

Identification and Functional Studies of Regulatory Variants Responsible for the Association of *NRG3* with a Delusion Phenotype in Schizophrenia

Mariela Zeledón^{a,c} Nicole Eckart^{a,b} Margaret Taub^d Hilary Vernon^a
Megan Szymanski^{a,b} Ruihua Wang^a Pei-Lung Chen^f Gerry Nestadt^c
John A. McGrath^{c,e} Akira Sawa^c Ann E. Pulver^{c,e} Dimitrios Avramopoulos^{a,c}
David Valle^a

^aMcKusick-Nathans Institute of Genetic Medicine, ^bPredoctoral Training Program in Human Genetics, and ^cDepartment of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, and ^dBiostatistics and ^eEpidemiology-Genetics Program, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Md., USA; ^fDepartment of Medical Genetics, National Taiwan University Hospital, Taipei City, Taiwan, ROC

Key Words

NRG3 · Neuregulin 3 · Delusion · Ashkenazi · Functional studies · Psychiatric disease

Abstract

We previously reported a genetic linkage for schizophrenia (SZ; nonparametric linkage score of 4.27) at 10q22 in an Ashkenazi Jewish (AJ) population. In follow-up fine-mapping, we found strong evidence for an association between 3 intronic single nucleotide variants (SNVs) in the 5' end of neuregulin 3 (*NRG3*) and the delusion factor score of our phenotypic principal component analysis. Two independent groups replicated these findings, indicating that variants in *NRG3* confer a risk for a delusion-rich SZ subtype. To identify the causative variants, we sequenced the 162-kb linkage disequilibrium block covering the *NRG3* 5' end in 47 AJ SZ patients at the extremes of the delusion factor quantitative trait distribution. Among the identified variants, we found 5 non-

coding SNVs which were present on the high delusion factor haplotype and significantly overrepresented in high delusion factor subjects. We tested these for regulatory effects and found that risk alleles of rs10883866 and rs60827755 decreased and increased, respectively, the expression of a reporter gene as compared to the reference allele. In postmortem brain mRNA quantification experiments, we found the same variants also perturb the relative expression of alternative *NRG3* isoforms. In summary, we have identified regulatory SNVs contributing to the association of *NRG3* with delusion symptoms in SZ.

© 2015 S. Karger AG, Basel

Introduction

Schizophrenia (SZ) is a chronic, disabling neuropsychiatric disease [1–3] with typical onset in young adulthood. The evidence for a genetic contribution to SZ is

strong [4–7], and while, initially, genetic linkage and association studies had produced few replicated positive results (in part because of phenotypic and genetic heterogeneity [2, 8–17]), genome-wide association studies have recently been more successful [18]. To minimize the challenges of phenotypic and genetic heterogeneity, we utilized rigorous phenotyping in a genetically homogeneous Ashkenazi Jewish (AJ) population [19, 20]. Previously, in this sample, we had identified a significant SZ linkage signal at 10q22 from 29 AJ pedigrees (nonparametric linkage score of 4.27) [19]. Our result was independently replicated in a study of 1,345 Han Chinese individuals [21]. During follow-up, we performed an association study in 1,199 SZ individuals across the linkage peak with 9 factors (heritability ranging from 0.27 to 0.60) derived from a principal component analysis of 73 phenotypic features [22]. We identified a strong association ($p = 1.6 \times 10^{-6}$ to 2.3×10^{-7}) of several *NRG3* single nucleotide variants (SNVs; rs6584400, rs10883866 and rs10748842) with the factor designated ‘delusion’ [23], corresponding to the constellation of symptoms known as ‘Schneiderian first-rank symptoms’ [24]. The association of these same *NRG3* SNVs with delusions was independently replicated by others in both Western European [25] and Anglo-Irish cohorts [26]. In the former study, the authors suggested a functional role for rs10748842, possibly by regulating the relative abundance of *NRG3* isoforms [25].

NRG3 is a strong biological candidate gene for SZ. It is one of 3 paralogs of *NRG1*, another replicated SZ susceptibility gene [27]. Like *NRG1*, *NRG3* undergoes extensive alternative splicing and alternate utilization of first exons (fig. 1). Expression of *NRG3* is limited to the brain, breast and testis and is higher early in development [25]. The protein product of *NRG3* is a single-pass type I membrane protein with an extracellular N-terminal EGF-like domain that is cleaved and binds ErbB4 receptors [28–30] encoded by *ErbB4*, another gene implicated in SZ [31–33]. *NRG3* appears to function as a chemorepellant signal in the patterning and migration of GABAergic interneurons in the developing cerebral cortex [34–36].

Here, we used in vitro and in vivo approaches to identify the functional variants underlying the association with the delusion factor in the 162-kb linkage disequilibrium (LD) block containing the associated SNVs. We found 3 variants that affect the regulation of *NRG3* expression and are strongly associated with the delusion factor phenotype.

Subjects, Materials and Methods

Subject Ascertainment and Factor Analysis

All subjects are participants in the JHU Epidemiology-Genetics Research Program in Psychiatry. The ascertainment and assessment of these patients and controls [19, 23], as well as factor analysis derivation, have been previously described [22].

Next-Generation Sequencing, Analysis Parameters and Software

We sequenced a total of 47 AJ SZ individuals (33 males and 14 females) from the extreme ends of the delusion factor distribution. This 2:1 sex ratio is as expected, based on our entire AJ SZ collection [23]. Additionally, we sequenced an unaffected female, identified as homozygous for the microdeletion at Hg19 chr10:83611664–83613507, to determine the background haplotype of the microdeletion.

To generate sequencing libraries, we amplified 19 overlapping ~10-kb amplicons to cover the entire 162-kb LD block using LA Taq (Kyoto, Japan). LD information was acquired from our data on Ashkenazi Jews [23], and the LD structure is very similar between AJ and European populations [37]. After amplification, we fragmented the amplicons to ~250 bp by sonication with a Covaris E210 device, indexed fragments with unique identifiers for each individual, and multiplexed samples for library preparation in groups of 12, each including 6 low delusion factor and 6 high delusion factor individuals. We sequenced each of the 4 libraries with 75-bp paired-end reads on either the Illumina Genome Analyzer or the Hi-Seq 2000 system. Sequencing primers and PCR conditions are available upon request.

We aligned the reads using the BWA aligner (version 0.5.9) [38] and postprocessed aligned reads with the Genome Analysis Toolkit (GATK version 1.0.5336) [39] to remove duplicate reads, to realign around small insertions or deletions (indels) and to recalibrate base-call quality scores. We excluded 1 high delusion factor subject from the analysis due to low coverage depth (55) relative to the other samples (average coverage of 850). The median fraction of reads that aligned was 90%. We called SNVs and indels using SAMtools mpileup [40] (version 0.1.17) on all samples simultaneously with parameters set to downweight reads with excessive mismatches and to allow for high depth coverage. We removed variants if their overall call quality score was <99 and set individual calls to missing if their call quality was <40. We assessed the quality and depth of coverage differences for each variant, comparing these quantities between low and high delusion factor individuals using a t test and removing variants with $p < 0.1$.

We validated a total of 790 known variants across the region in all subjects, using previously obtained SNV genotypes at 19 positions [23]. Of these, 773 were correctly called from our unfiltered SNV detection, for an accuracy of 97.8% (95.1% sensitivity, 99.5% specificity, using a variant presence/absence criterion). After setting all calls with an individual call quality <40 or a variant call quality <99 to missing, 6 incorrect calls remained. This accuracy of 99.2% reflects 100% specificity and 98.1% sensitivity after filtering.

We excluded 4 variants (GRCh37/hg19 chr10:83667530, 83616428, 83659377 and 83702217) from further analysis due to significant (two-sample t test $p < 0.1$) differences in call quality or coverage between high and low delusion factor groups. After filtering, we identified 620 SNVs. The number of positions with non-reference (non-REF) alleles in an individual ranged from 11 to 235 with a median of 98.

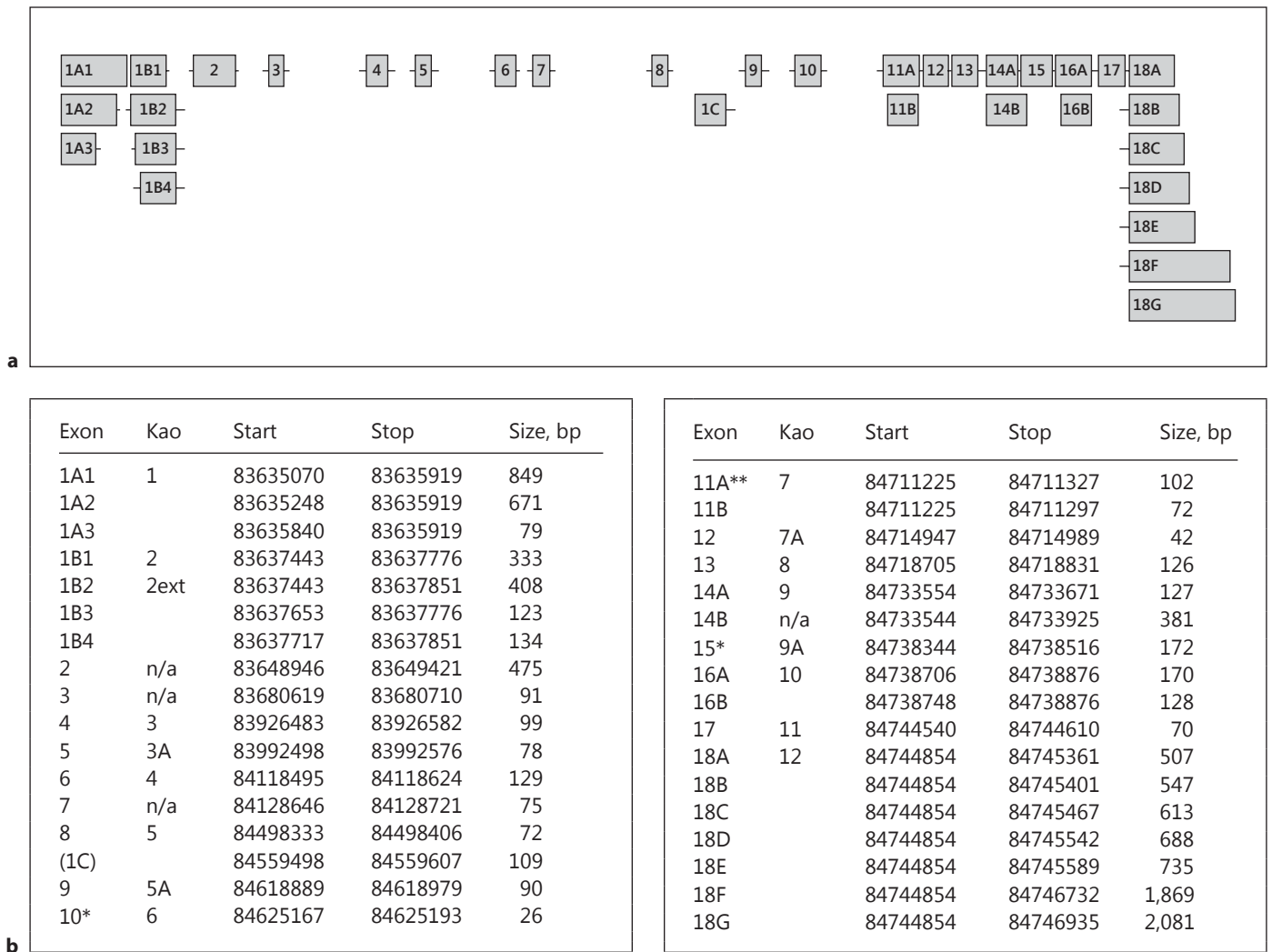


Fig. 1. Structure of *NRG3*. **a** Diagram of *NRG3* exons as well as a proposed numerical nomenclature for validated exons in *NRG3* (gray) located at 10q 23.1. Exons and introns are not to scale. **b** Table of the genomic coordinates on 10q (Start, Stop) and the size of each of the *NRG3* exons. The column labeled 'Kao' corre-

lates these exons with those described previously by Kao et al. [25]. * A shorter (3-bp) version of this exon has been observed [Vernon, unpubl. data]. ** A shorter (5-bp) version of this exon has been reported (EMBL:ABC69293.1.).

Microdeletion Studies

During the amplification step of DNA library preparation for next-generation sequencing, we identified a 1.8-kb microdeletion 23 kb 5' of the transcription start site, at Hg19 chr10:83611664–83613507. We genotyped the microdeletion by sizing the PCR product in 724 AJ SZ cases as described in Chen et al. [23]. For some individuals, we were not able to make a definitive genotype call from the PCR products. We genotyped these individuals, plus an additional 32 individuals from 10 unrelated pedigrees with SZ probands, using a 3-primer PCR assay where a deleted allele would yield a 249-bp band and a nondeleted allele would yield 406- and 2,092-bp bands. The primer sequences and PCR conditions are available upon request. We calculated the association of this

microdeletion with the delusion factor quantitative trait (QT) locus using the QT likelihood ratio χ^2 test implemented in the UNPHASED software [41].

Reporter Constructs

We generated dual-luciferase reporter assay (DLA) reporter constructs by subcloning PCR amplicons of selected genomic segments from our subjects. The amplicons were cloned using the pCR8/GW/TOPO TA Cloning Kit (Invitrogen, Carlsbad, Calif., USA), sequenced for confirmation and shuttled to pDSMA, a Gateway-modified pGL3-promoter plasmid [42] using LR Clonase Enzyme Mix (Invitrogen). They were then shuttled in genomic orientation (5' to 3') upstream of a constitutive promoter in the

pDSMA plasmid to determine the consequences on expression of the firefly luciferase reporter gene. We performed one plasmid preparation per construct by QIAfilter Plasmid Purification Midi Kit (Qiagen, Hilden, Germany) and sequence verified all constructs. We caution the reader that since all replicates come from a single plasmid preparation, it is possible for differences in preparation quality to influence the results of the DLA.

The REF amplicon was 3.8 kb and included the region of the microdeletion plus 1-kb flanks on both sides; the deletion (DEL) construct was 2 kb and included only the 1-kb regions flanking the microdeletion. To determine if the microdeletion had some effect based on spatial characteristics of the region, we also restored normal spacing by inserting a 1.8-kb segment of a zebrafish construct (ZNAS-17.9) with no enhancer activity in DLAs [43] at the site of the microdeletion. To produce constructs containing the SNVs, we amplified approximately 320 bp of the region of highest conservation surrounding the variant. Because rs10883862 and rs7919976 are only 68 bp apart, we identified 3 haplotypes and made single constructs for each. The primers and PCR conditions used to generate all amplicons are available upon request.

Dual-Luciferase Reporter Assays

We cultured HT22 mouse hippocampal [44], HEK 293 human embryonic kidney and Neuro2A mouse neuroblastoma cells in DMEM with GlutaMAX and 10% FBS. Primary cortical neuron cells from E18 C57BL/6 mice or Sprague-Dawley rats were cultured as described previously [45]. Due to the quantity of neuron cells required for the SNV series of DLAs, we chose to use rat primary neurons to minimize the number of animals sacrificed. It should be noted that based on data from Gene Expression Omnibus (GEO) [46], *NRG3* is not expressed in the commercially available cell lines (GEO accessions GSE45534, GDS4802 and GSE56275), but it is highly expressed in both mouse and rat primary cortical neurons (GEO accessions GDS4759 and GDS3832).

We cotransfected cells with the pDSMA and Renilla vectors at a 20:1 ratio using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions and lysed the cells 24 h after transfection (Promega, Madison, Wis., USA). We used 20 μ l of lysate and 50 μ l of each reagent for the assays. Firefly/Renilla luciferase expression ratios are normalized to the pDSMA promoter-only construct to facilitate comparisons between cell lines.

Human Brain mRNA Studies

We extracted DNA and total RNA from subjects without macroscopic pathology from 190 superior temporal gyrus (STG) samples from the Harvard Brain Tissue Resource Center (HBTRC) and 94 dorsolateral prefrontal cortex (DLPFC) samples from the National Institute of Child and Human Development (NICHD). The HBTRC samples were 77% male, averaged 62 years of age and had an average postmortem interval of 23.2 h. Ancestry data for the HBTRC samples were not available, but based on an ancestry SNP panel and comparison with Caucasian samples from the Johns Hopkins Brain Resource Center, there were only 2 genetic outliers in the HBTRC samples [47], which were removed from analysis. The NICHD samples were all reported Caucasian, were 66% male, averaged 36 years of age and had an average postmortem interval of 17.7 h.

We genotyped all samples for rs10748842 and rs60827755 by TaqMan genotyping assays (Invitrogen). The expression of *NRG3* transcripts initiating with exons 1a and 1b was measured sepa-

rately, using exon-specific primers for qPCR of cDNA reverse transcribed from brain RNA, and was quantified by comparison to a standard curve with normalization to the expression of two housekeeping genes, *MRIP* and *ACTB*, as previously described [47]. The data were log transformed to achieve a normal distribution, and outliers were removed.

Using a generalized linear model, we interrogated the expression data for correlations with genotypes at selected SNVs using age, sex, qPCR plate and postmortem interval as covariates. To examine changes in the relative abundance of classes of transcripts, we adjusted the expression of transcripts initiating with exon 1b for that of transcripts initiating with exon 1a.

Results and Discussion

Previously, we had examined 685 AJ SZ probands for an association between each of 9 QTs (phenotypic principal component factor scores) and selected genotypes around *NRG3*. We found a significant association of the delusion factor with the genotypes of SNVs in a 162-kb LD block covering the 5' end of *NRG3* [22]. Here, we sequenced SZ subjects stratified according to their scores on the delusion factor QT with the goal of finding the variant(s) responsible for the association signal. High and low delusion factor subjects differed primarily in the number of types of delusions that they exhibited, with high delusion factor subjects experiencing more kinds of delusions [22].

Sequencing of the NRG3 5' 162-kb LD Block

To identify all variants across the LD block, we sequenced subjects in 4 pooled sets of 12 samples (6 from the high and 6 from the low extremes of the delusion factor QT distribution). The minimum average depth for samples included in the analysis was 334, and 98.6% of the positions in our region of interest were covered by at least 15 reads [48]. There was no difference between the high and low delusion factor groups in average coverage (t test $p = 0.49$) or percentage of bases covered by at least 15 reads (t test $p = 0.20$), indicating that between-sample differences in these variables did not bias our subsequent analyses.

A Microdeletion Upstream of NRG3 Is Associated with the Delusion Factor but Lacks Functional Consequences

PCR and sequencing revealed a 1.8-kb microdeletion located 23 kb 5' of the predicted transcription start site of *NRG3* at Hg19 chr10:83611664–83613507. Figure 2 shows the amplification scheme, the location of the deletion and the corresponding evidence from PCR product

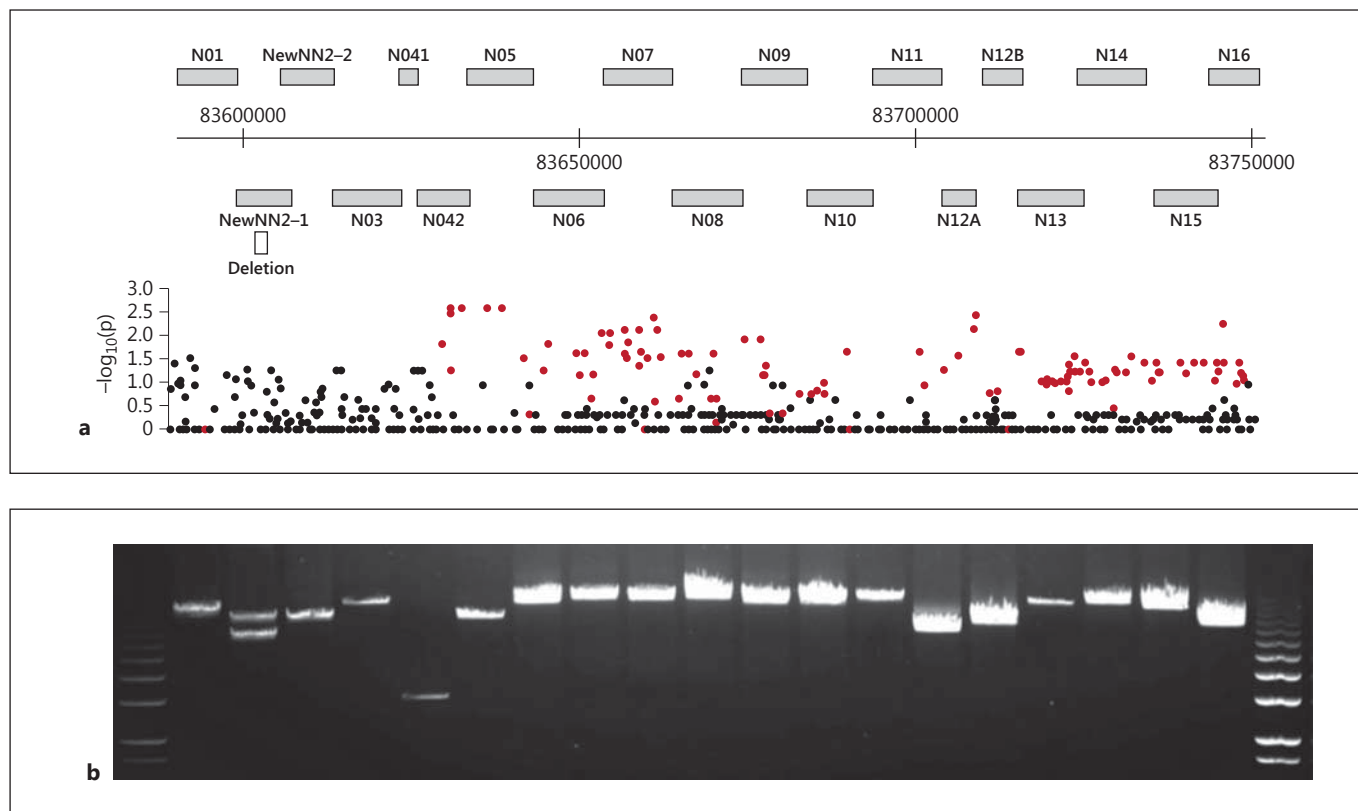


Fig. 2. a Diagram of *NRG3* amplicons used for sequencing (top) and negative $\log_{10} p$ values for differences in allele count between high and low delusion factor subjects (bottom). Variants with more alternative alleles in low delusion factor patients are in black and those with more alternative alleles in high delusion factor pa-

tients are in red. **b** Gel (0.75% agarose) electrophoresis of the *NRG3* amplicons shown in **a** flanked by a 1-kb ladder size marker from a subject heterozygous for a microdeletion (hg19 chr10:83611664–83613507).

electrophoresis. This microdeletion has not previously been reported by dbSNP or any tracks on the UCSC Genome Browser and is in a region with low sequence conservation [49, 50]. There are flanking clusters of transcription factor binding sites ~6 kb 5' and ~12 kb 3' of the microdeletion (hg18 HMR Conserved Transcription Factor Binding Sites track created by Weirauch and Raney) [49].

To determine if the microdeletion itself was a causative variant, we first measured its frequency in various populations. We found similar frequencies in our sample of AJ SZ (2.65%, $n = 641$), AJ bipolar (2.57%, $n = 564$) and AJ control (2.8%, $n = 642$) cohorts. However, when we tested SZ patients at both extremes of the delusion factor QT, we found a significant difference in frequencies (8/138 alleles or 5.8% in the high compared to 0/134 alleles or 0% in the low delusion factor patients; Fisher's exact test $p = 0.013$). To investigate this association with delusion factor scores

in a larger SZ cohort, we genotyped the microdeletion in 724 AJ SZ subjects across the range of scores and also found a strong association (likelihood ratio $\chi^2 p = 0.006895$). We did not detect the microdeletion in 346 outbred American European SZ subjects, suggesting that it is specific to Ashkenazim. We also genotyped the microdeletion in 10 unrelated families with heterozygous probands and found that it was inherited and had identical breakpoints in all. From these results, we reasoned that the microdeletion was either a causative variant or marked a haplotype containing (a) causative variant(s).

To investigate potential regulatory effects of the microdeletion, we performed DLAs in HEK 293 cells, HT22 cells and cultured mouse cortical neurons. We found no difference in the relative expression of the reporter gene between the REF and DEL constructs (fig. 3). We also corrected for possible spatial effects by introducing an inactive stuffer sequence into our deletion construct, which

Table 1. Variant identification and distribution amongst low and high delusion factor patients

rs ID	Position (Hg19)	REF	Alt.	Alt. allele frequency	Low alleles	High alleles	p value	On DEL haplotype
rs10883862 ^d	83642101	T	G	0.11	0/46	9/46	0.003	TRUE
rs10883866 ^b	83643639	C	G	0.11	0/46	9/46	0.003	TRUE
rs60827755 ^b	83647518	A	G	0.10	0/46	9/46	0.003	TRUE
rs10748842 ^b	83649739	T	C	0.11	0/46	9/46	0.003	TRUE
rs7919976 ^a	83642034	G	A	0.58	5/46	18/46	0.003	TRUE
rs74885547 ^a	83672780	T	C	0.56	8/44	22/46	0.004	TRUE
rs10786760 ^a	83664575	C	T	0.55	9/44	23/46	0.004	TRUE
rs1018484 ^a	83665670	T	C	0.56	8/44	20/42	0.005	TRUE
rs10884019 ^c	83757021	G	A	0.15	3/46	13/44	0.005	TRUE
rs11191757 ^a	83667910	G	T	0.56	9/46	22/46	0.008	TRUE
rs10786762 ^a	83670102	G	A	0.58	9/46	22/46	0.008	TRUE
rs10883921 ^a	83685802	A	C	0.57	8/46	20/46	0.012	TRUE
rs7075126 ^a	83688159	A	G	0.57	8/46	20/46	0.012	TRUE
rs10883898 ^a	83670126	C	T	0.53	8/44	20/46	0.012	TRUE
rs9943299 ^b	83640808	A	G	0.22	1/46	9/46	0.015	TRUE
rs6584400 ^b	83656526	G	A	0.22	1/46	9/46	0.015	TRUE
rs3918483 ^a	83668440	C	A	0.53	8/44	18/42	0.018	TRUE

Alt. = Alternative. ^{a-d} LD group in which member SNVs have an $r^2 > 0.8$ based on 1000 Genomes Project data (www.broadinstitute.org/mpg/snap/).

also failed to show an effect on reporter gene expression. The experimental constructs show reduced expression compared to the promoter-only vector, normalized to 1 on the y-axis (fig. 3), which is likely due to the presence of an additional sequence in the vector causing reduced transfection or translation efficiency. These in vitro experiments indicate that the microdeletion lacks regulatory consequences in the tested cell types; although one explanation might be that these cell types do not have the appropriate transcription machinery, it is also possible that the microdeletion itself is not causative but rather marks a haplotype containing (an)other causative variant(s).

Identification of Variation in cis with the Microdeletion

During the course of these studies, we identified 2 homozygotes for the microdeletion, a bipolar subject and a nonaffected relative of an SZ subject. The latter, a 66-year-old asymptomatic woman, has a brother and a nephew with SZ. Sequencing of the microdeletion homozygote revealed homozygosity for all SNVs across the entire LD block, including the risk alleles of previously identified delusion factor-associated SNVs [23]. She has 121 homozygous non-REF variants and no heterozygous variants, suggesting that the microdeletion haplotype extends across the entire sequenced region. Genotyping 5 other

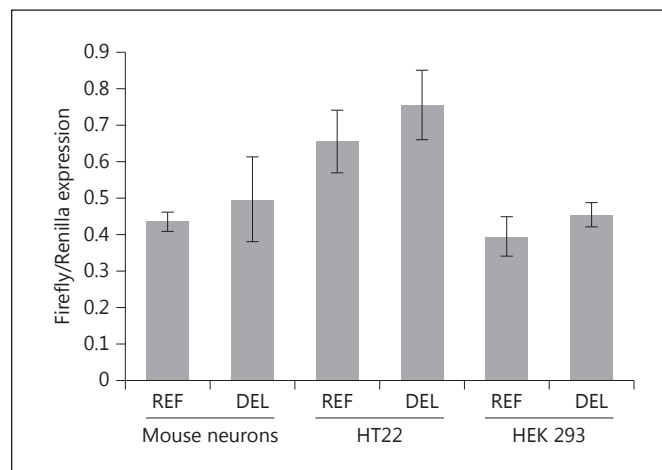


Fig. 3. Results from DLAs of the microdeletion region. Each construct was tested in 6 biological replicates per experiment and was normalized to the promoter-only vector. No statistically significant differences are observed between the DEL and the REF allele.

members in her family and 31 unrelated affected AJ microdeletion heterozygotes allowed us to define 3 distinct haplotypes across the LD block: 1 that is marked by the microdeletion and likely contains the risk variants and 2 others without the microdeletion that likely do not carry the risk variants. Sequencing the LD block in 24 individu-

als with high delusion factor QT scores, 23 with low delusion factor QT scores and the asymptomatic woman identified a total of 620 SNVs in the LD block. Of the 57 SNVs associated with a high delusion factor score ($p < 0.05$) in the present study, 54 were non-REF alleles homozygous in the asymptomatic woman, and we concluded these made up the microdeletion haplotype.

To assess differences between low and high delusion factor subjects at each of these 54 SNVs, we ranked them according to the p value as calculated by Fisher's exact test (table 1). While no results withstood correction for multiple testing, the 5 nominally most significant ($p = 0.003$ for all 5, minor allele frequency 0.104–0.424) were located in a ~8-kb cluster flanking and including *NRG3* exon 2. Of the 3 previously identified SNVs most associated with the delusion factor, 2 (rs10883866 and rs10748842) were among the 5 showing strongest statistical significance in this study [23]. The third previously identified SNV, rs6584400, had a p value of 0.015 and was also included in our functional studies of a total of 6 variants. The risk alleles of all of these variants ($p < 0.05$) are on the microdeletion haplotype.

Variants in Intronic Regions of NRG3 Affect NRG3 Expression in vitro and ex vivo

To determine if the variant genotype affected the regulatory function of the surrounding sequence, we performed DLAs testing ~320 bp surrounding either allele of our 6 candidate SNVs in HEK 293, HT22 and Neuro2A cell lines as well as in cultured rat primary cortical neurons.

We tested rs7919976 and rs10883862, which are only 68 bp apart ($r^2 = 0.329$), in a single construct and detected no discernible regulatory activity (fig. 4). Variant rs6584400 had a minimal and inconsistent effect regardless of the genotype. For these reasons, we did not perform additional studies on these variants. We observed modest effects for rs10748842, with the risk allele (C) increasing relative expression ~1.2-fold. For this reason, and because this variant was implicated by previous authors based on *in silico* computational analysis, we included it in additional studies. In contrast, we found that rs10883866, located 5 kb 5' of exon 2, had a strong and consistent effect in DLA experiments. The construct containing the REF variant (C) conferred a 2- to 3-fold increase in expression over the promoter-only vector in all cell lines tested (in primary rat cortical neurons, the effect was more modest but in the same direction), and this effect was completely abrogated by the risk variant (G). Similarly, rs60827755, located 1.5 kb 5' of exon 2, also

showed a consistent effect on expression in all the cell types tested. In this case, however, the risk allele (G) is associated with an increased expression. The risk allele (G) of rs10883862 and the REF allele (A) of rs60827755 have consistently a lower expression of the reporter gene compared to the promoter-only vector, which may indicate possible repressor activity or may be a result of reduced transfection efficiency due to the additional inactive sequence in the vector. We conclude that rs60827755 and rs10883866 are both functional, each perturbing the regulatory function of a DNA element in this genomic segment but with opposite effects on expression in these *in vitro* assays. Together, they generate 2 haplotypes with different regulatory properties that likely represent the causative genetic variation leading to changes in *NRG3* expression and in the delusion factor scores. While the minor alleles of rs10883866 and rs60827755 are both on the risk haplotype and drive reporter gene expression in opposite directions, *in vivo*, the effects of these variants will depend on the relative abundance of the *trans*-acting factors that bind these regions at any given time as well as the genomic context beyond the 320-bp inserts in the test plasmids.

SNVs in NRG3 Correlated with Alternate First Exon Use

NRG3 undergoes complex alternative splicing with transcripts that originate from at least 2 alternative first exons, exons 1a and 1b. We hypothesized that the SNVs associated with the delusion factor, which are all located in the 5' end of *NRG3*, might influence the abundance of *NRG3* transcripts that begin with exon 1a or 1b. To test this possibility, we asked if there was a correlation between the genotype and the expression of *NRG3* transcripts (see Subjects, Materials and Methods and fig. 1) [25] in 2 brain regions. We genotyped one of the functional variants from the DLAs, rs60827755, in addition to rs10748842, which has an r^2 of 1.0 with the other functional DLA variant, rs10883866 [51].

After controlling for qPCR plate, age, sex and post-mortem interval, we found that carriers of the risk allele (C) at rs10748842 have a decreased expression of transcripts containing exon 1b in the DLPFC ($p = 1.9 \times 10^{-7}$). In the STG, we observed a similar but somewhat weaker effect ($p = 0.026$) of the risk allele on exon 1b transcripts. We saw no significant correlation between genotype and the expression of transcripts containing exon 1a in either brain region. We also wanted to determine if the relative expression of transcripts containing exon 1b, compared to transcripts containing exon 1a, was correlated with

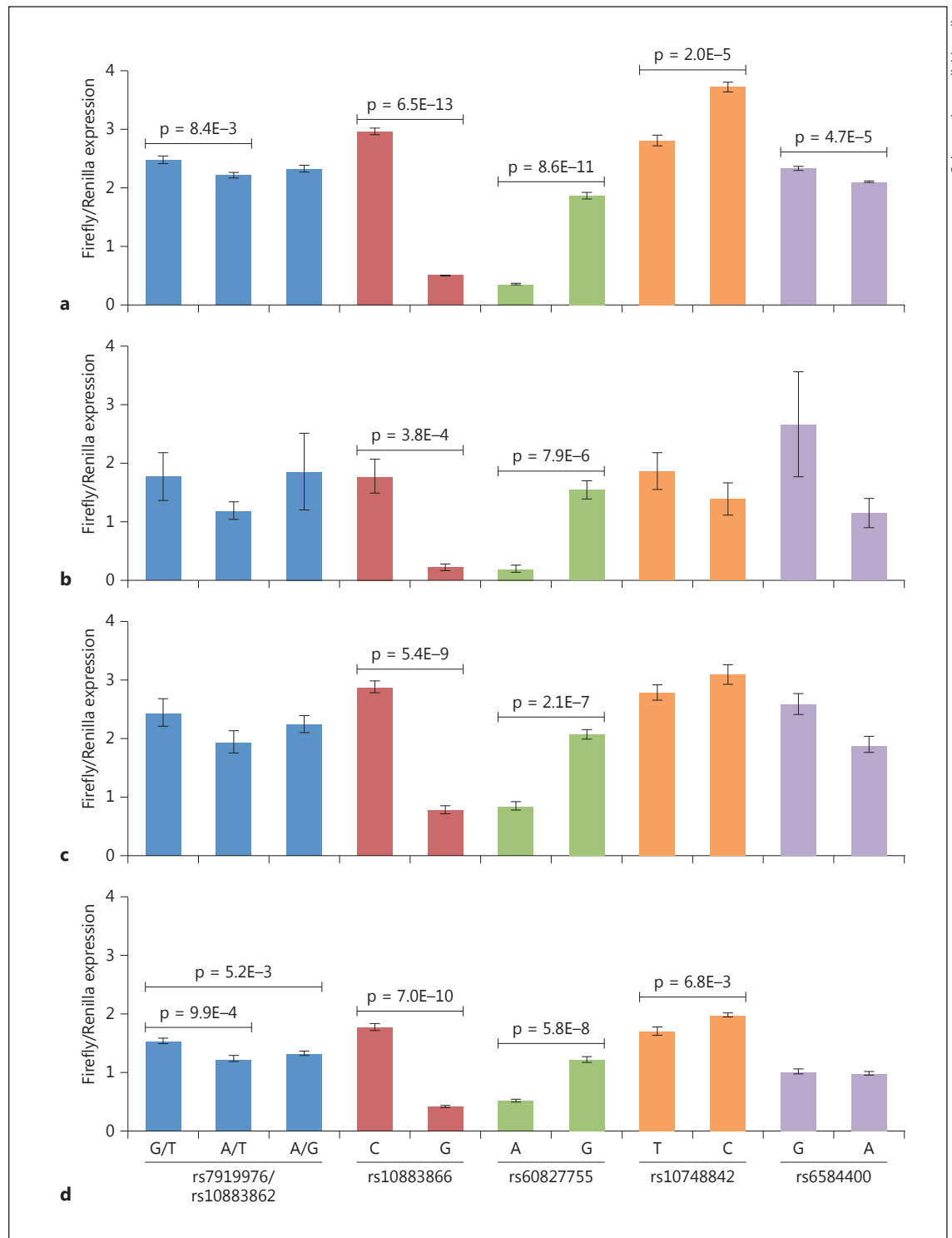


Fig. 4. DLAs of short ~320-bp sequence segments surrounding the indicated SNPs. The constructs were expressed in HEK 293 cells (**a**), cultured rat primary cortical neurons (**b**), HT22 cells (**c**) and Neuro2A cells (**d**). Statistically significant results from two-tailed t tests are shown. Each SNP (or pair of SNPs, in the case of

rs7919976 and rs10883862) is represented by a different color; the left bar in each set represents the REF allele and the right the risk allele. Each construct was tested in 6 biological replicates per experiment and normalized to the promoter-only vector.

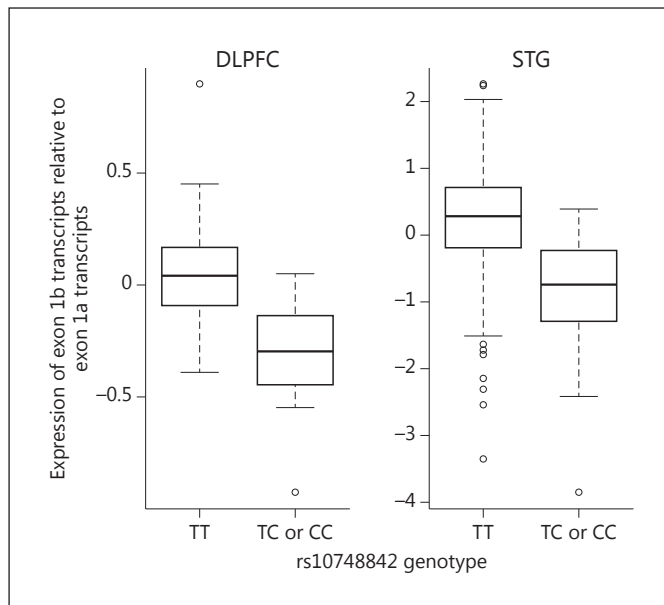


Fig. 5. Effect of rs10748842 on *NRG3* alternative transcripts. Box plots of levels of exon 1b transcripts relative to exon 1a transcripts for the different genotype groups and in the two brain regions examined. The homozygotes for the rare genotype CC were very few (2 and 1 for the STG and DLPFC, respectively) and grouped with the heterozygotes.

genotype. In the DLPFC, the risk allele is highly correlated with a decrease in the relative expression of exon 1b transcripts when adjusted for the expression of transcripts with exon 1a ($p = 4.7 \times 10^{-8}$). In the STG, the relative expression of exon 1b transcripts became highly significant when adjusting for exon 1a transcripts ($p = 5.11 \times 10^{-9}$; fig. 5).

As expected given the strong LD ($r^2 = 0.92$), the risk allele (G) of rs60827755 also showed a decreased expression of exon 1b transcripts ($p = 1.42 \times 10^{-7}$) and a decreased relative expression when adjusted for exon 1a transcripts ($p = 6.0 \times 10^{-8}$) in the DLPFC and in the STG ($p = 0.048$ and 1.9×10^{-7} , respectively).

From these data, we conclude that the effect of the delusion factor-associated variants involves the relative expression of various *NRG3* transcripts. This suggests a transcript-specific mechanism in which the risk variants affect the utilization of alternative *NRG3* first exons.

Current psychiatric diagnostic criteria, including the Diagnostic and Statistical Manual (DSM), focus on clinical reliability but not on etiologic, particularly genetic, validity, hampering a biological understanding of psychiatric illness [52]. To rectify this, the National Institute of

Mental Health (NIMH) launched the Research Domain Criteria (RDoC) project to create a new framework for research on the pathophysiology of mental illness [52]. RDoC describes a matrix in which the rows represent various constructs grouped hierarchically into domains of neuropsychiatric function, while the columns denote ascending levels of etiological analysis: the genetic, molecular and cellular levels; the systems level, and the levels of the individual, the family environment and social context. Our work indicates that genetic variations of a specific gene, *NRG3*, contribute to a functional domain (delusion).

In summary, we identified all variants in the 162-kb LD block previously shown to confer a risk for a constellation of phenotypic features characterized as ‘delusion factor’ in an AJ SZ cohort. We then identified a risk haplotype extending across the entire LD block and showed that alleles of 3 intronic SNVs strongly associating with the delusion factor QT also modulate gene expression and alternative first exon utilization. The molecular mechanism by which the single base pair changes exert their effect is unknown, although we surmise that the variants alter binding sites for transcription factors that interact indirectly with the promoter(s). The identity of the proteins that utilize these putative binding sites to control the regulation of particular transcripts of *NRG3*, as well as the functional consequences of various splice forms in specific cells types, will shed light on the molecular pathways involved in delusions and psychiatric disease as well as deepen our understanding of the regulation of this gene.

Acknowledgements

Human tissue was obtained from the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, Baltimore, Md., USA. The Bank is funded by NIH Contract No. #HHSN275200900011C, Ref. No. NO1-HD-9-0011. Tissue was also provided by the HBTRC. The authors thank all participants in the research program of the JHU Epidemiology-Genetics Program, Y. Taniguchi for preparation of primary neurons, Y. Lema for assistance preparing the manuscript, and A. McCallion and D. Gorkin for advice and assistance regarding DLA experiments as well the use of gateway-modified reporter vectors. This material is based upon work supported by a National Alliance for Research in Schizophrenia and Affective Disorders (NARSAD) award (D.A.), a National Science Foundation Graduate Research Fellowship (grant No. DGE-0707427 to M.Z.), the Predoctoral Training Program in Human Genetics (5T32GM07814 to M.S., M.Z. and N.E.), the Johns Hopkins Brain Science Institute (D.V.), a grant for ‘DPYSL2 (CRMP2) variants and risk for schizophrenia’ (5P50MH094268 to D.V.), NIMH R01s (MH092515 and

MH085018 to D.A.) and the NIH (MH-084018, MH-094268 Silvio O. Conte center, MH-069853, MH-085226, MH-088753, MH-092443, Stanley, RUSK, S-R foundations, NARSAD and the Maryland Stem Cell Research Fund to A.S.).

Disclosure Statement

The authors declare no conflict of interest.

References

- 1 Sawa A, Snyder SH: Schizophrenia: diverse approaches to a complex disease. *Science* 2002;296:692–695.
- 2 Insel TR: Rethinking schizophrenia. *Nature* 2010;468:187–193.
- 3 Eaton WW: Epidemiology of schizophrenia. *Epidemiol Rev* 1985;7:105–126.
- 4 Bilder RM, Mukherjee S, Rieder RO, Pandurangi AK: Symptomatic and neuropsychological components of defect states. *Schizophr Bull* 1985;11:409–419.
- 5 Gottesman II: Schizophrenia Genesis: The Origin of Madness. New York, Freeman, 1991.
- 6 Shih RA, Belmonte PL, Zandi PP: A review of the evidence from family, twin and adoption studies for a genetic contribution to adult psychiatric disorders. *Int Rev Psychiatry* 2004;16:260–283.
- 7 Tsuang M: Schizophrenia: genes and environment. *Biol Psychiatry* 2000;47:210–220.
- 8 Lencz T, Morgan TV, Athanasiou M, Dain B, Reder CR, Kane JM, Kucherlapati R, Malhotra AK: Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol Psychiatry* 2007;12:572–580.
- 9 O'Donovan MC, Craddock N, Norton N, et al: Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 2008;40:1053–1055.
- 10 Shifman S, Johannesson M, Bronstein M, Chen SX, Collier DA, Craddock NJ, Kendler KS, Li T, O'Donovan M, O'Neill FA, Owen MJ, Walsh D, Weinberger DR, Sun C, Flint J, Darvasi A: Genome-wide association identifies a common variant in the reelin gene that increases the risk of schizophrenia only in women. *PLoS Genet* 2008;4:e28.
- 11 Stefansson H, Ophoff RA, Steinberg S, et al: Common variants conferring risk of schizophrenia. *Nature* 2009;460:744–747.
- 12 Athanasiu L, Mattingsdal M, Kähler AK, Brown A, Gustafsson O, Agartz I, Giegling I, Muglia P, Cichon S, Rietschel M, Pietiläinen OPH, Peltonen L, Bramon E, Collier D, Clair DS, Sigurdsson E, Petursson H, Rujescu D, Melle I, Steen VM, Djurovic S, Andreassen OA: Gene variants associated with schizophrenia in a Norwegian genome-wide study are replicated in a large European cohort. *J Psychiatr Res* 2010;44:748–753.
- 13 Need AC, Ge D, Weale ME, et al: A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 2009;5:e1000373.
- 14 Kirov G, ZI, Georgieva L, Moskvina V, Nikolov I, Cichon S, Hillmer A, Toncheva D, Owen MJ, O'Donovan MC: A genome-wide association study in 574 schizophrenia trios using DNA pooling. *Mol Psychiatry* 2009;14:796–803.
- 15 Purcell SM, Wray NR, Stone JL, et al: Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009;460:748–752.
- 16 Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, Dudbridge F, Holmans PA, Whittemore AS, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Crowe RR, Oksenberg JR, Mirel DB, Kendler KS, Freedman R, Gejman PV: Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature* 2009;460:753–757.
- 17 Sullivan PF, Lin D, Tzeng J-Y, van den Oord E, Perkins D, Stroup TS, Wagner M, Lee S, Wright FA, Zou F, Liu W, Downing AM, Lieberman J, Close SL: Genomewide association for schizophrenia in the CATIE study: results of stage 1. *Mol Psychiatry* 2008;13:570–584.
- 18 Schizophrenia Working Group of the Psychiatric Genomics Consortium: Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014;511:421–427.
- 19 Fallin M, Lasseter V, Wolyniec P, McGrath J, Nestadt G, Valle D, Liang K, Pulver A: Genomewide linkage scan for schizophrenia susceptibility loci among Ashkenazi Jewish families shows evidence of linkage on chromosome 10q22. *Am J Hum Genet* 2003;73:601–611.
- 20 Fallin MD, Lasseter VK, Wolyniec PS, McGrath JA, Nestadt G, Valle D, Liang K-Y, Pulver AE: Genomewide linkage scan for bipolar-disorder susceptibility loci among Ashkenazi Jewish families. *Am J Hum Genet* 2004;75:204–219.
- 21 Wang Y-C, Chen J-Y, Chen M-L, Chen C-H, Lai I-C, Chen T-T, Hong C-J, Tsai S-J, Liou Y-J: Neuregulin 3 genetic variations and susceptibility to schizophrenia in a Chinese population. *Biol Psychiatry* 2008;64:1093–1096.
- 22 McGrath JA, Avramopoulos D, Lasseter VK, Wolyniec PS, Fallin MD, Liang K-Y, Nestadt G, Thornquist MH, Luke JR, Chen P-L, Valle D, Pulver AE: Familiality of novel factorial dimensions of schizophrenia. *Arch Gen Psychiatry* 2009;66:591–600.
- 23 Chen P-L, Avramopoulos D, Lasseter VK, McGrath JA, Fallin MD, Liang K-Y, Nestadt G, Feng N, Steel G, Cutting AS, Wolyniec P, Pulver AE, Valle D: Fine mapping on chromosome 10q22–q23 implicates *Neuregulin 3* in schizophrenia. *Am J Hum Genet* 2009;84:21–34.
- 24 Wing J, Nixon J: Discriminating symptoms in schizophrenia. A report from the international pilot study of schizophrenia. *Arch Gen Psychiatry* 1975;32:853–859.
- 25 Kao W-T, Wang Y, Kleinman JE, Lipska BK, Hyde TM, Weinberger DR, Law AJ: Common genetic variation in Neuregulin 3 (*NRG3*) influences risk for schizophrenia and impacts *NRG3* expression in human brain. *Proc Natl Acad Sci* 2010;107:15619–15624.
- 26 Morar B, Dragović M, Waters FAV, Chandler D, Kalaydjieva L, Jablensky A: Neuregulin 3 (*NRG3*) as a susceptibility gene in a schizophrenia subtype with florid delusions and relatively spared cognition. *Mol Psychiatry* 2011;16:860–866.
- 27 Zhang D, Sliwkowski MX, Mark M, Frantz G, Akita R, Sun Y, Hillan K, Crowley C, Brush J, Godowski PJ: Neuregulin-3 (*NRG3*): a novel neural tissue-enriched protein that binds and activates ErbB4. *Proc Natl Acad Sci USA* 1997;94:9562.
- 28 Buonanno A, Fischbach GD: Neuregulin and ErbB receptor signaling pathways in the nervous system. *Curr Opin Neurobiol* 2001;11:287–296.
- 29 Buonanno A: The neuregulin signaling pathway and schizophrenia: from genes to synapses and neural circuits. *Brain Res Bull* 2010;83:122–131.
- 30 Mei L, Xiong W-C: Neuregulin 1 in neural development, synaptic plasticity and schizophrenia. *Nat Rev Neurosci* 2008;9:437–452.
- 31 Silberberg G, Darvasi A, Pinkas-Kramarski R, Navon R: The involvement of ErbB4 with schizophrenia: association and expression studies. *Am J Med Genet B Neuropsychiatr Genet* 2006;141B:142–148.
- 32 Hatzimanolis A, McGrath JA, Wang R, Li T, Wong PC, Nestadt G, Wolyniec PS, Valle D, Pulver AE, Avramopoulos D: Multiple variants aggregate in the neuregulin signaling pathway in a subset of schizophrenia patients. *Transl Psychiatry* 2013;3:e264.
- 33 Law AJ, Kleinman JE, Weinberger DR, Weickert CS: Disease-associated intronic variants in the *ErbB4* gene are related to altered *ErbB4* splice-variant expression in the brain in schizophrenia. *Hum Mol Genet* 2007;16:129–141.
- 34 Assimacopoulos S, Grove EA, Ragsdale CW: Identification of a Pax6-dependent epidermal growth factor family signaling source at the lateral edge of the embryonic cerebral cortex. *J Neurosci* 2003;23:6399–6403.

- 35 Anton ES, Ghashghaei HT, Weber JL, McCann C, Fischer TM, Cheung ID, Gassmann M, Messing A, Klein R, Schwab MH, Lloyd KCK, Lai C: Receptor tyrosine kinase ErbB4 modulates neuroblast migration and placement in the adult forebrain. *Nat Neurosci* 2004;7:1319–1328.
- 36 Li H, Chou S-J, Hamasaki T, Perez-Garcia CG, O'Leary DD: Neuregulin repellent signaling via ErbB4 restricts GABAergic interneurons to migratory paths from ganglionic eminence to cortical destinations. *Neural Dev* 2012;7:10.
- 37 Bray SM, Mulle JG, Dodd AF, Pulver AE, Wooding S, Warren ST: Signatures of founder effects, admixture, and selection in the Ashkenazi Jewish population. *Proc Natl Acad Sci USA* 2010;107:16222–16227.
- 38 Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754–1760.
- 39 McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA: The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–1303.
- 40 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–2079.
- 41 Dudbridge F: Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered* 2008;66:87–98.
- 42 Grice EA, Rochelle ES, Green ED, Chakravarti A, McCallion AS: Evaluation of the RET regulatory landscape reveals the biological relevance of a HSCR-implicated enhancer. *Hum Mol Genet* 2005;14:3837–3845.
- 43 McGaughey DM, Vinton RM, Huynh J, Al-Saif A, Beer MA, McCallion AS: Metrics of sequence constraint overlook regulatory sequences in an exhaustive analysis at *phox2b*. *Genome Res* 2008;18:252–260.
- 44 Adamczyk A, Kaźmierczak A, Czapski GA, Strosznajder JB: α -Synuclein induced cell death in mouse hippocampal (HT22) cells is mediated by nitric oxide-dependent activation of caspase-3. *FEBS Lett* 2010;584:3504–3508.
- 45 Ishizuka K, Kamiya A, Oh EC, Kanki H, Seshadri S, Robinson JF, Murdoch H, Dunlop AJ, Kubo K, Furukori K, Huang B, Zeledon M, Hayashi-Takagi A, Okano H, Nakajima K, Houslay MD, Katsanis N, Sawa A: DISC1-dependent switch from progenitor proliferation to migration in the developing cortex. *Nature* 2011;473:92–96.
- 46 Edgar R, Domrachev M, Lash AE: Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002;30:207–210.
- 47 Szymanski M, Wang R, Bassett SS, Avramopoulos D: Alzheimer's risk variants in the clusterin gene are associated with alternative splicing. *Transl Psychiatry* 2011;1:e18.
- 48 Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, Schork NJ, Murray SS, Topol EJ, Levy S, Frazer KA: Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol* 2009;10:R32.
- 49 Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D: The human genome browser at UCSC. *Genome Res* 2002;12:996–1006.
- 50 Sherry ST, Ward M-H, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;29:308–311.
- 51 Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PIW: SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008;24:2938–2939.
- 52 Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, Sanislow C, Wang P: Research Domain Criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry* 2010;167:748–751.