Molecular Cloning of a Brain-specific, Developmentally Regulated Neuregulin 1 (NRG1) Isoform and Identification of a Functional Promoter Variant Associated with Schizophrenia*5

Received for publication, April 6, 2007, and in revised form, June 1, 2007 Published, JBC Papers in Press, June 12, 2007, DOI 10.1074/jbc.M702953200

Wei Tan6, Yanhong Wang3, Bert Gold4, Jingshan Chen6, Michael Dean1, Paul J. Harrison1, Daniel R. Weinberger3, and Amanda J. Law1

From the 1Department of Psychiatry, Warneford Hospital, University of Oxford, Oxford, Oxford OX3 7JX, United Kingdom, 2SAIC-Frederick and 3Laboratory of Genomic Diversity, NCI, National Institutes of Health, Frederick, Maryland 21702, and the 4Clinical Brain Disorders Branch, Intramural Research Program, National Institute of Mental Health, Bethesda, Maryland 20892-1385

Neuregulin 1 (NRG1) is essential for the development and function of multiple organ systems, and its dysregulation has been linked to diseases such as cancer and schizophrenia. Recently, altered expression of a novel isoform (type IV) in the brain has been associated with schizophrenia-related genetic variants, especially rs6994992 (SNP8NRG243177). Here we have isolated and characterized full-length NRG1 type IV cDNAs from the adult and fetal human brain and identified novel splice variants of NRG1. Full-length type IV spans 1.8 kb and encodes a putative protein of 590 amino acids with a predicted molecular mass of ~66 kDa. The transcript consists of 11 exons with an Ig-like domain, an epidermal growth factor-like (EGF) domain, a β-stalk, a transmembrane domain, and a cytoplasmic “a-tail,” placing it in the B1a NRG1 subclass. NRG1 type IV was not detected in any tissues except brain and a putative type IV NRG1 protein of 66 kDa was similarly brain-specific. Type IV transcripts are more abundantly expressed in the fetal brain, where, in addition to the full-length structure, two novel type IV variants were identified. In vitro luciferase-reporter assays demonstrate that the 5′ promoter region upstream of type IV is functional, with differential activity associated with genetic variation at rs6994992, and that promoter competition may impact on type IV expression. Our data suggest that type IV is a unique brain-specific NRG1 that is differentially expressed and processed during early development, is translated, and its expression regulated by a schizophrenia risk-associated functional promoter or single nucleotide polymorphism (SNP).

Neuregulin 1 (NRG1)2 is a signaling protein that mediates cell-cell interactions and plays critical roles in the growth of the nervous system, heart, breast, and other organ systems. The gene is located on 8p12.21, and ~15 “classic” NRG1 isoforms are generated through alternative promoter usage and splicing (1, 2). In terms of their structural organization, all NRG1 isoforms have an epidermal growth factor-like domain (EGF), which is necessary and sufficient for the biological activities of NRG1 (3, 4). Upstream of the EGF domain, NRG1s contain either an immunoglobulin-like (Ig) domain or a cysteine-rich domain (CRD). The Ig-like domain is thought to mediate binding to the extracellular matrix and potentiate the NRG1 signal (5, 6). Downstream of the EGF domain, C termini contain a transmembrane (TM) domain and a common cytoplasmic “a”-tail and a rare “b”-tail, or they are synthesized as soluble peptides (2, 4). Recently, additional 5′ exons have been identified in the NRG1 gene, giving rise putatively to novel families of NRG1, types IV, V, and VI (7). Genetic variation in NRG1 has been linked to risk for schizophrenia (for review see Ref. 8). The association with schizophrenia was first reported in an Icelandic population in which a NRG1 haplotype (HapA1) consisting of five single nucleotide polymorphisms (SNPs) and two microsatellites covering the 5′-end of the gene doubled the risk for the disorder (9). Four of these SNPs (SNP8NRG221132, SNP8NRG221533, SNP8NRG241930, and SNP8NRG243177 (rs6994992)) represent a 22-kb surrogate haplotype that resides in the 5′ flanking putative promoter region of NRG1 directly upstream of the novel E187 exon, which is unique to the NRG1 type IV isoform (7, 10). Follow-up studies in multiple ethnic populations and a meta-analysis have confirmed genetic association between schizophrenia and NRG1 using markers within the same core haplotype (8, 11) or with overlapping markers in the 5′ region, making NRG1 a leading schizophrenia susceptibility gene.

Recent studies have focused on possible molecular mechanisms of NRG1-mediated susceptibility for schizophrenia by

* This work was supported by the Medical Research Council (MRC), United Kingdom and by Federal funds from the Intramural Research Program of the National Institutes of Health, NCI, Center for Cancer Research and from SAIC-Frederick under Contract NO1-CO-12400. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† This article was selected as a Paper of the Week.

‡ The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. 1 and Table 1.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) EF372273–EF372277 and EF517295–EF517297.

1 An MRC, UK Career Development Fellow and NARSAD (National Alliance for Research on Schizophrenia and Depression) Young Investigator. To whom correspondence should be addressed: University of Oxford, Dept. of Psychiatry, Warneford Hospital, Oxford OX3 7JX, United Kingdom. Tel.: 301-4027810; Fax: 301-4807795; E-mail: amanda.law@psych.ox.ac.uk.

2 The abbreviations used are: NRG1, neuregulin 1; EGF, epidermal growth factor; SNP, single nucleotide polymorphism; TSS, transcription start site; RT-PCR, reverse transcription-PCR; qRT-PCR, quantitative real-time RT-PCR; PBGD, porphobilinogen deaminase; bis-Tris, 2-[bis(2-hydroxyethyl)-amino]-2-(hydroxymethyl)propane-1,3-diol; ORF, open reading frame; CTF, C-terminal fragment; CT, cycle at threshold; HEK293, human embryonic kidney cell line 293.
studying the regulation of NRG1 expression in the human brain. Hashimoto et al. (12) examined type I–III NRG1 transcripts in the dorsolateral prefrontal cortex and found increased NRG1 type I mRNA in schizophrenia, a finding that was subsequently confirmed in a separate and larger sample in the hippocampus (10). However, evidence for a relationship between genetic risk in the gene and increased expression of type I mRNA in the brain was inconclusive. Subsequently, Law et al. (10) identified that genetic variation in the Hapice risk region; particularly a single SNP, rs6994992, is associated with NRG1 type IV mRNA levels in the normal human brain and in individuals with schizophrenia, with individuals carrying the risk allele having higher NRG1 type IV levels. This finding suggests a novel molecular mechanism for the genetic association of the Hapice region of NRG1 with schizophrenia, implicating specifically the type IV isoform. Moreover, based on the location of rs6994992 (in the proximal promoter of NRG1, 1.2 kb upstream of the transcription start site (TSS)) and on the in silico prediction that the SNP maps to a serum response element, Law et al. (10) proposed that rs6994992 represents a functional promoter site involved in NRG1 type IV-specific transcriptional control. Subsequently, other studies have demonstrated that Hapice, and especially rs6994992, is linked to transcriptional control. Subsequently, other studies have demonstrated that Hapice, and especially rs6994992, is linked to transcriptional control. Finally, a single SNP, rs6994992, is associated with schizophrenia risk-associated SNP, rs6994992, is a functional promoter variant associated with schizophrenia genetic predisposition and NRG1 type IV expression.

**EXPERIMENTAL PROCEDURES**

Molecular Cloning and Characterization of NRG1 Type IV cDNA Clones: RT-PCR and cDNA Cloning and Sequencing—To isolate full-length cDNA clones encoding type IV splice isoforms of the human NRG1 gene, RT-PCR and primer-specific amplifications were performed using human brain and human fetal brain cDNAs. Marathon-Ready cDNA libraries were purchased from Clontech (Mountain View, CA). Adult human brain cDNA libraries were generated using total RNA from the hippocampus and prefrontal cortex (Clontech, BD Biosciences, and Ambion (Austin, TX), respectively). For construction of human cDNA libraries, 5 μg of total RNA was reverse-transcribed to cDNA in a total volume of 20 μl by using SuperScript™III (Invitrogen) primed with oligo(dT)20 according to manufacturer’s instructions. After reverse transcription, the cDNA product was digested with 2 units of RNase H (Invitrogen) at 37 °C for 30 min. Two microliters of the first-strand cDNA was employed in PCR amplification using Platinum TaqDNA Polymerase High Fidelity (Invitrogen). RT-PCR primers for full-length NRG1 type IV transcript amplification were designed specific to the unique 5′′ exon E187 of NRG1 type IV (E187_s3, 5′″GGCAGCAGCATGGGGAAAGGA-3″) overlapping the putative translation start site and reverse primers specific for the conserved termination exon of the NRG1 gene, the a-tail (E846_Atail_anti2, 5′″AGGTITTTATACAGCATTTTGGGC-3″). To identify potential transcripts containing a cytoplasmic b-tail (E778), we used a reverse primer located in the a-b-tail junction (a/bjunc anti 1, 5′″TAGCAGGGAAGCT-GTTACTGTCT-3″). Reverse primers for full-length cloning of type IV were designed specific to exon E846 based on previous data, suggesting that NRG1 type IV represents a class of Ig-containing NRG1 variants (7). In the vast majority of Ig-NRG1 variants (types I and II) E846 represents the conserved termination exon. To search for the potential existence of rare b-cytoplasmic tail variants (2), we used a reverse primer in the a-b-tail junction that would delineate a b-tail from the common a-tail. Long range PCR was performed as follows: 94 °C for 5 min followed by 35 cycles of 94 °C for 1 min, 55 °C for 2 min, and 68 °C for 6 min. After the last cycle, extension was conducted at 68 °C for 10 min. PCR products were resolved on 1% agarose gels in 0.5× Tris borate-EDTA buffer. Gels were stained with ethidium bromide, and the DNA bands were visualized with a Kodak EDAS 290 imaging setup that consists of an orange band pass filter, a Kodak DC290 Camera, and a 302 nm UVA transiluminator. A fragment of ~1.8 kb was excised from the gel for each individual sample and purified using the QIAquick gel extraction kit (Qiagen, Valencia, CA). The purified fragment was cloned into either pCR®-XL-TOPO or pCR®2.1-TOPO vector (Invitrogen) and subsequently transformed into TOP10F-competent cells (Invitrogen). Five-to-ten-well isolated colonies were picked up and grown at 37 °C, 225 rpm overnight. The plasmid miniprep DNAs were prepared from 1.5 ml of bacterial cultures by using QIAprep Spin miniprep kit (Qiagen). The insert in selected clones was bidirectionally sequenced by the BigDye terminator kit (Applied Biosystems, Foster City, CA) with M13 primers (M13forward, 5′″GTA-AAACGACGCCAG-3″; M13reverse, 5′″CAGGAAAACAGC-
TATGAC-3’). For deeper sequencing of the inserts and confirmation of clone sequences generated with M13 primers, additional purified plasmid DNAs were resequenced using a “primer walking” strategy with a set of NRG1 sequencing primers designed to the conserved exons within the NRG1 gene (supplemental data, Table 1). The combination of these primers covered the full-length coding region of NRG1 type IV. Sequence data were constructed using the Sequencher software (Gene Codes Corp., Ann Arbor, MI), and the exon structure of each clone was determined by aligning the cDNA sequence to the genomic DNA sequence of the human NRG1 gene (GenBank™ BK000383).

Nucleotide Sequence Accession—The cDNA sequences of the NRG1 type IV splice variants have been submitted to the GenBank™ database under accession numbers EF372273–EF372277 and EF517295–EF517297. The NRG1 genomic DNA sequence used in comparison with the NRG1 type IV cDNA sequences has the GenBank™ accession number BK000383.

Quantitative Real-time RT-PCR (qRT-PCR) Analysis of NRG1 Type IV Expression in Human Tissues—To investigate whether NRG1 type IV is expressed in a panel of selected human tissues, qRT-PCR assays were performed using cDNAs constructed from poly(A)+ mRNA or total RNA from human heart, skeletal muscle, breast tumor, liver, lung, testis, adult hippocampus, fetal brain (Clontech BD Biosciences), human B lymphoblast, and human monocyte cell lines. The human cell lines were derived from normal volunteers participating in studies at the National Institutes of Health. NRG1 type IV mRNA expression levels were measured by qRT–PCR using an ABI Prism 7900 sequence detection system with 384-well format (Applied Biosystems) as described previously (10). The forward primer for qRT-PCR amplification overlaps with that used for RT-PCR amplification of NRG1 type IV in the cloning experiments. Co-amplification and normalization of NRG1 type IV mRNA levels were to the internal control gene, porphobilinogen deaminase (PBGD). Fold changes in expression were calculated as 2–ΔΔCT (cycle at threshold) relative to adult human hippocampus. For comparison, NRG1 types I, II, and III were measured in the same panel of cDNAs as described previously (10).

Western Blot Analysis—Western blot analysis was performed using an antibody to the a-tail cytoplasmic domain of NRG1 (neuregulin-1a/b1/2 (C-20) (sc-348, Santa Cruz Biotechnology). 20 μg of protein from human dorsolateral prefrontal cortex, hippocampus, fetal brain, human heart, and skeletal muscle (Clontech, Protein Medleys) and lymphocytes and monocytes were washed five times (for 5 min each) with TBS-T. Horseradish peroxidase (sc-2004, Santa Cruz Biotechnology). Membranes were washed five times (for 5 min each) with TBS-T. Horseradish peroxidase-immunoreactive protein bands were detected by enhanced chemiluminescence (ECL) Western blotting analysis system (RPN2109, Amersham Biosciences) and exposed to Kodak scientific imaging film. Prestained Kaleidoscope standards were used for product size determination (Bio-Rad, 161-0324) Blots were stripped using Restore Western blot stripping buffer (21062, Pierce Biotechnology) and reprobed with 1:5000 anti-β-actin antibody (A5441, Sigma).

Construction of Promoter–Luciferase Reporter Plasmids, DNA Sequencing, Cell Culture, Transfection, and Measurement of NRG1 Promoter Activity—To investigate whether a mechanism of the genetic association of the 5’ region of NRG1 with schizophrenia involves transcriptional regulation, we used a combined approach of bioinformatic modeling and informed promoter-luciferase fusion experiments. To test whether the 5’ flanking region of the NRG1 gene upstream of E187 represents a functional promoter and rs6994992 represents a cis-acting regulatory element, a 1.5-kb fragment of the human NRG1 gene 5’ to the type IV transcription initiation site was amplified by PCR and used in the construction of luciferase fusion plasmids. PCR primers were designed based on human GenBank™ sequence BK000383 (forward primer F-1500, 5′-GCAGAGC- CATCAATGAGGCTCA-3′; reverse primer R2, 5′-CTGG- GAGTGAAGGTGACCTCA-3′). To examine the effects of genetic variation specifically at rs6994992, promoter fragments were amplified from separate control individuals based on genotype at this SNP. Each construct was compared in single experiments. The recombinants pGL4.SNPC and pGL4.SNPT represent the ancestral and derived alleles, respectively, at rs6994992. Furthermore, to confirm that rs6994992 per se is the functional SNP responsible for differential promoter activity, we used the 1.5-kb pGL4.SNPT construct as a template for site-directed mutagenesis of rs6994992. Single mutants were generated with the QuikChange® II site-directed mutagenesis kit (Strategene) with mutagenic primers to convert T to C at rs6994992 (supplemental Table 1). All mutagenic reactions were performed according to manufacturer’s instructions, and all constructs were verified by direct sequencing.

Because gene transcription can be regulated by promoter competition (15, 16), and the NRG1 type IV promoter may be comparatively inefficient (based on low expression levels in the brain (10)) compared with the neighboring promoter for the type II isoform, we sought insight into neighboring NRG1 promoter localization and activity through the use of a Markov model-based statistical tool for promoter detection (17). Based on bioinformatic predictions we also amplified promoter constructs containing DNA sequence upstream of E187 and into the type II exon TSS including: 1) 1.5 kb upstream of the type IV TSS combined with the downstream tandem promoter for the type II isoform (forward primer F-1500, 5′-GCAGAGC- CATCAATGAGGCTCA-3′; reverse primer −215B, 5′-CTTGCTTCGCTGTTGGACT-3′); 2) 2 kb upstream of the type IV TSS combined with the downstream tandem promoter for the type II isoform (forward primer F-2000, 5′-CTTGCTTCGCTGTTGGACT-3′; reverse primer −215B, 5′-CTTGCTTCGCTGTTGGACT-3′); and 3) 3 kb upstream of the type IV TSS combined with the downstream tandem promoter for the type II isoform (forward primer F-3000, 5′-GGAAGAGTCAAGTGAGGGA-3′; reverse primer −215B, 5′-CTTGCTTCGCTGTTGGACT-3′).
Type IV/II promoter fusion constructs were generated from the SNP-T individual above and were used for comparison with the NRG1 type IV-only promoter constructs to determine the relative strengths of these regions to drive gene expression in vitro.

All promoter PCR products were cloned into pCRII-TOPO vectors (Invitrogen) and then subcloned into the poly linker region of pGL4.10 (Promega) using standard molecular techniques. The promoter region sequences obtained by PCR were verified by direct sequence analysis using M13 forward and reverse primers. Orientations of all promoter sequence constructs were confirmed by restriction endonuclease digestion and DNA sequencing. Sequencing results were analyzed with BLAST. For transfection, a human embryonic kidney cell line (HEK293) was isolated and cultured in Dulbecco's modified Eagle's medium/F-12 (Invitrogen) plus 10% fetal bovine serum and 1% penicillin-streptomycin in a humidified 5% CO2 incubator at 37°C. Three 5 x 10⁵ cells/well were seeded into 6- or 12-well plates, respectively, for transfection. NRG1 promoter-luciferase fusion plasmids (4 μg, 6-well, and 1.6 μg, 12-well; and 400 ng, 6-well, and 160 ng, 12-well) of PRL-TK plasmid (Renilla luciferase expression vector, an internal control for transfection efficiency) were co-transfected with Lipofectamine 2000 (Invitrogen). Plasmid pGL4-SV40 was used as a positive control promoter vector. Negative controls comprised transfection with empty pGL4.10 and no vector-transfected cell controls. All experiments were performed in triplicate and then repeated. 48 h post-transfection, cells were lysed. Luciferase and Renilla activity was measured in 20-μl reactions using the Dual-Luciferase® reporter assay system kit (Promega) on a Lumat LB9507 luminometer. Levels of luciferase activity are expressed as relative light units. The ratio of firefly to Renilla relative light units was determined, and data are normalized to the pGL4.10 empty vector. To address the potential tissue and species specificity of NRG1 type IV promoter activity, we also transfected under similar conditions the following additional cell lines: Neuro 2A, SH-SY5Y, and SK-N-SH-neuroblastoma cell lines. Rat primary cortical neurons were cultured and transfected as described previously (18). qRT-PCR analysis of NRG1 type IV expression in these additional cell lines also was examined.

RESULTS

Cloning, Sequencing, and Characterization of Human Full-length NRG1 Type IV cDNAs—Human NRG1 type IV full-length cDNAs were cloned from adult and fetal human brain cDNA libraries. Sequencing of cDNA clones derived from the adult human brain (hippocampus and prefrontal cortex) revealed that they contained an ~1800-bp insert and an open reading frame (ORF) of 1770 bp encoding a putative NRG1 type IV proprotein of 590 amino acids with a calculated molecular mass of ~66 kDa. This variant corresponds to full-length NRG1 of the novel type IV family, and we termed this variant “type IV-β1a” (Fig. 1). We report that E187 has an AUG codon in the context of a Kozak sequence (CAGCATGG), potentially encoding 13 unique amino acids upstream of E178 (Fig. 1). Detailed characterization of the sequenced products aligned to the assembled genomic sequence (BK000383) revealed the presence of 11 exons (Fig. 1). Each of the 11 exons identified have been reported previously as part of the NRG1 gene and the 5’ exon structure upstream of the EGFc domain (E130) confirms identification of partial NRG1 type IV transcripts utilizing 5’-RACE (rapid amplification of cDNA ends) in the human brain (7). Characterization of the full-length NRG1 type IV transcript revealed that the 3’-exons downstream of the EGF-like domain (E130c) comprise E103, E127, E131, E207, and E846 with a termination codon, placing the transcript in the β1α’ family of NRG1 proteins (Fig. 1). These findings categorize the adult full-length type IV as a NRG1 proprotein belonging to the Ig class of EGF β-1 containing transmembrane variants with a cytoplasmic a-tail.

In the fetal brain, the main full-length β1α variant of NRG1 type IV is identical to that found in the adult human hippocampus and prefrontal cortex (Fig. 1, Fetal A, IV-β1a). In addition, two other full-length variants were identified with variability in the “spacer region” (s1/E51a and s2/E51b) downstream of the Ig-like domains (E178/122) with either the inclusion of exons E51a and E51b (Fig. 1, Fetal B) or just E51a alone (Fetal C). The ORF was maintained in both transcripts. These observations suggest that these variants represent full-length spliced versions of the adult and fetal NRG1 type IV-β1a. In addition, a transcript was identified (Fig. 1, Fetal D, IV-β1a) in the fetal brain that is identical to Fetal C but harbors a nonsense mutation (CAG/TAG) in the inverted repeat codon. This transcript is predicted to produce a truncated NRG1 type IV protein product of 206 amino acids.

Two novel NRG1 type IV splice variants that are unique with respect to any previously described NRG1 transcripts were identified (Fig. 1; Fetal E and F). These fetal brain-specific transcripts result from splicing of exons E59 and E24, creating two transcripts of ~1700–1751 bp consisting of exons E187, E178, E122, and E130 directly spliced to the TMc domain (E103) followed by a full-length cytoplasmic a-tail. These two variants also harbor a premature stop codon in exon E127, which may give rise to a truncated protein. Variability in the spacer region (s1/E51a) distinguishes the two variants.

In addition to the variants described above, other novel splice variants belonging to the NRG1 type IV family were identified in the adult brain (data not shown). A single splice variant was detected from prefrontal cortex cDNAs containing a novel 128-bp exon (termed E128) between E187 and E178 (data not shown). This exon is a perfect match to the NRG1 genomic sequence (nucleotides 359451–359578 in BK000383).1 In addition, E128 is flanked by an AG dinucleotide at the 5’-end, and a GT dinucleotide at the 3’-end, suggesting that E128 may be a novel alternative exon of the NRG1 gene. Although the exon E128 contains five ATG triplets, the ORF is not maintained, so it appears that none of them are used as translation initiation sites. Interestingly, an ATG in the downstream exon E178 may instead act as a translation initiation codon where it may initiate translation of an N-terminally truncated NRG1 type IV protein lacking 22 amino acids. Because E128 may not be translated, the presence of the E128 sequence in the transcript may affect mRNA stability or translation efficiency of the protein. A second splice variant was detected from adult hip-
Molecular Cloning and Characterization of Type IV NRG1

**FIGURE 1.** Organization of NRG1 type IV transcripts cloned from fetal and adult human brain. Exons are shown in the order in which they occur in the transcript. The length of each exon is not proportional to the number of amino acids encoded. The nomenclature describing each exon is listed in the Code box at the bottom of the figure. The top row is a compendium of all exon nomenclatures (see Refs. 2 and 7). The initiation and termination codons are indicated on each variant. Isoforms are named based on the tissue of origin (adult prefrontal cortex (PFC), adult hippocampus, and fetal brain), the categorical structure in terms of the genomic sequence of NRG1 (IV-β1a), and whether they are novel splice variants of the NRG1 gene. (The nucleotide sequences for the type IV variants have been deposited in the GenBank™ data base under accession numbers EF372273–EF372277 and EF517295–EF517297).

pocampus cDNAs containing a tandem repeat of the exon E178 immediately downstream of E187 (data not shown). This duplicated E178 sequence is not found in the NRG1 genomic sequence, and the ORF is not maintained from E187 or the duplicated exon. However, the ATG in the second E178 exon may be used as a translation initiation site. Interestingly, the variants from both the prefrontal cortex and the hippocampus have the potential to produce an identical putative protein of 570 amino acids, which has a perfect sequence match to the full-length NRG1 type IV protein lacking 20 N-terminal amino acids. Determining whether the presence of the novel E128 exon or E178 duplication has any functional implications will require further investigation. All exon deletions and insertions were confirmed by short range RT-PCR analysis of cDNA libraries, with the exception of E128.

It is important to note that forward primers used for amplification of NRG1 type IV that were designed 5' to the AGT of the E187 exon failed to specifically amplify NRG1 type IV and returned many false positives including nonspecific amplification of other gene transcripts. This may be in part because of the GC-rich nature of the 5'-region of the E187 exon. These observations are of critical importance for studies of NRG1 type IV at the transcript level.

Expression of NRG1 Type IV: qRT-PCR and Western Blot Analysis—NRG1 type IV mRNA expression was not detected outside of the human brain (Fig. 2A). In contrast, expression of NRG1 types I and II was detected in all tissues with the exception of lymphocytes and skeletal muscle, and NRG1 type III was detected in heart and lung as well as in the developing and adult brain (data not shown). These results provide compelling evidence that NRG1 type IV is a unique brain-specific isoform of the NRG1 gene. Furthermore, NRG1 type IV expression was ~3.5-fold higher in the fetal brain compared with that of the adult brain (2ΔΔCT; Fig. 2A).

As no information is available regarding the type IV NRG1 protein, we used an antibody specific to the C-terminal a-tail of NRG1 (sc-348, Santa Cruz) to identify NRG1 type IV. We identified a 66-kDa band exclusively expressed in the adult and fetal brain (Fig. 2B). This band is consistent with the predicted molecular mass of type IV-β1a (and no other known NRG1 isoform or product) and is absent in peripheral tissues where NRG1 type IV mRNA is not expressed. We propose that this ~66-kDa protein represents NRG1 type IV. We also identified NRG1 proteins at 110, 95, and 60 kDa (Fig. 2B), which are consistent with previous observations (19–21). These proteins are proposed to represent I-β1a and III-β1a proproteins and a C-terminal fragment (CTF) cleavage product, respectively. The ~60-kDa CTF was identified in the heart and muscle (at longer exposures; see Fig. 2B, inset), where NRG1 type IV mRNA was not expressed, thus confirming that the CTF fragment of NRG1 is distinct from the 66-kDa type IV protein. A 45-kDa band was identified in all tissues (Fig. 2B), which is reported as corresponding to NRG1-α (3).

Characterization of the 5’ Flanking Promoter Region of the Human NRG1 Gene—Using a bioinformatic method for examining eukaryotic polymerase II transcription start sites, we identified the presence of a putative promoter in the 5’ flanking region of NRG1 directly adjacent to the E187 exon followed by a second one ~500 bp downstream of the end of E187 representing the contiguous promoter for E1006, the exon specific to the type II isoform of NRG1 (Fig. 3, A and B).

First, we experimentally addressed whether the 5’ flanking region of E187 represents a functional promoter in vitro and
whether differential promoter activity is associated with rs6994992. Under basal conditions, the activity of the 1.5-kb human NRG1 type IV promoter-luciferase fusion gene was 65% higher in T allele constructs (pGL4.SNPT) compared with that of the C allele construct (pGL4.SNPC) in transfected HEK293 cells (Fig. 3C). These findings are consistent with previous observations in the human brain (10) whereby the allele associated with risk for schizophrenia (T), is associated with higher expression of the NRG1 type IV variant. To determine whether rs6994992 is singularly responsible for the differential promoter activity seen between the two sequences, we generated single point mutation of this site (T to C) within the 1.5-kb promoter activity seen between the two sequences, we generated

rs6994992 is singularly responsible for the differential promoter activity observed. Importantly, when comparable experiments were conducted in Neuro 2A, SH-SY5Y, and SK-N-SH-neuroblastoma cell lines and rat primary cortical neurons, activity of the NRG1 type IV promoter was not observed (data not shown). These observations suggest that the human NRG1 type IV promoter may be “silenced” or “inactivated” in a cell type (i.e. in cancer)- or species-specific (i.e. rodent) manner. These suggestions are based on the observations that NRG1 type IV gene expression was absent in the cell lines of cancer origin that we tested and that the NRG1 type IV promoter region and rs6994992 are not conserved across mouse, rat, and human species. Furthermore, transcript analysis confirmed that type IV mRNA was not expressed in neuroblastoma cell lines (compared with high levels of NRG1 type II; data not shown) suggesting that the native type IV promoter is inactive in these cells.

We identified that the NRG1 type II promoter is located 500 bp 3' of E187 and includes a TATA-like sequence (TAAATTTAAA) immediately upstream of the TSS (which is absent in the type IV promoter). Based on neighboring maxima predictions and expression levels in the brain, the NRG1 type II promoter may represent a stronger promoter in competition with the type IV promoter (Fig. 3A). Indeed, a greater than 100-fold induction of promoter activity was detected compared with the single NRG1 type IV promoter construct, when the 1.5-kb 5' flanking fragment was cloned in a natural physical context with the type II promoter (compare panels C and D in Fig. 3). Furthermore, serial 5' extension of the NRG1 type IV-II promoter construct resulted in either no alteration in promoter activity (2-kb 5' addition) or decreased activity (3-kb 5' addition) compared with the 1.5 kb-containing construct (Fig. 3, B and D). These observations suggest that repressor or insulator elements may be present 5' to the 2-kb promoter region. For NRG1 type IV-II promoters, similar results were seen in Neuro 2A cells and rat primary cortical neurons (data not shown).

**DISCUSSION**

Many human genes are known to produce more than one protein isoform through the use of alternative promoters and complex arrays of cis-regulatory elements that determine the efficient and accurate initiation of gene transcription (for review see Ref. 23). The use of multiple promoters is a frequent mechanism for the generation of isoforms that provide tissue or cell type specificity (i.e. spatial specificity) or developmental stage specificity (temporal specificity) and that allow for diverse functional properties of a gene. NRG1 is a classic example of
differential promoter usage whereby nine alternative promoters have been identified (2, 7). SNPs residing in promoters can cause or contribute to genetic diseases by creating entirely new transcriptional elements that change the level of gene transcription. A striking example of this is the set of mutations involved in α-thalassemia (24). Such mechanisms may also underlie the pathogenesis of complex heritable diseases such as schizophrenia, where disease-associated coding variants appear to be uncommon in comparison with noncoding SNPs residing in modulatory DNA (see Ref. 25 for review). Recently, a similar pathogenetic scenario has been described pertaining to a genetic variant (SNP8NRG243177; rs6994992) in the 5′-flanking region of the NRG1 gene (10), which is linked with risk for schizophrenia (8, 9, 11). This SNP, which is predicted to interfere with a cis-regulatory serum response element in the 5′ NRG1 promoter, is associated with increased transcript expression of a novel E187-containing type IV variant of NRG1 in the human brain, an effect that is exaggerated in schizophrenia (10).

We demonstrate for the first time the presence of a functional promoter for the human type IV transcript, 1.5 kb upstream of the E187 TSS, and show that allele-specific differences related to promoter activity are seen for rs6994992 with the same directionality seen in the human brain (10). Specifically, promoter activity of construct rs6994992-T, which represents the schizophrenia risk allele, is 65% higher than that of construct rs6994992-C, an observation confirmed by single point mutagenesis. These data provide evidence that rs6994992 represents an operative cis-acting regulatory element in the NRG1 5′ promoter and that the mechanism of association with this SNP involves regulation of type IV expression. We also demonstrate that a 2.2–2.7-kb region of DNA, 1.5–2 kb upstream of E187 and downstream into the type II promoter, represents a stronger promoter, suggesting that promoter com-
petition may occur between the type IV and II promoters \textit{in vivo}, a common phenomenon seen in other genes where tandem promoters are present (15, 16, 26). The two alternative promoters are located only 687 bp apart, and the type II promoter includes a TATA-like sequence (TAAATATAA) directly upstream of the TSS, which may explain the propensity for a stronger promoter.

Characterization of full-length novel NRG1 type IV variant in the fetal and adult human brain reveals that it exists as a 1.8-kb transcript that falls into the category of a transmembrane NRG1 of the Ig-\beta 1a class. The 5' region of the type IV transcript consists of a unique E187 leader exon, an Ig-like domain(s), and an EGFc domain, whereas the 3' end consists of a \beta-1 stalk sequence and a full transmembrane domain and cytoplasmic a-tail, categorizing it as similar in structure to the Ig-containing type I variant of NRG1 (1, 2). From its identified full-length sequence, NRG1 type IV is predicted to have an ORF encoding 590 amino acids initiated at an AUG methionine codon in E187; we have identified a 66-kDa NRG1 protein representing a putative “type IV proprotein,” which is exclusively expressed in the adult and fetal human brain.

In the fetal brain we describe two novel NRG1 type IV variants that have not been described previously in the context of any NRG1 isoform family. These variants arise from the alternative splicing of the type IV transcript resulting in the exclusion of the E59/E24 exons, which comprise the \beta 1 stalk region of the gene. Furthermore, these two variants contain premature termination sites in the E127 exon of the cytoplasmic tail and may thus give rise to truncated C-terminal variants, as reported previously in native NRG1 type I variants (27). Splicing of the novel type IV isoforms results in the deletion of regions implicated in EGF bioactivity (\beta-EGF domain, fifth and sixth cysteine) and the regulated cleavage of transmembrane NRG1s by matrix metalloproteases (\textit{i.e.} 1 stalk, otherwise referred to as the “linker” region, 2, 28–30). Such variability may confer distinct processing and binding affinities to these proteins and developmental functions. Equally, whether these variants are translated or functional \textit{in vivo} remains to be determined. Moreover, these variants are absent from the adult brain, suggesting that they play a specific role in early brain development. Given the proposed abnormality of neurodevelopment that is thought to occur in schizophrenia (31–34), these developmental variants may have particular relevance.

Quantitative expression profiling of a range of tissues demonstrates that NRG1 type IV mRNA is expressed exclusively in the human brain. This is unique compared with other known classes of NRG1 (types I–III) that exhibit certain degrees of cell type specificity but not tissue exclusivity. Furthermore, we demonstrate that NRG1 type IV expression is 3.5-fold higher in the fetal brain, again supporting the notion that the isoform may have an especially important functional role during early brain development, mediated potentially through the effects of NRG1 on neurogenesis, neuronal migration, cell differentiation, and synapse formation, as well as the regulation of neurotransmitter function, including \textit{N}-methyl-\textit{d}-aspartate, \textit{\gamma}-aminobutyric acid, \textit{\alpha}-7, and dopamine, all of which have been reported as biological functions of the Ig-like domain variants of the NRG1 gene (2, 8). The fact that its expression is specific to the brain may be valuable in terms of the possible therapeutic candidacy of NRG1 type IV in schizophrenia.

Finally, many studies show that NRG1 plays an important role in oncogenesis and is overexpressed in various cancer tissues (35). In this regard, it is notable that we did not detect NRG1 type IV in breast tumor and neuroblastoma cell lines. This absence, when juxtaposed with the increased NRG1 type IV mRNA associated with genetic risk for schizophrenia (10), may contribute to the established, but unexplained, reduced incidence of cancer in patients with schizophrenia and in their relatives (22, 36). Although further studies are needed to determine the function of NRG1 type IV in human cancer, this convergent pattern suggests that genetic regulation of type IV may have a dual effect of both protecting against cancer while increasing the risk for schizophrenia.

In conclusion, we have shown that type IV represents a novel, translated, developmentally regulated, brain-specific isoform that belongs to the Ig-\beta 1a NRG1 family. We have also shown that a NRG1 promoter, SNP, which impacts on schizophrenia risk and other brain phenotypes, is functional and selectively affects transcriptional levels of the type IV isoform, providing a mechanistic basis for the functional associations.

REFERENCES

1. Buonanno, A., and Fischbach, G. D. (2001) \textit{Curr. Opin. Neurobiol.} \textbf{11}, 287–296

2. Falls, D. L. (2003) \textit{J. Neurocytol.} \textbf{32}, 619–647

3. Holmes, W. E., Sliwkowski, M. X., Akiti, R. W., Henzel, W. I., Lee, J., Park, J. W., Yansura, D., Abadi, N., Raab, H., Lewis, G. D., Shepherd, H. M., Kuang, W. J., Wood, W. I., Goeddel, D. V., and Vanolen, R. L. (1992) \textit{Science} \textbf{256}, 1205–1210

4. Wen, D., Suggs, S. V., Karunagaran, D., Liu, N., Cuples, R. L., Luo, Y., Janssen, A. M., Ben Baruch, N., Trollinger, D. B., Jacobsen, V. L., Meng, S. Y., Lu, H. S., Hu, S., Chang, D., Yang, W., Yanigahara, D., Koshi, R. A., and Yarden, Y. (1994) \textit{Mol. Cell. Biol.} \textbf{14}, 1909–1919

5. Loeb, J. A., and Fischbach, G. D. (1995) \textit{J. Cell Biol.} \textbf{130}, 127–135

6. Rimer, M. (2003) \textit{J. Neurocytol.} \textbf{32}, 665–675

7. Steinhorsdottir, V., Stefansson, H., Ghosh, S., Birgisdottir, B., Bjornsdot- tir, S., Fasquel, A. C., Olafsson, O., Stefansson, K., and Gulcher, J. R. (2004) \textit{Gene} \textbf{342}, 97–105

8. Harrison, P. J., and Law, A. J. (2006) \textit{Biol. Psychiatry} \textbf{60}, 132–140

9. Stefansson, H., Sigurdsson, E., Steinhorsdottir, V., Bjornsdotir, S., Sig- mundsson, T., Ghosh, S., Brynjolfsson, I., Gunnarsdottir, S., Ivarsson, O., Chou, T. T., Hjaltaison, O., Birgisdottir, B., Jonsson, H., Guadnottid, V. G., Gudmundsdottir, E., Bjornsson, A., Ingvarsson, B., Ingason, A., Sigfusson, S., Hardardottir, H., Harvey, R. P., Lai, D., Zhou, M., Brunner, D., Mutel, V., Gonzalez, A., Lemke, G., Sainz, J., Johannesson, G., Andresson, T., Gud- hjartsson, D., Manolescu, A., Frigge, M. L., Gurney, M. E., Kong, A., Gulcher, J. R., Petursson, H., and Stefansson, K. (2002) \textit{Am. J. Hum. Genet.} \textbf{71}, 877–892

10. Law, A. J., Lipska, B. K., Weickert, C. S., Hyde, T. M., Straub, R. E., Hashimoto, R., Harrison, P. J., Kleinman, J. E., and Weinberger, D. R. (2006) \textit{Proc. Natl. Acad. Sci. U. S. A.} \textbf{103}, 6747–6752

11. Li, D., Collier, D. A., and He, L. (2006) \textit{Hum. Mol. Genet.} \textbf{15}, 1995–2002

12. Hashimoto, R., Straub, R. E., Weickert, C. S., Hyde, T. M., Kleinman, J. E., and Weinberger, D. R. (2004) \textit{Mol. Psychiatry} \textbf{9}, 299–307

13. Hall, J., Whalley, H. C., Job, D. E., Baig, B. J., McIntosh, A. M., Evans, K. L., Thomson, P. A., Porteous, D. J., Cunningham-Owens, D. G., Johnstone, E. C., and Lawrie, S. M. (2004) \textit{Neurosci. Lett.} \textbf{37}, 1477–1478

14. Stefanis, N. C., Trikalinos, T. A., Avramopoulos, D., Smyrnis, N., Evdokimidis, I., Ntzani, E. E., Ioannidis, J. P., and Stefanis, C. N. (2007) \textit{Biol. Psychiatry}, in press

15. Choi, O. R., and Engel, J. D. (1988) \textit{Cell} \textbf{55}, 17–26
Molecular Cloning and Characterization of Type IV NRG1

16. Conte, C., Dastugue, B., and Vaury, C. (2002) EMBO J. 21, 3908–3916
17. Ohler, U. (2006) Nucleic Acids Res. 34, 5943–5950
18. Sala, C., Piech V., Wilson, N. R., Passafaro M., Liu G., and Sheng, M. (2001) Neuron 31, 115–130
19. Wang, J. Y., Miller, S. J., and Falls, D. L. (2001) J. Biol. Chem. 276, 2841–2851
20. Frenzel, K. E., and Falls, D. L. (2001) J. Neurochem. 77, 1–12
21. Bao, J., Wolpowitz, D., Role, L. W., and Talmage, D. A. (2003) J. Cell Biol. 161, 1133–1141
22. Levav, I., Lipshitz, I., Novikov, I., Pugachova, I., Kohn, R., Barchana, M., Ponizovsky, A., and Werner, H. (2007) Br. J. Psychiatry 190, 156–161
23. Ayoubi, T. A., and Van de Ven, W., J. (1996) FASEB J. 10, 453–460
24. De Gobbi, M., Viprakasit, V., Hughes, J. R., Fisher, C., Buckle, V. J., Ayyub, H., Gibbons, R. J., Vernimmen, D., Yoshinaga, Y., de Jong, P., Cheng, J. F., Rubin, E. M., Wood, W. G., Bowden, D., and Higgs, D. R. (2006) Science 312, 1215–1217
25. Harrison, P. J., and Weinberger, D. R. (2005) Mol. Psychiatry 10, 40–68
26. Schibler, U., and Sierra, F. (1987) Annu. Rev. Genet. 21, 237–257
27. Zhang, Z., Prentiss, L., Heitzman, D., Stahl, R. C., DiPino, F., Jr., and Carey, D. J. (2006) J. Neurosci. Res. 84, 1–12
28. Lu, H. S., Hara, S., Wong, L. W.-J., Jones, M. D., Katta, V., Trail, G., Zou, A., Brankow, D., Cole, S., Hu, S., and Wen, D. (1995) J. Biol. Chem. 270, 4775–4783
29. Montero, J. C., Yuste, L., Díaz-Rodriguez, E., Esparis-Ogando, A., and Pandiella, A. (2000) Mol. Cell. Neurosci. 16, 631–648
30. Montero, J. C., Rodriguez-Barrueco, R., Yuste, L., Juanes, P. P., Borges, J., Esparis-Ogando, A., and Pandiella, A. (2007) Mol. Biol. Cell 18, 380–393
31. Weinberger, D. R. (1986) in The Neurology of Schizophrenia (Nasrallah, H. A., ed) pp. 397–406, Elsevier, Amsterdam
32. Weinberger, D. R. (1987) Arch. Gen. Psychiatry 44, 660–669
33. Murray, R. M., and Lewis, S. W. (1987) Br. Med. J. 295, 681–682
34. Lewis, D. A., and Levitt, P. (2002) Annu. Rev. Neurosci. 25, 409–432
35. Armin, D. N., Tuck, D., and Stern, D. F. (2005) Exp. Cell Res. 309, 12–23
36. Jablensky, A., and Lawrence, D. (2001) Arch. Gen. Psychiatry 58, 579–580