Functional Effects of a Tandem Duplication Polymorphism in the 5′Flanking Region of the DRD4 Gene

Ursula M. D'Souza, Carsten Russ, Eda Tahir, Jonathan Mill, Peter McGuffin, Philip J. Asherson, and Ian W. Craig

Background: *Several polymorphisms have been identified in the 5′ flanking region of the human dopamine D_4 receptor gene (DRD4), including a tandem duplication polymorphism. This comprises a 120-base-pair repeat sequence that is known to have different allele frequencies in various populations around the world. Furthermore, various studies have revealed evidence of linkage to attention-deficit/hyperactivity disorder and association with schizophrenia and methamphetamine abuse. The location of the polymorphism in the 5′ regulatory region of the DRD4 gene and the fact that it consists of potential transcription factor binding sites suggest that it might confer differential transcriptional activity of the alleles.*

Methods: *We investigated the functional effects of this gene variant with transient transfection methods in four human cell lines and then assessed transcriptional activity with luciferase reporter gene assays.*

Results: *The longer allele has lower transcriptional activity than the shorter allele in SK-N-MC, SH-SY5Y, HEK293, and HeLa cell lines.*

Conclusions: *This evidence suggests that the duplication might have a role in regulating the expression of the DRD4 gene and provides an understanding of the biological mechanisms underlying the etiology of neuropsychiatric disorders such as ADHD, schizophrenia, and metamphetamine abuse.*

Key Words: Dopamine receptors, DRD4, ADHD, tandem duplication, polymorphism, regulatory

The human dopamine D_4 receptor gene (DRD4), which codes for a seven transmembrane domain protein belonging to the super family of G-protein-coupled receptors, was originally cloned by Van Tol et al (1991). The gene has been implicated in several neuropsychiatric and neurobehavioral disorders, and pharmacologic studies have revealed that the atypical antipsychotic drug clozapine binds with high affinity to the dopamine D_4 receptor. High expression of DRD4 messenger ribonucleic acid (mRNA) was detected in the medulla, frontal cortex, midbrain, and amygdala, with lower levels of the gene expressed in the basal ganglia (Van Tol et al 1991). Multiple polymorphic variants of the DRD4 gene in the human population are known to exist as two-, four-, or seven-repeat sequences of a 48-base-pair (bp) region in exon 3, which codes for the third putative cytoplasmic loop of the receptor (Van Tol et al 1992). Van Tol et al also reported that the seven-repeat variants had different binding profiles with clozapine and spiperone compared with the two- and four-repeat variants, with respect to sodium chloride sensitivity. Similarly, the same group later found that the polymorphic repeat sequences showed small differences in DRD4 pharmacologic binding properties (Asghari et al 1994; Jovanovic et al 1999) and also in functional characteristics to inhibit cyclic adenosine monophosphate (Asghari et al 1995). Therefore, it was concluded from the evidence that there was no direct relationship between length of the polymorphism and changes in these particular activities. The DRD4 variants were also capable of coupling to several G protein ($G_{i\alpha}$) subtypes, but no evidence of any quantitative difference in G protein coupling related to the number of repeats was observed (Kazmi et al 2000). Additionally, evidence suggests that the seven-repeat variant originated as a rare mutational event that increased to high frequency in human populations by positive selection (Ding et al 2002; Wang et al 2004).

The DRD4 gene has been implicated in other neurobehavioral disorders, such as novelty seeking, attention-deficit/hyperactivity disorder (ADHD), and substance abuse. Additionally, reports have described the association of the personality trait of novelty seeking (which comprises impulsive, exploratory, and sensation-seeking behavior) with the allele having seven repeats in both Israeli and American populations (Benjamin et al 1996; Ebstein et al 1996). More recently, many findings demonstrate the association and linkage between the seven-repeat allele in DRD4 and ADHD (Faraone et al 1999, 2001; Holmes et al 2000; LaHoste et al 1996; Langley et al 2004; Mill et al 2001; Rowe et al 1998; Smalley et al 1998; Swanson et al 1998; Tahir et al 2000); however, a smaller number of studies have reported no association of the seven-repeat allele with ADHD (Castellanos et al 1998; Eisenberg et al 2000; Hawi et al 2000; Kotler et al 2000; Todd et al 2001). Attention-deficit/hyperactivity disorder is a common syndrome that begins in childhood and is characterized by inattention, hyperactivity, and impulsivity and is also associated with school failure, peer-relationship problems, early use of nicotine and alcohol, and an increased frequency of accidents (Todd and O'Malley 2001). The seven-repeat allele was also reported to be associated with opiate dependence (Kotler et al 1997) and with heroin abuse in a Han Chinese sample (Li et al 1997).

The cloning and characterization of the 5′flanking region of the human DRD4 gene (Kamakura et al 1997) enabled the search for novel polymorphisms. One of these was the −521CT polymorphism, which was identified in the promoter region of DRD4 and showed weak association to schizophrenic Japanese patients (Okuyama et al 1999). Functional effects of this promoter polymorphism were analyzed in human retinoblastoma cells (Y-79 cells), and transient transfection methods demonstrated that the T allele reduced transcriptional efficiency by 40%

From the Medical Research Council Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, United Kingdom.

Address reprint requests to Ursula M. D'Souza, Social, Genetic and Developmental Psychiatry Centre, Box Number P082, Denmark Hill, London SE5 8AF, United Kindom; E-mail: spjgumd@iop.kcl.ac.uk.

Received February 26, 2004; revised July 20, 2004; accepted August 25, 2004.

0006-3223/04/$30.00
doi:10.1016/j.biopsych.2004.08.008

compared with the C allele. The same functional polymorphism was also found to be significantly associated with the human novelty-seeking personality trait (Okuyama et al 2000). A further study also demonstrated association of this promoter polymorphism with novelty seeking (Ronai et al 2001).

A different polymorphism was later identified in the 5′flanking region of the DRD4 gene (Seaman et al 1999) and found to be a tandem duplication of 120 bp that was located 1.2 kilobases (kb) upstream from the initiation codon. The frequency of the duplicated allele ranged from .40 to .81 in 11 populations around the world. Further information on the allele frequency of this polymorphism is currently available from Kidd's laboratory, through a database known as ALFRED (the Allele FREquency Database; http://alfred.med.yale.edu). This database stores frequencies of alleles at human polymorphic sites and of haplotypes for multiple populations and has revealed frequencies of the duplicated allele ranging from .4 to .9 in 33 populations. Furthermore, haplotype data of the tandem duplication with single nucleotide polymorphisms in the promoter region (−521CT) and the 5′untranslated region of the DRD4 gene is also available from this database. Interestingly, after the identification of the tandem duplication polymorphism, a report was subsequently published that showed linkage of this polymorphism with children having ADHD (McCracken et al 2000). Recently, further evidence was reported supporting linkage of this polymorphism to ADHD (Mill et al 2003) and from association studies to schizophrenia (Xing et al 2003), methamphetamine abuse (Li et al 2004), and novelty-seeking in bipolar and alcoholic families (Rogers et al 2004). There are consensus-binding sequences for several transcription factors within this duplication that is situated approximately 850 bp upstream of the start of transcription and upstream from the promoter region and silencer regions of the gene. This suggests that this particular polymorphism might have a role in regulating the transcriptional activity of the gene. It was hypothesized that the duplication could create a risk for ADHD by decreasing expression of the DRD4 gene (McCracken et al 2000). Therefore, we set out to investigate the functional effects of this duplication, using transfection methods in mammalian cell lines and reporter gene technology. The results reported here indeed reveal that the tandem duplication lowers transcription activity in various cell lines compared with the short variant.

Methods and Materials

Subjects

A sample of Turkish children with ADHD was collected as described by Tahir et al (2000), and samples 542 and 622 were used for cloning in the reporter gene vectors because they were homozygous for the long and short allele, respectively.

Oligonucleotides for Polymerase Chain Reaction and Plasmid Constructs

The primers used in the polymerase chain reaction (PCR) for genotyping were designed according to Seaman et al (1999). The numbering was based on the nucleotide sequence given by Kamakura et al (1997) (GenBank accession number U95122). The forward primer was D4upstrFor2 (5′GTT GTC TGT CTT TTC TCA TTG TTT CCA TTG3′, −1607/–1578), and the reverse primer was D4upstrRev3 (5′GAA GGA GCA GGC ACC GTG AGC3′, −1179/−1199). The primers used for the cloning of the inserts into the reporter gene vector were as follows: D4upF/XhoI: 5′GGG **CTC GAG** CTG GGA GAG AAG AAA CTT CCA C 3′ (bold type represents the XhoI restriction enzyme site, which preceded bases −1389/−1368); D4upR/HindIII: 5′GGG **AAG CTT** ACA GGA CAA GGT CAC CGG 3′ (bold type represents the HindIII restriction enzyme site, which precedes bases −1203/−1220). The duplicated and nonduplicated versions of the polymorphism were amplified with the above primers and cloned into the TA cloning vector pCR 2.1 TOPO, released by HindIII and XhoI and subcloned into the same sites of pGL3-Basic vector, which lacks eukaryotic promoter and enhancer sequences (Promega UK, Southhampton, Hampshire, United Kingdom). Orientations of the inserts were verified by fluorescent sequencing using the ABI 3100 DNA Sequencer (Applied Biosystems, Foster City, California).

Cell Culture, Transient Transfection, and Luciferase Assays

The human neuroepithelioma cell line SK-N-MC, the human embryonic kidney cell line HEK293, the human neuroblastoma cell line SH-SY5Y, and HeLa cells (human adenocarcinoma) were all purchased from American Type Culture Collection (ATCC, Manassas, Virginia). All cell lines were cultured in six-well tissue culture plates containing minimum essential medium Eagle (obtained from ATCC) and supplemented with 10% fetal bovine serum (Life Technologies, Rockville, Maryland) and grown at 37°C in a humidified atmosphere containing 5% CO_2. Transfections of SK-N-MC and SH-SY5Y cells were carried out with the Calcium Phosphate Transfection Kit (Invitrogen, Paisley, United Kingdom). Transfection efficiencies were normalized by cotransfection with the Renilla vector, pRL-SV40 (Promega). Each transfection of the SK-N-MC and SH-SY5Y cells used 5 μg of deoxyribonucleic acid (DNA) together with .25 μg of Renilla control vector. The HEK293 and the HeLa cells were transiently transfected with Lipofectamine 2000 reagent (Life Technologies), with 1 μg of luciferase DNA constructs together with 1:20 (50 ng) Renilla control vector. Both pGL3-Basic (lacks the SV40 promoter and enhancer) and pGL3-Control (contains the SV40 promoter and enhancer sequences, which results in strong expression of the luciferase gene) were separately transfected into the cells, which served as negative and positive controls, respectively, within the assay.

After the transfection period, the cells were allowed to grow for 48 hours. Then, the cells were washed with Dulbecco's phosphate-buffered saline and 500 μL of lysis buffer, added to each well of cells. The firefly luciferase and Renilla luciferase assays were carried out with the Dual-Luciferase® Reporter (DLR™) Assay System (Promega). Briefly, 10–20 μL of cell lysate was mixed with 50–100 μL of the luciferase assay reagent II for 3 sec and the firefly luciferase activity quantified for 12 sec. Then, 50–100 μL of Stop and Glo® Reagent (Promega) was added and mixed for 3 sec and the Renilla luciferase activity quantified for 12 sec. The transfection efficiencies were also qualitatively assessed with a vector containing green flourescent protein (GFP) under the control of the cytomegalovirus (CMV) promoter, and cells expressing the protein were observed under a fluorescence microscope.

Fluorogenic 5′Nuclease Assay

Total RNA was first prepared from each of the four cell lines (SK-N-MC, SH-SY5Y, HEK293, and HeLa) with the RNeasy Mini Kit (Qiagen Ltd UK, Crawley, West Sussex, United Kingdom) and then treated with DNase I (Qiagen) to remove any genomic DNA contamination. A 2-μg aliquot of total RNA was used to prepare complementary (c)DNA with random hexamers and Taqman reversal transcription reagents (Applied Biosystems, Foster City,

Table 1. Luciferase Transcriptional Activity of the DRD4 Constructs in Cell Lines

Luciferase/Renilla Ratio	SK-N-MC Cells	HEK293 Cells	SH-SY5Y Cells	HeLa Cells
pGL3-Basic	.00204 ± .0004	.071 ± .0077	.01504 ± .0023	.00028 ± .00004
pGL3-Basic DRD4S	.026 ± .00257	.295 ± .0545	.197 ± .030	.00692 ± .00042
pGL3-Basic DRD4L	.019 ± .00133	.167 ± .0142	.124 ± .0173	.00476 ± .00035
pGL3-Control	2.22 ± .11	2.61 ± .42	5.95 ± .92	.42 ± .08

The transcriptional activity results are expressed as a ratio of the firefly luciferase versus the renilla luciferase values and represented as mean ± SEM ($n = 9$). Three independent experiments were carried out for each cell line, with each experiment performed in triplicate.

DRD4S, one-repeat allele variant of the 120-bp polymorphism in the DRD4 gene; DRD4L, two-repeat allele variant of the 120-bp polymorphism in the DRD4 gene.

California). All cDNA samples were tested for genomic DNA contamination by PCR amplification of a nontranscribed sequence. The samples were stored at −70°C before further use in TaqMan assays. The TaqMan assays were all carried out in triplicate in 384-well plates with 1 μL of cDNA per well and TaqMan Universal Mastermix (Applied Biosystems) primers and probe for the DRD4 gene designed with the program Primer Express (Applied Biosystems) and mRNA sequence accession number NM_000797. The primers selected out include the forward primer 5′ CTG GGC TAC GTC AAC AGC G 3′, 1162/1180, the reverse primer 5′ CGT TGC GGA ACT CGG C 3′, 1213/1228, and the probe designed as follows: 5′ FAM CCT CAA CCC CGT CAT CTA CAC TGT CTT CA 3′, 1182/1210. Two TaqMan predeveloped assay reagent probes were used as endogenous controls: human β-actin (human ACTB) and human GAPD (GAPDH) (glyceraldehydes-3-phosphate-dehydrogenase) were commercially supplied from Applied Biosystems. These housekeeping genes help to normalize the levels of gene expression and are designed so that they do not detect any genomic DNA. The TaqMan assays were performed on an ABI Prism 7900HT Sequence Detector (Applied Biosystems). The expression data produced were analyzed and converted into threshold cycle values (Ct values) with the software program SDS 2.0 (Applied Biosystems). Qualitative levels of expression were calculated by subtracting the Ct values obtained for DRD4 from the Ct values obtained for the housekeeping genes.

Results

The transcriptional luciferase activity of the long and short alleles of the DRD4 tandem duplication were calculated by taking the ratio of the firefly luciferase activity versus the Renilla luciferase activity (Table 1). These ratios were compared with that obtained for the pGL3-Basic vector to produce fold values representing increments in transcriptional activity over basic for the short and long construct, respectively, for cell line as follows: SH-SY5Y: 13.1 ± 2.0, 8.2 ± 1.2; SK-N-MC: 12.7 ± 1.3, 9.3 ± .7; HeLa: 24.7 ± 1.5, 17.0 ± 1.3; and HEK293: 4.2 ± .8, 2.4 ± .2 (Figure 1). Significant differences in luciferase activity were observed between the long and short constructs at $p < .001$ in SK-N-MC, SH-5Y5Y, and HeLa cells and at $p = .02$ in HEK293 cells. Significant differences in luciferase activity were also observed between the basic plasmid and the short construct at $p < .001$ in all cells lines and between the basic plasmid and the long construct at $p < .001$ in SK-N-MC, SH-5Y5Y, and HeLa

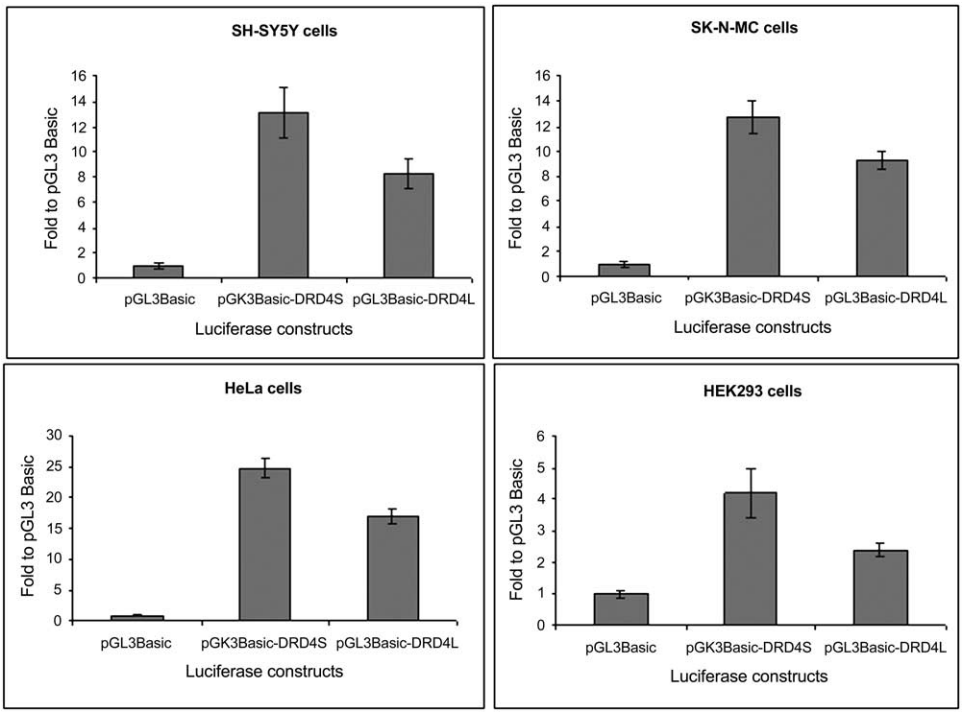

Figure 1. Transcriptional activity of the DRD4 tandem duplication polymorphism in cells. The data are expressed as luciferase activity relative to pGL3-Basic and shown as means ± SEM ($n = 9$).

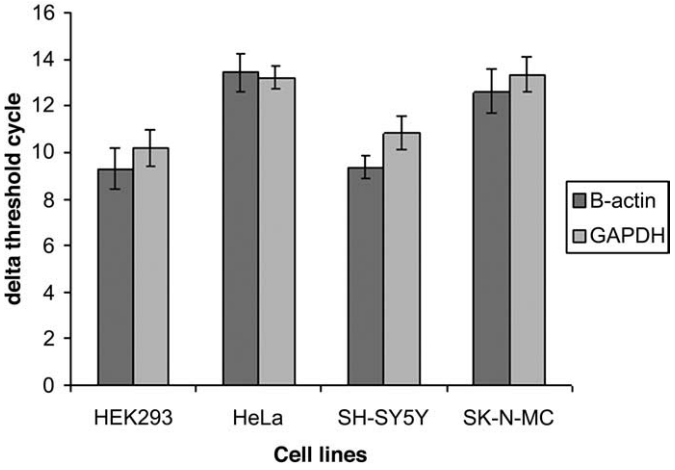

Figure 2. Expression of DRD4 messenger ribonucleic acid (mRNA) in mammalian cell lines according to real-time reverse-transcriptase polymerase chain reaction. The DRD4 mRNA expression is represented as threshold cycle values relative to expression of a housekeeping gene (deltaCT) obtained by subtracting the housekeeping gene value (number of amplification cycles for the product to reach a fixed threshold value) from obtained for DRD4. The light gray bars are relative to glyceraldehyde-3-phosphate-dehydrogenase (GAPDH) expression and the darker shaded bars relative to β-actin. The means ($\pm$ SD) were calculated from three experiments, each performed in triplicate ($n = 3$).

cell lines and at $p = .071$ in HEK293 cells. The statistical analysis was done with the analysis of variance test in a general linear model in the SPSS program (SPSS, Chicago, Illinois).

The target DRD4 mRNA expression normalized to an endogenous reference was calculated by generating a ΔCt. This value is the difference in threshold cycles for the target and reference genes. The reference genes included the housekeeping genes, β-actin and GAPDH. The ΔCt DRD4 values relative to β-actin and GAPDH for each of the cell lines are as follows (means $\pm$ SD): SK-N-MC: 12.6 $\pm$ .94, 13.32 $\pm$ .77; SH-SY5Y: 9.35 $\pm$ .49, 10.86 $\pm$ .70; HEK293: 9.3 $\pm$ .87, 10.18 $\pm$ .78; HeLa: 13.41 $\pm$ .85, 13.21 $\pm$ .47 (Figure 2). These results show that all cell lines used in this study express the DRD4 gene endogenously (at detectable levels), with the HEK293 and SH-SY5Y cell lines showing the highest DRD4 expression relative to β-actin and GAPDH. The lower the Ct values, the higher the expression of the gene.

Discussion

In the present study, we performed functional analysis of the tandem duplication polymorphism in the DRD4 gene by using transient transfection in mammalian cell lines and reporter gene (luciferase) assays. This 120-bp duplication is located in the 5′ flanking region of the gene, approximately 1.2 kb upstream of the initiation codon (Seaman et al 1999). Functional analysis of the 5′ flanking region of the DRD4 gene by a transient expression method in IMR32 and Y-79 cells demonstrated promoter activity between −591 bp and −123 bp and a silencer between −770 bp and −679 bp relative to the initiation codon (Kamakura et al 1997). Interestingly, the tandem duplication polymorphism is present upstream of these two regulatory regions and is known to contain predicted consensus binding sites for several transcription factors (e.g., MEP-1 [metal element protein-1] and Sp1, a transcription factor that was purified with sephacryl and phosphocellulose columns), suggesting that the different alleles might differ in transcriptional activity (Seaman et al 1999). Our results show the longer allele to have consistently lower luciferase transcriptional activity than the shorter allele in four cell lines: SK-N-MC (human neuroepithelioma), which was originally used to clone the DRD4 gene (Van Tol et al 1991), SH-SY5Y (human neuroblastoma), HEK293 (human embryonic kidney), and HeLa (human adenocarcinoma), which all endogenously express the DRD4 gene, as demonstrated by real-time reverse transcriptase (RT)-PCR TaqMan assays. The assays detected only mRNA because all traces of any genomic DNA contamination in the total RNA preparations from the various cell lines were removed with a DNase I treatment and then confirmed with PCR amplification with primers that amplify an intron sequence. No PCR products were present in all the total RNA samples, indicating no genomic DNA contamination (data not shown). The cloning of the human DRD4 gene was achieved with SK-N-MC cells by Van Tol et al (1991), thus the decision to use this particular cell line for the functional studies. Interestingly, the other mammalian cell lines that were tested also generated results similar to those obtained with SK-N-MC cells, with all cell lines endogenously expressing the DRD4 gene, as indicated from TaqMan assays. These data suggest that these cells contain the essential transcriptional genetic machinery for the regulation of the DRD4 gene. To date, this is the first study in which the expression of the DRD4 gene in mammalian cell lines has been determined with the TaqMan real-time RT-PCR methodology. This assay is more sensitive and accurate compared with traditional PCR and measures the kinetics of the reaction in the early phases of the reaction (Bustin 2000). Therefore, even though Kamakura et al (1997), using standard RT-PCR, did not detect any DRD4 expression in the HeLa cells, it is possible that the conditions they used were not sensitive enough to identify the gene. A similar argument would also account for the detection of endogenous DRD4 expression in the HEK293 cells seen from our studies. In various studies, HEK293 cells have been used for the heterologous expression of proteins, including dopamine receptors, suggesting no presence of dopaminergic activity in these cells. Interestingly, however, this particular cell line has been used recently to demonstrate the functional effects of a polymorphism in the regulatory region of the dopamine transporter gene (Miller and Madras 2002), suggesting endogenous expression of dopaminergic genes in these cells. Moreover, HEK293 cells are derived from human embryonic kidney tissue, and there are reports on the expression of the D_2 subfamily of dopamine receptor genes in kidney tissue (Gao et al 1994; O'Connell et al 1998). This adds further support for the endogenous expression of dopaminergic genes observed in the HEK293 cells.

The results obtained here are in agreement with the hypothesis that the tandem duplication might result in decreased receptor levels as a result of altered DRD4 transcription (McCracken et al 2000). McCracken et al (2000) used the transmission disequilibrium test and reported that the long allele of the polymorphism transmitted preferentially in 371 children with ADHD. The exploratory analyses of the inattentive phenotypic subtype of ADHD seemed to strengthen the evidence for association of the long allele with this feature. It is also important to bear in mind that there was a trend for the transmission of the allele with one 120-bp repeat, but this was not significant in a mixed European Caucasian sample of ADHD children and their families collected in Toronto, Canada (Barr et al 2001); however, no association of the polymorphism was observed in a population of twins (Todd et al 2001). Thus, the strongest findings of linkage of the 120-bp repeat polymorphism with ADHD have been observed by McCracken et al (2000) and recently by Mill et

al (2003). The latter study indicated a haplotype of markers in the 5′regulatory region of the DRD4 gene (including the long allele of the 120-bp repeat polymorphism) that conferred susceptibility to ADHD in a large clinical sample and their families. More recently it was found that the long allele of the 120-bp repeat polymorphism was significantly associated with schizophrenic patients of Han Chinese origin who were recruited from the Shanghai Mental Health Centre (Xing et al 2003). Additionally, analysis of the 120-bp variable number of tandem repeat (VNTR) polymorphism and the exon 3 VNTR in the dopamine D_4 receptor as a haplotype showed significant association with methamphetamine abuse, which gave an empirical p value of .0034 for a heterogeneity model (Li et al 2004). The above evidence points to the relevance of the tandem duplication polymorphism in behavioral phenotypes and thus the significance in investigating the functional effects of this particular polymorphism.

There are speculations as to the location of the DRD4 promoter region itself and thus as to the exact position of the tandem duplication polymorphism relative to the transcription initiation region, as initially discussed by Todd and O'Malley (2001). They suggested that the promoter region characterized by Kamakura et al (1997) could be in an intron region. This was due to a discrepancy in the size of the DRD4 mRNA that had been previously reported. The first report of cloning of the human DRD4 gene described the size of the mRNA as 5.3 kb in the human neuroblastoma cell line SK-N-MC, in rat striatum, in rat brain area comprising hypothalamus, mesencephalon, and medulla and in monkey brain areas (olfactory tubercle, hippocampus, medulla, amygdala, cerebellum, midbrain, frontal cortex, and striatum) by Northern blot analysis (Van Tol et al 1991). The size of the DRD4 mRNA, however, was later found to be 1.5 kb and present at higher levels in human corpus callosum, spinal cord, medulla, and subthalamic nucleus, as well as in the mesolimbic system and related regions (amygdala, hippocampus, and hypothalamus) (Matsumoto et al 1996). The studies by Matsumoto et al revealed lower DRD4 mRNA levels in the striatum (caudate nucleus and putamen) and barely detectable levels in the cerebral cortex. Thus, the discrepancy in the mRNA size suggests that the DRD4 gene might have alternative transcription start sites. A similar finding has been revealed for the catechol-O-methyltransferase (COMT) gene, which is known to have alternative transcription start sites resulting in the generation of two distinct specific transcripts detected in various human tissues and cell lines (Tenhunen et al 1994). Additionally, in view of the discrepancy of the size of the DRD4 mRNA, Todd and O'Malley (2001) suggested that there is also the possibility that a large intron could separate the coding exons of the DRD4 gene from the untranslated exons, as observed for the other D_2-like receptors (e.g., in the D_2 and D_3 receptor genes [D'Souza et al 2001; Minowa et al 1992; O'Malley et al 1990; Valdenaire et al 1994]). This suggests that the tandem duplication could be located in an intron region. More recently, bioinformatics on the University of California-Santa Cruz browser (http://genome.ucsc.edu) has revealed that this specific polymorphism is indeed located in intron 1 of the gene, according to GENESCAN predictions. Further detailed characterization of the 5′ flanking region of the DRD4 gene would need to be carried out in future experiments to clearly define the core promoter region and transcription initiation sites.

It is important to emphasize here that the aim in this study was very specific in nature and focused on determining whether any functional differences exist between the long and short alleles of the DRD4 tandem duplication in four different cell lines. Therefore, because the regulatory region of the DRD4 gene comprising this polymorphism has not been well established and the possibility that it might lie in an intron region, we set out to test its effects in a vector without any endogenous promoter and therefore used the pGL3-Basic vector in the cloning experiments. The results reported here clearly demonstrate that there are consistent functional differences between the one- and two-repeat alleles of the 120-bp polymorphism in four cell lines tested. Interestingly, our data suggest that the DRD4 gene has an alternative promoter region, because the region in which the tandem duplication polymorphism is located has demonstrated promoter activity. This is in agreement with previous findings in other dopamine receptor genes (e.g., an additional promoter exists in intron 1 of the human D_1 dopamine receptor gene [Lee et al 1996]). Furthermore, there is increasing evidence of the functional effects of intronic regions, as seen with the microsatellite repeat polymorphisms located in the intron regions of the human tyrosine hydroxylase (Albanese et al 2001; Meloni et al 1998) and serotonin transporter genes (Fiskerstrand et al 1999; Klenova et al 2004; MacKenzie and Quinn 1999).

Therefore, after consideration of the above comments, the results can be interpreted meaningfully as the basic vector on its own producing very low background activity and the control vector (having both the simian virus 40 early promoter and enhancer regions) showing the expected high transcriptional activity in all cell lines, as indicated in Table 1. The data reported here clearly demonstrate that significant differences exist in the transcriptional activity of the long and short alleles of this polymorphism. Lower levels of transcription activity observed with the long form of the polymorphism would result in lower levels of expression of the DRD4 gene. This in turn could have an effect on the levels of dopamine in the synapse and therefore could represent a plausible molecular mechanism underlying the etiology of ADHD, because the long form of the polymorphism is preferentially transmitted in children with ADHD. The location of this polymorphism was found to be upstream from the promoter and silencer regions of the DRD4 gene that were previously determined in IMR32 and Y-79 cells (Kamakura et al 1997).

Our data reveal that the tandem duplication polymorphism has differential transcriptional activity in the four cell lines tested when luciferase is used as the reporter gene. It is important to note that there is a discrepancy in the findings obtained with HeLa cells because previously no activity of DRD4 constructs in this cell line was seen (Kamakura et al 1997); however, this disagreement could be explained by the use of a different reporter gene, chloramphenicol acetyltransferase (CAT). It is known that the sensitivity of luciferase assays is in the subattomole range, approximately 30–1000 times greater than the sensitivity of CAT assays (Pazzagli et al 1992). This indicates that the CAT assays used by Kamakura et al (1997) were not sensitive enough to detect transcriptional activity within the DRD4 5′flanking region.

The tandem duplication of 240 bp is an exact 120-bp repeat with no mismatches, and when used in the Basic Local Alignment Search Tool (BLAST) database search only the short allele (120 bp) was found to have 98%, 97%, and 96% homology with the a common chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), and orangutan (*Pongo pygmaeus*) DRD4 sequences, respectively, with no homology seen with rat and mouse sequences. This is in agreement with evolution of exon 1 of the DRD4 gene in primates. In humans, this region of the gene comprises a

number of polymorphic variants and thus evolutionary history of the exon examined in thirteen other primate species (Seaman et al 2000). The results showed different patterns of variation in monkeys compared with apes and were suggestive of a strong purifying selective pressure and a relaxed selection pressure on the exon in the monkeys and apes, respectively. Further studies will be required to establish the precise genomic organization of control and coding regions of the human DRD4 locus.

Other tandem repeat sequences (twofold repeats) have also been found in the human dopamine D_4 receptor gene, where a polymorphic 12-bp repeat was identified in exon 1 that codes for the extracellular N-terminal part of the receptor (Catalano et al 1993). The rarer shorter allele (A2 allele) was present at a higher frequency than the more common twofold repeat variant (A1 allele) in patients having delusional disorder. Additionally, there is a report on the existence of twofold repeats in transmembrane channel proteins and the potential effects of these on their tertiary structures (Wistow et al 1991).

We have demonstrated the enhanced transcriptional activity of the short allele compared with the long (120-bp duplication) at the human DRD4 locus. The functional effects of this tandem duplication might have significant implications in behavioral phenotypes and be relevant to the study by McCracken et al (2000), which showed a significant preferential transmission of the long allele from parents to their ADHD children. This is in agreement with the findings described here, whereby the longer allele has lower transcriptional activity in cells compared with the shorter allele, thus suggesting differential expression levels of the two alleles. The McCracken study also showed that the 120-bp repeat was in linkage disequilibrium with the seven-repeat allele in exon 3 in the coding region of the DRD4 gene. There is much evidence for association of the seven-repeat allele in exon 3 with behavioral phenotypes such as the personality trait of novelty seeking (Ebstein et al 1996) and also ADHD (Faraone et al 1999; Holmes et al 2000; LaHoste et al 1996; Swanson et al 1998; Tahir et al 2000). The experimental strategy for advancing evidence of a functional role for the DRD4 tandem duplication polymorphism would be to use gel shift assays to investigate any interactions with transcription factors and to assess the functional effect of haplotypes of the DRD4 polymorphisms in the 5′flanking region. Furthermore, multiple combinations of the 5′tandem duplication and exon 3 VNTR alleles could exist where each haplotype could be associated with disease/disorder and have a functional characteristic.

The data reported here demonstrate that the tandem duplication polymorphism in the 5′flanking region of the DRD4 gene affects cellular expression. Duplications are known to effect expression and also preserve the function of genes (Bailey et al 2002), and there is increasing evidence of their prevalence in the human genome and of their significance in human disease and physiology (Mazzarella and Schlessinger 1998). The evidence shown from the studies described here supports the above statements and reveals an understanding of the functional importance of repeat sequence polymorphisms that are associated with behavioral phenotypes.

The research study was supported by the Medical Research Council (MRC) with an MRC Component Grant (#G0000197). We thank Dr. Sabine Landau at the Biostatistics Department of the Institute of Psychiatry, London, for her assistance with the statistical analysis of the data.

Albanese V, Biguet NF, Kiefer H, Bayard E, Mallet J, Meloni R (2001): Quantitative effects on gene silencing by allelic variation at a tetranucleotide microsatellite. *Hum Mol Genet* 10:1785–1792.

Asghari V, Sanyal S, Buchwaldt S, Paterson A, Jovanovic V, Van Tol HMM (1995): Modulation of intracellular cyclic AMP levels by different human dopamine D_4 receptor variants. *J Neurochem* 65:1157–1165.

Asghari V, Schoots O, Van Kats S, Ohara K, Jovanovic V, Guan HC, et al (1994): Dopamine D_4 receptor repeat: Analysis of different native and mutant forms of the human and rat genes. *Mol Pharmacol* 46:364–373.

Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, et al (2002): Recent segmental duplications in the human genome. *Science* 297:1003–1007.

Barr CL, Feng Y, Wigg KG, Schachar R, Tannock R, Roberts W, et al (2001): 5′-Untranslated region of the dopamine D_4 receptor gene and attention-deficit hyperactivity disorder. *Am J Med Genet* 105:84–90.

Benjamin J, Li L, Patterson C, Greenberg BD, Murphy DL, Hamer DH (1996): Population and familial association between the D_4 dopamine receptor gene and measures of novelty seeking. *Nat Genet* 12:81–84.

Bustin SA (2000): Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 25: 169–193.

Castellanos FX, Lau E, Tayebi N, Lee P, Long RE, Giedd JN, et al (1998): Lack of an association between a dopamine-4 receptor polymorphism and attention deficit/hyperactivity disorder: Genetic and brain morphometric analyses. *Mol Psychiatry* 3:431–434.

Catalano M, Nobile M, Novelli E, Nothen MM, Smeraldi E (1993): Distribution of a novel mutation in the first exon of the human dopamine D_4 receptor gene in psychotic patients. *Biol Psychiatry* 34:459–464.

Ding Y-C, Chi H-C, Grady DL, Morishima A, Kidd JR, Kidd KK, et al (2002): Evidence of positive selection acting at the human dopamine receptor D4 gene locus. *Proc Natl Acad Sci U S A* 99:309–314.

D'Souza UM, Wang W, Gao D-Q, Kanda S, Lee G, Junn E, et al (2001): Characterization of the 5′ flanking region of the rat D_3 dopamine receptor gene. *J Neurochem* 76:1736–1744.

Ebstein RP, Novick O, Umansky R, Priel B, Osher Y, Blaine D, et al (1996): Dopamine D_4 receptor (D4DR) exon III polymorphism associated with the human personality trait of novelty seeking. *Nat Genet* 12:78–80.

Eisenberg J, Zohar A, Mei-Tal G, Steinberg A, Tarkakovsky E, Gritsenko I, et al (2000): A haplotype relative risk study of the dopamine D_4 receptor (DRD4) exon III repeat polymorphism and attention deficit hyperactivity disorder (ADHD). *Am J Med Genet* 96:258–261.

Faraone SV, Biederman J, Weiffenbach B, Keith T, Chu MP, Weaver A, et al (1999): Dopamine D_4 gene 7-repeat allele and attention deficit hyperactivity disorder. *Am J Psychiatry* 156:768–770.

Faraone SV, Doyle AE, Mick E, Biederman J (2001): Meta-analysis of the association between the 7-repeat allele of the dopamine D_4 receptor gene and attention deficit hyperactivity disorder. *Am J Psychiatry* 158: 1052–1057.

Fiskerstrand CE, Lovejoy EA, Quinn JP (1999): An intronic polymorphic domain often associated with susceptibility to affective disorders has allele dependent differential enehancer activity in embryonic stem cells. *FEBS Lett* 458:171–174.

Gao DQ, Canessa LM, Mouradian MM, Jose PA (1994): Expression of the D_2 subfamily of dopamine receptor genes in kidney. *Am J Physiol* 266:F646–F650.

Hawi Z, McCarron M, Kirley A, Daly G, Fitzgerald M, Gill M (2000): No association of the dopamine DRD4 receptor (DRD4) gene polymorphism with attention deficit hyperactivity disorder (ADHD) in the Irish population. *Am J Med Genet* 96:268–272.

Holmes J, Payton A, Barrett JH, Hever T, Fitzpatrick H, Trumper AL, et al (2000): A family-based and case-control association study of the dopamine D_4 receptor gene and dopamine transporter gene in attention deficit hyperactivity disorder. *Mol Psychiatry* 5:523–530.

Jovanovic V, Guan HC, Van Tol HH (1999): Comparative pharmacological and functional analysis of the human dopamine D4.2 and D4.10 receptor variants. *Pharmacogenetics* 9:561–568.

Kamakura S, Iwaki A, Matsumoto M, Fukumaki Y (1997): Cloning and characterisation of the 5′ flanking region of the human dopamine D_4 receptor gene. *Biochem Biophys Res Commun* 235:321–326.

Kazmi MA, Snyder LA, Cypess AM, Graber SG, Sakmar TP (2000): Selective reconstitution of human D_4 dopamine receptor variants with $G_{i\alpha}$ subtypes. *Biochemistry* 39:3734–3744.

Klenova E, Scott AC, Roberts J, Shamsuddin S, Lovejoy EA, Bergmann S, et al (2004): YB-1 and CTCF differentially regulate the 5-HTT polymorphic intron 2 enhancer which predisposes to a variety of neurological disorders. *J Neurosci* 24:5966–5973.

Kotler M, Cohen H, Segman R, Gritsenko I, Nemanov L, Lerer B, et al (1997): Excess dopamine D_4 receptor (D4DR) exon III seven repeat allele in opioid-dependent subjects. *Mol Psychiatry* 2:251–254.

Kotler M, Manor I, Sever Y, Eisenberg J, Cohen H, Ebstein RP, Tyano S (2000): Failure to replicate an excess of the long dopamine D_4 exon III repeat polymorphism in ADHD in a family-based study. *Am J Med Genet* 96: 278–281.

LaHoste GJ, Swanson JM, Wigal SB, Glabe C, Wigal T, King N, Kennedy JL (1996): Dopamine D_4 receptor gene polymorphism is associated with attention deficit hyperactivity disorder. *Mol Psychiatry* 1:121–124.

Langley K, Marshall L, van den Bree M, Thomas H, Owen M, O'Donovan M, Thapar A (2004): Association of the dopamine D_4 receptor gene 7-repeat allele with neuropsychological test performance of children with ADHD. *Am J Psychiatry* 161:133–138.

Lee S-H, Minowa MT, Mouradian MM (1996): Two distinct promoters drive transcription of the human D_{1A} dopamine receptor gene. *J Biol Chem* 271:25292–25299.

Li T, Chen C-K, Hu X, Ball D, Lin S-K, Chen W, et al (2004): Association analysis of the DRD4 and COMT genes in methamphetamine abuse. *Am J Med Genet* 129B(1):120–124.

Li T, Xu K, Deng H, Cai G, Liu J, Liu X, et al (1997): Association analysis of the dopamine D_4 gene exon III VNTR and heroin abuse in Chinese subjects. *Mol Psychiatry* 2:413–416.

MacKenzie A, Quinn J (1999): A serotonin transporter gene intron 2 polymorphic region correlated with affective disorders has allele dependent differential enhancer-like properties in the mouse embryo. *Proc Natl Acad Sci U S A* 96:15251–15255.

Matsumoto M, Hidaka K, Tada S, Tasaki Y, Yamaguchi T (1996): Low levels of mRNA for dopamine D_4 receptor in human cerebral cortex and striatum. *J Neurochem* 66:915–919.

Mazzarella R, Schlessinger D (1998): Pathological consequences of sequence duplications in the human genome. *Genome Res* 8:1007–1021.

McCracken JT, Smalley SL, McGough JJ, Crawford L, Del'Homme M, Cantor RM, et al (2000): Evidence for linkage of a tandem duplication polymorphism upstream of the dopamine D_4 receptor gene (DRD4) with attention deficit hyperactivity disorder (ADHD). *Mol Psychiatry* 5:531–536.

Meloni R, Albanese V, Ravassard P, Treilhou F, Mallet J (1998): A tetranucleotide polymorphic microsatellite located in the first intron of the tyrosine hydroxylase gene acts as a transcription regulatory element in vitro. *Hum Mol Genet* 7:423–428.

Mill J, Curran S, Kent L, Richards S, Gould A, Virdee V, et al (2001): Attention deficit hyperactivity disorder (ADHD) and the dopamine D_4 receptor gene: Evidence of association but no linkage in a UK sample. *Mol Psychiatry* 6:440–444.

Mill J, Fischer N, Curran S, Richards S, Taylor E, Asherson P (2003): Polymorphisms in the dopamine D_4 receptor gene and attention-deficit hyperactivity disorder. *Neuroreport* 14:1463–1466.

Miller GM, Madras BK (2002): Polymorphisms in the 3′untranslated region of human and monkey dopamine transporter genes affect reporter gene expression. *Mol Psychiatry* 7:44–55.

Minowa T, Minowa MT, Mouradian MM (1992): Analysis of the promoter region of the rat D_2 dopamine receptor gene. *Biochemistry* 31:8389–8396.

O'Connell DP, Vaughan CJ, Aherne AM, Botkin SJ, Wang ZQ, Felde RA, Carey RM (1998): Expression of the dopamine D_3 receptor protein in the rat kidney. *Hypertension* 32:886–895.

Okuyama Y, Ishiguro H, Nankai M, Shibuya H, Watanabe A, Arinami T (2000): Identification of a polymorphism in the promoter region of DRD4 associated with the human novelty seeking personality trait. *Mol Psychiatry* 5:64–69.

Okuyama Y, Ishiguro H, Toru M, Arinami T (1999): A genetic polymorphism in the promoter region of DRD4 associated with expression and schizophrenia. *Biochem Biophys Res Commun* 258:292–295.

O'Malley KL, Mack KJ, Gandelman KY, Todd RD (1990): Organization and expression of the rat D_{2A} receptor gene: Identification of alternative transcripts and a variant donor splice site. *Biochemistry* 29:1367–1371.

Pazzagli M, Devine JH, Peterson DO, Baldwin TO (1992): Use of bacterial and firefly luciferases as reporter genes in DEAE-dextran-mediated transfection of mammalian cells. *Anal Biochem* 204:315–323.

Rogers G, Joyce P, Mulder R, Sellman D, Miller A, Allington M, et al (2004): Association of a duplicated repeat polymorphism in the 5′-untranslated region of the DRD4 gene with novelty seeking. *Am J Med Genet* 126B: 95–98.

Ronai Z, Szekely A, Nemoda Z, Lakatos K, Gervai J, Staub M, Sasvari-Szekely M (2001): Association between novelty seeking and the –521 C/T polymorphism in the promoter region of the DRD4 gene. *Mol Psychiatry* 6:35–38.

Rowe DC, Stever C, Giedinghagen LN, Gard JMC, Cleveland HH, Teris ST, et al (1998): Dopamine DRD4 receptor polymorphism and attention deficit hyperactivity disorder. *Mol Psychiatry* 3:419–426.

Seaman MI, Chang FM, Quinones AT, Kidd KK (2000): Evolution of exon 1 of the dopamine D_4 receptor (DRD4) gene in primates. *J Exp Zool* 288: 32–38.

Seaman MI, Fischer JB, Chang FM, Kidd KK (1999): Tandem duplication polymorphism upstream of the dopamine D_4 receptor gene (DRD4). *Am J Med Genet* 88:705–709.

Smalley SL, Bailey JN, Palmer CG, Cantwell DP, McGough JJ, Del'Homme MA, et al (1998): Evidence that the dopamine D_4 receptor is a susceptibility gene in attention deficit hyperactivity disorder. *Mol Psychiatry* 3: 427–430.

Swanson JM, Sunohara GA, Kennedy JL, Regino R, Fineberg E, Wigal T, et al (1998): Association of the dopamine receptor D_4 (DRD4) gene with a refined phenotype of attention deficit hyperactivity disorder (ADHD): A family-based approach. *Mol Psychiatry* 3:38–41.

Tahir E, Yazgan Y, Cirakoglu B, Ozbay F, Waldman I, Asherson PJ (2000): Association and linkage of DRD4 and DRD5 with attention deficit hyperactivity disorder (ADHD) in a sample of Turkish children. *Mol Psychiatry* 5:396–404.

Tenhunen J, Salminen M, Lundstrom K, Kiviluoto T, Savolainen R, Ulmanen (1994): Genomic organization of human catechol-O-methyltransferase gene and its expression from two distinct promoters. *Eur J Biochem* 223:1049–1059.

Todd RD, Neuman RJ, Lobos EA, Jong Y-J, Reich W, Heath AC (2001): Lack of association of the dopamine D_4 receptor gene polymorphisms with ADHD subtypes in a population sample of twins. *Am J Med Genet* 105: 432–438.

Todd RD, O'Malley KL (2001): The dopamine receptor DRD4 gene: Are duplications distracting? *Trends Pharmacol Sci* 22:55–56.

Valdenaire O, Vernier P, Maus M, Dumas Milne Edwards JB, Mallet J (1994): Transcription of the rat dopamine-D2-receptor gene from two promoters. *Eur J Biochem* 220:557–84.

Van Tol HHM, Bunzow JR, Guan HC, Sunahara RK, Seeman P, Niznik HB, Civelli O (1991): Cloning of the gene for a human dopamine D_4 receptor with high affinity for the antipsychotic clozapine. *Nature* 350:610–614.

Van Tol HHM, Wu CM, Guan HC, Ohara K, Bunzow JR, Civelli O, et al (1992): Multiple dopamine D_4 receptor variants in the human population. *Nature* 358:149–152.

Wang E, Ding Y-C, Flodman P, Kidd JR, Kidd KK, Grady DL, et al (2004): The genetic architecture of selection at the human dopamine receptor D_4 (DRD4) gene locus. *Am J Med Genet* 74:931–944.

Wistow GJ, Pisano MM, Chepelinsky AB (1991): Tandem sequence repeats in transmembrane channel proteins. *Trends Biol Sci* 16:170–171.

Xing Q-H, Wu S-N, Lin Z-G, Li H-F, Yang J-D, Feng G-Y, et al (2003): Association analysis of polymorphisms in the upstream region of the human dopamine D_4 receptor gene in schizophrenia. *Schizophr Res* 65:9–14.