A novel gene codes for a putative G protein-coupled receptor with an abundant expression in brain

Brian F. O'Dowd[a,b,*], Tuan Nguyen[a], Kevin R. Lynch[e], Lee F. Kolakowski, Jr.[f], Miles Thompson[b], Regina Cheng[b], Adriano Marchese[b], Gordon Ng[b], Henry H.Q. Heng[d], Susan R. George[a,b,c]

[a]*Addiction Research Foundation, 33 Russell St, Toronto, Ont. M5S 2S1, Canada*
[b]*Department of Pharmacology, University of Toronto, Toronto, Ont. M5S 1A8, Canada*
[c]*Department of Medicine, University of Toronto, Toronto, Ont. M5S 1A8, Canada*
[d]*SeeDNA Biotech Inc., Farquharson Bldg., 4700 Keele St., Downsview, Ont. M3J 1P3, Canada*
[e]*Department of Pharmacology, University of Virginia Health Sciences Center, 1300 Jefferson Park Ave, Charlottesville, VA 22908, USA*
[f]*Department of Pharmacology, University of Texas Health Science Center at San Antonio, San Antonio, TX 78284, USA*

Received 7 July 1996; revised version received 2 August 1996

Abstract Following the cloning of the dopamine receptors we continued a search of the human genome for related genes. We searched an EST data base and discovered cDNA fragments encoding novel G protein-coupled receptor genes. The available GenBank sequence of one of these EST fragments showed that it encoded a receptor with closest similarity to the D2 dopamine and adrenergic receptors. This cDNA was used to isolate the gene (GPR19), and the encoded receptor also demonstrated similarity with the neuropeptide Y receptor. The gene was mapped to chromosome 12, in region p13.2–12.3. Northern blot analysis revealed expression of GPR19 in peripheral regions, and brain regions significantly overlapping with the D2 receptor gene expression. A sequence of the rat orthologue of GPR19 was obtained and in situ hybridization analysis demonstrated a very abundant expression in rat brain.

Key words: Polymerase chain reaction; Intronless gene; Polymorphism; Chromosome

1. Introduction

Following our reports on the cloning of the D1, D5 and the related pseudogenes we continued a broad search for additional and related members of this G protein-coupled receptor (GPCR) gene family. We have utilized several technical strategies to isolate novel members of this gene family, belonging to many diverse classes. During this process we reported the characterization of many human genes encoding GPCRs, namely APJ [1], GPR1, GPR2, and GPR3 [2], GPR4, GPR5, and GPR6 [3], GPR7, and GPR8 [4], GPR9, GPR10, and GPR14 [5], GPR15 [6], GPR20, 21, 22 and 23 [7]. In this report, we describe a novel human gene (presently named GPR19) encoding a putative receptor which has significant amino acid sequence identity with the D2 dopamine and neuropeptide Y families of receptors. This gene, intronless in the coding region, also has a distribution in brain overlapping that of the D2 dopamine receptor gene, and is located on chromosome 12.

2. Materials and methods

2.1. Searching the data base of expressed sequence tags

We queried the data base of expressed sequence tags (EST) maintained by the National Center for Biotechnology Information (NCBI), with different rhodopsin family G protein-coupled receptor amino acid sequences using the TBLASTN algorithm [8]. The dbEST sequences returned that had statistically significant scores (E values < 0.05) were searched manually to determine whether highly conserved amino acid sequence motifs (e.g. N/DPXXY) were present in the translated sequences. The conceptualized amino acid sequences were then used to query our G protein-coupled receptor data base to eliminate from further consideration those EST sequences that represented known G protein-coupled receptors. The amino acid sequences thus filtered were used to query the SwissProt (release 31) data base using the FASTA algorithm (BLOSSUM 50 matrix, ktup = 1) [9,10] and those sequences that scored significantly (E value < 0.01) against more than one known G protein-coupled receptor were retained. The EST cDNAs representing sequences that met these criteria were requested from the I.M.A.G.E. Consortium. The reported nucleotide sequences were verified by sequencing of both ends of the cDNA.

2.2. Human library screening and PCR amplification of rat genomic DNA

The EST fragment (I.M.A.G.E. I.D. #45231), a 1.4 kb cDNA inserted in the lafmid vector at the *Hin*dIII/*Not*I sites, was used to screen a human λEMBL3 SP6/T7 (Clontech) genomic library as previously described [2]. The probe binding phage clones were purified, digested with restriction enzymes, and subjected to Southern blot hybridization with the same probe used to screen the library. The DNA fragments were isolated, subcloned into pBluescript and sequenced. Rat genomic DNA was amplified by PCR using degenerate oligonucleotides (5′-ATGGAATTAAGTGAGGAGCAC and 5′-AATGAGTCATAGATCGAGTC) designed based on the sequence of the human EST fragment. Rat genomic DNA (1 μg) was amplified with these primers, under the following conditions: 1.5 min at 95°C, 2 min at 50°C, and 4 min at 73°C for 30 cycles. The PCR product was subcloned as described previously [2].

2.3. Northern blot analysis

RNA samples from different brain regions and tissues were extracted as previously described [11]. Briefly, total RNA was extracted by the method of Chomczynski and Sacchi [12] and polyA$^+$ RNA was isolated using oligo-dT cellulose spin columns (Pharmacia). RNA was denatured and size fractionated on a 1% formaldehyde agarose gel, transferred onto nylon membrane and immobilized by UV crosslinking. The blots were probed with a ^{32}P-labeled fragment, washed with 2×SSC, 0.1% SDS at 50°C for 20 min and with 0.1×SSC, 0.1% SDS and exposed to X-ray film for 3 days at −70°C with an intensifying screen.

2.4. In situ hybridization analysis

Male rats (Charles River, ~200–250 g) were killed by decapitation and brains removed in 30 s and frozen in crushed dry ice. Frozen

*Corresponding author. Fax: (1) (416) 978-2733.
E-mail: bodowd@hookup.net

PII S0014-5793(96)00901-5

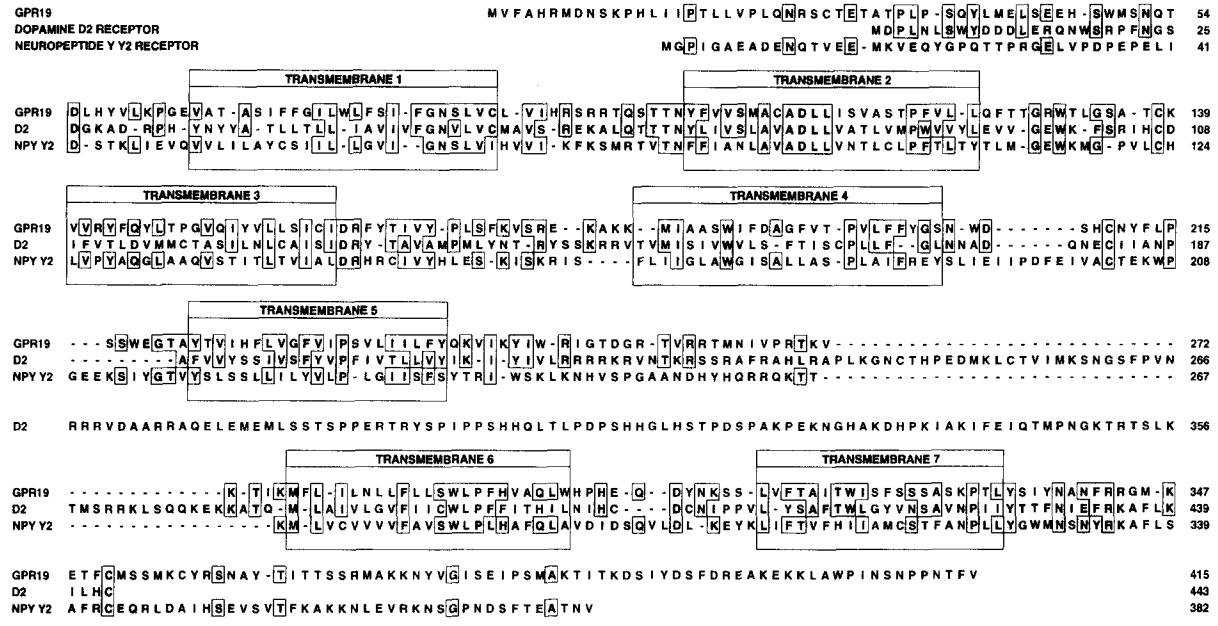

Fig. 1. Amino acid comparison of the putative receptor encoded by GPR19 with related G protein-coupled receptors. The protein encoded by GPR19 is compared with the human D2 receptor, and neuropeptide Y2 receptor. Amino acids identical with the protein encoded by GPR19 are boxed and shaded. The predicted seven-transmembrane domains are indicated. Gaps (−) have been introduced to maximize the alignments among the sequences. The GenBank Accession numbers for human and rat GPR19 are U64871 and U65417, respectively.

brains were sectioned at 14 μm thickness on a Reichert-Jung cryostat at −20°C and thaw-mounted onto microscope slides. Sections were fixed in freshly prepared 4% paraformaldehyde in 0.02% DEPC water for 20 min at 4°C in an ice bath and then washed for 5 min in cold phosphate-buffered saline, pH 7.4 before dehydration in an alcohol series, as described previously [11]. Fixed sections were stored at −70°C until use.

The PCR derived rat orthologue of GPR19 was labeled by random priming using [^{35}S]dCTP (NEN, Dupont). Rat brain sections were prehybridized for 2 h in buffer containing 50% deionized formamide, 0.6 M NaCl, 10 mM Tris-HCl pH 7.5, 10% dextran sulfate, 1% polyvinyl pyrrolidone, 2% SDS, 100 mM dithiothreitol, 200 μg/ml herring sperm DNA, and hybridized with the labeled probe (10^6 cpm/slice) for 16 h, and washed in conditions of increasing temperature and decreasing ionic strength. The hybridized sections were dehydrated in a graded alcohol series and were exposed to X-ray film (Dupont MRF-34) for 4 weeks at −70°C and developed. For use as controls, adjacent sections were hybridized following treatment with RNase, to confirm the specificity of hybridization.

2.5. FISH detection

Phage probes were biotinylated with dATP using the BRL Bionick labeling kit. For in situ hybridization and FISH detection, lymphocytes were cultured and synchronized as described [13]. The cells were harvested and slides were made. The procedure for FISH was performed as previously described [13].

2.6. Single stranded conformational polymorphism (SSCP) analysis

The SSCP method was used to screen GPR19 for polymorphic variants. The intronless gene was divided into five overlapping segments using overlapping primer pairs and the amplified gene fragments were obtained from the PCR. The reaction was conducted using 0.2 mM of each primer, 1.5 mM $MgCl_2$ and 2.5 units Taq polymerase. The PCR product was denatured for 5 min at 95°C and soaked for 5 min at 4°C prior to loading the SSCP gel. The denaturing running buffer consisted of 40% denaturing solution (0.12% SDS and 10 mM EDTA) and 60% stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol).

The SSCP was conducted using Novex 4–20% precast gels run in the Novex Xcell II mini-cell system. The gels were typically run at 50 volts for 14 h overnight and developed using the Novex Silver Xpress staining kit.

2.7. Radioligand binding assays

Using primers flanking the coding region of GPR19 we amplified by PCR this fragment and subcloned this DNA into the expression vector pcDNA3 (Invitrogen). The DNA construct was used to transiently transfect COS cells, and membranes from these cells were tested for their ability to bind a variety of ligands. Expression of GPR19 in COS cells was verified by Northern blot analysis.

3. Results and discussion

3.1. Cloning of the gene GPR19

Using a customized search procedure of a human EST data base, several cDNA sequences that partially encoded novel GPCRs were identified. One of these EST fragments obtained from the I.M.A.G.E Consortium (I.D. #45231; GenBank accession number H07970) is 1.4 kb long and on further sequencing it was determined to contain the entire coding sequence of a novel GPCR, with 35 base pairs in the 5′ untranslated region. To obtain the gene encoded by this EST fragment, a human genomic library was screened and 6 positive clones were obtained. One of these phage was purified and after restriction and Southern blot analysis, a 5-kb fragment was subcloned and sequenced. This genomic clone, which we have called GPR19, contained an intronless open reading frame of 1245 nucleotides (identical to the cDNA clone), encoding a 415 amino acid protein (Fig. 1). PCR analysis of the rat genome using two oligonucleotides specific for the coding region of the human gene, also identified a DNA fragment of the predicted size. The DNA fragment was subcloned and sequenced and demonstrated 85% nucleotide identity with human. This rat genomic clone of GPR19 was used to map the distribution of mRNA transcripts in the rat brain.

Hydrophobic analysis of the deduced amino acid receptor sequence encoded by GPR19, demonstrated the seven TM regions characteristic of all GPCRs. A search of the GenBank data base, followed by a manual alignment, revealed that overall the protein sequence had significant sequence similar-

ity both with the dopamine D2 receptor family, and the neuropeptide Y receptors. A comparison of the TM regions of the receptor sequence encoded by GPR19 with the D2 dopamine and neuropeptide Y receptor, revealed that there were 101 and 95 amino acids identical, respectively (Fig. 1). Furthermore a comparison with all three members of the D2 receptor family (D2, D3 and D4) revealed that overall 130 amino acids were identical. To a lesser extent similarity was also revealed, by the GenBank search with the cholecystokinin receptor and two other aminergic muscarinic and adrenergic receptors.

The novel receptor contains three putative glycosylation sites, consensus sequences for phosphorylation by protein kinase C in the intracellular loops, and a cysteine for potential modification by palmitoylation in the carboxy tail. The carboxy tail contains a stretch of 6 amino acids of which five are either serines or threonines and may be targets for phosphorylation by GPCR kinases. Significantly, an aspartic acid residue that forms part of the ligand binding site in the catecholaminergic receptors in TM3 was absent, indicating that the endogenous ligand is likely not aminergic.

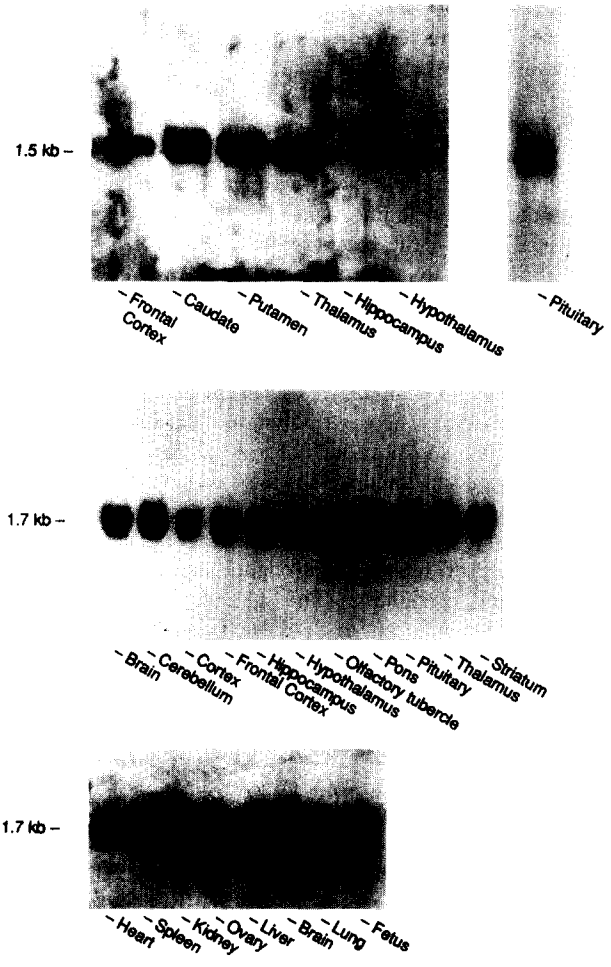

Fig. 2. Northern blot analysis of the tissue distribution of GPR19 mRNA in human brain (top panel), rat brain (middle panel), and rat peripheral tissues (lower panel). Each lane contained 5 μg of poly(A)$^+$ RNA isolated from various human and rat tissues. The human blot was probed with a 1.4 kb fragment isolated from the coding region of GPR19. The rat blots were probed with the rat orthologue of GPR19.

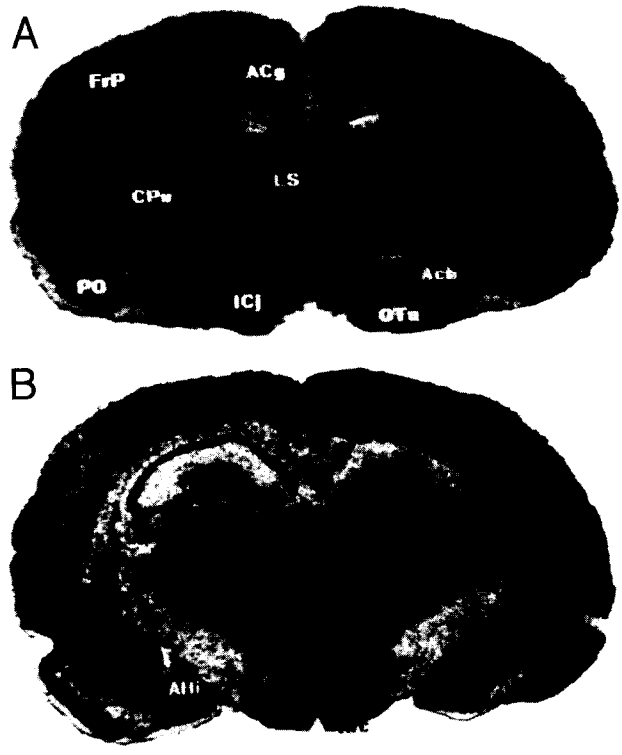

Fig. 3. Darkfield autoradiograms of coronal sections of rat brain showing the localization of GPR19 receptor mRNA. Shown are representative sections at levels relative to the bregma at −0.0 mm (A) and −3.3 mm (B), according to the coordinates [14]. Acb, nucleus accumbens; ACg, anterior cingulate cortex; AHi, amygdala; Arc, arcuate nucleus; CA1 and CA2, hippocampus regions; CPu, caudate putamen; DG, dentate gyrus; FrP, frontoparietal cortex; ICj, islands of Calleja; LS, lateral septum; MHb, medial habenula; OTu, olfactory tubercle; PO, piriform area; VMH, ventromedial nucleus of hypothalamus; ZI, zona incerta.

3.2. Expression of GPR19

Northern blot analysis studies using poly(A)$^+$ RNA isolated from multiple human and rat brain and peripheral regions revealed a single mRNA transcript for gene GPR19. In human brain the signal was strongest in the caudate nucleus and putamen, followed by thalamus. Fainter signals were noted in frontal cortex and hypothalamus with two transcripts observed in pituitary (Fig. 2). In rat tissues a single mRNA transcript was visualized in all regions examined. There was relatively greatest abundance of expression of GPR19 mRNA in brain (overlapping D2 receptor expression), with successively lesser amounts in kidney, heart, spleen, liver and ovary. Only a faint signal was observed in lung tissue, and a moderate level of expression was present in rat fetus. In the Northern blot analysis of rat brain regions the greatest abundance was in the pituitary and cerebellum, followed by olfactory tubercle and thalamus. Significant GPR19 mRNA expression was also present in striatum, hippocampus, frontal cortex, hypothalamus and medulla pons.

In rat brain sections, the distribution of GPR19 mRNA showed discrete localization in many areas. As shown in Fig. 3, mRNA was detected in several cortical regions such as the somatosensory areas, anterior limbic area, piriform area and the posterior limbic area including the retrosplenium. GPR19 mRNA was densely localised to the olfactory tubercle, islands of Calleja, hippocampal areas and den-

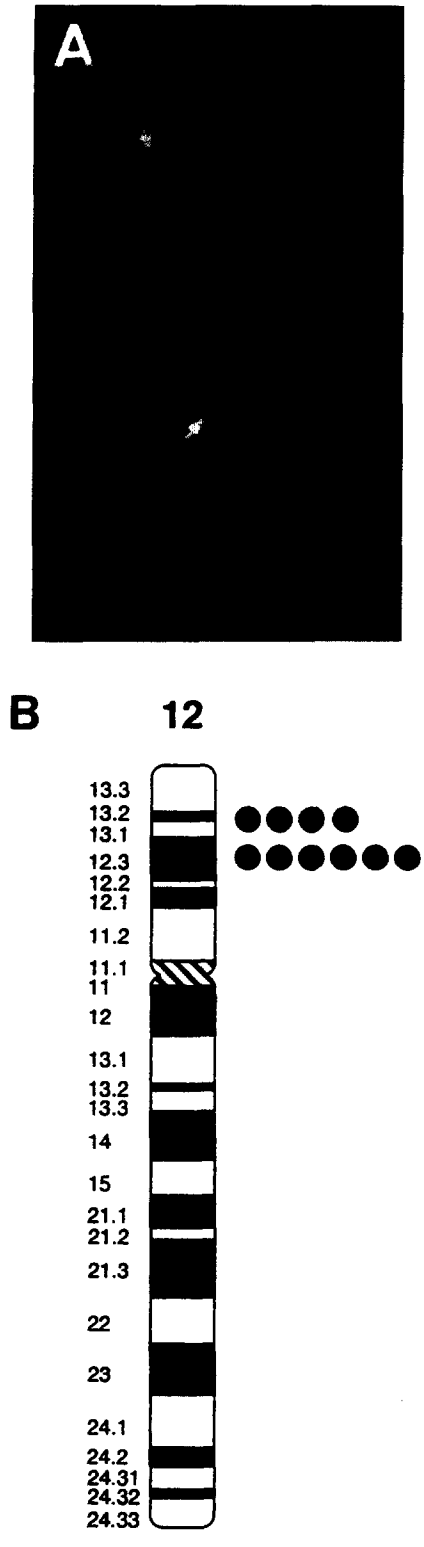

Fig. 4. FISH analysis of GPR19. A: Results of metaphase spread chromosomes probed with a phage clone encoding GPR19. Arrows point to the FISH signals on a pair of chromosomes. B: A summary of the FISH analysis, each dot represents the location of a fluorescent signal on the chromosome using phage GPR19 as aprobe.

tate gyrus. mRNA was also detected in several amygdaloid nuclei and hypothalamic nuclei, and lower levels were present in ventral tegmental area, substantia nigra compacta and central gray.

3.3. Chromosomal assignment of GPR19

Fluorescence in situ hybridization (FISH) of lymphocyte metaphase spread chromosomes was used to identify the chromosomal location of the gene GPR19. Under the conditions used, the hybridization efficiency of the phage probe containing GPR19 was 95%; among 100 checked mitotic figures, 96 showed signals on one pair of the chromosomes. To assign the signal to a specific chromosome the DAPI banding pattern was examined. These results indicate that GPR19 is located on chromosome 12 in region p13.2–12.3, the data are summarized in Fig. 4.

3.4. Pharmacological characterization of the receptor encoded by gene GPR19

Membranes from COS cells transiently transfected with GPR19 were tested for their ability to bind a variety of ligands. No saturable binding was detected with tritium-labeled haloperidol, spiperone, nemonapride, mesulergine, naloxone, butaclamol, clozapine, DTG, flupenthixol, or LSD. Membranes from COS cells transfected with GPR19 were also examined for activation of adenylyl cyclase activity in the presence and absence of forskolin, following the addition of a variety of ligands. No effects on adenylyl cyclase activity were seen with dopamine, serotonin, epinephrine, (−)sulpiride, HVA, 2-hydroxyestradiol, 4-hydroxyestrone and 4-hydroxy-3-methoxyphenyl glycol.

3.5. Polymorphic variants of GPR19

SSCP analysis was used to identify three polymorphic variants of GPR19 and each was confirmed by sequencing. The first resulted in an amino acid change of isoleucine 289 to valine in TM 4 of the receptor, as a result of a single base substitution. The frequency of this allele was 6.5% in Caucasians ($n = 200$), and occurred more commonly in blacks (15%; $n = 88$). A second polymorphism with a frequency of 1.5%, was located in TM 2 of the receptor and represented an amino acid change from valine 116 to isoleucine as a result of a single base change. The third polymorphism found in a single individual was a substitution (G for A) at base pair 1062. The first polymorphism noted by virtue of its significant incidence may facilitate linkage analysis of neuropsychiatric diseases.

In summary, we have identified a novel gene encoding a putative receptor with extensive and abundant distribution in brain and periphery. The abundant distribution of GPR19 mRNA in brain regions such as olfactory tubercle, islands of Calleja, caudate-putamen and pituitary are very similar to the expression pattern of the D2 receptor gene. Further work is necessary to identify the endogenous ligand which will likely have important neurotransmitter function.

Acknowledgements: We thank Dr. Philip Seeman for providing some of the COS cell binding data. This research was supported by grants from the National Institutes of Drug Abuse, the Medical Research Council of Canada, and the Smokeless Tobacco Research Council to B.F.O. and S.R.G.

References

[1] O'Dowd, B.F., Heiber, M., Chan, A., Heng, H.H.Q., Tsui, L.-C., Kennedy, J.L., Shi, X., Petronis, A., George, S.R. and Nguyen, T. (1993) Gene 136, 355–360.
[2] Marchese, A., Docherty, J.M., Nguyen, T., Heiber, M., Cheng, R., Heng, H.H.Q., Tsui, L.-C., Shi, X., George, S.R. and O'Dowd, B.F. (1994) Genomics 23, 609–618.

[3] Heiber, M., Docherty, J.M., Shah, G., Nguyen, T., Cheng, R., Heng, H.H.Q., Marchese, A., Tsui, L.-C., Shi, X., George, S.R. and O'Dowd, B.F. (1995) DNA Cell Biol. 14, 25–35.
[4] O'Dowd, B.F., Scheideler, M.A., Nguyen, T., Cheng, R., Rasmussen, J.S., Marchese, A., Zastawny, R., Heng, H.H.Q., Tsui, L.-C., Shi, X., Asa, S., Puy, L. and George, S.R. (1995) Genomics 28, 84–91.
[5] Marchese, A., Heiber, M., Nguyen, T., Heng, H.H.Q., Saldivia, V.R., Cheng, R., Murphy, P.M., Tsui, L.-C., Shi, X., Gregor, P., George, S.R., O'Dowd, B.F. and Docherty, J.M. (1995) Genomics 29, 335–344.
[6] Heiber, M., Marchese, A., Nguyen, T., Heng, H.H.Q., George, S.R. and O'Dowd, B.F. (1996) Genomics 32, 462–465.
[7] O'Dowd, B., Nguyen, T., Jung, B., Marchese, A., Cheng, R., Heng, H., Kolakowski, F.L., Lynch, K.R., and George, S. (1996) Gene, in press.
[8] Atschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[9] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
[10] Pearson, W.R. (1995) Prot. Sci. 4, 1145–1160.
[11] Zastawny, R.L., George, S.R., Nguyen T., Cheng, R., Tsatsos, J., Briones-Urbina, R. and O'Dowd, B.F. (1994) J. Neurochem. 62, 2099–2105.
[12] Chomczynski, P. and Sacchi, N. (1987). Anal. Biochem. 162, 156–159.
[13] Heng, H.H.Q. and Tsui, L.-C. (1994) in: Methods in Molecular Biology: In Situ Hybridization Protocols (Choo, K.H.A., Ed.), pp. 35–49, Humana Press, Clifton, NJ.
[14] Paxinos, G. and Watson, C. (1982) The Rat Brain. Academic Press, New York.