Vol. 207, No. 1, 1995 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
February 6, 1995 Pages 467-475

POLYMORPHIC TANDEM REPEATS IN DOPAMINE D4 RECEPTOR ARE SPREAD OVER PRIMATE SPECIES

Mitsuyuki Matsumoto*, Kazuyuki Hidaka, Shoko Tada, Yoshikazu Tasaki and Tokio Yamaguchi

Neuroscience & Gastrointestinal Research Lab.,
Institute for Drug Discovery Research,
Yamanouchi Pharmaceutical Co., Ltd.,
21 Miyukigaoka, Tsukuba, Ibaraki 305, Japan

Received December 27, 1994

The human dopamine D4 receptor has polymorphic tandem repeats in the third cytoplasmic loop. However, these repeats are not present in the rat counterpart. To determine whether the tandem repeats are specific to humans or not, we analyzed genomic DNA sequences for the D4 receptor of six primate species (human, chimpanzee, gorilla, orangutan, macaque, marmoset). Sequencing data revealed that all primates have the 48-bp tandem repeats in the D4 receptor gene. This finding suggests that these repeats originated before the separation of the New World monkey lineage from the Old World monkey and ape-human lineages.

Recent molecular cloning techniques have revealed five dopamine receptors (D1,D2,D3,D4,D5). The cDNA of the human dopamine D4 receptor has been isolated from neuroblastoma SK-N-MC cells (1). This receptor has high affinity for the antipsychotic clozapine, which does not produce extrapyramidal side effects and improves positive and negative symptoms of schizophrenia (1,2).

The structure of the human D4 receptor gene is quite unique. This gene has polymorphic tandem repeats in the third cytoplasmic loop region and the alleles vary not only in the number of repeats but also in the sequence of repeats (3,4). At present, the human D4 gene is the only example of the variable number of tandem repeats (VNTR) polymorphism in the coding region for catecholamine receptors, and

* To whom correspondence should be addressed.

0006-291X/95 $5.00

the number of repeats is reported to affect the ligand binding properties of the receptor (3,5). These observations have provoked many studies on the relationship between the number of repeats and psychotic illness, especially schizophrenia (6-11). On the other hand, the origin of these tandem repeats is not clear. The only fact is that the repeats in the human D4 receptor are not found in the rat counterpart (12). Assuming that the tandem repeats are connected with the function of the D4 receptor, the pharmacological and physiological nature of the human D4 receptor might be different from that of D4 receptor without the repeats. It is important for understanding of the physiological role of tandem repeats in the human D4 receptor to make clear whether they are specific to humans.

In the present study, we cloned and sequenced genomic DNAs for the D4 receptor of various primate species to determine whether the tandem repeats in the dopamine D4 receptor are specific to humans or spread over primate species.

MATERIALS AND METHODS

Genomic DNAs and Oligonucleotide Primers

Genomic DNAs of human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), orangutan (*Pongo pygmaeus*), macaque (*Macaca fasciculans*) and marmoset (*Leontopithecus saguinus*) were obtained from BIOS Laboratories. Oligonucleotide primers were synthesized on a model 394 (ABI) DNA synthesizer. The names and sequences of the primers were,
D4EVO1 (3020-3043, sense)
5'-AAAATCTAGATTCTTCCTACCCTGCCCGCTCATG-3'
D4EVO2 (3452-3475, antisense)
5'-AAAATCTAGAGACCACCACCGGCAGGACCCTCAT-3'
The underlined sequences show artificial XbaI sites. The position numberings refer to HUMD4G in the EMBL database.

Genomic DNA Cloning by Improved Polymerase Chain Reaction (PCR)

For cloning of the genomic DNAs for the third cytoplasmic loop of the D4 receptor from primates, formamide and *Pfu* DNA polymerase (Stratagene) in PCR amplification were used as previously described (13). The reaction sample contained 0.5 μg of genomic DNA of each primate, 10 pmol of the primers D4EVO1 and D4EVO2, 200 μM each of dNTP (dATP, dTTP, dGTP, dCTP), 20 mM Tris-HCl, pH 8.2, 10 mM KCl, 6 mM ammonium sulfate, 2 mM $MgCl_2$, 0.1% Triton X-100, 10 ng/μl BSA, 5% formamide, and 5 units of *Pfu* DNA polymerase in a final volume of 50 μl. The reaction mixture was amplified for 35 cycles in a Perkin Elmer Cetus thermal cycler model 9600. The amplification profile consisted of denaturation at 98 °C for 1 minute, primer annealing at 65 °C for 1 minute, and extension at 74 °C for 4 minutes. The amplified products were digested with restriction endonuclease XbaI and ligated into XbaI site of Bluescript phagemid (Stratagene).

Sequencing analysis was performed on both strands using 373A Autosequencer (ABI) by the dideoxynucleotide chain-termination method. For confirmation of the amplified genomic DNA sequence, two or three clones from independent PCR amplifications were sequenced for each primate.

RESULTS

Because the usual PCR method using *Taq* DNA polymerase without denaturant was unsuccessful in the amplification of the D4 gene, which has an extraordinarily high GC content, we cloned genomic DNAs for the D4 receptor of various primate species by an improved PCR method using formamide and *Pfu* DNA polymerase as previously described (13).

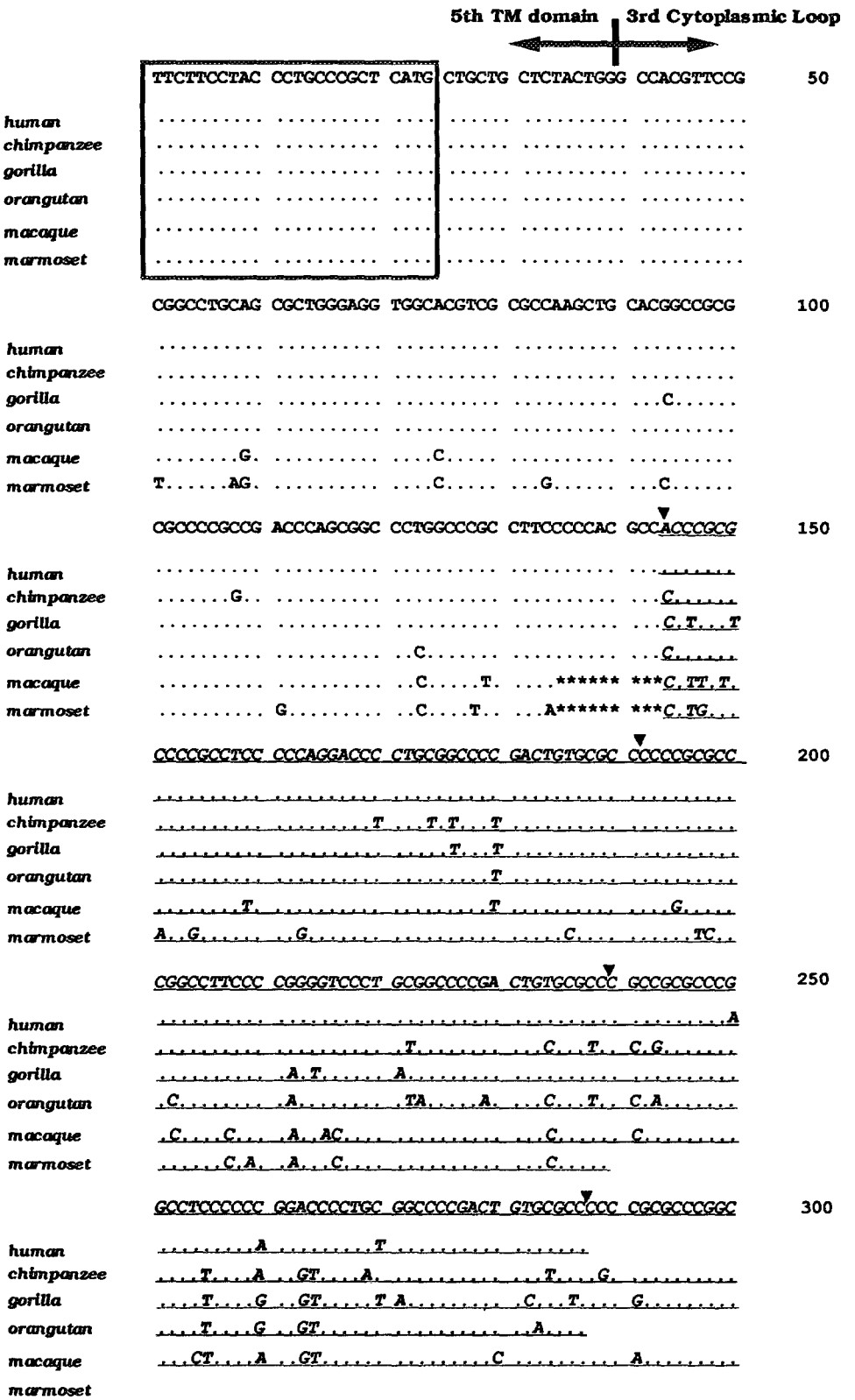
The primers for PCR amplification were selected from the human D4 transmembrane domains which were completely conserved at amino acid level between the human and rat D4 receptors. With the primers D4EVO1 on the fifth transmembrane domain and D4EVO2 on the sixth transmembrane domain, genomic DNAs for the third cytoplasmic loop of the D4 receptor were amplified from all primates tested. The nucleotide sequences of the amplified fragments were determined after being subcloned into the sequencing vector. The nucleotide sequences and deduced amino acid sequences are shown in figure 1 and figure 2, respectively. The sequencing data confirmed the existence of tandem repeats in the third cytoplasmic loop of the D4 receptor of human (four repeat units), chimpanzee (five repeat units), gorilla (five repeat units), orangutan (four repeat units), macaque (five repeat units) and marmoset (three repeat units). As the number of tandem repeat units differed among these primates, the repeat units were aligned in both figures in

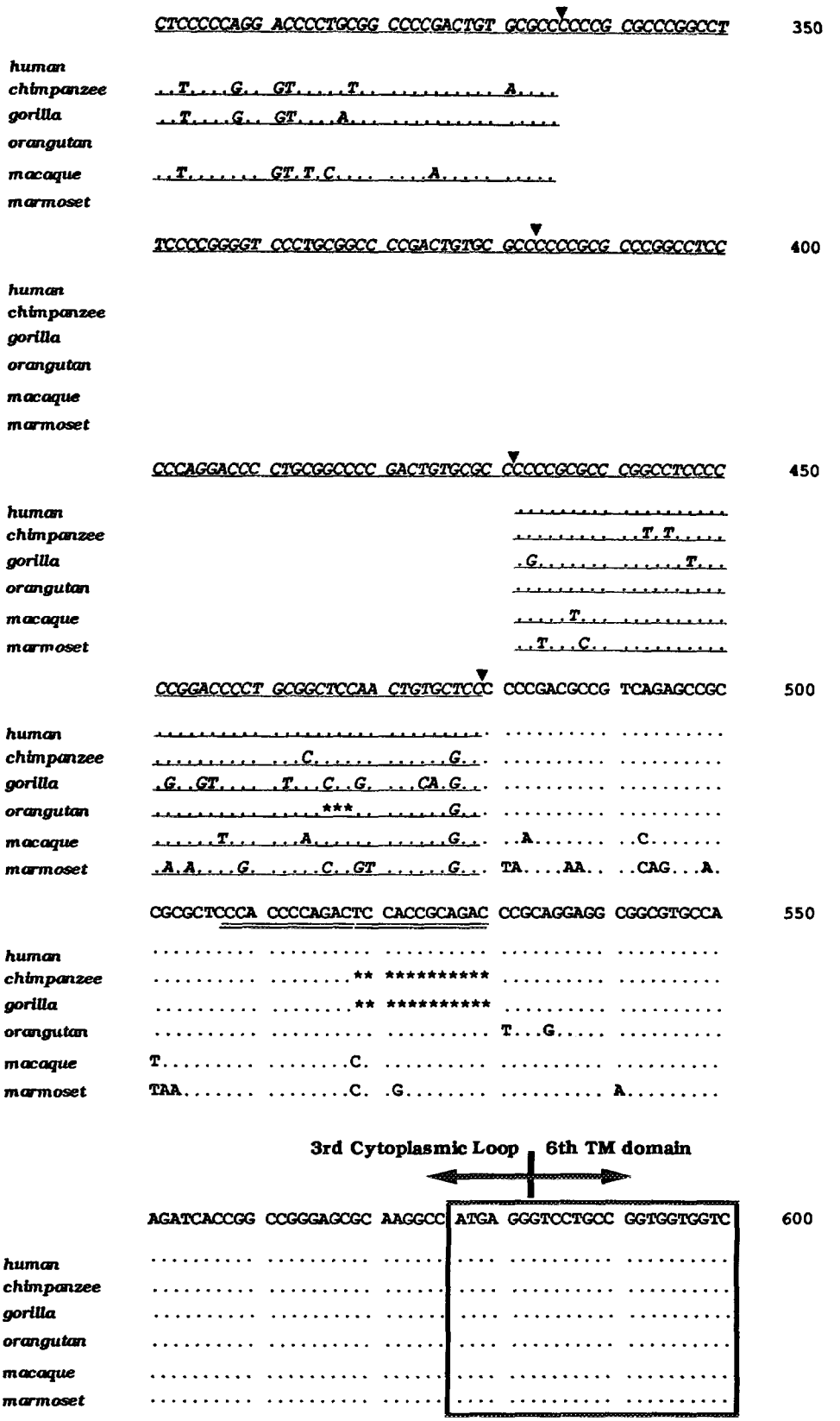
---

Figure 1.

Nucleotide sequences for the third cytoplasmic loop of the dopamine D4 receptor of human, chimpanzee, gorilla, orangutan, macaque and marmoset.

The nucleotide sequence at the top is quoted from the published human D4 receptor cDNA sequence with seven repeat units, HUMD4C, in the Genbank data base. A period (.) denotes an identical nucleotide with the top sequence, and an asterisk (*) denotes a gap in the sequence. Sequences of the tandem repeats are underlined, and the ends of each repeat unit are marked with arrowheads (▼). Double underlines show the duplicated sequence described in "Results". Sequences of the primers for the PCR amplification are boxed. Borders between the transmembrane (TM) domain and the cytoplasmic loop region are indicated in the figure. The sequence of the last repeat of all primates is aligned with that of HUMD4C at the top.


5th TM domain
3rd Cytoplasmic Loop
TTCTTCCTAC CCTGCCCGCT CATG CTGCTG CTCTACTGGG CCACGTTCCG 50
human
chimpanzee
gorilla
orangutan
macaque
marmoset
CGGCCTGCAG CGCTGGGAGG TGGCACGTCG CGCCAAGCTG CACGGCCGCG 100
CGCCCCGCCG ACCCAGCGGC CCTGGCCCGC CTTCCCCCAC GCCACCCGCG 150
CCCCGCCTCC CCCAGGACCC CTGCGGCCCC GACTGTGCGC CCCCCGCGCC 200
CGGCCTTCCC CGGGGTCCCT GCGGCCCCGA CTGTGCGCCC GCCGCGCCCG 250
GCCTCCCCCC GGACCCCTGC GGCCCCGACT GTGCGCCCCC CGCGCCCGGC 300


CTCCCCCAGG ACCCCTGCGG CCCCGACTGT GCGCCCCCCG CGCCCGGCCT 350
human
chimpanzee ..T...G.. GT.....T. ..... A....
gorilla ..T...G.. GT....A. ....
orangutan
macaque ..T..... GT.T.C.... ....A..... .....
marmoset
TCCCCGGGGT CCCTGCGGCC CCGACTGTGC GCCCCCCGCG CCCGGCCTCC 400
human
chimpanzee
gorilla
orangutan
macaque
marmoset
CCCAGGACCC CTGCGGCCCC GACTGTGCGC CCCCCGCGCC CGGCCTCCCC 450
human ....
chimpanzee ....T.T.....
gorilla .G..... .....T...
orangutan ....
macaque .....T. .....
marmoset ..T...C. .....
CCGGACCCCT GCGGCTCCAA CTGTGCTCCC CCCGACGCCG TCAGAGCCGC 500
human ....
chimpanzee .....C..... .....G.. ....
gorilla .G..GT.... .T...C..G. ...CA.G.. ....
orangutan .....***.. .....G.. ....
macaque .....T... ...A..... .....G.. ..A..... ..C.....
marmoset .A.A....G. .....C..GT .....G.. TA....AA.. ..CAG...A.
CGCGCTCCCA CCCCAGACTC CACCGCAGAC CCGCAGGAGG CGGCGTGCCA 550
human ....
chimpanzee ....** ********** ....
gorilla ....** ********** ....
orangutan .... T...G..... .....
macaque T..... .....C. ....
marmoset TAA..... .....C. .G..... ..... A.....
3rd Cytoplasmic Loop
6th TM domain
AGATCACCGG CCGGGAGCGC AAGGCC ATGA GGGTCCTGCC GGTGGTGGTC 600
human ....
chimpanzee ....
gorilla ....
orangutan ....
macaque ....
marmoset ....

order from the first unit, but the last unit of all primates was aligned with that of the published human D4 receptor shown at the top. The tandem repeats of each primate consisted of 48-bp units encoding 16 amino acids. The repeat units in each species were not completely identical in nucleotide sequence or amino acid sequence, which resulted in imperfect tandem repeats in all primates. Comparison of the sequence of humans with that of the other primates showed several short in-frame deletions in the flanking region of the tandem repeats and within the repeat, as shown by asterisks in figures 1 and 2. An identical 9-bp deletion upstream of the tandem repeats was found in macaque and marmoset, an identical 12-bp deletion downstream of the repeats in chimpanzee and gorilla, and a 3-bp deletion within the last repeat of orangutan. Interestingly, the 12-bp deletion in chimpanzee and gorilla occurred on duplicated 12-bp nucleotides, which resulted in duplicated four amino acids in other species (double underlined in figures 1 and 2).

DISCUSSION

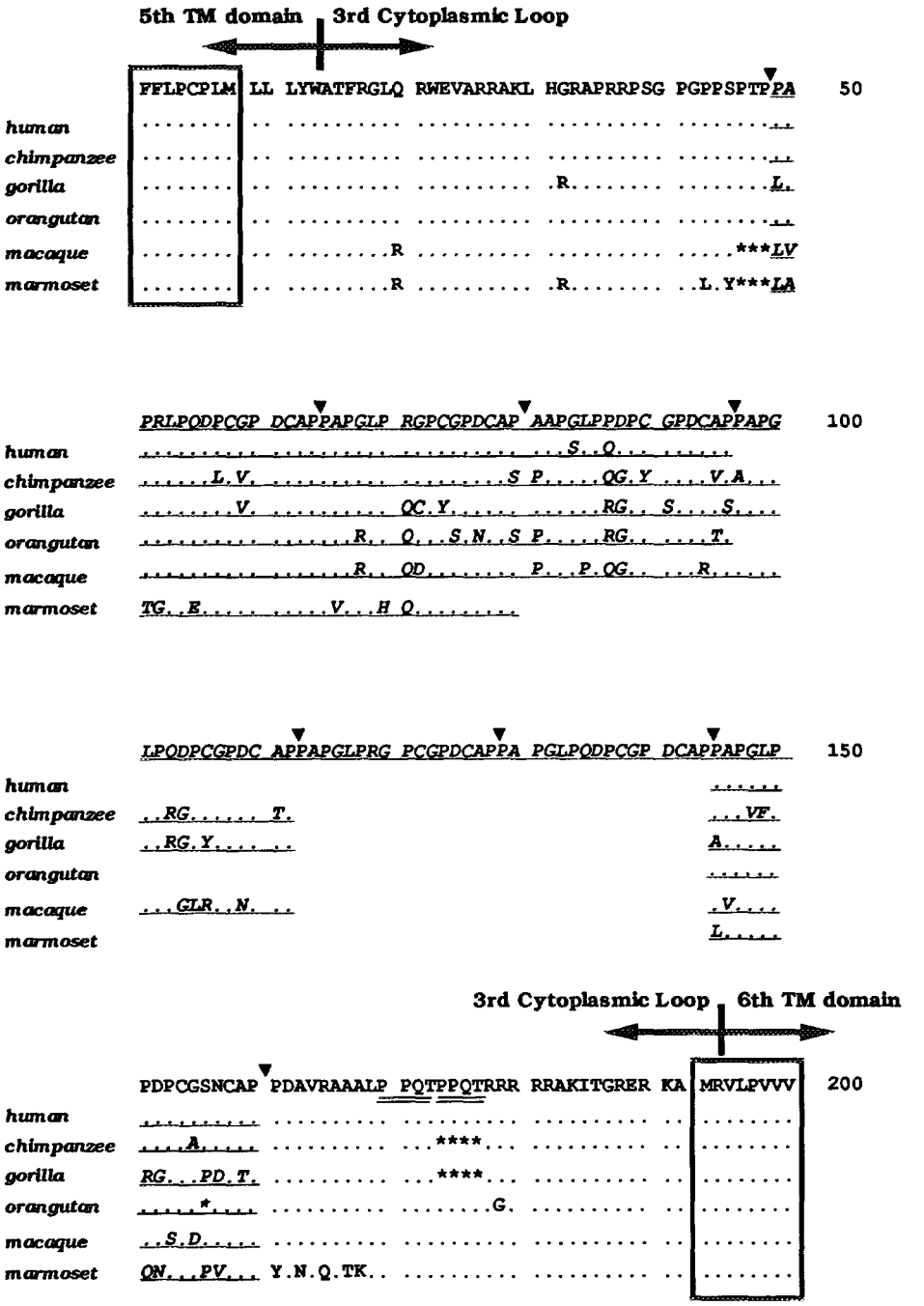
In this study, we demonstrated that the tandem repeats in the dopamine D4 receptor are not specific to humans and that these repeats are spread over various primate species. The existence of the repeats in all primates tested, especially in marmoset, a species of New World monkey, suggests that the repeats originated before the separation of the New World monkey lineage from the Old World monkey and ape-human lineages. As we were unable to obtain a large number of individual genomic DNA from the primates, it is not clear

---

Figure 2.

Deduced amino acid sequences for the third cytoplasmic loop of the dopamine D4 receptor of human, chimpanzee, gorilla, orangutan, macaque and marmoset.

The amino acid sequence at the top is deduced from the published *human D4 receptor cDNA sequence* with seven repeat units, HUMD4C, in the Genbank data base. A period (.) denotes an identical amino acid with the top sequence, and an asterisk (*) denotes a gap in the sequence. Sequences of the tandem repeats are underlined, and the ends of each repeat unit are marked with arrowheads (▼). Double underlines show the duplicated sequence described in "Results". Amino acid sequences corresponding to the primers for the PCR amplification are boxed. Borders between the transmembrane (TM) domain and the cytoplasmic loop region are indicated in the figure. The sequence of the last repeat of all primates is aligned with that of HUMD4C at the top.



whether the tandem repeats of primates show the hypervariable polymorphism observed in the human population (4,5). However, considering the high rate of mutation on the number of tandem repeat units by deletion and insertion through processes such as unequal

crossover or replication slippage (14), the presence of such polymorphic variation is suspected. Our sequence data will be useful in investigations on possible variations in the tandem repeats of the primates.

Our data show not only an alternation of the number of tandem repeat units but also that a small deletion and single base alternation occurred on the third cytoplasmic loop of the D4 receptor during the evolution of primates. Although the variability of the repeat unit itself in each primate makes it difficult to compare one sequence of tandem repeat region with another, we note that the rate of single base alternation in tandem repeat regions is quite high compared to the flanking regions among primates.

Concerning VNTR polymorphism of the human D4 receptor, linkage and association studies between the number of repeats and schizophrenia have been negative (7-11). Recently, Asghari et al. reported that the number of repeats affected ligand binding properties, but this effect was relatively small (5). Although these data suggest the possibility that the tandem repeat region is not essential for the function of the receptor, the third cytoplasmic loop region of G protein coupled receptors has been shown to be involved in signal transduction. Further studies are needed to determine the effect of the repeats upon physiological function of the D4 receptor. Our present data showing the presence of tandem repeats in primates will be useful in clarifying the role and origin of tandem repeats of the D4 receptor.

Note: During the preparation of this manuscript, the polymorphic nature of the D4 receptor gene of nonhuman primates was reported in abstract form by Adamson M.D. et al. Soc. Neurosci. Abstr. 20:20.

REFERENCES

1. Van Tol, H.H.M., Bunzow, J.R. Guan, H.C., Sunahara, R.K., Seeman, P., Niznik, H.B., and Civelli, O. (1991) Nature 350, 610-614
2. Lahti, R.A., Evans, D.L., Stratman, N.C., and Figur, L.M. (1993) Eur. J. Pharmacol. 236, 483-486
3. Van Tol, H.H.M., Wu, C.M., Guan, H.C., Ohara, K., Bunzow, J.R., Civelli, O., Kennedy, J., Seeman, P., Niznik, H.B., and Jovanovic, V. (1992) Nature 358, 149-152
4. Lichter, J.B., Barr, C.L., Kennedy, J.L., Van Tol, H.H.M., Kidd, K.K., and Livak, K.J. (1993) Hum. Mol. Genet. 2, 767-773
5. Asghari, V., Schoots, O., Van Kats, S., Ohara, K., Jovanovic, V., Guan, H., Bunzou, J.R., Petronis, A., and Van Tol, H.H.M. (1994) Mol. Pharmacol. 46, 364-373

6. George, S.R., Cheng, R., Nguyen, T., Israel, Y., and O'Dowd B.F. (1993) Biochem. Biophys. Res. Commun. 196, 107-114
7. Sommer, S.S., Lind, T.J., Heston, L.L., and Sobell, J.L. (1993) Am. J. Med. Genet. 48, 90-93
8. Shaikh, S., Collier, D., Kerwin, R.W., Pilowsky, L.S., Gill, M., Xu, W., and Thornton, A. (1993) Lancet 341, 116
9. Nanko, S., Hattori, M., Ikeda, K., Sasaki, T., Kazamatsuri, H., and Kuwata, S. (1993) Lancet 341, 689-690
10. Macciardi, F., Smeraldi, E., Marino, C., Cavallini, M.C., Petronis, A., Van Tol, H.H.M., and Kennedy, J.L. (1993) Arch. Gen. Psychiatry 51, 288-293
11. Barr, C.L., Kennedy, J.L., Lichter, J.B., Van Tol, H.H.M., Wetterberg, L., Livak, K., and Kidd, K.K. (1993) Am. J. Med. Genet. 48, 218-222
12. O'Malley, K.L., Harmon, S., Tang, L., and Todd, R.D. (1992) New Biol. 4, 137-146
13. Matsumoto, M., Hidaka, K., Tada, S., Tasaki, Y., and Yamaguchi, T. Mol. Brain Res. in press
14. Jeffreys, A.J., Royle, N.J., Wilson, V., and Wong, Z. (1988) Nature 332, 278-281