Echoes of D4 receptor repeats

SIR — Van Tol *et al.* have reported polymorphic variation of the dopamine D4 receptor in the human population[1]. This variation is due to the presence in the gene of different numbers of a 48-base-pair repeat, which results in the expression of receptor molecules containing between two and seven copies of a 16-amino-acid motif in the third intracellular loop. To date, such an arrangement appears to be unique for proteins within the seven transmembrane domain family; its significance is not yet clear.

Human D4 repeat consensus 744–791, 792–839, etc.	A G A T A AT T A T CC GCG CCC CGC CT CCC CCG G CCC TGC GGC CC AC TGT GC CC C C G C G GC C G G	Homology to repeat consensus (%)
Human D4 696–743	C CgC GCG CCC CGC Cga CCC aGc GGC CCt gGC ccg CCt tcC ccc aCG CC	65
Rat D4 675–722	C CgC GCG CCg CGC Cga CCC aGc GGC CCg gGC ccg **CCg Gtg Tcg Gac CC**	63
Rat D4 708–755	**g CCg GtG tCg GaC CcT** aCt CAG GGT CCC ctC ttC TCa GAt TGT cCG CC	67
Rat D4 756–803	t CCC tCa CCC AGC CTC Cgg aCG aGC CCC acC GtC TCC AgC a**Ga cCa ga**	65
Rat D4 798–845	**A CCa Ga**G tCa GaC CTC tCt CAG aGC CCC TGC aGC CCC Ggg TGT ctG Ct	71

Alignment of the consensus sequence for the repeats in the human dopamine D4 receptor gene and similar regions in the rat D4 gene and elsewhere in the human D4 gene. Nucleotides which are the same as in the consensus sequence are shown in upper case. The numbers refer to the co-ordinates given for the human and rat sequences[2,3]. Nucleotides 708–722 and 798–803 are in the two overlap regions and are shown in bold. The alignment used the program Multalin[4] (Cherwell Scientific), which searched the complete rat D4 sequence for homologies with the human repeats. The four regions with the greatest degree of similarity were found to be immediately adjacent to each other. Optimal alignment between the human and rat sequences showed that the rat region 675–722 is very similar to human region 696–743, immediately upstream of the repeats.

Van Tol *et al.* stated that the rat homologue of the human D4 receptor does not contain this repeat region. This is not strictly true. At the DNA level, using the published rat D4 sequence of O'Malley *et al.*[2], I have identified four adjacent and partly overlapping regions which are 63–71% similar to the 48-bp repeat sequence of the human gene (see figure). All four regions are in the part of the rat sequence encoding the third intracellular loop and in a position analogous to that of the repeats in the human sequence. At the amino-acid level, the similarity is 38–50%, suggestive of poor functional conservation.

Despite their high DNA similarity to the human repeats, these regions in the rat sequence are not repeats themselves. Their levels of similarity are too low, varying from 33 to 60% for each pair. This compares with 90–96% among the human repeats. In addition to its close relationship to the human repeats, the 5′ 36-bp of the first of the rat sequences shares 95% DNA identity (100% amino-acid identity) with a region in the human sequence[3] immediately upstream of the repeats (696–743), and which appears to be a minor repeat.

These sequence relationships strongly suggest that the origin of the repeats in the human sequence predates the divergence of mammalian species as different as rat and man. In view of the presence in the rat of non-repeat sequences homologous to the human repeats, it seems most unlikely that the repeats are a more recent insertion into the human genome. The repeats therefore seem to have been conserved in the human population for a very long time, albeit with different numbers of copies. By contrast, the repetitive nature of the sequences has been lost in the rat. This argues for an important role for this part of the third intracellular loop of the human dopamine D4 receptor.

A. J. Makoff
Department of Cell Biology,
Wellcome Foundation,
Beckenham, Kent BR3 3BS, UK

LICHTER *ET AL.* REPLY — Makoff compares the polymorphic DNA repeat sequences in the human dopamine D4 receptor gene (DRD4)[1,3] with a similar region in the rat DRD4 gene[2] and argues for an evolutionary origin of the repeat sequence that predates the divergence of rats and humans. We believe that these conclusions cannot yet be substantiated, and now have additional data that bear on the issue.

We find a tremendous amount of variation in the nucleotide sequence of the repeat region, above and beyond the variation we described[1]. Sequence analysis of DNA from 87 unrelated individuals shows that there are at least three times more haplotypes than the seven forms described previously[3]. Despite the variation among repeats in humans, we find that the first and last repeats are invariant, which suggests that they have a functional role, particularly if one considers that we have not found an individual with less than two repeats. In a related study (as yet unpublished) of several species of non-human primates, we confirm the presence of a repeat sequence with similarities at around 90% to the human repeat sequence.

Makoff's interpretation is not the only plausible explanation for the origin of the repeat sequence. (1) It could be argued that the repeat sequence found in humans became inserted (possibly through homologous recombination) after rat and human divergence. In this respect it is important that in an optimal alignment of the entire rat sequence with human D4 only the 36 base pairs immediately upstream to the human repeat sequence show good similarity with the rat sequence (100% at the amino-acid level), in contrast to the successive sequences (approximately 25%). (2) A primordial core sequence, made by a tandem duplication of what we call the flanking region, could have occurred in an ancestral mammal. On divergence of mammalian lineages the rat lineage may have lost this primordial sequence, while it was maintained in the human lineage and eventually elaborated to its current form.

Finally, it remains surprising that the rat has only a mere 63–71% similarity with the human D4 repeat consensus sequence, as given by Makoff — especially if one considers that outside this region there is much higher similarity (approximately 90%) between the two species. As a comparison, the human and rat D2 receptors display approximately 90% nucleic acid sequence similarity over their entire coding regions[5]. For the identification of the genetic mechanism responsible for the generation of the repeat sequence and/or divergence of the rat sequence, data from other species are needed.

Jay B. Lichter
Kenneth J. Livak
Du Pont Merck Pharmaceutical Co.,
PO Box 80400,
Wilmington, Delaware 19880-0400, USA
Jeffery Rogers
Southwest Foundation for Biomedical Research,
Department of Genetics,
PO Box 28147,
San Antonio, Texas 78228, USA
James L. Kennedy
Hubert H. M. Van Tol
Clarke Institute of Psychiatry & University of Toronto,
250 College Street,
Toronto, Ontario M5T 1R8, Canada

1. Van Tol, H. H. M. *et al. Nature* **358**, 149–152 (1992).
2. O'Malley, K. L. *et al. New Biol.* **4**, 137–146 (1992).
3. Van Tol, H. H. M. *et al. Nature* **350**, 610–614 (1991).
4. Corpet, F. *Nucleic Acids Res.* **16**, 10881 (1988).
5. Grandy, D. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 9762–9766 (1989).