Identification and Characterization of a Tandem Repeat in Exon III of the Dopamine Receptor D4 (DRD4) Gene in Cetaceans

Line Mogensen, Carl Christian Kinze, Thomas Werge, and Henrik Berg Rasmussen

From the Research Institute of Biological Psychiatry, H:S Sct. Hans Hospital, Copenhagen University, 2 Boserupvej, DK-4000 Roskilde, Denmark; and CC Konsult, 35 Falkoner Allé, DK-2000 Frederiksberg, Denmark.

Address correspondence to H. B. Rasmussen at the address above, or e-mail: henrik.berg.rasmussen@shh.hosp.dk.

A large number of mammalian species harbor a tandem repeat in exon III of the gene encoding dopamine receptor D4 (DRD4), a receptor associated with cognitive functions. In this study, a DRD4 gene exon III tandem repeat from the order Cetacea was identified and characterized. Included in our study were samples from 10 white-beaked dolphins (Lagenorhynchus albicrissus), 10 harbor porpoises (Phocoena phocoena), eight sperm whales (Physeter macrocephalus), and five minke whales (Balaenoptera acutorostrata). Using enzymatic amplification followed by sequencing of amplified fragments, a tandem repeat composed of 18-bp basic units was detected in all of these species. The tandem repeats in white-beaked dolphin and harbor porpoise were both monomorphic and consisted of 11 and 12 basic units, respectively. In contrast, the sperm whale harbored a polymorphic tandem repeat with size variants composed of three, four, and five basic units. Also the tandem repeat in minke whale was polymorphic; size variants composed of 6 or 11 basic units were found in this species. The consensus sequences of the basic units were identical in the closely related white-beaked dolphin and harbor porpoise, and these sequences differed by a maximum of two changes when compared to the remaining species. There was a high degree of similarity between the cetacean basic unit consensus sequences and those from members of the horse family and domestic cow, which also harbor a tandem repeat composed of 18-bp basic units in exon III of their DRD4 gene. Consequently, the 18-bp tandem repeat appears to have originated prior to the differentiation of hoofed mammals into odd-toed and even-toed ungulates. The composition of the tandem repeat in cetaceans differed markedly from that in primates, which is composed of 48-bp repeat basic units.

Introduction

Dopamine is a neurotransmitter implicated in the control of intellectual skills, motivational behavior, language, and motor activity (Nicoullon and Coquerel 2003). The importance of this transmitter in terms of cognition and conceptual skills has nourished the idea that development of human intelligence has been dependent on the expansion of the dopamine system (Previc 1999).

The dopamine receptor D4 (DRD4) is expressed at high levels in the prefrontal cortex, that is, a region of the brain associated with cognitive abilities (Oak et al. 2000; Tarazi and Baldessarini 1999). Observations with DRD4 knockout mice have associated this receptor with the trait termed novelty seeking (Dulawa et al. 1999). More recently, DRD4 has been implicated in working memory in rats (Zhang et al. 2004).

Exon III of the DRD4 gene harbors a tandem repeat in several mammalian species. This tandem repeat is composed of 48-bp basic units in humans (Lichter et al. 1993; Van Tol et al. 1992), nonhuman primates (Livak et al. 1995; Matsumoto et al. 1995), and prosimians (Inoue-Murayama et al. 1998). A DRD4 gene exon III tandem repeat consisting of 18-bp or 36-bp basic units has been identified in domestic cow (Larsen et al., 2005) and in members of the horse family (Hasegawa et al. 2002). The repeat structure is complex in the dog family (Canidae) because it consists of basic units of different sizes, namely 39, 27, and 12 bp (Niimi et al. 2001). Recently, tandem repeats composed of 18-bp basic units were identified in exon III of the DRD4 gene in other members of the mammalian order Carnivora, including domestic cat (Felis catus), polar bear (Ursus maritimus), Asiatic bear (Ursus thibetanus), and common raccoon (Procyon lotor), implying that this repeat size is not restricted to hoofed animals (Larsen et al., 2005). A repeat structure has not been identified in the DRD4 gene of mouse (Fishburn et al. 1995).

It is possible that there are functional implications of size variation of the tandem repeat in exon III in the DRD4 gene. Recent observations suggest that the length of the polymorphic repeat modulates the level of expression of the human...
DNA was extracted from the skin samples using the DNeasy kit (Qiagen GmbH, Hilden, Germany). Skin samples were obtained from 10 white-beaked dolphins (Lagenorhynchus albirostris), 10 harbor porpoises (Phocoena phocaena), eight sperm whales (Physeter macrocephalus), and five minke whales (Balaenoptera acutorostrata). Two of the harbor porpoises originated from Greek waters. All other animals had been found dead along the coasts of Denmark. Some of these animals came ashore in small groups. Genomic DNA was extracted from the skin samples using the DNeasy kit (Qiagen GmbH, Hilden, Germany).

Primers [5’-CTACTCGTCCGTGTGCTCCT-3’ (forward)] and [5’-ACCACAGGGACTCTCA-3’ (reverse)] for amplification of a fragment of exon III of the cetacean DRD4 gene were designed on the basis of a sequence from white-beaked dolphin deposited under accession number DQ132799 in GenBank. Amplification mixtures contained 60 mM Tris- SO₄ (pH 9.1), 18 mM (NH₄)₂SO₄, 400 nM of both upstream and downstream primers, 200 µM of each deoxynucleoside triphosphate, 1.5 mM MgCl₂, 5% dimethyl sulphoxide, 1 unit Elongase Enzyme Mix (Invitrogen®, Carlsbad, CA, USA), and 50 ng genomic DNA. Initially, the amplification mixtures were heated to 94°C for 3 min. Then followed 35 cycles, each consisting of 30 s at 94°C, 45 s at 55°C, and 45 s at 68°C. After the last cycle, elongation was allowed to proceed for 10 min at 72°C.

Amplified fragments were subjected to agarose gel electrophoresis. Fragments of interest were excised from the gel, purified using the Geneclean® III Kit (Qiogene, Inc., Carlsbad, CA, USA), and sequenced on both strands. Some of the amplified samples produced two fragments of nearly the same size. In these cases, excision from the gel was omitted to avoid cross-contamination of one of the polymerase chain reaction fragments with the other. Instead, the amplified fragments were cloned by insertion into the pDrive Cloning Vector (Qiagen GmbH, Hilden, Germany) followed by transformation using NovaBlue Singles™ Competent cells (Merck Biosciences, Darmstadt, Germany). Plasmids were isolated from the bacterial clones, and their inserts were sequenced.

Sequences were analyzed by the TANDEM REPEAT FINDER (TRF), a program that identifies tandem repeats (Benson 1999). The weight for match is +2 in this program, while three different weights can be chosen for mismatch and indels, namely −3, −5, and −7. In this study, the alignment parameters +2, −3, and −7 (match, mismatch, and gap) were used. Lower numbers permit the detection of tandem copies with more mismatches or indels; higher weight numbers increase the search stringency. If successful in identifying a tandem repeat, TRF reports its compositional characteristics, including the size, number of basic units, consensus sequence of the basic units, percentage of matches between adjacent basic units (a measure of their degree of identity), and percent GC. Often the number of basic units reported by TRF is not a whole number. This may reflect an accumulation of nucleotide substitutions in ends of the tandem repeat, resulting in a partial disappearance of the repeat motif in these portions of the tandem repeat. For example, a tandem repeat, which is reported to be composed of 18-bp basic units in an array of 4.7 copies, has four full-size 18-bp basic units followed by a partial basic unit of 13 bp (13/18 ~ 0.7). The consensus sequence of the basic units in a tandem repeat is determined by the majority rule and is the sequence, which produces the best alignment. All full-length and partial basic units are utilized in this alignment.

For comparative purposes, sequences of DRD4 exon III from domestic cow (Bos taurus), wild horse (Equus przewalskii), wild ass (Equus hemionus), and mountain zebra (Equus zebra) with the GenBank accession numbers AB086028, AB086029, AB080628, AB080631, and AB080635, respectively, were also analyzed by TRF. Although these sequences previously have been subjected to analysis by TRF (Larsen et al., 2005), they have not been analyzed using the stringency applied in the present study. A consensus sequence of the basic units of DRD4 exon III tandem repeats from various primate species (Livak et al. 1995) was included for comparative purposes.

The start of a tandem repeat as determined by TRF may vary with several nucleotides between different species. This complicates a comparison of the consensus sequences. To compensate for this, consensus sequences from different tandem repeats were subjected to cyclic alignment, which allows every position in one of two sequences to be the first.

Results

Fragments of DRD4 exon III from four different cetacean species were amplified and analyzed. The approximate sizes of the amplified fragments from white-beaked dolphin and harbor porpoise were 480 and 500 bp, respectively. Analysis of amplification products from the sperm whales revealed three different fragment sizes ranging from about 320 to 370 bp. Fragment sizes of about 390 and 490 bp were detected on analysis of amplified samples from the minke whales.

The sequences of the amplified fragments were determined and deposited in GenBank under the accession
Analysis of these sequences revealed tandem repeats composed of 18-bp basic units in all four cetacean species. Important characteristics of these tandem repeats are shown (Table 1). Suggestions for basic unit sizes of 36 or 51 bp were also provided by TRF, but these alternatives achieved inferior scores and were not considered further.

The number of 18-bp basic units was 11.3 in all 20 of the white-beaked dolphin tandem repeat sequences examined. All these repeat sequences were completely identical (only one unique allele was found), implying that the white-beaked dolphin tandem repeat was monomorphic. The harbor porpoise tandem repeat was composed of 12.3 of 18-bp basic units. Also this tandem repeat was monomorphic. In contrast, the sperm whale tandem repeat was polymorphic with respect to the number of 18-bp basic units; three size variants composed of 3.1, 4.1, and 5.1 basic units were found in this species. On analysis of the minke whale sequences, two size variants with 6.6 and 11.6 basic units were identified.

All tandem repeats were imperfect, that is, the sequences of the 18-bp basic units differed within each of the tandem repeats (Figure 1). Further examination revealed that none of the 18-bp basic units were present in a number higher than two in any of the tandem repeats. The degree of identity between adjacent copies of the 18-bp basic units ranged from 69% in white-beaked dolphin to 91% in sperm whale. Comparison of the two tandem repeat variants from minke whale revealed that the short variant differed from its longer counterpart by the lack of the five 3\# terminal basic units. The relationship between the alleles of the sperm whale was more complex as the 3\# terminal basic unit in the two shortest alleles was absent in the long allele.

Several of the 18-bp basic units were present in two or more cetacean species, including three consecutive 18-bp basic units in white-beaked dolphin and harbor porpoise. Overall, the tandem repeats from these two closely related species had seven 18-bp basic unit sequences in common. Comparison of the 18-bp basic unit consensus sequences from the four different cetacean species revealed that three from harbor porpoise and white-beaked dolphin were identical. These consensus sequences differed in those of minke whale and sperm whale at one and two nucleotide positions, respectively (Table 1).

Comparison of the 18-bp basic unit consensus sequences from domestic cow and members of the horse family (Table 1) revealed differences at three or four nucleotide positions. In contrast, the 18-bp basic units were almost identical in the four cetaceans examined. The number of basic units was lower in the tandem repeats from sperm whale than in those from the cetacean species. The GC content of the various tandem repeats ranged from 0.74 to 0.89. The highest value was found in the tandem repeat from sperm whale.

The number of 18-bp basic units was 11.3 in all 20 of the white-beaked dolphin tandem repeat sequences examined. All these repeat sequences were completely identical (only one unique allele was found), implying that the white-beaked dolphin tandem repeat was monomorphic. The harbor porpoise tandem repeat was composed of 12.3 of 18-bp basic units. Also this tandem repeat was monomorphic. In contrast, the sperm whale tandem repeat was polymorphic with respect to the number of 18-bp basic units; three size variants composed of 3.1, 4.1, and 5.1 basic units were found in this species. On analysis of the minke whale sequences, two size variants with 6.6 and 11.6 basic units were identified.

Table 1. Properties of the tandem repeat in exon III of the DRD4 gene from cetaceans and modern ungulates

<table>
<thead>
<tr>
<th>Species</th>
<th>Number of 18-bp basic units</th>
<th>Percent matches&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Percent GC</th>
<th>Heterozygosity&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Consensus sequence of 18-bp basic units</th>
<th>GenBank accession numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>White-beaked dolphin (Lagenorhynchus albirostris)</td>
<td>11.3</td>
<td>69</td>
<td>78</td>
<td>0</td>
<td>CCGGCCGCGAGCCGCACTC</td>
<td>AY615861</td>
</tr>
<tr>
<td>Harbor porpoise (Phocoena phocoena)</td>
<td>12.3</td>
<td>73</td>
<td>80</td>
<td>0</td>
<td>CCGGCCGCGAGCCGCACTC</td>
<td>AY615862</td>
</tr>
<tr>
<td>Sperm whale (Physeter macrocephalus)</td>
<td>3.1</td>
<td>86</td>
<td>88</td>
<td>0.53</td>
<td>CCGCCCCCGAGCG/CACCC</td>
<td>AY615863</td>
</tr>
<tr>
<td>Minke whale (Balaenoptera acutorostrata)</td>
<td>5.1</td>
<td>81</td>
<td>87</td>
<td></td>
<td></td>
<td>AY615865</td>
</tr>
<tr>
<td>Domestic cow (Bos taurus)</td>
<td>5.6</td>
<td>68</td>
<td>86</td>
<td>No data</td>
<td>CCGGCCGAGCCGCCGCACTC</td>
<td>AB069668</td>
</tr>
<tr>
<td>Wild horse (Equus przewalskii)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>19.1</td>
<td>68</td>
<td>74</td>
<td>No data</td>
<td>CCGGCCGCGAGCCGCCA</td>
<td>AB080628</td>
</tr>
<tr>
<td>Wild ass (Equus hemionus)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>16.7</td>
<td>69</td>
<td>75</td>
<td>No data</td>
<td>CCGGCCGCGAGCCCA</td>
<td>AB080631</td>
</tr>
<tr>
<td>Mountain zebra (Equus zebra)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>19.1</td>
<td>69</td>
<td>75</td>
<td>No data</td>
<td>CCGGCCGCGAGCCCA</td>
<td>AB080635</td>
</tr>
</tbody>
</table>

<sup>a</sup> Percent matches is an overall measure of the degree of sequence identity between adjacent basic units in a tandem repeat.

<sup>b</sup> Heterozygosity is the probability that a given individual is heterozygous.

<sup>c</sup> Nucleotide number 14 is a G in the consensus sequence of the short allele and a C in that from the medium-sized and long alleles.

<sup>d</sup> Alignment of the equine basic unit consensus sequences with those from cetaceans changed the start position and other properties of the equine tandem repeats, including percent matches.
On alignment of the 18-bp dolphin consensus sequence with the 48-bp consensus sequence deduced on the basis of DRD4 exon III tandem repeats from various primate species (Livak et al. 1995), a consecutive stretch of seven identical nucleotides was found. No other similarities were found.

**Discussion**

This is the first study to identify and characterize a tandem repeat in exon III of the DRD4 gene within the order Cetacea. The species included in the study derived from four different cetacean families with both suborders being represented, that is, Odontoceti (white-beaked dolphin, harbor porpoise and sperm whale), and Mysticeti (minke whale). Imperfect tandem repeats composed of 18-bp basic units were detected in all species. These tandem repeat were GC rich, a property characteristic of some hypermutable tandem repeats (Denoeud et al. 2003).

The consensus sequences of the tandem repeats in the four species were identical or differed slightly. The finding
of an identical consensus sequence in white-beaked dolphin and harbor porpoise probably reflects the close evolutionary relationship between these two species. The consensus sequence from minke whale (a Mysticeti) appeared to be more similar to that identified in dolphin and harbor porpoise than the consensus sequence from sperm whale. Probably, this reflects limitations adherent to the use of these short sequences for comparative purposes.

The tandem repeat appeared to be monomorphic in the white-beaked dolphin and harbor porpoise. Because several of the dolphin and harbor porpoise samples derived from strandings of single individuals, the lack of genetic variation within these two species is probably not a result of a family relationship. Moreover, the harbor porpoises originated from two different geographical localities, namely, Danish and Greek waters, supporting the notion that the tandem repeat in this species is truly monomorphic.

Different size variants were detected in sperm whale and minke whale. The tandem repeat of sperm whale exhibited a relatively high degree of heterozygosity with three different allele sizes. Presence of several basic units with identical or nearly identical sequences in a tandem repeat is an indication that there has been several rounds of expansions and that the tandem repeat in question exists in multiple allelic forms (Benson 1999). In agreement with this, the degree of identity between the basic units in the sperm whale tandem repeat (polymorphic with three size variants) was higher than in the three other species, especially the two monomorphic species. The presence of a basic unit with a unique nucleotide sequence in the two short alleles of sperm whale and a basic unit with a sequence specific for the long allele suggests that the short alleles are evolutionarily more closely related to each other than they are to the long allele. The long allele of the minke whale, the sole representative of the suborder Mysticeti, differed from the short allele in this species by the presence of five consecutive basic units in the 3′ end, suggesting a marked degree of polarity in expansion/retraction events.

Molecular evidence places cetaceans and their terrestrial ancestors within the order of even-toed ungulates (Graur and Higgins 1994). This order also includes domestic cow, while horses belong to another order, the odd-toed ungulates. The indication that the consensus sequences from cetaceans were more closely related with those identified in the horse family than they were with the consensus sequence from domestic cow was surprising. However, only a single tandem repeat from even-toed ungulates was available, and whether this tandem repeat is typical is not known.

We also compared the basic unit consensus sequence of the *DRD4* exon III tandem repeat from dolphin with that from primates assuming that the *DRD4* tandem repeat modulates cognitive abilities. If this assumption holds, structural similarities should exist between the tandem repeats from the cetacean and primate lineages because members of both lineages possess advanced cognitive skills (Previc 1999). Similarities in gene sequences from different phylogenetic lineages reflecting shared functional properties have been observed previously. For example, the langur stomach lysozyme, which is involved in degradation of microorganisms that pass from the foregut to the stomach, has acquired sequence similarity to that of cow after the langur acquired its foregut fermentation ability (Stewart et al. 1987). The observations in the present study revealed that the composition of the tandem repeat in cetaceans differed profoundly from that in primates, which is composed of 48-bp basic units (Lichter et al. 1993; Van Tol et al. 1992). Consequently, there is no evidence that the advanced conceptual skills in cetaceans are reflected in a composition of the tandem repeat in the *DRD4* gene similar to that of primates.

### Acknowledgments
The samples were kindly provided by Drs. Hans Baagøe, Natural History Museum of Denmark, University of Copenhagen, Denmark, Aimilia Drougas, Department of Civil Engineering, Technological Education Institute of Piraeus, Greece, and Anastasia Komenou, Department of Clinical Sciences, Aristotle University of Thessaloniki, Greece. The study was financed by grants from Copenhagen Hospital Corporation Research Fund and The Danish National Psychiatric Research Foundation.

### References


Received January 20, 2005
Accepted December 9, 2005

Corresponding Editor: William Modi