
Identification of V1R-like Putative Pheromone Receptor Sequences in Non-human Primates. Characterization of V1R Pseudogenes in Marmoset, a Primate Species that Possesses an Intact Vomeronasal Organ

Dominique Giorgi and Sylvie Rouquier

IGH, CNRS UPR 1142, Montpellier, France

Correspondence to be sent to: Sylvie Rouquier, IGH, CNRS UPR 1142, rue de la Cardonille, 34396 Montpellier cédex 5, France.
e-mail: rouquier@igh.cnrs.fr

Abstract

The vomeronasal organ (VNO) is responsible in terrestrial vertebrates for the sensory perception of some pheromones, chemicals that elicit characteristic behaviors among individuals of the same species. Two multigene families (V1R, V2R) that encode proteins with seven putative transmembrane domains that are expressed selectively in different neuron subsets of the VNO have been described in rodents. Pheromone-induced behaviors and a functional VNO have been described in a number of mammals, but this sensory organ seems absent in adult catarrhines and apes, including humans. Until now, only pseudogenes have been isolated in humans, except one putative V1R (hV1RL1) sequence expressed in the main olfactory epithelium. We sought to isolate V1R-like genes in a New World monkey species, the marmoset *Callithrix jacchus*, that possesses an intact VNO and for which pheromone-induced behavior has been well documented. Using library screening approaches, we have identified five different sequences that exhibit characteristic features of V1R sequences, but that are non-functional pseudogenes. In an attempt to sort out functional V1R genes, we next cloned by polymerase chain reaction (PCR) the primate orthologues of hV1RL1. This approach was successful for gorilla, chimpanzee and orangutan, but not for the other species, including marmoset, probably because these species are too divergent from humans. Chimpanzee and orangutan V1RL1 genes are pseudogenes, whereas the gorilla counterpart is potentially functional. These observations raise the possibility that the V1R family has evolved in such a manner in mammals that every species that relies on a VNO-mediated sensory function possesses its own set of functional vomeronasal genes.

Introduction

Chemoreception is an ancient sense system that enables an organism to detect chemicals in its environment. Being thus able to sense molecules in their environment allows organisms to alter their behavior accordingly. Chemosensation in vertebrates is achieved by at least two distinct nasal organs: the main olfactory epithelium (MOE) and the vomeronasal organ (VNO) (Bargmann, 1997). Sensory neurons of the MOE detect volatile odorants and are therefore responsible for the perception of smell. Those of the VNO are believed to respond mainly to pheromones, eliciting a characteristic array of innate reproductive and social behaviors, along with neuroendocrine responses. However, only a few pheromones are known to stimulate vomeronasal sensory neurons and some pheromones appear to be detected by the main olfactory system.

In mammals, the olfactory receptors (ORs) of the MOE are encoded by a large family of ~1000 genes. They belong to the superfamily of seven-transmembrane-domain (7TM) G-protein-coupled receptors (GPCR) (Mombaerts, 1999a). In rodents, differential screenings of cDNA libraries have led to the isolation of two large independent superfamilies

of putative pheromone receptor genes (VRs), the V1Rs and V2Rs (Dulac and Axel, 1995; Herrada and Dulac, 1997; Matsunami and Buck, 1997; Ryba and Tirindelli, 1997). V1Rs are restricted to $G\alpha i2$ -expressing neurons, and V2Rs are restricted to $G\alpha 0$ -expressing neurons in the VNO. They both constitute novel families of 7TM receptors with no sequence similarity with ORs.

During the course of this work, a third family of pheromone receptors has been reported as the V3R family (Pantages and Dulac, 2000), but recent complementary data indicate that V3R correspond to one of the 12 V1R families (Rodriguez *et al.*, 2002). In mouse, the V1R family comprises ~300 receptor genes, distributed in 12 families, with pseudogenes accounting for ~53% of the total. Indeed, besides the four gene families that were first described (V1ra, V1rb, V1rc and V3R) (Dulac and Axel, 1995; Del Punta *et al.*, 2000; Dulac, 2000; Pantages and Dulac, 2000), eight new families have been characterized by database mining of the mouse genome, revealing a large sequence diversity among the V1Rs (Rodriguez *et al.*, 2002). V2Rs contain a large extracellular N-terminal domain and share similarity to

extracellular calcium-sensing receptors and metabotropic glutamate receptors. The rodent genome is estimated to contain ~100 V2R-like genes, including a large number of pseudogenes (Dulac, 2000).

In mammals, nerve fibers from the VNO converge to the accessory olfactory bulb (AOB), that projects to areas of the brain that are involved in hormonal and reproductive functions. Although the existence of a functional VNO is still debated (Meredith, 2001), a study reporting anatomical, histological and immunohistochemical data indicates that in human adults, the vomeronasal structure is a remnant of the VNO and probably cannot function as a sensory organ (Trotier *et al.*, 2000). Indeed, it appears that a VNO-like structure is present during early human embryogenesis and regresses after birth to become vestigial in adults (Rodriguez *et al.*, 2000). Furthermore, there is no AOB in humans. Although different studies argue in favor of a pheromone-like communication in human, as for example the synchronization of menstrual cycles among women living together (Stern and McClintock, 1998), the proof that this is mediated by the VNO is lacking. In other catarrhines the VNO is also vestigial, with no obvious thick sensory epithelium, while it is present in New World monkey platyrrhines (Stoddart, 1980; Taniguchi *et al.*, 1992).

Based on known and potentially functional rodent genes, genomic and bioinformatics approaches were used to explore the possibility that humans possess functional V1R-like sequences; only pseudogenes were found (Giorgi *et al.*, 2000). Nevertheless, a single V1R/V3R (V1RL1) sequence with an open-reading frame and expressed in the olfactory mucosa has been recently isolated (Pantages and Dulac, 2000; Rodriguez *et al.*, 2000). On the other hand, despite the mouse V1R repertoire recently having been characterized by database mining, no data are available in other species due to the difficulty of cloning these receptors by conventional approaches. To shed light on the evolution of the V1R family, we chose to investigate the genome of a primate species that possesses an intact VNO, i.e. the marmoset (*Callithrix jacchus*, New World monkey), in an attempt to isolate potentially functional V1R genes. Because many primates are endangered species from which it is nearly impossible to obtain VNO tissue suitable for RNA isolation, we used a genomic approach. The absence of introns in the coding region of V1Rs (Dulac and Axel, 1995) permits the determination of the coding sequence directly from genomic clones. In addition, we sought to clone the primate orthologues of the human V1RL1 gene, the only potentially functional primate V1R gene described to date.

Materials and methods

Probes

Human V1R clones Ph2 and PhB4C5 (Giorgi *et al.*, 2000) were amplified by polymerase chain reaction (PCR) using internal V1R-sequence primers. Twenty-five-microliter

PCRs were performed using 10 ng of recombinant plasmid DNA and primer sets CL2-1F (5'-GAT GAG GGA CTC TCC ATC TGC-3') and CL2-2R (5'-AAG GGC TCA CAG TGG CAT AGG) for clone Ph2 and BC5'F (5'-TAT TTG CAC AAG GTG ATG AGG) and BCR (5'-GAA AGC ACA CAG TTC TTG GC-3') for clone PhB4C5. An initial denaturation step at 94°C for 2 min was followed by 30 cycles (94°C for 15 s, 66°C for Ph2 or 55°C for PhB4C5 for 30s, 72°C for 45 s) and a final extension at 72°C for 10 min. The rat VN1 receptor sequence (Dulac and Axel, 1995) was cloned by PCR using 100 ng of genomic rat DNA and primer sets rVN1-F (5'-AGA ACA GCA GAC TCT ACA CTG-3') and rVN1-R (5'-ACA TGG ACC TCA AAG AGT TAA C-3'). Cycling conditions were as indicated above, except that the annealing temperature was 58°C and the duration of the extension step 90 s. The marmoset probe was PCR-generated from subclone Ms7E9 (issued from cos Ms7) containing the V1R sequence (see Results) with primers Ms7E9-2F (5'-GAA TAA CTT CAT CTG TCA GGC-3') and Ms7E9-2R (5'-ATG TAT GCA CTG AGA GCA GC-3'), using an annealing temperature of 58°C and an extension step of 45 s. PCR products were then purified by electrophoresis on low-melting-point agarose and directly radiolabeled by random hexamer priming (Feinberg and Vogelstein, 1983) using [α -³²P]dCTP.

Cosmid library screening

A *C. jacchus* cosmid library gridded on high density filters (marmoset genome coverage of ~1.5 \times) was purchased from the Resource Center of the German Human Genome Project (RZPD, Berlin, Germany).

The different screenings were performed separately at low stringency with human, rat and marmoset V1R clones as probes. Filter hybridization was carried out overnight at 55°C in 6 \times SSC, 5 \times Denhardt's, 0.5% SDS and 100 μ g/ml sonicated herring sperm DNA. Following the hybridization, filters were washed twice in 2 \times SSC and 0.1% SDS at 55°C. Positive clones were identified by autoradiography after overnight or longer exposure at -80°C.

Cosmid clone analysis

Cosmid DNAs were isolated by alkaline lysis and purification using a High-Pure Plasmid Isolation kit (Roche Diagnostic) or Qiagen tip-100 (Qiagen) columns, following the procedure recommended by the manufacturers. DNAs were digested with various restriction enzymes in buffers supplied by the manufacturer (New England Biolabs). The restriction digests were analysed by electrophoresis on 0.8% agarose gels. The gels were alkali blotted as described previously (Gemmill *et al.*, 1987) onto nylon membranes (Hybond N+; Amersham) and used for hybridization as described above. Because of their extreme instability, some cosmid DNAs were rescued by *in vitro* packaging and transfected in the bacterial host DH5 α MCR (Yokobata *et al.*, 1990) using Gigapack III Gold Packaging extract from

Stratagene and following the instructions of the manufacturer.

PCR with degenerate primers

Degenerate PCR primers were designed in the most conserved regions of the V1R receptor sequences isolated from rat (Dulac and Axel, 1995) and from human (Giorgi *et al.*, 2000). V1R-2F is located in the fourth transmembrane domain [5'-TC(A/T) ATT TGT C(C/T)T T(C/T)A (A/G) (C/T)A GT(G/T) (A/G)C] and V1R-2R is located between transmembrane domains V and VI [5'-T(C/T) (C/T) (A/C/T) (C/T)A (C/T)A(C/T) GCT (C/T)GT (C/T)GT GCT GTG-3']. These primers were used in PCR assays to isolate V1R-like sequences from cosmids issued from library screenings and confirmed as positives by Southern-blot hybridization. The cycling conditions consisted of an initial denaturation at 94°C for 2 min, followed by 30 cycles (94°C for 30 s, 50°C for 40 s, 72°C for 45 s) and a final extension step at 72°C for 10 min. PCR products were subcloned in a PCR cloning system (TA cloning kit; Invitrogen). Recombinant clones were analysed by PCR and sequenced using M13-vector primers and dye-labelled dideoxy terminator chemistry.

Sequencing of the V1R fragments

Directional sequencing with custom-made primers was performed directly on cosmid DNA that yielded a positive V1R band with degenerate primers V1R-2F/2R, following the instructions of the manufacturer.

Cosmids that did not amplify with the V1R primers, but that produced a positive signal by Southern-blot hybridization, were analyzed using conventional methods. Fragments containing putative V1R sequences were subcloned into pBluescript KS⁻ (Stratagene). Sequencing was carried out by a combination of vector and custom-made primers using dye-labelled dideoxy terminator chemistry.

Multiple alignment was performed using the PileUp program from the GCG package (University of Wisconsin) and the CLUSTAL W v. 1.8 software package (Thompson *et al.*, 1994). The percentage of identity for pairwise comparisons was calculated by using the GAP program (Wisconsin Package v. 8, GCG).

Analysis of cos Ms7 subclone

PCR primers surrounding the junction between the V1R sequence and the non-V1R sequence were designed: Ms7E9-C1F (5'-CTT GCA GGA GGT AAA TAA TTG-3') and Ms7E9-C1R (5'-ATT ACA CTC AGG ATG CAG AC-3'). PCR reaction was performed as indicated above, with an annealing temperature of 55°C and an extension step of 45 s.

PCR cloning of primate orthologues of the hV1RL1 gene

PCR primers (PM1F, 5'-ATG GTT GGA GAC ACA TTA AAA C-3' and PM1R, 5'-TCA TGG CAT GAC AAC CAG

ATT AG-3') spanning both ends of human V1RL1 (GenBank AF255342) were used to amplify genomic DNA from different primate species kindly provided by Dr A. Blancher: human (*Homo sapiens*); chimpanzee (*Pan troglodytes*); gorilla (*Gorilla gorilla*); orangutan (*Pongo pygmaeus*); gibbon (*Hylobates lar*); macaque (*Macaca sylvanus*); baboon (*Papio papio*); marmoset (*Callithrix jacchus*); squirrel-monkey (*Saimiri boliviensis*); and lemur (*Eulemur fulvus*). The cycling conditions consisted of an initial denaturation at 94°C for 2 min, followed by 30 cycles (94°C for 15 s, 55°C for 30 s, 72°C for 90 s) and a final extension step at 72°C for 10 min. PCR products were subcloned and recombinant clones were analysed as described above.

Results

Screening of a marmoset cosmid library with human probes

Given the phylogenetic proximity of marmoset and human, we first screened at low stringency a *C. jacchus* cosmid library with probes corresponding to two human V1R coding regions (Ph2 and PhB4C5) belonging to two different subfamilies (Giorgi *et al.*, 2000). Nineteen positive clones that produced medium or weak signals were isolated. Out of them, nine clones were confirmed by hybridization of a Southern blot of the EcoRI-digested DNAs with the mixed probe used for the library screening. The cosmids were analyzed by restriction-digest fingerprinting and were grouped on the basis of the similarity of their restriction pattern (not shown). Duplicate clones were eliminated and analysis was pursued on five clones: Ms7 (ICRF MPMGc160-K02359Q3); Ms8 (ICRF MPMGc160E14338Q3); Ms11 (ICRF MPMGc160I04322Q3); Ms18 (ICRF MPMGc160-H0890Q3); and Ms19 (ICRF MPMGc160N01207Q3). Ms7 was positive with both probes Ph2 and PhB4C5; Ms8 was positive only with Ph2; Ms11, Ms18 and Ms19 were positive only with PhB4C5.

Analysis of the positive clones

All these cosmid clones showed an extreme instability and were deleting. We rescued them by *in vitro* packaging and transfection into the bacterial host *Escherichia coli* DH5 α -MCR. Fragments testing positive with the human V1R probes were subcloned and sequenced. To our surprise, except Ms7 that contained a V1R-like sequence, all the remaining clones contained the calossin gene (GenBank Y17920), a calmodulin-binding protein of unknown function conserved throughout animal phylogeny. Pairwise comparisons between nucleotide sequences of the calossin gene and the human V1R probes did not reveal any stretch of identity or common repetitive sequences that could explain this result. Analysis of the Ms7E9 subclone revealed 67.8 and 70.1% nucleotide sequence identity (NSI) with human V1R genes Ph2 and PhB4C5, respectively, and

55.6% NSI with VN6, the most homologous rat V1R gene (Dulac and Axel, 1995). During the course of this work, other workers (Del Punta *et al.*, 2000) determined the sequences of the coding region of 25 potentially functional mouse V1R genes and seven pseudogenes. These sequences are distributed into three distinct groups—a, b, and c—with group c defining a novel group of V1R genes. In addition to V3Rs (V1rd), eight supplementary families have been identified very recently (Rodriguez *et al.*, 2002). The Ms7E9 sequence presents 66.9% NSI with the closest V1Rc sequence (V1rc3), while the closest homologues in groups a and b are V1Ra2 (57% NSI) and V1Rb3 (56.5% NSI).

These results clearly indicate that Ms7 belongs to this new group c (Figure 1). However, the homology with known V1R sequences starts at the end of TMII, the 5' end containing numerous stop codons indicating that this gene is either a truncated V1R-sequence or has accumulated so many mutations in the 5' moiety that the amino-acid identity with V1Rs is not evident anymore. To verify that this sequence was a faithful representation of the marmoset genome and not an artefact due to the deletion of the cosmid, we designed PCR primers surrounding the junction between the

V1R sequence and the unknown genomic sequence. The PCR reactions were carried out on marmoset genomic DNA, cosmid DNA and plasmid subclone DNA. In the three cases, the same band at the expected size of 250 bp was obtained, confirming that this sequence is a true pseudo-gene, not a cloning artefact and therefore does not code for a full-length V1R protein (not shown).

Screening of the marmoset library with rat probes

We next decided to screen the cosmid library at low stringency with a rat coding V1R sequence (VN1). Indeed, given the result of the first library screening, we thought that even though marmoset is phylogenetically closer to human than rat, functional V1R sequences could be closer to potentially functional rat sequences than to human V1R pseudogenes. This approach was unsuccessful. Twenty-four weakly positive clones were selected, but were not confirmed to contain V1R sequences by Southern-blot hybridization.

Screening of the marmoset library with the marmoset Ms7E9 clone

We performed a last screening at low stringency with the

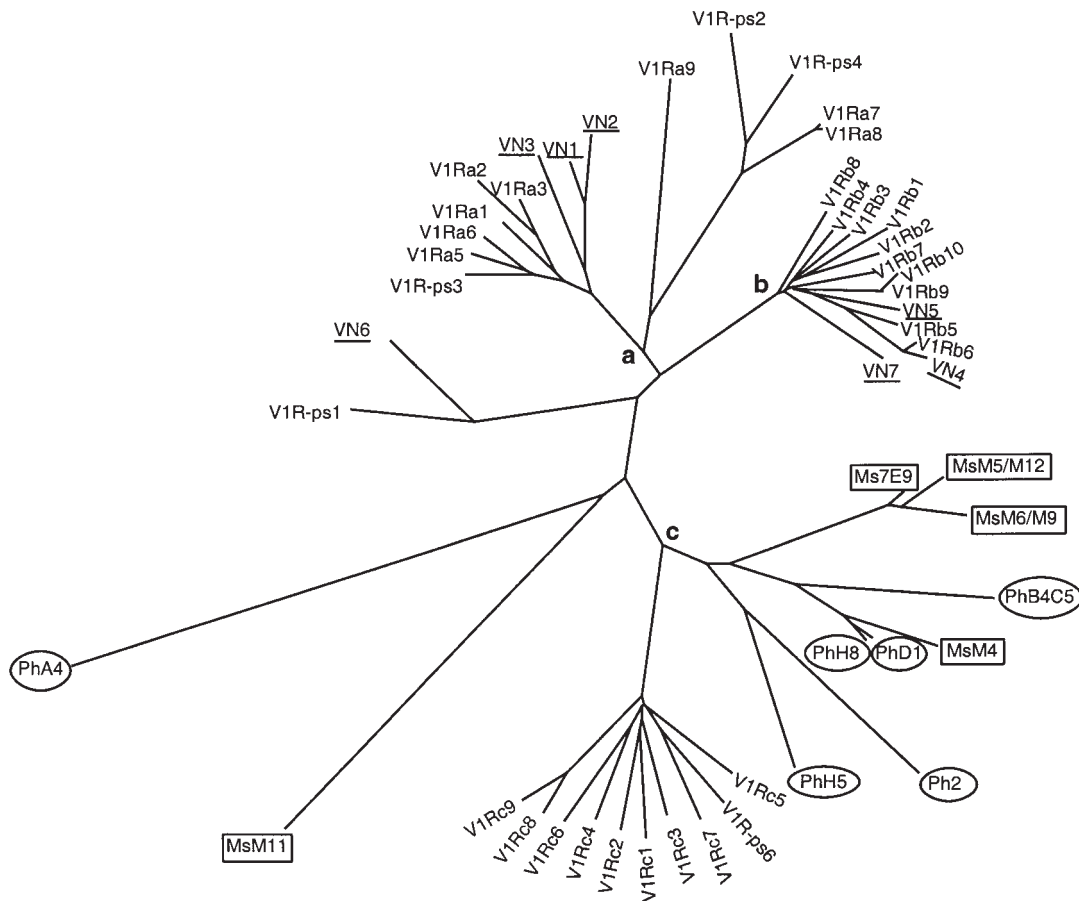


Figure 1 Unrooted phylogenetic tree placing the marmoset V1R-like nucleotide sequences relative to rat V1R sequences VN1–VN7 (Dulac, 1995), mouse V1R sequences groups a, b, c (Del Punta *et al.*, 2000) and human V1R sequences (Giorgi *et al.*, 2000). Sequences were aligned using CLUSTAL W v. 1.8. Marmoset sequences are boxed, rat sequences are underlined and human sequences are circled.

Ms7E9 V1R-like sequence. Four calossin-containing clones were again isolated. Twelve other clones yielded a positive signal ranging from weak to strong and were analyzed by Southern blot (Figure 2). Out of them, one clone did not grow, four produced a weak signal (MsM2, ICRF MPMGc160P0934; MsM4, ICRF MPMGc160M09317; MsM10, ICRF MPMGc160G16333Q2; MsM11, ICRF MPMGc160M1118Q2) and four clones gave a medium to strong signal (MsM5, ICRF MPMGc160O2195Q2; MsM6, ICRF MPMGc160O19149Q2; MsM9, ICRF MPMGc160B15280Q2; MsM12, ICRF MPMGc160N02356Q2) comparable to the positive control Ms7. Clones MsM1, MsM7 and MsM8 were negative. Clones MsM2 and MsM10 were further shown to contain repeats but no V1R-like sequences.

Analysis of the marmoset V1R-like clones

In parallel, degenerate PCR primers chosen in conserved domains of rat and human V1Rs were tested on these cosmids as well as on the cosmids isolated during the two other screening steps. Primer pair V1R-2F/2R was able to amplify products only from cosmids Ms7, MsM4, MsM5, MsM6, MsM9, MsM11 and MsM12 (not shown). These PCR products were subcloned and sequenced. Sequences obtained from clones M5 and M12 were identical, as well as those from MsM6 and MsM9. We established the full-length sequences of all these clones, except MsM4 (for which ~16 amino acids are missing on the 5' end), by primer-walking on cosmid DNA. Ms7 presents 94 and 93.7% NSI with MsM5/M12 and MsM6/M9, respectively. Furthermore, as observed in Ms7, the homology of MsM5/M12 and MsM6/M9 with rodent V1R sequences starts at the beginning of domain II, with a very diverged 5' end.

MsM4 and MsM11 present, respectively, 73.9 and 48.7% NSI with Ms7. When compared to groups a, b and c (Figure 1), MsM4 clearly belongs to group c (68.7% NSI with V1Rc3), while MsM11 presents an average NSI of ~45% with the sequences of the three groups (44% with V1ra1, 41.4% with V1Rb1 and 49.4% with V1Rc3). When compared to human sequences, MsM4 and MsM11 present, respectively, 69 and 49% NSI with Ph2 and 73 and 44.9% NSI with PhB4C5. As in the case of Ms7, the four additional V1R-like sequences obtained from these six cosmid clones contained stop codons and/or frameshifts, indicating that they correspond to pseudogenes.

Although the proteins deduced from these marmoset V1R-like sequences are similar to V1Rs, the inferred amino-acid sequences were uncertain in a few regions of some clones due to many frameshift-causing indels (insertion/deletion events). We therefore present in Figure 1 the relationships among these sequences in the form of a phylogenetic tree based on the nucleotide sequences. The tree includes the seven rat V1R sequences (Dulac and Axel, 1995), the mouse V1R genes and pseudogenes belonging to the three different groups a, b and c recently characterized

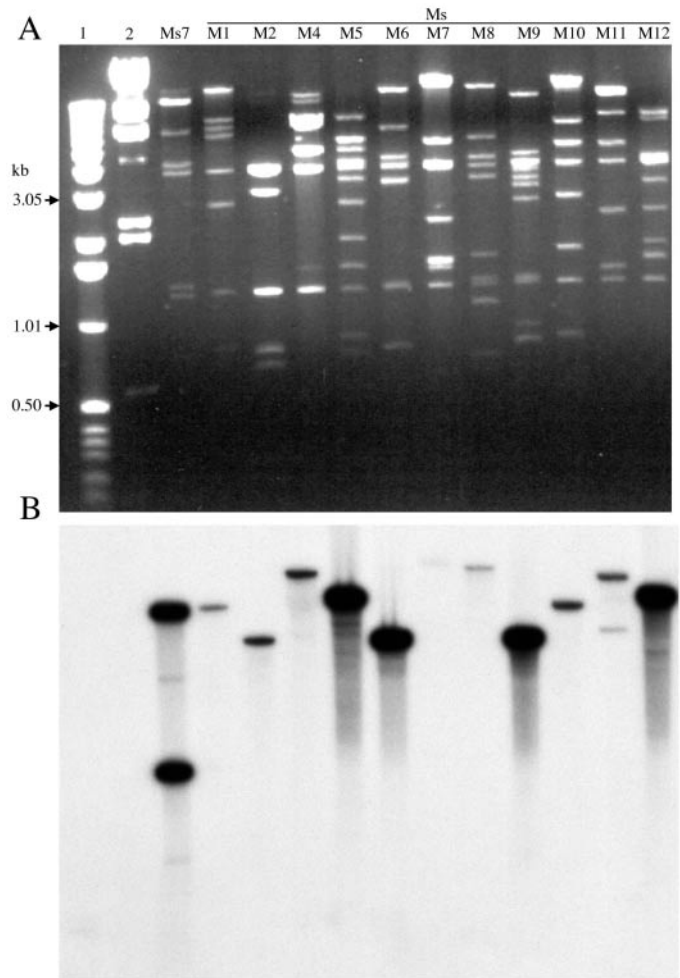


Figure 2 Southern-blot analysis of the marmoset V1R-like cosmid clones. **(A)** EcoRI restriction of the 11 cosmid clones isolated with the Ms7E9 probe. Cosmid Ms7 was loaded as a positive control. DNAs were run on a 0.6% agarose gel. Two mol. wt markers (1 kb ladder and Lambda/Hind III; Life Technologies) were run in parallel. The gel was then alkali blotted on nylon membrane and hybridized with the Ms7E9 probe. **(B)** Autoradiogram of the Southern blot.

(Del Punta *et al.*, 2000) and the seven human pseudogenes that we previously described (Giorgi *et al.*, 2000). Except for PhA4 and MsM11 that are more distantly related, marmoset and human V1R-like sequences belong to group c initially defined by mouse V1R sequences, but form a sub-group distinct from rodent group c sequences. MsM4 groups with PhH8 and PhD1 that localize human chromosome 7, and sequences Ms7E9, MsM5/M12 and MsM6/M9 form another sub-group. MsM11 may constitute a new group that is not a, b or c.

The marmoset sequences show a NSI ranging from 44.5 to 94%, with M11 being the most divergent, leading in all cases to an amino-acid sequence identity (ASI) > 40%. One classification proposed for multigene families (Nebert *et al.*, 1991; Ben-Arie *et al.*, 1994) defines 40% ASI as a threshold for belonging to the same family. According to this definition,

all marmoset sequences are members of the same family. Except for MsM11, the group of marmoset sequences is very close to the mouse group c sequences that are considered as a new sub-group (Del Punta *et al.*, 2000). Families are then subdivided in sub-families (ASI > 60%). Accordingly, only M11 belongs to a distinct sub-family with an NSI with the other V1R genes ranging from 44 to 48%. As shown in Figure 3, assembling of the marmoset protein sequences by correcting numerous frameshifts revealed that they contain the amino-acid residues described previously (Del Punta *et al.*, 2000) that are specific to the 7TM-GPCR superfamily (N28, C88, C176 and F254) and the V1R family (G24, L31, R99, L109, K128, N163 and L250).

Characterization of primate orthologues of the hV1RL1 gene

In an attempt to characterize potentially functional V1R genes in primates, we took advantage of the ORF (open reading frame)-containing human gene hV1RL1 (GenBank AF255342) to characterize by PCR the orthologues in various species. Human hV1RL1 is the only gene reported to date to encode a potentially functional VR-type receptor in primates. PCR products were obtained with DNA from human, chimpanzee, gorilla and orangutan, but not from the other species (gibbon, macaque, baboon, marmoset, squirrel-monkey and lemur), suggesting that these latter sequences are too divergent from their human counterpart.

The human sequence was 100% identical to the published hV1RL1 sequence (Pantages and Dulac, 2000; Rodriguez *et al.*, 2000). Protein sequence comparison (Figure 4) indicates that human, chimpanzee and gorilla share ~97% of amino-acid sequence identity (ASI), whereas orangutan, that is more distantly related on the evolutionary tree, shares ~92% ASI with the three other species. Surprisingly, only human and gorilla contain an ORF. Chimpanzee and orangutan contain obvious pseudogenes due to frameshift mutations interrupting the ORF. The orangutan sequence contains five frameshifts and the chimpanzee sequence contains one frameshift associated with a motif change in conserved transmembrane domain TM7 (CIH instead of VSL in the other species). In addition, the four species present 38 different amino-acid substitutions, a number of which are non-synonymous (Figure 4).

Discussion

The VNO is a chemosensory organ located at the base of the nasal cavity. It is distinct from the olfactory epithelium and is present in most amphibians, reptiles and non-primate mammals, but is absent in birds, adult catarrhine monkeys and apes (Stoddart, 1980). It plays an important role in intra-specific chemical (pheromone) communication (Halpern, 1987; Keverne, 1999; Holy *et al.*, 2000; Leinders-Zuffall *et al.*, 2000). Two main families of putative pheromone receptors expressed in rodent VNO have been described, i.e.

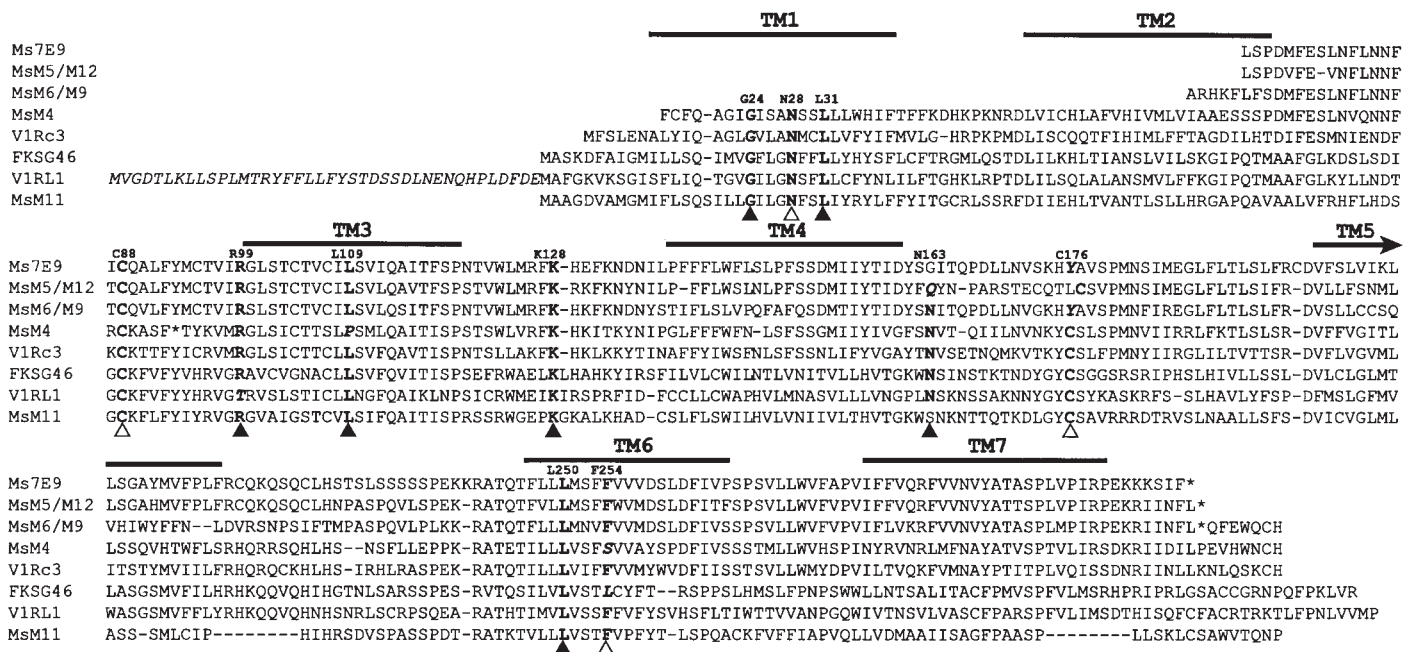


Figure 3 Alignment of predicted amino-acids sequences of marmoset pseudogenes and their closest potentially coding homologues in the databases, i.e. V1Rc3 (GenBank AF291499), FKSG46 (AF336873) and V1RL1 (AF255342). Due to many mutations and frameshifts, marmoset protein sequences were decrypted, when possible, using blastx (AGC) and assembled by eye. Sequence comparison and alignment were performed using CLUSTAL W v. 1.8 software. The seven putative transmembrane domains are represented as horizontal lines. Amino acids that are conserved in 7TM-GPCR and V1R proteins (Del Punta *et al.*, 2000) are in bold, numbered and indicated by open and closed arrowheads, respectively. Amino acids that diverged from the consensus for one nucleotide in the codon are indicated in bold italic. The NH2 extension of the V1RL1 long-form is indicated in italics.

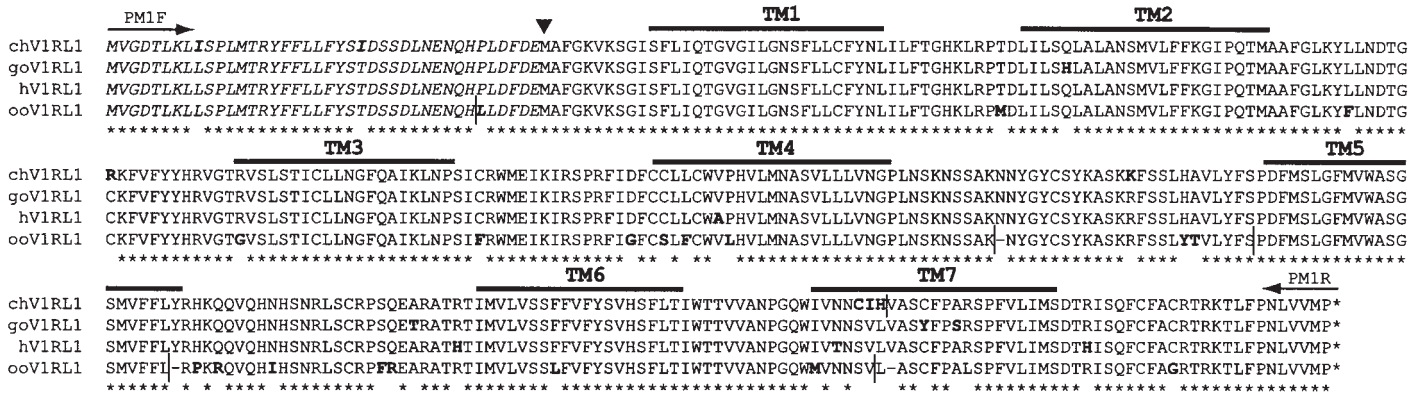


Figure 4 Alignment of predicted amino-acids sequences of V1RL1 primate orthologues. Sequence comparison and alignment were performed using CLUSTAL W v. 1.8 software. The position of PCR primers PM1F and PM1R is indicated by arrows. The seven putative transmembrane domains are represented as horizontal lines. The NH2 extension of the V1RL1 long-form is indicated in italic. Amino-acids identical in the four species are indicated by asterisks under the sequences. The putative start methionine is indicated by an arrowhead. Non-synonymous substitutions are indicated in bold. Frameshifts are represented by vertical bars. Chimpanzee, ch; gorilla, go; human, h; orangutan, oo.

V1R/V3R and V2R, each comprising ~300 and 100–150 members, respectively. Pseudogenes have been reported for the mouse V1R (50%) and V2R (~66%) families. V1R sequences have been cloned in human in three separate studies. In the three analyses, the human V1R-like family appears to be much smaller than the rodent family since only seven, eight and 34 different sequences, respectively, were characterized (Dulac and Axel, 1995; Giorgi *et al.*, 2000; Kouros-Mehr *et al.*, 2001). So far, only human V1R pseudogenes have been cloned. However, one potentially functional human sequence, V1RL1 (GenBank AF255342) described (Pantages and Dulac, 2000; Rodriguez *et al.*, 2000) and expressed in the MOE has been isolated. This is the only example of a potentially functional primate VR gene reported to date.

Platyrrhines are the last primates to diverge from the human evolutionary lineage that have the VNO morphology of most other mammals. We thought that a primate species with an intact and functional VNO would retain a high proportion of functional V1R genes, as in rodents. We therefore sought to isolate putative marmoset V1R homologs to approach the question of the relationship between an intact VNO and potentially functional V1R sequences. As a result, we have only identified five different marmoset V1R-like pseudogenes using an approach based on genomic library screenings at low stringency with V1R probes derived either from humans or rodents. These results are similar to those obtained in humans using rodent probes, i.e. a small number of different clones containing only V1R pseudogenes were isolated (Giorgi *et al.*, 2000). Marmosets possess an intact VNO and the role of olfactory communication in social behaviors such as recognition of group members (Epple *et al.*, 1993) and reproductive status (Smith and Abbott, 1998) has been documented. Despite these observations, our approach did not allow us to isolate functional marmoset V1R genes. Similarly, an approach conducted in pig, another

species with an intact VNO and whose behavior related to pheromones is well established, resulted in the identification of pseudogenes only (S. Rouquier *et al.*, unpublished results).

Nevertheless, the release of the ORF-containing hV1RL1 gene (Rodriguez *et al.*, 2000) allowed us to isolate by PCR the orthologues of this gene in different primate species, in an attempt to pull out potentially functional primate genes. Among the different species tested and in addition to humans, we were able to amplify only chimpanzee, gorilla and orangutan sequences, the other species, including marmoset, presumably being too divergent. Similarly, a pig library screening using human V1RL1 as a probe did not allow us to select for a putative pheromone receptor gene (S. Rouquier *et al.*, unpublished results). Chimpanzee and orangutan V1RL1 genes are pseudogenes and, despite the fact that human and gorilla contain an ORF, these observations raise the question of whether the gene is functional in these two last species. Indeed, from an evolutionary point of view, knowing the regression of the olfactory structures (VNO, MOE) and the increasing pseudogenization level of the genes involved in the corresponding chemosensory functions from mouse to humans, and even more precisely from New World monkeys to hominoids (Rouquier *et al.*, 1998a; Giorgi *et al.*, 2000; Kouros-Mehr *et al.*, 2001), the notion that the human V1RL1 gene became functional again, or retained a function, while it is inactivated in chimpanzees and orangutans is questionable. Alternatively, due to its expression pattern (olfactory mucosa), it is possible that V1RL1 was primarily a VNO receptor which evolved into an odorant receptor with pheromone-like ligand properties (Kouros-Mehr *et al.*, 2001). This question will be ultimately addressed by a functional test.

Our results lead us to hypothesize that during evolution, the V1R gene repertoire diversified to generate different sets of functional genes that are specific of each species. Indeed,

using classical heterologous screening strategies, we were able to select only for vestigial genes, whereas the presumed functional pheromone receptor genes that are specific for a given species may be too divergent to be isolated by these approaches. These observations suggest that the species-specific nature of pheromone communication makes this system quite different from regular olfaction, since functional olfactory receptor genes are quite conserved in vertebrates, especially in mammals, and can be isolated from one species to another using various strategies (Mombaerts, 1999b). This hypothesis is reinforced by data published while this work was under submission (Lane *et al.*, 2002). The authors performed a sequence analysis of mouse V1R gene clusters and were unable to identify the human orthologues, suggesting that these genes have been lost during primate evolution. In the absence of extensive genome sequence data, it is likely that cDNA approaches using complete VNO or single-cell PCR would be more appropriate to characterize putative pheromone receptor genes of a particular species.

The results reported here indicate that V1R genes with sequences likely to code for functional proteins are weakly homologous between species and therefore suggest alternative approaches for their identification. Indeed, the characterization of the V1R repertoire is rather slow and very few papers have been published for other species than rodents. On the other hand, the identification of a human V1R/V3R ORF that is expressed in the MOE suggests that the VNO should not be considered as the exclusive site for VR type receptors, as previously described in goldfish (Cao *et al.*, 1998).

Overall, these observations lead to several reflections on the evolution of chemical communication which are not mutually exclusive, as follows.

1. The pseudogenization process parallels the regression of the structures (VNO, AOB, MOE) devoted to chemical communication and the need for this function (Giorgi *et al.*, 2000; Rouquier *et al.*, 2000; Kouros-Mehr *et al.*, 2001), i.e. in absence of selective pressure there would be a parallel degeneration of organs and genes, meaning that primates (at least hominoids) would not extensively use chemicals to communicate as rodents do.
2. One may wonder why so few VR genes are identified by using heterologous (rodent) probes in primates. Two explanations are possible. Either numerous genes were deleted over time, or they evolved rapidly in a species-specific fashion leading to a loss of the homology necessary to ensure a specific hybridization.
3. In animals that possess a VNO and for which chemical communication via pheromones has been well established (pig, marmoset), we see the same type of results, i.e. very few genes were identified using heterologous probes and all of them were pseudogenes. Starting from the hypothesis that these species have functional VR genes (GPCR), we may hypothesize that, during evolution, each species acquired its own set of functional VR genes and that each set does not overlap with those of other species, except perhaps in the case of very close species such as rat and mouse, or human and chimpanzee or gorilla. This situation evokes the case of olfactory receptors of vertebrates and insects, i.e. vertebrate sequences could not be used to pull out insect olfactory receptor genes.
4. Alternatively, primates or any species communicating via pheromones might use receptors radically different from those used as probes in this study.
5. Finally, it is also possible that in absence of a functional VNO, other genes different from the VR genes and expressed in other tissues may function as pheromone receptor genes. Olfactory receptors could play this role. For example, gene OR912-93 which is potentially functional in chimpanzee, orangutan and gibbon (Rouquier *et al.*, 1998b), binds the odorant 2-heptanone (Gaillard *et al.*, 2002), a chemical which displays pheromone activity in mouse (Leinders-Zufall *et al.*, 2000). However, it should be demonstrated that the signal transmission in the nervous system is in accordance with a pheromone stimulus inducing modifications of the physiology and/or behavior.

In most species that rely on pheromone communication, the receptors responsible for the function remain to be discovered. The study of the evolution of the VR genes in primates is an important step in discovering whether humans possess functional VR genes.

Note

Sequence data from this article have been deposited with the DDBJ/EMBL/GenBank Data Libraries under accession Nos AF397897–AF397901 and AF426106–AF426108

Acknowledgements

We are grateful to S. Dubel for critical reading of the manuscript, and to Prof. A. Blancher for providing primate DNAs.

References

- Bargmann, C. (1997) *Olfactory receptors, vomeronasal receptors, and the organization of olfactory information*. Cell, 90, 585–587.
- Ben-Arie, N., Lancet, D., Taylor, C., Khen, M., Walker, N., Ledbetter, D.H., Carrozzo, R., Patel, K., Sheer, D., Lehrach, H. and North, M.A. (1994) *Olfactory receptor gene cluster on human chromosome 17: possible duplication of an ancestral receptor repertoire*. Hum. Mol. Genet., 3, 229–235.
- Cao, C., Oh, B. and Stryer, L. (1998) *Cloning and localization of two receptor multigene families in goldfish olfactory epithelium*. Proc. Natl Acad. Sci. USA, 95, 11987–11992.
- Del Punta, K., Rothman, A., Rodriguez, I. and Mombaerts, P. (2000) *Sequence diversity and genomic organization of vomeronasal receptor genes in the mouse*. Genome Res., 10, 1958–1967.
- Dulac, C. (2000) *Sensory coding of pheromone signals in mammals*. Curr. Opin. Neurobiol., 10, 511–518.

- Dulac, C. and Axel, R.** (1995) *A novel family of genes encoding putative pheromone receptors in mammals*. *Cell*, 83, 195–206.
- Epple, G., Belcher, A.M., Knderling, I., Zeller, U., Scolnick, L., Greenfield, K.L. and Smith, A.B.** (1993) *Making sense out of scents: species differences in scent glands, scent-making behavior, and scent-mark composition in Callitrichidae*. In Rylands, A.B. (ed.), *Marmosets and Tamarins: Systematics, Behavior, and Ecology*. Oxford University Press, Oxford, pp. 123–151.
- Feinberg, A. and Vogelstein, B.** (1983) *A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity*. *Anal. Biochem.*, 132, 6–13.
- Gaillard, I., Rouquier, S., Pin, J.-P., Mollard, P., Richard, S., Barnabé, C., Demaille, J. and Giorgi, D.** (2002) *A single olfactory receptor specifically binds a set of odorant molecules*. *Eur. J. Neurosci.*, 15, 409–418.
- Gemmill, R.M., Coyle-Morris, J.F., McPeck, F.D., Ware-Urbe, L.F. and Hecht, F.** (1987) *Construction of long-range restriction maps in human DNA using pulsed field gel electrophoresis*. *Gene Anal. Technol.*, 4, 119–131.
- Giorgi, D., Friedman, C., Trask, B.J. and Rouquier, S.** (2000) *Characterization of nonfunctional V1R-like pheromone receptor sequences in human*. *Genome Res.*, 10, 1979–1985.
- Halpern, M.** (1987) *The organization and function of the vomeronasal system*. *Annu. Rev. Neurosci.*, 10, 325–362.
- Herrada, G. and Dulac, C.** (1997) *A novel family of putative pheromone receptors in mammals with a topographically organized and sexually dimorphic distribution*. *Cell*, 90, 763–773.
- Holy, T.E., Dulac, C. and Meister, M.** (2000) *Responses of vomeronasal neurons to natural stimuli*. *Science*, 289, 1569–1572.
- Keverne, E.B.** (1999) *The vomeronasal organ*. *Science*, 286, 716–720.
- Kouros-Mehr, H., Pintchovski, S., Melnyk, J., Chen, Y.J., Friedman, C., Trask, B. and Shizuya, H.** (2001) *Identification of non-functional human VNO receptor genes provides evidence for vestigiality of the human VNO*. *Chem. Senses*, 26, 1167–1174.
- Lane, R.P., Cutforth, T., Axel, R., Hood, L. and Trask, B.J.** (2002) *Sequence analysis of mouse vomeronasal receptor gene clusters reveals common promoter motifs and a history of recent expansion*. *Proc. Natl Acad. Sci. USA*, 99, 291–296.
- Leinders-Zufall, T., Lane, A.P., Puche, A.C., Ma, W., Novotny, M.V., Shipley, M.T. and Zufall, F.** (2000) *Ultrasensitive pheromone detection by mammalian vomeronasal neurons*. *Nature*, 405, 792–796.
- Matsunami, H. and Buck, L.** (1997) *A putative family encoding a diverse array of putative pheromone receptors in mammals*. *Cell*, 90, 775–784.
- Meredith, M.** (2001) *Human vomeronasal organ function: a critical review of best and worst cases*. *Chem. Senses*, 26, 433–445.
- Mombaerts, P.** (1999a) *Molecular biology of odorant receptors in vertebrates*. *Annu. Rev. Neurosci.*, 22, 487–509.
- Mombaerts, P.** (1999b) *Seven-Transmembrane Proteins as Odorant and Chemosensory Receptors*. *Science* 286, 707–711.
- Nebert, D.W., Nelson, D.R., Coon, M.J., Estabrook, R.W., Feyereisen, R., Fujii-Kuriyama, Y., Gonzalez, F.J., Guengerich, F.P., Gunsalus, I.C., Johnson, E.F., et al.** (1991) *The P450 superfamily: update on new sequences, gene mapping, and recommended nomenclature*. *DNA Cell Biol.*, 10, 1–14.
- Pantages, E. and Dulac, C.** (2000) *A novel family of candidate pheromone receptors in mammals*. *Neuron*, 28, 835–845.
- Rodriguez, I., Greer, C.A., Mok, M.Y. and Mombaerts, P.** (2000) *A putative pheromone receptor gene expressed in human olfactory mucosa*. *Nat. Genet.*, 26, 18–19.
- Rodriguez, I., Punta, K.D., Rothman, A., Ishii, T. and Mombaerts, P.** (2002) *Multiple new and isolated families within the mouse superfamily of V1r vomeronasal receptors*. *Nat. Neurosci.*, 5, 134–140.
- Rouquier, S., Taviaux, S., Trask, B., Brand-Arpon, V., van den Engh, G., Demaille, J. and Giorgi, D.** (1998a) *Distribution of olfactory receptor genes in the human genome*. *Nat. Genet.*, 18, 243–250.
- Rouquier, S., Friedman, C., Delettre, C., van den Engh, G., Blancher, A., Crouau-Roy, B., Trask, B. and Giorgi, D.** (1998b) *A gene recently inactivated in human defines a new olfactory receptor family in mammals*. *Hum. Mol. Genet.*, 7, 1337–1345.
- Rouquier, S., Blancher, A. and Giorgi, D.** (2000) *The olfactory receptor gene repertoire in primates and mouse: evidence for reduction of the functional fraction in primates*. *Proc. Natl Acad. Sci. USA*, 97, 2870–2874.
- Ryba, N. and Tirindelli, R.** (1997) *A new multigene family of putative pheromone receptors*. *Neuron*, 19, 371–379.
- Smith, T.E. and Abbott, D.H.** (1998) *Behavioral discrimination between circumgenital odor from peri-ovulatory dominant and anovulatory female common marmosets (Callithrix jacchus)*. *Am. J. Primatol.*, 46, 265–284.
- Stern, K. and McClintock, M.** (1998) *Regulation of ovulation by human pheromones*. *Nature*, 392, 177–179.
- Stoddart, D.** (1980) *The Ecology of Vertebrate Olfaction*. Chapman & Hall, New York.
- Taniguchi, K., Matsusaki, Y., Ogawa, K. and Saito, T.R.** (1992) *Fine structure of the vomeronasal organ in the common marmoset (Callithrix jacchus)*. *Folia Primatol.*, 59, 169–176.
- Thompson, J.D., Higgins, D.G. and Gibson, T.J.** (1994) *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice*. *Nucl. Acids Res.*, 22, 4673–4680.
- Trotier, D., Eloit, C., Wassef, M., Talmain, G., Bensimon, J.L., Doving, K.B. and Ferrand, J.** (2000) *The vomeronasal cavity in adult humans*. *Chem. Senses*, 25, 369–380.
- Yokobata, K., Trenchak, B. and de Jong, P.** (1990) *Rescue of unstable cosmids by in vitro packaging*. *Nucleic Acids Res.*, 19, 403–404.

Accepted April 12, 2002