Chem. Senses 26: 1157–1166, 2001

New GPCRs from a Human Lingual cDNA Library

J.-C. Gaudin, L. Breuils and T. Haertlé

Laboratoire d'Etude des Interactions des Molécules Alimentaires, Institut National de la Recherche Agronomique, BP 71627, 44316 Nantes, Cedex 3, France

Correspondence to be sent to: T. Haertlé, Laboratoire d'Etude des Interactions des Molécules Alimentaires, Institut National de la Recherche Agronomique, BP 71627, 44316 Nantes, Cedex 3, France. e-mail: haertle@nantes.inra.fr

Abstract

Sweet and bitter taste perception involve G protein coupled receptors (GPCRs) present at the taste receptor cell surface. It is likely that various mechanisms are active and various families of GPCRs are involved in the perception of these tastes. The expression of GPCRs in human tongue was studied using degenerated primers corresponding to transmembrane domains 2 or 3 (for 5′ primer), 6 or 7 (for 3′ primer) of olfactory-like receptors in reverse transcription-polymerase chain reaction experiments. It was demonstrated that four previously identified, eight new olfactory-like receptor genes, three previously known and eight new olfactory-like receptor pseudogenes, mostly located on chromosome 11, are expressed in adult tongue and/or in fetal tongue. Previously identified genes include HGMP07I, HTPCR06, TPCR120 and TPCR85 whose cDNAs were originally isolated from male germinal cells. New genes were named JCG1, JCG2, JCG3, JCG4, JCG5, JCG6, JCG9 and JCG10. HGMP07I, HTPCR06, TPCR120, JCG3 and JCG5 are also expressed in the epithelium of adult tongue, whereas all these genes are expressed in fetal tongue. Although functional studies are needed before definitive conclusions are made, the obtained results imply that lingual olfactory-like receptors could be involved in taste perception.

Introduction

Mammals are able to discriminate between five basic tastes: sweet, salty, sour, bitter and umami. Taste receptor cells (TRCs) are localized in small specialized organs called taste buds distributed on the surface of the tongue and palate, mainly in circumvallate, fungiform and foliate papillae. It is well known that salty and sour perception involve membrane ion channels of TRCs. In rats, umami taste, the taste elicited by sodium glutamate, is mediated by a truncated form of the metabotropic glutamate receptor 4 (mGluR4) (Chaudhari *et al.*, 2000) and probably also by ionotropic glutamate receptors (Chaudhari and Roper, 1998; Lin and Kinnamon, 1999). mGluR4 is a G protein coupled receptor (GPCR). Although still not very precisely defined, it is generally accepted that at least a part of sweet and bitter taste signals is also transduced by GPCRs. Recent and independent cloning, in two laboratories, of a family of candidate receptors for bitter and sweet compounds confirms this hypothesis (Adler *et al.*, 2000; Matsunami *et al.*, 2000). Among this family of receptors forming a new class of GPCRs, three members named mT2R-5 and mT2R-8 in mice and hT2R-4 in humans function as bitter taste receptors in transfected cells (Chandrashekar *et al.*, 2000). Good candidate receptors for sweet taste named Tas1r3 (Max *et al.*, 2001) or T1R3 (Kitagawa *et al.*, 2001; Montmayeur *et al.*, 2001; Sainz *et al.*, 2001) have been cloned in four independent laboratories recently. This receptor is also a GPCR belonging to another class of GPCRs containing T1R1 and T1R2 taste receptors (Hoon *et al.*, 1999).

Before the cloning of members of these classes of GPCR, it had been proposed that receptors for sweet and bitter molecules may display a high degree of homology with olfactory receptors. Abe *et al.* showed using reverse transcription-polymerase chain reaction (RT-PCR) strategy that a family of olfactory-like receptor (OLR) mRNAs is expressed in rat tongue (Abe *et al.*, 1993a). One of these mRNAs, coding a protein named GUST27, is present in epithelial cells of rat tongue, including taste buds (Abe *et al.*, 1993b). Additionally, GUST27 was found to be expressed in a taste bud area where α-gustducin, a taste-specific G protein, is also present. These data suggested that GUST27 could be involved in taste transduction (Kusakabe *et al.*, 1996). A RT-PCR approach also allowed Matsuoka *et al.* (Matsuoka *et al.*, 1993) to obtain cDNA clones highly similar to olfactory receptors from bovine taste tissue.

The present study reports the cloning of several cDNAs corresponding to OLR mRNAs expressed in human tongue and the identification of the corresponding genes.

Material and methods

Total RNA extraction

Total RNA was prepared from a sample of adult human

tongue epithelium. The sample originated from a 52-year-old Caucasian male subject to tongue ablation. It was quickly washed in PBS and frozen in liquid nitrogen. Epithelial and sub-epithelial layers were separated from the muscular layer by dissection using a scalpel. Tissue sample still frozen (~100 mg) was broken in liquid nitrogen with a freezer/mill Spex 6700 apparatus. The powder obtained was suspended in 350 µl of lytic solution (Qiagen SA, Courtaboeuf, France) and homogenized for 30 s using a polytron apparatus (Kinematica, Luzern, Switzerland). Total RNA was extracted using Qiagen RNeasy extraction kit and *DNase I* treatment (Qiagen SA). The integrity of RNA was checked by 1.2% agarose gel electrophoresis.

RT-PCR

Reverse transcription (RT) was performed on 1–2 µg of total RNA incubated 1 h at 37°C in 40 µl of a reaction mixture containing 1× buffer, 0.625 µM each dNTP, 3.5 µM oligo-dT primer, 1 µM random nonamer primer, 0.5 units of RNase inhibitor (Eurogentec, Seraing, Belgium) and 4 units of Omniscript Reverse Transcriptase™ (Qiagen).

PCR primers were synthesized by Cybergene (Evry, France). Degenerated primers *RS1* 5′-CA(AGCT)AC(AGCT)CC(AGCT)ATGTA(CT)(ACT)(AGCT)TT(CT)(CT)T-3′, *RS2* 5′-ATGGC(AGCT)TA(CT)GA(CT)(AC)G(AGCT)TA(CT)GT(AGCT)GC-3′, *RAS3* 5′-(GC)(CT)(AG)CA(AGCT)GT(AGCT)(GC)(AT)(AG)AA(AGCT)GC(CT)TT-3′, *RAS4* 5′-TA(AGT)AT(AGCT)A(AG)(AGCT)GG(AG)TT(AGCT)A(AG)CAT(AGCT)GG-3′ were chosen according to the amino-acid sequences (H/Q)TPMY(F/L/I)FL, MAYDRYVA(I/V), KAFSTC(G/T/A), PMLNP(F/L)IY(S/T), which are highly conserved among the olfactory receptors family and situated close to transmembrane domains 2, 3, 6 and 7, respectively (Horn *et al.*, 1998). *RS1* and *RS2* were used as 5′ primers, *RAS3* and *RAS4* as 3′ primers.

Specific primers were: *JCG1–5* 5′-ATGGGGACTGGAAATGA-3′ and *JCG1–3* 5′-TCAAGAAAATATTTTTATTCTAAG-3′ for full length JCG1 cDNA amplification; *JCG2–5* 5′-ATGGCTACTTCAAACCATTCTTC-3′ and *JCG2–3* 5′-TCAGGATGACTGCCTTCCC-3′ for full length JCG2 cDNA amplification; *JCG3–5* 5′-ATGAATTCCCTGAAGGACG-3′ and *JCG3–3* 5′-CTATGTAATATCATTATTTGAAGTTC-3′ for full length JCG3 cDNA amplification; *JCG5–5* 5′-ATGATGTGGGAAAACTGG-3′ and *JCG5–3* 5′-TCATAGTTTCTGAGAGCC-3′ for full length JCG5 cDNA amplification; *JCG6–5* 5′-ATGGCTATAGGAAACTGG-3′ and *JCG6–3* 5′-CTATGGGATACAGTTTCTG-3′ for full length JCG6 cDNA amplification; *JCG9–5* 5′-ATGACCATGGAAAATTATTCTA-3′ and *JCG9–3* 5′-TCATTTTCCTACTAAGACCT-3′ for full length JCG9 cDNA amplification; *OR1E1–5* 5′-ATGATGGGACAAAATCAAAC-3′ and *OR1E1–3* 5′-TCAGAGAGAGAAGAAAGTT-3′ for full length HGMP07I/OR1E1 cDNA amplification; *JCG8–5* 5′-ATGGCTGCTGAGAATTC-3′ and *JCG8–3* 5′-TCAGGAGAATGCATTTTTG-3′ for full length TPCR85 cDNA amplification; *HTPCR-5* 5′-ATGCAAGGAGAAAACTTCAC-3′ and *HTPCR-3* 5′-TCAGAGATGTTCGTGTGTTT-3′ for full length HTPCR06 cDNA amplification, *GAPDH-5* 5′-GAAATCCCATCACCATCT-3′ and *GAPDH-3* 5′-TCCACAGTCTTCTGGGTG-3′ for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) cDNA amplification.

PCRs were performed on either 2 µl of RT reaction mixture, 1 µl of human fetal tongue Gene Pool™ (Invitrogen, Groningen, The Netherlands), 0.05–0.1 µg of total RNA from human adult tongue epithelium or 100 ng of human placenta genomic DNA. The PCR mixture (final volume 50 µl) contained 1× PCR buffer, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.3 µM of each specific primer (1 µM in the case of degenerated primers *RS1*, *RS2*, *RAS3* and *RAS4*) and 2.5 U HotStartTaq™ DNA polymerase (Qiagen SA). An initial step 15 min at 95°C was followed by 35 or 40 cycles: 30 s at 94°C for denaturation, 1 min 30 s at 55°C (45°C in the case of degenerated primers *RS1*, *RS2*, *RAS3* and *RAS4*) for annealing and 2 min at 72°C for elongation. After this, a final step of 7 min at 72°C was carried out. Thirty-five cycles were applied in PCR using degenerated primers and 40 cycles were applied in PCR using specific primers. In few cases, 2 µl of the first PCR reaction product were subject to 30 additional cycles. According to the sample, 5–20 µl of PCR products were analysed on 1.5% agarose gel.

Cloning and sequencing

According to the instructions of the supplier (Promega, Madison, WI, USA), PCR fragments were inserted into pGEM®-T Easy vector and characterized by restriction analysis. Restriction enzymes were purchased from Eurogentec (Seraing, Belgium). Plasmid DNA of selected clones was purified using Plasmid Midi™ columns (Qiagen SA) and sequenced from T7 and SP6 primers by E.S.G.S. (Groupe Cybergene, Evry, France).

Sequence homology searches were done with BLAST 2.1 (NCBI, National Library of Medicine, Bethesda, MD, USA) in either *nr* database (‘all non-redundant GenBank CDS translations + PDB + SwissProt + PIR + PR’) or *htgs* [‘Unfinished High Throughput Genomic Sequences: phases 0, 1 and 2 (finished, phase 3 HTG sequences are in *nr* database)’], or with FASTA (Pearson and Lipman, 1988) in the Human Olfactory Receptor Data Exploratorium (HORDE) database (Weizmann Institute of Science, Rehovot, Israel). Sequence alignments were done with ClustalW (EMBL-EBI, Cambridge, UK). Data from the ClustalW alignment were treated with the NEIGHBOR program from the PHYLIP package (http://evolution.genetics.washington.edu/phylip.html) to construct a phylogenetic tree.

Results

Cloning of partial olfactory-like receptors cDNA from human tongue

In order to investigate the expression of GPCRs in human tongue, degenerated primers *RS1* or *RS2* as 5′ primer and *RAS3* or *RAS4* as 3′ primer were used in attempt to amplify OLR cDNAs in RT-PCR experiments. Taking advantage of the fact that all known OLR genes are intronless, different couples of primers (*RS1*/*RAS3*, *RS1*/*RAS4*, *RS2*/*RAS3* and *RS2*/*RAS4*) were tested first on human genomic DNA. The best results were obtained with *RS2*/*RAS4*, which allowed the amplification of a single-band amplicon product of ~520 bp (data not shown), a size corresponding to that which could be expected. Therefore, primers *RS2* and *RAS4* annealing with transmembrane domain 3 (TM3)/intracellular loop II junction and with TM7 coding sequence, respectively, were used for further experiments. PCR experiments were carried out on both human fetal tongue Gene Pool™ (ready-to-use RT reaction realized on polyA$^+$ RNAs) provided by Invitrogen and on RT reaction realized on total RNA from human adult lingual epithelial and sub-epithelial layers.

Because all the OLR genes known are intronless, it was important to verify the absence of contaminating genomic DNA. PCR using *RS2*/*RAS4* realized directly on adult tongue RNA without reverse transcription step did not show any product (Figure 1D, line 1), even after two successive rounds of PCR (Figure 1D, line 3). This result showed that the RNA extracted from adult tongue was DNA-free. In the case of fetal tongue, a Gene Pool™ provided by Invitrogen was used. This product is the result of a reverse transcription reaction made on polyA$^+$ RNAs by the supplier, therefore it was impossible to perform a PCR without reverse transcription as a control. The absence of contaminating genomic DNA in the Invitrogen Gene Pool™ was double-checked by amplification of β-actin, clathrin and GAPDH cDNAs. DNA-containing introns was absent for these three genes. Additionally, we used amplification of the ubiquitously expressed housekeeping gene GAPDH as another test for genomic DNA presence/absence on both cDNA sources. GAPDH cDNA was amplified using primers *GAPDH-5* and *GAPDH-3*. These primers were chosen because they allow the amplification of a band of 355 bp in the case of cDNA and a band of 855 bp in the case of contamination with genomic DNA. These two bands were observed when a PCR was performed on human genomic DNA (data not shown). A single band of ~350 bp was amplified from both fetal and adult tongue after 35 PCR cycles (Figure 1A,C). These results confirmed the absence of contaminating genomic DNA in both cases.

PCR using *RS2*/*RAS4* realized on human fetal tongue Gene Pool™ allowed the amplification of a single band of ~520 bp after 35 cycles (Figure 1B) whereas PCR realized on total RNA from human adult tongue epithelium led to the amplification of two major bands at ~520 and 900 bp and two very thin bands at 380 and 700 bp (Figure 1D, line 2). In the second case, because the intensity of the bands was low, the PCR product was subject to 30 additional PCR cycles using the same primers. The products corresponding to all four bands were slightly re-amplified (Figure 1D, line 4) and were cloned separately into pGEM®-T Easy vector as well as the product obtained from the fetal tongue.

Sequencing of 380, 700 and 900 bp inserts and the screening of BLAST *nr* database showed that they constitute

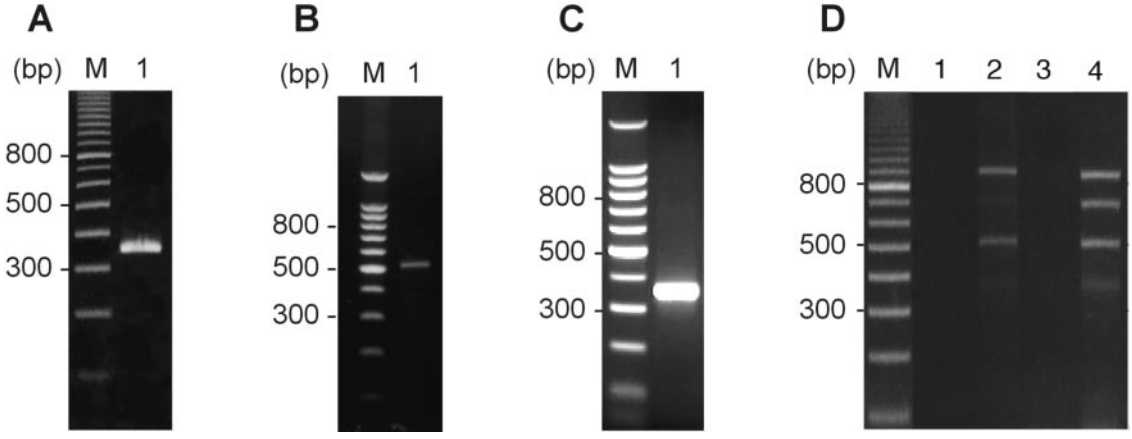

Figure 1 RT-PCR on human fetal tongue Gene Pool™ and human adult tongue epithelium RNA. Reverse transcription was performed on total RNA extracted from adult tongue epithelium with both random nonamer and oligo-dT primers. PCRs were carried out on both fetal tongue Gene Pool™ **(A, B)** and RT reaction products **(C, D)**. (A) and (C) PCR (35 cycles) using primers *GAPDH-5* and *GAPDH-3* specific for GAPDH (used as a control), 5 µl of PCR products were loaded on the gel. (B) and (D) (lanes 1 and 2): PCR using degenerated primers *RS2* and *RAS4* (35 cycles), 10 µl (B) or 20 µl (D) of PCR products were loaded on the gel. PCR products of (D) lanes 1 and 2 were subject to additional amplification (30 cycles) and 20 µl were loaded in lanes 3 and 4. Lanes 1 and 3 represent control reactions in which PCR was performed directly on RNA (without RT). The molecular weight marker (M) was a 100 bp ladder.

artefactual amplification due to the use of degenerated primers.

Inserts of size neighbouring 520 bp were analysed by restriction digestions. Their analysis demonstrated the presence of several different fragments. A total of 42 independent clones contained 35 different inserts (23 partial cDNA sequences from fetal tongue and 12 from adult tongue) were sequenced. Four inserts were present in two clones and one insert was present in four clones The sequences were translated according to six coding frames.

OLR prototype sequence was defined using an alignment of 88 members of the olfactory receptor family found in the GPCR Database (GPCRDB) (Horn *et al.*, 1998). The frequency of amino acid residues at a given position was analysed. This prototype sequence is shown in Figure 2 where conserved residues are reported and non-conserved residues are indicated by a star. Homology search of nucleotidic and deduced peptide sequences with BLAST in *nr* database and comparison with the prototype OLR sequence revealed that among cloned sequences, 12 (35%) were sequences unrelated to OLR (data not shown), 7 (20%) contained already published OLR sequences (cDNAs, genes or pseudogenes) and 16 (45%) were new cDNA sequences (Figure 2). Presence of artefactual fragments (e.g. KIAA-0782 protein, DNA polymerase alpha subunit, plakoglobin partial cDNAs) was mainly due to the use of degenerated primers leading to mispriming and amplification of non-OLR-related sequences. It was found that the sequence of clone RC86 was identical to partial HGMP07I cDNA sequence (Parmentier *et al.*, 1992) and to a part of the complete OR1E1 gene sequence (Glusman *et al.*, 2000); clone JCG8 sequence was identical to partial TPCR85 cDNA sequence (Vanderhaeghen *et al.*, 1997); clone RC254 sequence was identical to partial HTPCR06 cDNA sequence (Parmentier *et al.*, 1992); clone RC212 sequence was identical to partial TPCR120 cDNA sequence (Vanderhaeghen *et al.*, 1997); clone RC70 was identical to partial TPCR24 pseudogene cDNA sequence (Vanderhaeghen *et al.*, 1997); clone RC95 sequence was identical to partial OR7E13P pseudogene sequence (Buettner *et al.*, 1998) and clone RC183 sequence was identical to partial OR7-86 pseudogene sequence (Rouquier *et al.*, 1998). All five sequences of TPCR24, TPCR85, TPCR120, HGMP07I and HTPCR06 were originally cloned by RT-PCR from male germinal cell mRNA whereas OR7E13P and OR7-86 sequences were originally cloned or identified from chromosome 11 genomic DNA. All the other clones named JCG1, JCG2, JCG3, JCG4, JCG5, JCG6, JCG9, JCG10, PJCG1, PJCG3, PJCG4, PJCG5, PJCG6, PJCG7, PJCG8 and PJCG9 contained previously unreported new partial cDNA sequences. [GenBank accession number of the sequences: HGMP07I/OR10E1 (AF087916, X64994); HTPCR06 (X64977); TPCR120 (X89669); TPCR24 (AF309699, AF309702); OR7-86 (U86282); OR7E13P (AF065855, AF238487); JCG1 (AF158377); JCG2 (AF162668, AF162669); JCG3 (AF173006); JCG4 (AF220494); JCG5 (AF209506); JCG6 (AF065874); TPCR85 (AF238488); JCG9 (AF238488); JCG10 (AF308814); PJCG1 (AF220493); PJCG3 (AF309700, AF309701); PJCG4 (AF359415); PJCG5 (AF209507); PJCG6 (AF356416); PJCG7 (AF359417); PJCG8 (AF359418) and PJCG9 (AF359419).]

Partial OLR peptide sequences obtained by translation of *RS2*/*RAS4* cDNA fragments were compared to the OLR prototype sequence. They could be classified in two categories (Figure 2A): (i) sequences containing uninterrupted open reading frames (ORF) homologous to OLR prototype: TPCR24, HGMP07I/OR1E1, OR7E13P, TPCR85, OR7-86, TPCR120, HTPCR06, JCG1, JCG2, JCG3, JCG4, JCG5, JCG6, JCG9, JCG10 and PJCG1; and (ii) sequences containing interrupted ORF or ORF with changed reading frame in respect to the defined OLR prototype: PJCG6, PJCG4, PJCG5, PJCG6, PJCG7, PJCG8 and PJCG9. The former probably constitute parts of expressed pseudogenes. It is important to highlight that, in spite of their classification in the first category, sequences OR7E13P and OR7-86 were already reported as pseudogenes (Buettner *et al.*, 1998; Rouquier *et al.*, 1998). Indeed, stop codons were found outside of the cloned portion of their cDNA.

Homology searches in *htgs* and HORDE databases were used in order to identify full-length sequences of the genes or pseudogenes corresponding to partial cDNA fragments. The genes corresponding to JCG4, JCG10, PJCG8 and PJCG9 (all of them new) could not be found. PJCG8 and PJCG9 are pseudogenes very similar to previously described OR2-53, OR2-75, OR2-4 and OR2-52 (Rouquier *et al.*, 1998). All these pseudogenes are very similar to each other (98–99% identity). Partial JCG4 is very similar to partial JCG3, showing 96% identity at the protein level. However, full-length coding sequences of the genes or pseudogenes corresponding to all the other cloned fragments were identified.

Full-length coding sequences of HGMP07I/OR1E1, TPCR85, HTPCR06; JCG1, JCG2, JCG3, JCG5, JCG6 and JCG9 cDNAs were cloned from the fetal tongue Gene Pool™. They ranged from 927 bp to 969 bp, so the deduced protein sequences ranged from 308 amino acids (aa) to 322 aa (Figure 2A). All found sequences agreed well with the theoretical sequences extracted from human genome sequencing data. Nevertheless, in the case of HGMP07I/OR1E1 a point mutations was observed between the sequences described in this paper and the sequences available in databanks.

The full-length coding sequence of HGMP07I/OR1E1 (HORDE name OR1E1) is 945 bp long. The deduced protein sequence is 314 aa long. The cDNA sequence cloned from fetal tongue contains one point mutation at position 771 in respect to genomic one (OR1E1 gene and clone RP11-587F22). Because this was found in partial cDNA clones from embryonic tongue, in full-length cDNA clones and by direct sequencing of PCR products from adult

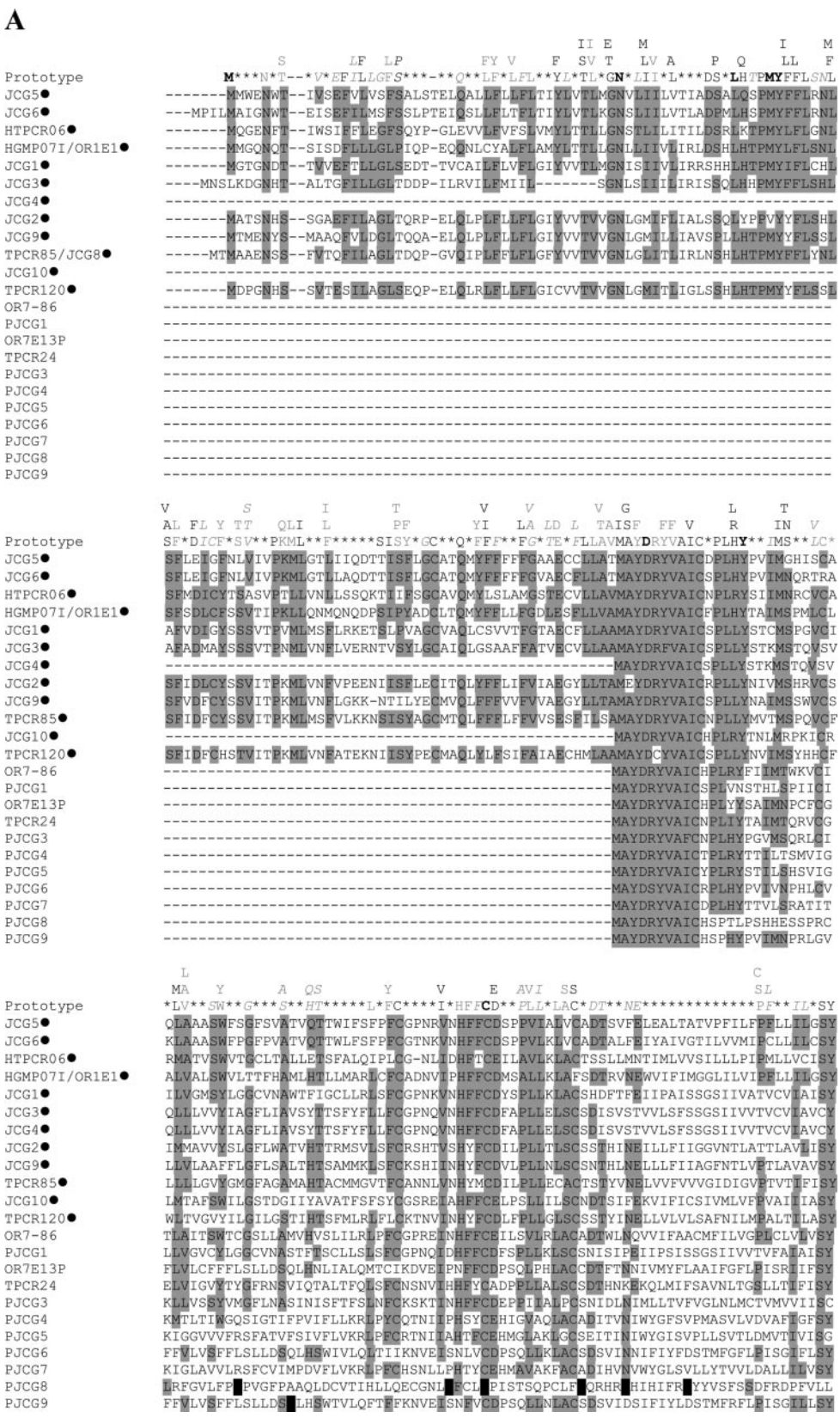
A
Prototype
JCG5●
JCG6●
HTPCR06●
HGMP07I/OR1E1●
JCG1●
JCG3●
JCG4●
JCG2●
JCG9●
TPCR85/JCG8●
JCG10●
TPCR120●
OR7-86
PJCG1
OR7E13P
TPCR24
PJCG3
PJCG4
PJCG5
PJCG6
PJCG7
PJCG8
PJCG9

Figure 2 (A)

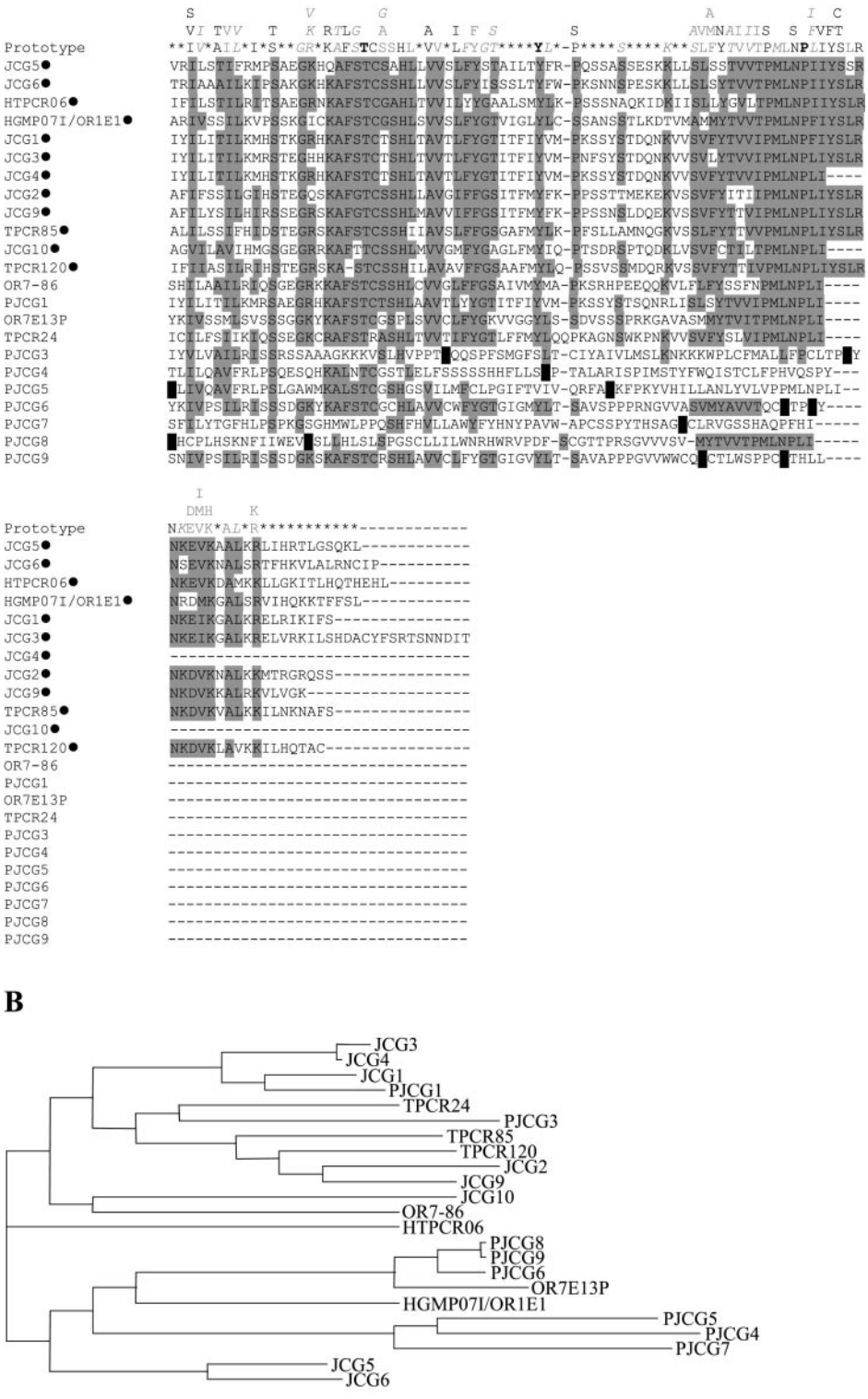

Figure 2 Comparison of deduced protein and cDNA ORL sequences found in human tongue. **(A)** Alignment of protein sequences deduced from partial and full-length coding cDNA sequences cloned from fetal and adult tongue with an OLR prototype defined according to alignment of 88 sequences found in GPCRdb. Bold letters of the OLR prototype sequence are 100% conserved residues on a position, black letters are 95–99% conserved residues or residues for which the total of representation is 95–100%, grey letters are 80–94% conserved residues or residues for which the total of representation is 80–94%, grey italic letters are 50–79% conserved residues or residues for which the total of representation is 60–79%, and stars are non-conserved residues. Grey boxes indicate amino acid residues conserved as compared with conserved residues found in OLR prototype sequence. Black boxes indicate stop codons. Gene names are labelled with a black circle, pseudogene names are not labelled. **(B)** Comparison of full-length or partial OLR genes and pseudogenes sequences expressed in human tongue. The dendrogram was established by pairwise sequence alignment using ClustalW followed by data treatment with the NEIGHBOR program from the PHYLIP package (http://evolution.genetics.washington.edu/phylip.html).

tongue it can be considered to be a result of polymorphism. Additionally, this silent mutation does not modify the sequence of the corresponding protein.

The full-length coding sequence of TPCR85 (HORDE name OR8B8) is 936 bp long. It is identical to the genomic sequence contained in clones RP11-728D14 and RP11-164A10. Its deduced protein sequence is 311 aa long. Nevertheless, its partial cDNA clone obtained by Vanderhaeghen and co-workers contains two point mutations when compared to the genomic sequence leading to $Gly_{204} \rightarrow Ser$ conversion.

The full-length coding sequence of HTPCR06 (HORDE name OR2K2) is 951 bp long. Its deduced protein sequence is 316 aa long. Its cDNA sequence cloned from fetal tongue is identical to the genomic sequence contained in clones RP11-17E20 and RP11-386D8, whereas the partial cDNA sequence cloned by Parmentier *et al.* contains one point mutation at position 632 ($Leu_{211} \rightarrow Pro$).

The full-length coding sequence of JCG1 (HORDE name OR5P3) is 936 bp long. Its deduced protein sequence is 311 aa long. Its full length cDNA sequence is identical to those found in genomic clones RP11-799H15, RP11-494M8 and RP11-399N15.

The full-length coding sequence of JCG2 (HORDE name OR8D2) is 936 bp long. Its deduced protein sequence is 311 aa long. cDNA sequence cloned from fetal tongue is identical to genomic sequence found in clones pDJ9j14 and RP11-164A10.

The full-length coding sequence of JCG3 (HORDE name OR5P2) is 969 bp long. Its deduced protein sequence is 322 aa long. The full-length cDNA sequence is identical to the sequence found in clones RP11-799H15, RP11-494M8 and RP11-399N15. There are 12 differences at the nucleotide level and six differences at the protein level between JCG4 partial cDNA sequence and JCG3 sequence in an area that does not include the sequence of the degenerated primers. Even if point mutations could be introduced by the use of a non-proofreading Taq-polymerase, the introduction of 12 mutations in a 472 bp sequence is very surprising. Because PCR fragments containing full-length JCG4 were not found, it can be suspected that differences between JCG4 and JCG3 are probably due to PCR artefacts. Nevertheless, it can not be completely excluded that JCG4 could be a different gene whose genomic sequence is still unknown.

The full-length coding sequence of JCG5 (HORDE name OR10A4) is 948 bp long. Its deduced protein sequence is 315 aa long. The full-length cDNA clone sequence contains one point mutation at position 880 in respect to the sequence of genomic clone RP11-560B16. By direct sequencing of the PCR fragment, two bases, A and C, were detected at this position in similar amounts. Consequently, the difference observed at position 880 is probably due to polymorphism, two different alleles being expressed. Nevertheless, this difference represents silent polymorphism because the protein sequence is unchanged. At position 617 the base found in genomic clone pDJ610i20 is different than in all other sequences leading to $Leu_{206} \rightarrow Pro$ conversion. If the sequence of clone PDJ610i20 is correct, a point mutation at position 617 can be assigned to polymorphism at this position as well. As well as this difference between the two genomic clones, clone RP11-560B16 contains deletion at position 351 when compared to PDJ610i20. This deletion is probably due to a sequencing mistake.

The full-length coding sequence of JCG6 (HORDE name OR10A5) is 954 bp long. The deduced protein sequence is 317 aa long. This sequence is identical to genomic sequence found in genomic clones RP11-560B16 and PDJ610i20.

The full-length coding sequence of JCG9 (HORDE name OR8D1) is 927 bp long. Its deduced protein sequence is 308 aa long. The full-length cDNA sequence is identical to genomic sequence contained in clone RP11-164A10.

All these sequences, when compared to the OLR prototype, could be classified in three groups (Figure 2A):

1. Previously identified genes encoding OLR proteins: TPCR85, HGMP07I/OR1E1, HTPCR06 and TPCR120 and new OLR genes: JCG1, JCG2, JCG3, JCG5, JCG6 and JCG9.
2. OLR pseudogenes sequences presenting the same reading frame as OLR genes but containing one or several stop codons located upstream from the stop codon normally used: TPCR24, PJCG5 and OR7-86.
3. OLR pseudogenes containing insertions or deletions leading to changes of reading frame: OR7E13P, PJCG1, PJCG3, PJCG4, PJCG6, PJCG7, PJCG8 and PJCG9.

Remarkably, 15 out of 19 genes or pseudogenes corresponding to cloned cDNAs are located on chromosome 11, some very close to each other. For instance, JCG2, TPCR85, JCG9 and TPCR120 are located on 11q25 between 137.68 megabases (Mb) and 137.96 Mb. JCG2, TPCR85 and JCG9 are localized in the same genomic clone RP11-164A10. The interval between JCG2 and JCG9 amounts to only 8640 bp. A phylogenetic tree based on nucleotide sequence conservation showed that they are closely related and derive from the same common ancestor gene (Figure 2B). JCG1, JCG3, JCG5, JCG6, TPCR24, PJCG1, PJCG4, PJCG5 and PJCG7 are located between 2.78 Mb and 6.93 Mb on chromosome 11. JCG1, JCG3 and TPCR24 are found in the same genomic clones (RP11-799H15, RP11-494M8 and RP11-399N15). This could indicate that one or several clusters of OLR genes located on chromosome 11 are expressed in tongue epithelium. OR13E7P and PJCG3 are also located on this chromosome.

Expression of OLR genes in adult and fetal human tongue

Expression of TPCR85, HGMP07I/OR1E1, HTPCR06, JCG1, JCG2, JCG3, JCG5, JCG6 and JCG9 was studied by RT-PCR on both human fetal and human adult tongue (Figure 3). For each gene, specific primers allowing the amplification of the complete theoretical coding sequence

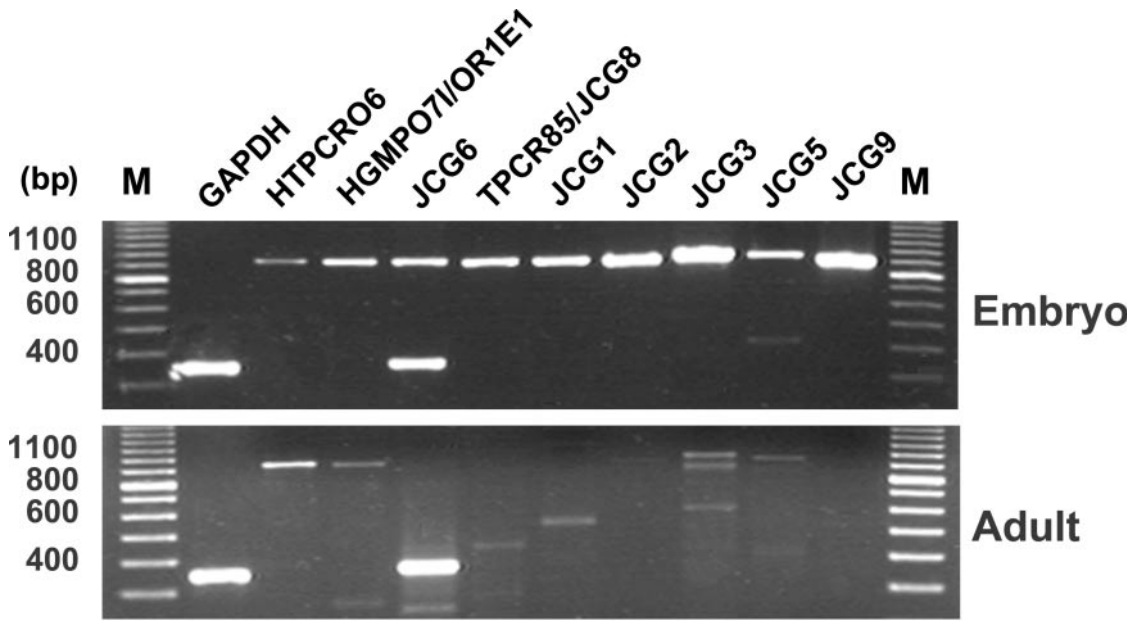

Figure 3 Detection of OLR transcripts by RT-PCR in fetal and adult tongue. RT-PCR was performed using specific primers for partial GAPDH cDNA and for full-length coding sequence of HGMP07I/OR1E1, TPCR85, HTPCR06, JCG1, JCG2, JCG3/JCG4, JCG5, JCG6 and JCG9 genes on both fetal and adult human tongue mRNA. In the case of fetal tongue 10 μl of PCR product were loaded, in the case of adult tongue 20 μl were loaded, except for GAPDH where 5 μl were loaded. The molecular weight marker (M) was a 100 bp ladder.

were used (the 5′ primer included the start codon and the 3′ primer included the stop codon). Because JCG3 was found to be very similar to JCG4 and because the JCG4 full-length coding sequence was not identified, primers chosen for JCG3 amplification would also have allowed the amplification of the full-length coding sequence. The sizes of the expected bands range from 927 to 969 bp. GAPDH was used as a positive control of the PCR.

During the experiment, PCRs allowed the amplification of fragments representing the expected size. However, in the case of JCG5 with fetal mRNA, TPCR85/JCG8, JCG1, JCG3 with adult mRNA and JCG6 with both mRNA, additional bands of lower molecular weight were observed. Cloning, restriction analysis and sequencing of these products showed that they are artefactual PCR fragments due to mispriming (e.g. myosin heavy chain in the case of JCG6).

It was found that all the tested genes are expressed in fetal tongue whereas only HTPCR06, OR1E1/HGMP07I, JCG3/JCG4 and JCG5 are expressed in epithelium of adult human tongue. The bands amplified from adult tissues were very thin, indicating that these mRNAs are probably very poorly expressed. The band for HTPCR06 had the lowest intensity on the gel corresponding to fetal tissues and the highest intensity on the gel corresponding to the adult tissues, suggesting that HTPCR06 is preferentially expressed in adults.

In order to confirm that the bands observed between 900 and 1000 bp markers in the previous experiment correspond well to cDNA of the genes of interest, PCR fragments obtained from fetal tongue were cloned into pGEM®-T Easy vector and sequenced. The cDNA sequences obtained were compared to both genomic and partial cDNA corresponding sequences. The results confirmed that the obtained PCR fragments correspond to the genes of interest. In the case of the amplification of JCG3/JCG4, only full-length coding sequence of JCG3 was found among three independent clones and by direct sequencing of the PCR product.

Discussion

The present paper describes the cloning of a subset of previously described and new OLR cDNAs expressed in tongue. It was shown that several OLR genes and pseudogenes are expressed in human fetal and adult tongue. Nevertheless, it was found that in adult human tongue, the number of these genes or pseudogenes expressed is much lower than in fetal lingual tissue. One possible reason for the observed differences could be differentiation of gene expression during maturation of this tissue. It was proposed that olfactory receptors might have other functions in addition to simple odorant detection [for a review see Dryer and Berghard (Dryer and Berghard, 1999)]. One of these functions could be axon guidance. Such a phenomenon could occur in tongue tissues during the establishment of neuronal connections between fetal taste receptors. However, this does not exclude the possibility that OLR proteins could also play a role in the recognition of sapid compounds.

It is also possible that the observed differences are related to the origins of the samples used for RNA extraction. Available fetal tongue Gene Pool™ was prepared from total

tongue, hence it was not the optimal source to search for genes expressed specifically in epithelium or in a precise area of the tongue. So, it is also possible that some of the cloned genes are expressed, for instance, in the muscular layers underneath the epithelial layer of the tongue, in its laropharyngeal part or in distinct areas of the lingual epithelium. Available sample from adult tongue was histologically much more precise since it contained only epithelial and sub-epithelial tissues sampled at the extremity of the tongue.

Surprisingly, some of the OLR mRNAs described in this paper found in tongue (HGMP07I, HTPCR06, TPCR24, TPCR85, TPCR120) were isolated for the first time from mammalian male germinal cells (Parmentier *et al.*, 1992; Vanderhaeghen *et al.*, 1997). Expression of the OLR gene in both taste and reproductive tissues was also reported in rats (Thomas *et al.*, 1996). This is also the case of adenylate cyclase type 3 (AC3) and the olfactory G protein subunit Gαolf (Defer *et al.*, 1998). Both are involved in olfactory signalling pathway. Even if several elements of the olfactory signalling pathway are present in male germ cells of several mammalian species, the function of OLR proteins in these cells remains unclear.

Point mutations were observed between new, partial or full-length cloned sequences and previously cloned sequences or genomic sequences—this is the case for HGMP-07I/OR10E1, TPCR85, HTPCR06 and JCG5. In the case of JCG5, the difference observed is most probably due to genetic polymorphism whose intensity was already reported within a cluster of 15 olfactory receptor genes located on chromosome 17 (Gilad *et al.*, 2000; Sharon *et al.*, 2000). Another illustration of such polymorphisms is that, in some cases, the same gene contained in two genomic clones exhibited one or several point mutations. In the case of HGMP07I/OR10E1, TPCR85 and HTPCR06, their sequences found are identical to the genomic one, whereas they vary slightly from previously cloned partial cDNA sequences. These point mutations can also be due to PCR artefacts.

The results presented show that some OLR pseudogenes are expressed in human tongue. Expression of pseudogenes is a rather uncommon phenomenon, which, however, has already been reported in the case of OLR. An OLR pseudogene located on chromosome 17 was found to be expressed in human olfactory tissue (Crowe *et al.*, 1996). TPCR24 from male germ cells mRNA first cloned in 1997 by Vanderhaeghen *et al.* (Vanderhaeghen *et al.*, 1997) and described as an OLR gene is in fact a pseudogene. One of the models of OLR regulation proposes that entire clusters of OLR genes could be regulated by common *cis*-acting elements. Because OLR pseudogenes are included mostly in such clusters, it is conceivable that the expression of OLR pseudogenes can be maintained by such a mechanism of regulation.

Among the genes and pseudogenes cloned in this study, 15 (eight genes and seven pseudogenes) are located on chromosome 11. They are probably parts of one or of several clusters. The fact that they are expressed together in fetal tongue could support this hypothesis. Nevertheless, out of all of them, only expression of JCG3 and JCG5 was detected in adult lingual tissues. This suggests that other regulation mechanisms could occur and/or that the expression levels of the other genes were too low to be detected in applied experimental conditions.

Recently, a new family of GPCRs expressed specifically in lingual tissue and containing receptors for bitter taste in humans was described (Hoon *et al.*, 1999; Adler *et al.*, 2000; Chandrashekar *et al.*, 2000; Matsunami *et al.*, 2000). Lingual OLRs identified in this study do not display significant homologies with any member of this class of bitter taste receptors. Nevertheless, the possibility of involvement of OLR proteins in taste perception can not be excluded. More intensive work would be needed confirm such a possibility. A histological localization study of the OLR transcripts, especially in taste bud cells, would be of great help in assessing their true functional relevance.

References

Abe, K., Kusakabe, Y., Tanemura, K., Emori, Y. and **Arai S.** (1993a) *Multiple genes for G protein-coupled receptors and their expression in lingual epithelia.* FEBS Lett., 316, 253–256.

Abe, K., Kusakabe, Y., Tanemura, K., Emori, Y. and **Arai, S.** (1993b) *Primary structure and cell-type specific expression of a gustatory G protein-coupled receptor related to olfactory receptors.* J. Biol. Chem., 268, 12033–12039.

Adler, E., Hoon, M.A., Mueller, K.L., Chandrashekar, J., Ryba, N.J. and **Zuker, C.S.** (2000) *A novel family of mammalian taste receptors.* Cell, 100, 693–702.

Buettner, J.A., Glusman, G., Ben-Arie, N., Ramos, P., Lancet, D. and **Evans, G.A.** (1998) *Organization and evolution of olfactory receptor genes on human chromosome 11.* Genomics, 53, 56–68.

Chandrashekar, J., Mueller, K.L., Hoon, M.A., Adler, E., Feng, L., Guo, W., Zuker, C.S. and **Ryba, N.J.** (2000) *T2Rs function as bitter taste receptors*. Cell, 100, 703–711.

Chaudhari, N. and **Roper, S.D.** (1998) *Molecular and physiological evidence for glutamate (umami) taste transduction via a G protein-coupled receptor*. Ann. N.Y. Acad. Sci., 855, 398–406.

Chaudhari, N., Landin, A.M. and **Roper, S.D.** (2000) *A metabotropic glutamate receptor variant functions as a taste receptor.* Nat. Neurosci., 3, 113–119.

Crowe, M.L., Perry, B.P. and **Connerton, I.F.** (1996) *Olfactory receptor-encoding genes and pseudogenes are expressed in humans.* Gene, 169, 247–249.

Defer, N., Marinx, O., Poyard, M., Lienard, M.O., Jegou, B. and **Hanoune, J.** (1998) *The olfactory adenylyl cyclase type 3 is expressed in male germ cells.* FEBS Lett., 424, 216–220.

Dryer, L. and **Berghard, A.** (1999) *Odorant receptors, a plethora of G-protein-coupled receptors.* Trends Pharmacol. Sci., 20, 413–417.

Gilad, Y., Segre, D., Skorecki, K., Nachman, M. W., Lancet, D. and **Sharon, D.** (2000) *Dichotomy of single-nucleotide polymorphism*

haplotypes in olfactory receptor genes and pseudogenes. Nat Genet., 26, 221–224.

Glusman, G., Sosinsky, A., Ben-Asher, E., Avidan, N., Sonkin, D., Bahar, A., Rosenthal, A., Clifton, S., Roe, B., Ferraz, C., Demaille, J. and **Lancet, D.** (2000) *Sequence, structure, and evolution of a complete human olfactory receptor gene cluster.* Genomics, 63, 227–245.

Hoon, M.A., Adler, E., Lindemeier, J., Battey, J.F., Ryba, N.J. and **Zuker, C.S.** (1999) *Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity.* Cell, 96, 541–551.

Horn, F., Weare, J., Beukers, M.W., Hörsch, S., Bairoch, A., Chen, W., Edvardsen, O., Campagne, F. and **Vriend, G.** (1998) *GPCRDB, an information system for G protein-coupled receptors.* Nucleic Acids Res., 26, 277–281.

Kitagawa, M., Kusakabe, Y., Miura, H., Ninomiya, Y. and **Hino, A.** (2001) *Molecular genetic identification of a candidate receptor gene for sweet taste.* Biochem. Biophys. Res. Commun., 283, 236–242.

Kusakabe, Y., Abe, K., Tanemura, K., Emori, Y. and **Arai, S.** (1996) *GUST27 and closely related G-protein-coupled receptors are localized in taste buds together with Gi-protein alpha-subunit.* Chem. Senses, 1, 335–340.

Lin, W. and **Kinnamon, S.C.** (1999) *Physiological evidence for ionotropic and metabotropic glutamate receptors in rat taste cells.* J. Neurophysiol., 82, 2061–2069.

Matsunami, H., Montmayeur, J.P. and **Buck, L.B.** (2000) *A family of candidate taste receptors in human and mouse.* Nature, 404, 601–604.

Matsuoka, I., Mori, T., Aoki, J., Sato, T. and **Kurihara, K.** (1993) *Identification of novel members of G-protein coupled receptor superfamily expressed in bovine taste tissue.* Biochem. Biophys. Res. Commun., 194, 504–511.

Max, M., Shanker, Y.G., Huang, L., Rong, M., Liu, Z., Campagne, F., Weinstein, H., Damak, S. and **Margolskee, R.F.** (2001) Tas1r3*, encoding a new candidate taste receptor, is allelic to the sweet responsiveness locus* Sac*.* Nat. Genet., 28, 58–63.

Montmayeur, J.P., Liberles, S.D., Matsunami, H. and **Buck, L.B.** (2001) *A candidate taste receptor gene near a sweet taste locus.* Nat. Neurosci., 4, 492–498.

Parmentier, M., Libert, F., Schurmans, S., Schiffmann, S., Lefort, A., Eggerickx, D., Ledent, C., Mollereau, C., Gerard, C., Perret, J. ***et al.*** (1992) *Expression of members of the putative olfactory receptor gene family in mammalian germ cells.* Nature, 355, 453–455.

Pearson, W.R. and **Lipman, D.J.** (1988) *Improved tools for biological sequence comparison.* Proc. Natl Acad. Sci. USA, 85, 2444–2448.

Rouquier, S., Taviaux, S., Trask, B.J., Brand-Arpon, V., van den Engh, G., Demaille, J. and **Giorgi, D.** (1998) *Distribution of olfactory receptor genes in the human genome.* Nat. Genet., 18, 243–250.

Sainz, E., Korley, J.N., Battey, J.F. and **Sullivan, S.L.** (2001) *Identification of a novel member of the T1R family of putative taste receptors.* J. Neurochem., 77, 896–903.

Sharon, D., Gilad, Y., Glusman, G., Khen, M., Lancet, D. and **Kalush, F.** (2000) *Identification and characterization of coding single-nucleotide polymorphisms within a human olfactory receptor gene cluster.* Gene, 260, 87–94.

Thomas, M.B., Haines, S.L. and **Akeson, R.A.** (1996) *Chemoreceptors expressed in taste, olfactory and male reproductive tissues.* Gene, 178, 1–5.

Vanderhaeghen, P., Schurmans, S., Vassart, G. and **Parmentier, M.** (1997) *Specific repertoire of olfactory receptor genes in the male germ cells of several mammalian species.* Genomics, 39, 239–246.

Accepted July 14, 2001