

The Molecular Basis for Odor Recognition

Stuart Firestein,^{1,*} Charles Greer,² and Peter Mombaerts³

¹Department of Biological Sciences, Columbia University, New York, NY 10027, USA

²Department of Neuroscience, Yale University, New Haven, CT 06520-8082, USA

³Max Planck Research Unit for Neurogenetics, Frankfurt 60438, Germany

*Correspondence: stuart@biology.columbia.edu

We are pleased to present a series of Annotated Classics celebrating 40 years of exciting biology in the pages of *Cell*. This installment revisits “A Novel Multigene Family May Encode Odorant Receptors: A Molecular Basis for Odor Recognition” by Linda Buck and Richard Axel. Here, Stuart Firestein, Charles Greer, and Peter Mombaerts comment on many discoveries foreshadowed by the identification of odorant receptors as GPCRs, the technical challenges for working with them, and as yet unanswered questions in the field.

Each Annotated Classic offers a personal perspective on a groundbreaking *Cell* paper from a leader in the field with notes on what stood out at the time of first reading and retrospective comments regarding the longer term influence of the work.

To see Stuart Firestein and colleagues's thoughts on different parts of the manuscript, just download the PDF and then hover over or double-click the highlighted text and comment boxes. You can also view the annotation by opening the Comments navigation panel in Acrobat.

A Novel Multigene Family May Encode Odorant Receptors: A Molecular Basis for Odor Recognition

Linda Buck* and Richard Axel

*Department of Biochemistry and Molecular Biophysics
†Howard Hughes Medical Institute
College of Physicians and Surgeons
Columbia University
New York, New York 10032

Summary

The mammalian olfactory system can recognize and discriminate a large number of different odorant molecules. The detection of chemically distinct odorants presumably results from the association of odorous ligands with specific receptors on olfactory sensory neurons. To address the problem of olfactory perception at a molecular level, we have cloned and characterized 18 different members of an extremely large multigene family that encodes seven transmembrane domain proteins whose expression is restricted to the olfactory epithelium. The members of this novel gene family are likely to encode a diverse family of odorant receptors.

Introduction

In vertebrate sensory systems, peripheral neurons respond to environmental stimuli and transmit these signals to higher sensory centers in the brain where they are processed to allow the discrimination of complex sensory information. The delineation of the peripheral mechanisms by which environmental stimuli are transduced into neural information can provide insight into the logic underlying sensory processing. Our understanding of color vision, for example, emerged only after the observation that the discrimination of hue results from the blending of information from only three classes of photoreceptors (Rushton, 1955, 1965; Wald et al., 1955; Nathans et al., 1986). The basic logic underlying olfactory sensory perception, however, has remained elusive. Mammals possess an olfactory system of enormous discriminatory power (for reviews see Lancet, 1986; Reed, 1990). Humans, for example, are thought to be capable of distinguishing among thousands of distinct odors. The specificity of odor recognition is emphasized by the observation that subtle alterations in the molecular structure of an odorant can lead to profound changes in perceived odor.

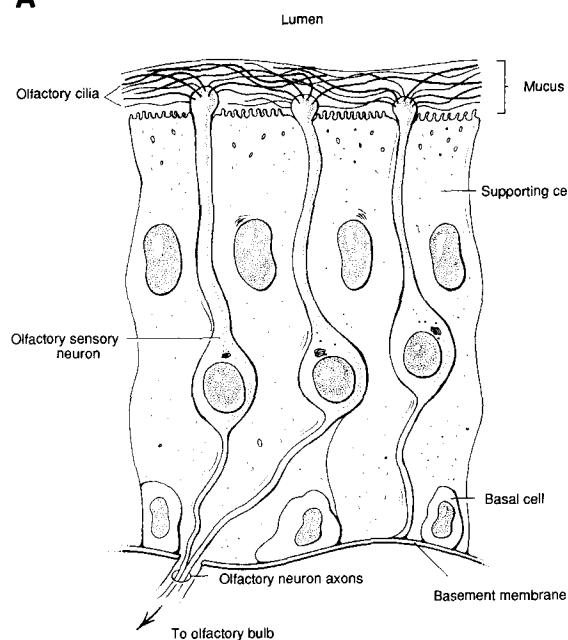
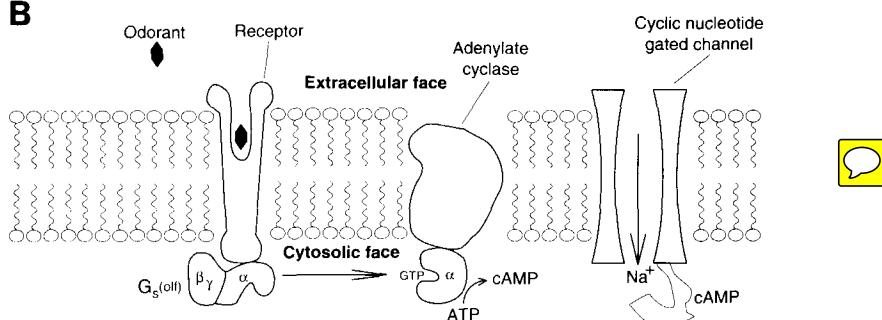
How are the diversity and specificity of olfactory perception accomplished? The detection of chemically distinct odorants presumably results from the association of odorous ligands with specific receptors on olfactory neurons, which reside in a specialized epithelium in the nose. Since these receptors have not been identified, it has been difficult to determine how odor discrimination might be achieved. It is possible that olfaction, by analogy with color vision, involves only a few odor receptors, each capable of interaction with multiple odorant molecules. Alternatively,

the sense of smell may involve a large number of distinct receptors each capable of associating with one or a small number of odorants. In either case, the brain must distinguish which receptors or which neurons have been activated to allow the discrimination between different odorant stimuli. Insight into the mechanisms underlying olfactory perception is likely to depend upon the isolation of the odorant receptors and the characterization of their diversity, specificity, and patterns of expression.

The primary events in odor detection occur in a specialized olfactory neuroepithelium located in the posterior recesses of the nasal cavity. Three cell types dominate this epithelium (Figure 1A): the olfactory sensory neuron, the sustentacular or supporting cell, and the basal cell, which is a stem cell that generates olfactory neurons throughout life (Moulton and Beidler, 1967; Graziadei and Monti Graziadei, 1979). The olfactory sensory neuron is bipolar: a dendritic process extends to the mucosal surface, where it gives rise to a number of specialized cilia that provide an extensive, receptive surface for the interaction of odors with the cell. The olfactory neuron also gives rise to an axon that projects to the olfactory bulb of the brain, the first relay in the olfactory system. The axons of the olfactory bulb neurons, in turn, project to subcortical and cortical regions where higher-level processing of olfactory information allows the discrimination of odors by the brain.

The initial events in odor discrimination are thought to involve the association of odors with specific receptors on the cilia of olfactory neurons. Selective removal of the cilia results in the loss of olfactory responses (Bronshtein and Minor, 1977). Moreover, in fish, whose olfactory system senses amino acids as odors, the specific binding of amino acids to isolated cilia has been demonstrated (Rhein and Cagan, 1980, 1983). The cilia are also the site of olfactory signal transduction. Exposure of isolated cilia from rat olfactory epithelium to numerous odorants leads to the rapid stimulation of adenylyl cyclase and elevations in cyclic AMP (an elevation in inositol trisphosphate in response to one odorant has also been observed) (Pace et al., 1985; Sklar et al., 1986; Breer et al., 1990; Boekhoff et al., 1990). The activation of adenylyl cyclase is dependent on the presence of GTP and is therefore likely to be mediated by receptor-coupled GTP-binding proteins (G proteins) (Jones and Reed, 1989). Elevations in cyclic AMP, in turn, are thought to elicit depolarization of olfactory neurons by direct activation of a cyclic nucleotide-gated, cation-permeable channel (Nakamura and Gold, 1987; Dhallan et al., 1990). This channel is opened upon binding of cyclic nucleotides to its cytoplasmic domain, and can therefore transduce changes in intracellular levels of cyclic AMP into alterations in the membrane potential.

These observations suggest a pathway for olfactory signal transduction (Figure 1B) in which the binding of odors to specific surface receptors activates specific G proteins. The G proteins then initiate a cascade of intracellular signaling events leading to the generation of an action potential that is propagated along the olfactory sensory axon

A**B**

to the brain. A number of neurotransmitter and hormone receptors that transduce intracellular signals by activation of specific G proteins have been identified. Gene cloning has demonstrated that each of these receptors is a member of a large superfamily of surface receptors that traverse the membrane seven times (for reviews see O'Dowd et al., 1989b; Strader et al., 1989). The pathway of olfactory signal transduction (Figure 1B) predicts that the odorant receptors might also be members of this superfamily of receptor proteins. The detection of odors in the periphery is therefore likely to involve signaling mechanisms shared by other hormone or neurotransmitter systems, but the vast discriminatory power of the olfactory system will require higher-order neural processing to permit the perception of individual odors. To address the problem of olfactory perception at a molecular level, we have cloned and characterized 18 different members of an extremely large multigene family that encodes seven transmembrane domain proteins whose expression is restricted to the olfactory epithelium. The members of this novel gene family are likely to encode the individual odorant receptors.

Figure 1. The Olfactory Neuroepithelium and a Pathway for Olfactory Signal Transduction
(A) The olfactory neuroepithelium. The initial events in odor perception occur in the nasal cavity in a specialized neuroepithelium that is diagrammed here. Odors are believed to interact with specific receptors on the cilia of olfactory sensory neurons. The signals generated by these initial binding events are propagated by olfactory neuron axons to the olfactory bulb.
(B) A pathway of olfactory signal transduction. In this scheme the binding of an odorant molecule to an odor-specific transmembrane receptor leads to the interaction of the receptor with a GTP-binding protein ($G_s(\text{olf})$). This interaction in turn leads to the release of the GTP-coupled α subunit of the G protein, which then stimulates adenylyl cyclase to produce elevated levels of cAMP. The increase in cAMP opens cyclic nucleotide-gated cation channels, thus causing an alteration in membrane potential.

Results

Experimental Strategy

The experimental design we employed to isolate genes encoding odorant receptors was based on three assumptions: First, the odorant receptors are likely to belong to the superfamily of receptor proteins that transduce intracellular signals by coupling to GTP-binding proteins. Second, the large number of structurally distinct odorous molecules suggests that the odorant receptors themselves should exhibit significant diversity and are therefore likely to be encoded by a multigene family. Third, expression of the odorant receptors should be restricted to the olfactory epithelium.

To identify molecules in the olfactory epithelium that resemble members of the seven transmembrane domain superfamily, homologs of this gene superfamily were amplified from olfactory epithelium RNA using the polymerase chain reaction (PCR). We then asked whether any of the PCR products we obtained consisted of a mixture of DNA sequences, consistent with the amplification of mem-

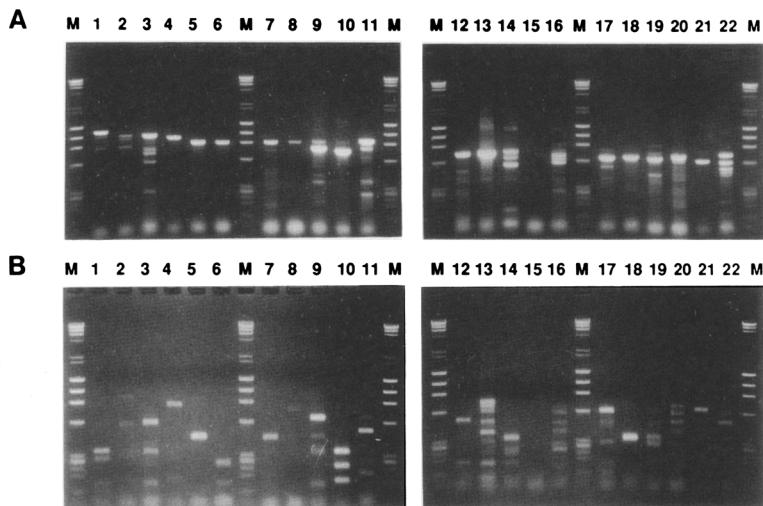


Figure 2. A PCR Amplification Product Containing Multiple Species of DNA

cDNA prepared from olfactory epithelium RNA was subjected to PCR amplification with a series of different primer oligonucleotides; the DNA products of appropriate size were isolated, further amplified by PCR, and size fractionated on agarose gels (A) (for details see text). Each of these semipurified PCR products was digested with the restriction enzyme *Hinf*I and analyzed by agarose gel electrophoresis (B). Lanes M contain size markers of 23.1, 9.4, 6.6, 4.4, 2.3, 2.0, 1.35, 1.08, 0.87, 0.60, 0.31, 0.28, 0.27, 0.23, 0.19, 0.12, and 0.07 kb. Twenty-two of the 64 PCR products that were isolated and digested with *Hinf*I are shown here. Digestion of one of these, PCR 13, yielded a large number of fragments whose sizes summed to a value much greater than that of the undigested PCR 13 DNA, indicating that PCR 13 [highlighted in yellow] contain multiple species of DNA that are representatives of a multigene family.

bers of a multigene family. We reasoned that restriction digestion of an individual PCR product consisting of a single species of DNA would generate a set of DNA fragments whose molecular weights sum to the molecular weight of the original PCR product. On the other hand, if an individual PCR product consisted of several different DNA sequences (consistent with the amplification of members of a multigene family), restriction endonuclease digestion should generate a large array of fragments whose molecular weights sum to far greater than the molecular weight of the original PCR product. In this manner, we identified a multigene family that encodes a large family of proteins belonging to the seven transmembrane domain receptor superfamily and whose expression is restricted to the olfactory epithelium.

Cloning the Gene Family

In initial experiments we designed a series of degenerate oligonucleotides that could anneal to conserved regions of members of the superfamily of G protein-coupled seven transmembrane domain receptor genes [highlighted in yellow]. Five degenerate oligonucleotides (A1–A5; see Experimental Procedures) matching sequences within transmembrane domain 2 and six degenerate oligonucleotides (B1–B6) matching transmembrane domain 7 were used in all combinations in PCR reactions to amplify homologous sequences in cDNA prepared from rat olfactory epithelium RNA. The amplification products of each PCR reaction were then analyzed by agarose gel electrophoresis. Multiple bands were observed with each of the primer combinations. The PCR products within the size range expected for this family of receptors (600 to 1300 bp) were subsequently picked and amplified further with the appropriate primer pair in order to isolate individual PCR bands. Sixty-four PCR bands isolated in this fashion revealed only one or a small number of bands upon agarose gel electrophoresis. Representatives of these isolated PCR products are shown in Figure 2A.

We next asked whether any of these discrete PCR products consisted of multiple DNA sequences reflecting the amplification of a large family of genes. The isolated PCR products were digested with *Hae*III or *Hinf*I, which recognize four base restriction sites and cut DNA at frequent intervals. In most instances, digestion of the PCR product with *Hinf*I generated a set of fragments whose molecular weights sum to the size of the original DNA (Figure 2B). These PCR bands are therefore each likely to contain a single DNA species. In some cases, however, restriction digestion yielded a series of fragments whose molecular weights sum to a value greater than that of the original PCR product. The most dramatic example is shown in Figure 2B, where the PCR 13 DNA (710 bp) is cleaved by *Hinf*I to yield a very large number of restriction fragments whose sizes sum to a value 5- to 10-fold greater than that of the original PCR product. These observations indicated that PCR product 13 consists of a number of different species of DNA, each of which could be amplified with the same pair of primer oligonucleotides. In addition, when PCR experiments similar to those described were performed using cDNA library DNAs as templates, a 710 bp PCR product was obtained with the PCR 13 primer pair (A4/B6) with DNA from olfactory cDNA libraries, but not from a glioma cDNA library. Moreover, digestion of this 710 bp PCR product also revealed the presence of multiple DNA species. In other cases (see PCR product 20, for example), digestion yielded a series of restriction fragments whose molecular weights also sum to a size greater than the starting material. Further analysis, however, revealed that the original PCR product consisted of multiple bands of similar but different sizes.

To determine whether the multiple DNA species present in PCR 13 encode members of a family of seven transmembrane domain proteins, PCR 13 DNA was cloned into the plasmid vector Bluescript, and five individual clones were subjected to DNA sequence analysis. Each of the five clones exhibited a different DNA sequence, but each

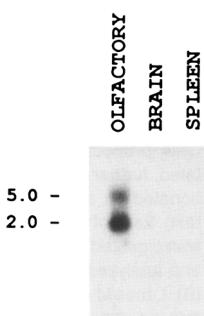


Figure 3. Northern Blot Analysis with a Mixture of 20 Probes
One microgram of poly(A)⁺ RNA isolated from rat olfactory epithelium, brain, or spleen was size fractionated in formaldehyde-agarose, blotted onto a nylon membrane, and hybridized with a ³²P-labeled mixture of segments of 20 cDNA clones. The DNA segments were obtained by PCR using primers homologous to transmembrane domains 2 and 7.

encoded a protein that displayed conserved features of the superfamily of seven transmembrane domain receptor proteins. In addition, the proteins encoded by all five clones shared distinctive sequence motifs not found in other superfamily members, indicating they were all members of a new family of receptors.

To obtain full-length cDNA clones, cDNA libraries prepared from olfactory epithelium RNA or from RNA of an enriched population of olfactory sensory neurons were screened. The probe used in these initial screens was a mixture of PCR 13 DNA as well as DNA obtained by amplification of rat genomic DNA or DNA from two olfactory cDNA libraries with the same primers used to generate PCR 13 (A4 and B6 primers). Hybridizing plaques were subjected to PCR amplification with the A4/B6 primer set, and only those giving a PCR product of the appropriate size (approximately 710 bp) were purified. The frequency of such positive clones in the enriched olfactory neuron cDNA library was approximately 5 times greater than the frequency in the olfactory epithelium cDNA library. The increased frequency of positive clones observed in the olfactory neuron library is comparable to the enrichment in olfactory neurons generally obtained in the purification procedure.

The original pair of primers used to amplify PCR 13 DNA was then used to amplify coding segments of 20 cDNA clones. A mixture of these PCR products was labeled and used as probe for further cDNA library screens. This mixed probe was also used in a Northern blot (Figure 3) to determine whether the expression of the gene family is restricted to the olfactory epithelium. The mixed probe detects two diffuse bands centered at 2 and 5 kb in RNA from olfactory epithelium; no hybridization can be detected in brain or spleen. (Later experiments, which examined a larger number of tissue RNAs with a more restricted probe, will be shown below.) Taken together, these data indicate that we have identified a novel multigene family encoding seven transmembrane domain proteins that are expressed in olfactory epithelium, and could be expressed predominantly or exclusively in olfactory neurons.

The Protein Sequences of Numerous, Olfactory-Specific Members of the Seven Transmembrane Domain Superfamily

Numerous clones were obtained upon screening cDNA libraries constructed from olfactory epithelium or olfactory neuron RNA at high stringency (see Experimental Procedures). Partial DNA sequences were obtained from 36 clones; 18 of these cDNA clones are different, but all of them encode proteins that exhibit shared sequence motifs indicating that they are members of the family identified in PCR 13 DNA. A complete nucleotide sequence was determined for the coding regions of ten of the most divergent clones (Figure 4). The deduced protein sequences of these cDNAs defines a new multigene family that shares sequence and structural properties with the seven transmembrane domain superfamily of neurotransmitter and hormone receptors. This novel family, however, exhibits features different from any other member of the receptor superfamily thus far identified.

Each of the ten sequences contains seven hydrophobic stretches (19–26 amino acids) that represent potential transmembrane domains. These domains constitute the regions of maximal sequence similarity to other members of the seven transmembrane domain superfamily (see legend to Figure 4). On the basis of structural homologies with rhodopsin and the β -adrenergic receptors (O'Dowd et al., 1989b), it is likely that the N-termini of the olfactory proteins are located on the extracellular side of the plasma membrane and the C-termini in the cytoplasm. In this scheme, three extracellular loops alternate with three intracellular loops to link the seven transmembrane domains (see Figure 5). Analysis of the sequences in Figure 4 demonstrates that the olfactory proteins, like other members of the receptor superfamily, display no evidence of an N-terminal signal sequence. As in several other superfamily members, a potential N-linked glycosylation site is present in all ten proteins within the short N-terminal extracellular segment. Other structural features conserved with previously identified members of the superfamily include cysteine residues at fixed positions within the first and second extracellular loops, which are thought to form a disulfide bond. Finally, many of the olfactory proteins reveal a conserved cysteine within the C-terminal domain that may serve as a palmitoylation site anchoring this domain to the membrane (O'Dowd et al., 1989a). These features, taken together with several short, conserved sequence motifs (see legend to Figure 4), clearly define this new family as a member of the superfamily of genes encoding the seven transmembrane domain receptors.

There are, however, important differences between the olfactory protein family and the other seven transmembrane domain proteins described previously, and these differences may be relevant to a proposed function of these proteins in odor recognition. Structure-function experiments involving *in vitro* mutagenesis suggest that adrenergic ligands interact with this class of receptor molecule by binding within the plane of the membrane (Kobilka et al., 1988; Strader et al., 1989). Not surprisingly, small receptor families that bind the same class of ligands, such as the adrenergic and muscarinic acetylcholine receptor

	I	II	
F3	MDSSNRTRVSEFLLLGPFENKDLQPLIYGLFLSMYLTVTIGNISIIIVAIIDPCLHTPMYFFLSNLSPVDFICFISTTVPKML		82
F5	MSSTNQSSVTEFLLLGPSRQQQQQQLFLIPIMVLTAVLGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		82
F6	MWASTGQNLSLTPGPFIILSFPGRPSMRIGLFLLFLVNVLLTTLVGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		85
F12	MESGNSTRRFSSFFLICFTENPQLHFLIFALFLSMYLTVLGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		83
I3	MN-NQTFITOFLLLGLPIPEEHQHLYFALFLVNLTTILGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		80
I7	MERRNHSGRYSEEVLLGFPAPPLRVLLFLSESLTVLVLGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		83
I8	MN-NKTVITHFLLLGLPIPEHQOFFALFLVNLTTILGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		80
I9	MTRRNQTAISQFFLGLIPFPEYQHLYFALFLAMLTLLGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		82
I14	MTCNNQTLILEFLLGLIPSEYHLYFALFLAMLTLLGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		82
I15	MTEENQTVISQFLLFLIPSEHQHVFYALFLSMYLTVLGNLTLAIGTDSRLHTPMYFFLSNLSPVDFICFISTTVPKML		82
	III	IV	
F3	---VNIQTQNNVITYAGCITQIYFFLLVEVDNFELITIAYDRYVAICHPMNYTIVINMYKLCGFVLYVSNIVSVLHALFQSLLM		163
F5	---ANHILGSOAISFSGLCTGLYFLAVFGNMDFLAVMSYDGFNAICHLRNYTTKMTROLEVLLVVGSNVANMNCLLHILLM		163
F6	---ATFAPRGYVISLAGCATOMYFVFSLCTEYFLAVAYDRYLAICLPLRYGGGINTPGLAMRLAIGSHLCGFSAITVPATLI		166
F12	---VNIYTQSKSITYEDCISQMCVFLVFAELGNFLAVAYDRYVAICHPCLCYTIVVNHRCLLCLLLELSMVISIFHAFIOSLIV		164
I3	---QNMRSQDTSPYGGCLACTYFHMVFDMESFLYVAFHAYDRYVAICHPCLCYTIVVNHRCLLCLLLELSMVISIFHAFIOSLIV		161
I7	AGFIGSKENHGLQISFEACMQLYFFLGLGCTECVLLAVAYDRYVAICHPCLCYTIVVNHRCLLCLLLELSMVISIFHAFIOSLIV		168
I8	---QNIQSQVPSISYAGCITQIFFFLFGYLGNGFLVAVAYDRYVAICHPCLCYTIVVNHRCLLCLLLELSMVISIFHAFIOSLIV		161
I9	---QNMQSQVPSIPYAGCICLADYFFLFLDGNFLVAVAYDRYVAICHPCLCYTIVVNHRCLLCLLLELSMVISIFHAFIOSLIV		163
I14	---QNMQSQVPSISYTGCLQLYFFMVFQDMESFLVAVAYDRYVAICHPCLCYTIVVNHRCLLCLLLELSMVISIFHAFIOSLIV		163
I15	---QNMQSQVPSIPFAGCCTQLYFYLYFADLESFLVAVAYDRYVAICHPCLCYTIVVNHRCLLCLLLELSMVISIFHAFIOSLIV		163
	V	VI	
F3	LALPFCETHLEIPHVFCEPNQVIQLTCSDAFLNDLVYFTLVLLATVPLAGIFYSYFKIVSSICAISSVHGKYKAFCSTCGSHLSVV		248
F5	ARKSFCADNMIPHFFCDGTPLLKSCSDTHLNEMLTEGAIVVMVTPFCVILISVIIHITCAVLRVSSPRGGWKSFSTCGSHLSVV		248
F6	ARLSFCGSRVINHFFCDSPHWVLSCTDQVVELVSGIAFCVILSCGILVSYFVTKIPSGARGRHRAFSTCGSHLSVV		251
F12	LOLTFCGDVKIPHFFCELNQLSQETCSNDPFSHLIMMNPVMLAAISFGILYVFKIVSSIHSISTVQGKYKAFCSTCGSHLSVV		249
I3	ARLSFCENNVVNLNNFCDLVLLKLACSBTYINELMIFIMSTLLIIIPPFLLIVMSYARISSILKVPSTQGICKVFSTCGSHLSVV		246
I7	SRLSYCGPTINHFFCDVSPLNLSCTDMSTAELTDFLVLAIFILLGFLSVTGASYMAITGAVMRIPSAAGRHKAFSTCGSHLSVV		253
I8	ARLSFCENNVVNLNNFCDLVLLKLACSDTYVNELMIMHMGVIVIIPVFLIVISYAKIISSLKVPSTQSIHKVFSTCGSHLSVV		246
I9	ARLSFCEDSYVIPHFFCDMSTLKKVLSLHNDELAIFIQGGPIVVLPPFLIIIVSYARISSIFKVPSSQSIHKAFSTCGSHLSVV		248
I14	ARLSFCEKNNVILHFFCDISPLLKSCSDTHVNELVIFVMGLVIVIPVFLIIIVSYARVVASILKVPSPVRGIHKIFSTCGSHLSVV		248
I15	ARLSFCADNMIPHFFCDISPLLKSCSDTHVNELVIFVMGLVIVIPVFLIIIVSYARVVASILKVPSPVRGIHKIFSTCGSHLSVV		248
	VII		
F3	SLFYCTGLGSVLLSSAANNSQASATASVHYTVVTPHVNPFYIYSLRKDKVSVLKTLCEEVIRSPPSLLHFFLVLCHLPCFIFCY		333
F5	CLFYGTIVIAYVFNPSSSHLAGRDMAAVHYAYVTPHNLNPFIYSLRNSDMKAALRKVLAMRFPSKQ		313
F6	L1WYGSTIFLHVRTSVESSLDLTKAITVLTNTIYTPVLPNPFYIYSLRNLKDVKRALERLLEGNCKVHHWTG		311
F12	SLFYSTGLGVVYSSAVVQSSHSAASASVHYTVVTPHNLNPFIYSLRNLKDVKRALRRTLHQAQDQEANTNKGSKIG		317
I3	SLFYGTIIGLYLCPAGNNSTVKEMVMAMMYTVVTPHNLNPFIYSLRNRDMKRALIRVICSMKITL		310
I7	IIIFYAASIFIYARPKALSAFTDNKLVSYLVAIVVPLFPNPIIYCLRNODVKRALRRTLHQAQDQEANTNKGSKIG		327
I8	SLFYGTIIGLYLCPGDNFSLKGSAMAMMYTVVTPHNLNPFIYSLRNRDMKRALIRVCSSKIKSLPW		312
I9	SLFYGTIVIGLYLCPSSANSTVKETVMSLHYTMVTPHNLNPFIYSLRNRDDIKELEMCKQIPSFL		314
I14	TLFYGTIFGIGLYLCPGDNFSLKGSAMAMMYTVVTPHNLNPFIYSLRNRDMKRALIRVICTKKISL		312
I15	SLFYGTIIGLYLCPSSANSTVKETVMSLHYTMVTPHNLNPFIYSLRNRDMKRALIRVICTKKISL		314

Figure 4. The Protein Sequences Encoded by Ten Divergent cDNA Clones

Ten divergent cDNA clones were subjected to DNA sequence analyses, and the protein sequence encoded by each was determined. Amino acid residues conserved in 60% or more of the proteins are shaded. The presence of seven hydrophobic domains (I–VII), as well as short conserved motifs shared with other members of the superfamily, demonstrates that these proteins belong to the seven transmembrane domain protein superfamily. Motifs conserved among members of the superfamily and the family of olfactory proteins include the GN in TM1 (transmembrane domain 1), the central W of TM4, the Y near the C-terminal end of TM6, and the NP in TM7. In addition, the DRY motif C-terminal to TM3 is common to many members of the G protein-coupled superfamily. However, all of the proteins shown here share sequence motifs not found in other members of this superfamily and are clearly members of a novel family of proteins. The nucleotide sequences from which these protein sequences were derived have been deposited in GenBank.

families, exhibit maximum sequence conservation (often over 80%) within the transmembrane domains. In contrast, the family of receptors we have identified shows striking divergence within the third, fourth, and fifth transmembrane domains (Figure 4). The variability in the three central transmembrane domains is highlighted schematically in Figure 5. The divergence in potential ligand-binding domains is consistent with the idea that the family of molecules we have cloned is capable of associating with a large number of odorants of diverse molecular structures.

Receptors that belong to the superfamily of seven transmembrane domain proteins interact with G proteins to gen-

erate intracellular signals. In vitro mutagenesis experiments indicate that one site of association between receptor and G protein resides within the third cytoplasmic loop (Kobilka et al., 1988; Hamm et al., 1988). The sequence of this cytoplasmic loop in 18 different clones we have characterized is shown in Figure 6A. This loop, which is often quite long and of variable length in the receptor superfamily, is relatively short (only 17 amino acids) and of fixed length in the 18 clones examined. Interestingly, 11 of the 18 different clones exhibit the sequence motif (K/R)IVSSI (or a close relative) at the N-terminus of this loop. Two of the cDNA clones reveal a different motif, HIT(C/W)AV, at this site. If this short loop is a site of contact

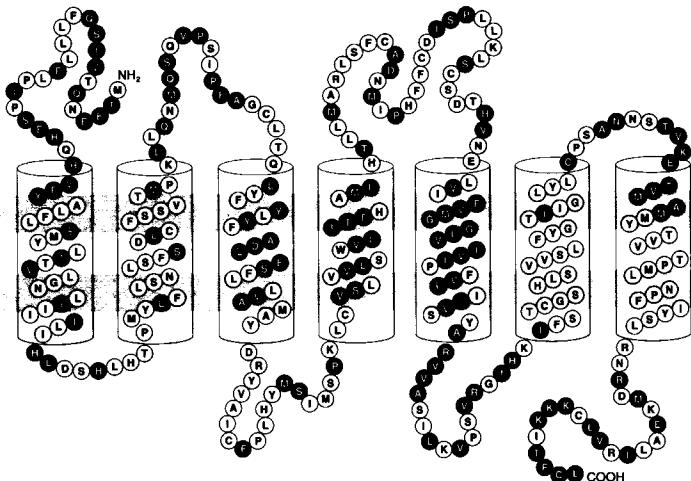


Figure 5. Positions of Greatest Variability in the Olfactory Protein Family

In this diagram the protein encoded by cDNA clone I15 is shown traversing the plasma membrane seven times, with its N-terminus located extracellularly and its C-terminus intracellularly. The vertical cylinders delineate the seven putative α helices spanning the membrane. Positions at which 60% or more of the 10 clones shown in Figure 4 share the same residue as I15 are shown as white balls. More variable residues are shown as black balls. The high degree of variability encountered in transmembrane domains III, IV, and V is evident in this schematic.



with G proteins, it is possible that the conserved motifs may reflect sites of interaction with different G proteins that activate different intracellular signaling systems in response to odors. In addition, the putative receptors we have cloned reveal several conserved serine or threonine residues within the third cytoplasmic loop. By analogy with

other G protein-coupled receptors, these residues may represent sites of phosphorylation for specific receptor kinases involved in desensitization (Bouvier et al., 1988).

Subfamilies within the Multigene Family

Figure 6A displays the sequences of the fifth transmem-

		V
A.	F2	RVNE VVIFIVVSILFLVLVPFALIIMSYY RIVSSILKVPSSQGIYK
	F3	FLND LVIFYFTLVLLATVPLAGIFYSYF KIVSSICAISSSVHGKYK
	F5	HLNE LMILTEGAVVMVTPEFCILISYI HITCAVLRVSSPRGGWK
	F6	QVVE LVSFGIAFCVILGSCGITALVSYA YIITTIKIKPSARGRHR
	F7	HVNE LVIFVMGGIILVIFPVLIIVSYV RIVSSILKVPSSQGIYK
	F8	FPSH LTMHLPVPMVLAISLSSGILSYF KIVSSIRSMSSVQGKYK
	F12	FPSH LIMNLVPVMLAAISLSSGILSYF KIVSSISHSISTVQGKYK
	F13	FPSH LIMNLVPVMLAAISLSSGILSYF KIVSSIRSVSSVKGKYK
	F23	FLND VIMYFALVLLAVVPLLGILSYS KIVSSIRRAISTVQGKYK
	F24	HEIE MIIYLVAAFNLISSLVVLYSYL FILIATLRMNSAEGRRK
	I3	YINE LMIFIMSTLLIIIPFFLIVMSYA RISSILKVPSTQGICK
	I7	STAE LTDVFVLAIFIILLGPLSVTGAASYM AITGAVMRIPSAAGRHK
	I8	YVNE LMIHIMGVIIIVIPFVLLIVISYK KIISSLKVPSTQSIHK
	I9	HDNE LAIFILGGPIVVLPPFLLIIIVSYA RIVSSIFKVPSSQSIHK
	I11	HLNE LMILTEGAVVMVTPEFCILISYI HITWAVLRVSSPRGGWK
	I12	FPSH LIMNLVPVMLGAISLSSGILSYF KIVSSVRSSVQGKH
	I14	YVNE LMIYILGGIILIPFLLIVMSYY RIFFSILKFPSIQQDIYK
	I15	HVNE LVIFVMGGIILVIFPVLIIVSYA RVVASILKVPSSVQGKH
B.	F12	FPSH LIMNLVPVMLAAISLSSGILSYF KIVSSISHSISTVQGKYK
	F13	FPSH LIMNLVPVMLAAISLSSGILSYF KIVSSIRSVSSVKGKYK
	F8	FPSH LTMHLPVPMVLAISLSSGILSYF KIVSSIRSMSSVQGKYK
	I12	FPSH LIMNLVPVMLGAISLSSGILSYF KIVSSVRSSVQGKH
	F23	FLND VIMYFALVLLAVVPLLGILSYS KIVSSIRRAISTVQGKYK
	F3	FLND LVIFYFTLVLLATVPLAGIFYSYF KIVSSICAISSSVHGKYK
C.	F7	HVNE LVIFVMGGIILVIFPVLIIVSYV RIVSSILKVPSSQGIYK
	I15	HVNE LVIFVMGGIILVIFPVLIIVSYA RVVASILKVPSSVQGKH
	I3	YINE LMIFIMSTLLIIIPFFLIVMSYA RISSILKVPSTQGICK
	I8	YVNE LMIHIMGVIIIVIPFVLLIVISYK KIISSLKVPSTQSIHK
	I9	HDNE LAIFILGGPIVVLPPFLLIIIVSYA RIVSSIFKVPSSQSIHK
	I14	YVNE LMIYILGGIILIPFLLIVMSYY RIFFSILKFPSIQQDIYK
D.	F5	HLNE LMILTEGAVVMVTPEFCILISYI HITCAVLRVSSPRGGWK
	I11	HLNE LMILTEGAVVMVTPEFCILISYI HITWAVLRVSSPRGGWK

Figure 6. The Presence of Subfamilies in a Divergent Multigene Family

Partial nucleotide sequences and deduced protein sequences were obtained for 18 different cDNA clones. Transmembrane domain V along with the flanking loop sequences, including the entire cytoplasmic loop between transmembrane domains V and VI, is shown here for each protein. Amino acid residues found in 60% or more of the clones in a given position are shaded (A). This region of the olfactory proteins (particularly transmembrane domain V) appears to be highly variable (see Figure 4). These proteins, however, can be grouped into subfamilies (B-D) in which the individual subfamily members share considerable homology in this divergent region of the protein.

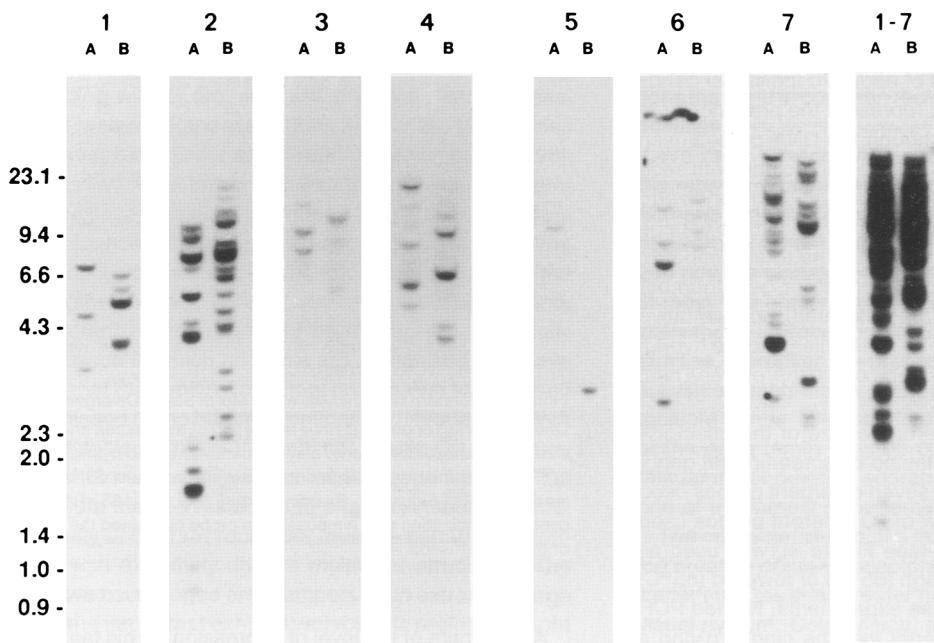


Figure 7. Southern Blot Analyses with Non-Cross-Hybridizing Fragments of Divergent cDNAs

Five micrograms of rat liver DNA was digested with EcoRI (A) or HindIII (B), electrophoresed in 0.75% agarose, blotted onto a nylon membrane, and hybridized with the 32 P-labeled probes indicated. The probes used were PCR-generated fragments of: 1, clone F9 (identical to F12 in Figure 4); 2, F5; 3, F6; 4, I3; 5, I7; 6, I14; or 7, I15. The lane labeled "1–7" was hybridized to a mixture of the seven probes. The probes used showed either no cross-hybridization or only trace cross-hybridization with one another. The size markers on the left correspond to the four blots on the left (1–4), whereas the marker positions noted on the right correspond to the four blots on the right (5–7 and "1–7").

brane domain and the adjacent cytoplasmic loop encoded by 18 of the cDNA clones we have analyzed. As a group, the 18 sequences exhibit considerable divergence within this region. The multigene family, however, can be divided into subfamilies such that the members of a given subfamily share significant sequence conservation. Analysis of the sequences in Figure 6A defines at least three subfamilies of related sequences (Figures 6B–6D). Subfamily B, for example, consists of six closely related sequences in which pairs of sequences can differ from one another at only four of 44 positions (91% identity) (see F12 and F13). The sequences encoded by clones F5 and I11 (subfamily D), which differ at only one residue, differ from F12 and F13 (subfamily B) at 34–36 of the 44 positions within this region and clearly define a separate subfamily. It is possible that the divergent subfamilies encode receptors that bind odorants of widely differing molecular structures. Members of the individual subfamilies could therefore recognize more subtle differences between molecules that belong to the same structural class of odorant molecules.

The Size of the Multigene Family

We have performed genomic Southern blotting experiments and have screened genomic libraries to obtain an estimate of the sizes of the multigene family and the member subfamilies encoding the putative odor receptors. DNAs extending from the 3' end of transmembrane domain 3 to the middle of transmembrane domain 6 were synthesized by PCR from DNA of seven of the divergent

cDNA clones (Figure 4). In initial experiments, these DNAs were labeled and hybridized to each other to define conditions under which minimal cross-hybridization would be observed among the individual clones. At 70°C the seven DNAs showed no cross-hybridization, or cross-hybridized only very slightly. The trace levels of cross-hybridization observed are not likely to be apparent upon genomic Southern blot analysis, where the amounts of DNA are far lower than in the test cross.

Probes derived from these seven DNAs were annealed under stringent conditions, either individually or as a group, to Southern blots of rat liver DNA digested with the restriction endonuclease EcoRI or HindIII (Figure 7). Examination of the Southern blots reveals that all but one of the DNAs detects a relatively large, distinctive array of bands in genomic DNA. Clone I15 (probe 7), for example, detects about 17 bands with each restriction endonuclease, whereas clone F9 (probe 1) detects only about 5–7 bands with each enzyme. A single band is obtained with clone I7 (probe 5). PCR experiments using nested primers (TM2/TM7 primers followed by primers to internal sequences) and genomic DNA as template indicate that the coding regions of the members of this multigene family, like those of many members of the G protein-coupled superfamily, may not be interrupted by introns. This observation, together with the fact that most of the probes encompass only 400 nucleotides, suggests that each band observed in these experiments is likely to represent a different gene. These data suggest that the individual probes

we have chosen are representatives of subfamilies that range in size from a single member to as many as 17 members. We detect a total of about 70 individual bands in this analysis, which could represent the presence of at least 70 different genes. Although the DNA probes used in these blots did not cross-hybridize appreciably with each other, it is possible that a given gene might hybridize to more than one probe, resulting in an overestimate of gene number. However, it is probable that the total number of bands reflects only a minimal estimate of gene number since it is unlikely that we have isolated representative cDNAs from all of the potential subfamilies and the hybridizations were performed under conditions of very high stringency.

A more accurate estimate of the size of the olfactory-specific gene family was obtained by screening rat genomic libraries. The mix of the seven divergent probes used in Southern blots, or the mix of 20 different probes used in our initial Northern blots (see Figure 3), was used as hybridization probe under high (65°C) or lowered (55°C) stringency conditions in these experiments. Nested PCR (see above) was used to verify that the clones giving a positive signal under low-stringency annealing conditions were indeed members of this gene family. We estimate from these studies that there are between 100 and 200 positive clones per haploid genome. The estimate of the size of the family we obtain from screens of genomic libraries again represents a lower limit. Given the size of the multigene family, we might anticipate that many of these genes are linked such that a given genomic clone may contain multiple genes. Thus the data from Southern blotting and screens of genomic libraries indicate that the multigene family we have identified consists of 100 to several hundred genes that can be divided into multiple subfamilies.

It should be noted that the cDNA probes we have isolated may not be representative of the full complement of subfamilies within the larger family of olfactory proteins. The isolation of cDNAs, for example, relies heavily on PCR with primers from transmembrane domains 2 and 7 and biases our clones for homology within these regions. Thus, estimates of gene number as well as subsequent estimates of RNA abundance should be considered as minimal. We anticipate that further characterization of cDNA clones will identify representatives of new subfamilies and that this family of putative odorant receptors may be extremely large.

Expression of the Members of This Multigene Family

We have performed additional Northern blot analyses to demonstrate that expression of the members of this gene family is restricted to the olfactory epithelium (Figure 8). Northern blot analysis with a mixed probe consisting of the seven divergent cDNAs used above reveals two diffuse bands about 5 and 2 kb in length in olfactory epithelium RNA. This pattern is the same as that seen previously with the mix of 20 DNAs. No annealing is observed to RNA from the brain or retina or other, nonneuronal tissues, including lung, liver, spleen, and kidney.

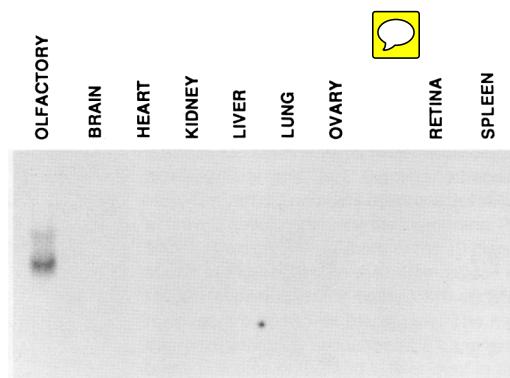


Figure 8. Northern Blot Analysis with a Mix of Seven Divergent Clones

One microgram of poly(A)⁺ RNA from each of the tissues shown was size fractionated, blotted onto a nylon membrane, and hybridized with a ³²P-labeled mixture of segments of seven divergent cDNA clones (see legend to Figure 7). Examination of 28S and 18S rRNAs and control hybridization with a mouse actin probe confirmed the integrity of the RNAs used.

An estimate of the level of expression of this family can be obtained from screens of cDNA libraries. The frequency of positive clones in cDNA libraries made from olfactory epithelium RNA suggests that the abundance of the RNAs in the epithelium is about one in 20,000. The frequency of positive clones is approximately 5-fold higher in a cDNA library prepared from RNA from purified olfactory neurons (in which approximately 75% of the cells are olfactory neurons). The increased frequency of positive clones obtained in the olfactory neuron cDNA library is comparable to the enrichment we obtain upon purification of olfactory neurons. These observations suggest that this multigene family is expressed largely, if not solely, in olfactory neurons and may not be expressed in other cell types within the epithelium. If each olfactory neuron contains 10⁵ mRNA molecules, from the frequency of positive clones we predict that each neuron contains only 25–30 transcripts derived from this gene family. Since the family of olfactory proteins consists of a minimum of 100 genes, a given olfactory neuron could maximally express only a proportion of the many different family members. These values thus suggest that olfactory neurons will exhibit significant diversity at the level of expression of these olfactory proteins.

Discussion

The mammalian olfactory system can recognize and discriminate a large number of odorous molecules. Perception in this system, as in other sensory systems, initially involves the recognition of external stimuli by primary sensory neurons. This sensory information is then transmitted to the brain, where it is decoded to permit the discrimination of different odors. Elucidation of the logic underlying olfactory perception is likely to require the identification of the specific odorant receptors, the analysis of the extent of receptor diversity and receptor specificity, as well as an understanding of the pattern of receptor expression in the olfactory epithelium.

What are the expected properties of odorant receptors,

and what is the evidence that the multigene family we have identified encodes these receptors? First, the odorant receptors are thought to transduce intracellular signals by interacting with G proteins, which activate second messenger systems (Pace et al., 1985; Sklar et al., 1986; Jones and Reed, 1989; Breer et al., 1990; Boekhoff et al., 1990). Although we have not demonstrated that the olfactory proteins we have identified can indeed trigger G protein-coupled responses, these proteins are clearly members of the family of G protein-coupled receptors, which traverse the membrane seven times (O'Dowd et al., 1989b). Second, the odorant receptors should be expressed specifically in the tissue in which odorants are recognized. We have shown that the family of olfactory proteins we have cloned is expressed in the olfactory epithelium. Hybridizing RNA is not detected in brain or retina, nor in a host of nonneuronal tissues. Moreover, expression of this gene family in the epithelium may be restricted to olfactory neurons. Third, the family of odorant receptors must be capable of interacting with extremely diverse molecular structures. The genes we have cloned are members of an extremely large multigene family that exhibits variability in regions thought to be important in ligand binding. The possibility that each member of this large family of seven transmembrane proteins is capable of interacting with only one or a small number of odorants provides a plausible mechanism to accommodate the diversity of odor perception. Finally, the odorant receptors must bind odorants and transduce an intracellular signal leading to the activation of second messenger systems. This criterion, at present, is difficult to satisfy experimentally because the diversity of odorants and the large number of individual receptors make it difficult to identify the appropriate ligand for a particular receptor. Nonetheless, the properties of the gene family we have identified suggest that this family is likely to encode a large number of distinct odorant receptors.

How Large Is the Multigene Family?

How many structurally distinct odors can an organism detect, and how large is the receptor gene family? The size of the receptor repertoire is likely to reflect the range of detectable odors and the degree of structural specificity exhibited by the individual receptors. It is difficult to assess the discriminatory capacity of the olfactory system accurately, but it has been estimated that humans can identify over 10,000 structurally distinct odorous ligands. However, this does not necessarily imply that humans possess an equally large repertoire of odorant receptors. For example, binding studies in lower vertebrates suggest that structurally related odorants may activate the same receptor molecules. In fish that smell amino acids, the binding of alanine to isolated cilia can be competed by other small polar residues (threonine and serine) but not by the basic amino acids lysine and arginine (Rhein and Cagan, 1983). These data suggest that individual receptors are capable of associating with several structurally related ligands, albeit with different affinities. Stereochemical models of olfactory recognition in mammals (Amoore, 1982) (based largely on psychophysical rather than biophysical data) have suggested the existence of several primary odor

groups including camphoraceous, musky, pepperminty, ethereal, pungent, and putrid. In such a model, each group would contain odorants with common molecular configurations that bind to common receptors and share similar odor qualities.

We have provided a minimum estimate of the size of the repertoire of the putative odorant receptors in the rat. Screens of genomic libraries with mixed probes consisting of divergent family members detect approximately 100 to 200 positive clones per genome. The present estimate of at least 100 genes provides only a lower limit since it is likely that our probes do not detect all of the possible subfamilies. Moreover, it is probable that many of these genes are linked such that a given genomic clone may contain multiple genes. We therefore expect that the actual size of the gene family may be considerably higher and that this family of putative odorant receptors could constitute one of the largest gene families in the genome.

The characterization of a large multigene family encoding putative odorant receptors suggests that the olfactory system utilizes a far greater number of receptors than the visual system. Color vision, for example, allows the discrimination of several hundred hues but is accomplished by only three different photoreceptors (Rushton, 1955, 1965; Wald et al., 1955; Nathans et al., 1986). The photoreceptors each have different, but overlapping, absorption spectra that cover the entire spectrum of visible wavelengths. Discrimination of color results from comparative processing of the information from these three classes of photoreceptors in the brain. Whereas three photoreceptors can absorb light across the entire visible spectrum, our data suggest that a small number of odorant receptors cannot recognize and discriminate the full spectrum of distinct molecular structures perceived by the mammalian olfactory system. Rather, olfactory perception probably employs an extremely large number of receptors each capable of recognizing a small number of odorous ligands.

Diversity within the Gene Family and the Specificity of Odor Recognition

The olfactory proteins we have identified are clearly members of the superfamily of receptors that traverse the membrane seven times. Analysis of the proteins encoded by the 18 distinct cDNAs we have cloned reveals structural features that may render this family particularly well suited for the detection of a diverse array of structurally distinct odorants. Experiments with other members of this class of receptors suggest that ligand binds to its receptor within the plane of the membrane such that the ligand contacts many, if not all, of the transmembrane helices (Strader et al., 1989; Kobilka et al., 1988). The family of olfactory proteins can be divided into several different subfamilies that exhibit significant sequence divergence within the transmembrane domains. Nonconservative changes are commonly observed within blocks of residues in transmembrane regions 3, 4, and 5 (Figures 4–6); these blocks could reflect the sites of direct contact with odorous ligands. Some members, for example, have acidic residues in transmembrane domain 3, which in other families are thought to be essential for binding aminergic ligands.

(Strader et al., 1987), while other members maintain hydrophobic residues at these positions. This divergence within transmembrane domains may reflect the fact that the members of the family of odorant receptors must associate with odorants of widely different molecular structures.

These observations suggest a model in which each of the individual subfamilies encodes receptors that bind distinct structural classes of odorants. Within a given subfamily, however, the sequence differences are far less dramatic and are often restricted to a small number of residues. Thus, the members of a subfamily may recognize more subtle variations among odor molecules of a given structural class. At a practical level, individual subfamilies may recognize grossly different structures such that one subfamily may associate, for example, with the aromatic compound benzene and its derivatives, whereas a second subfamily may recognize odorous, short-chain aliphatic molecules. Subtle variations in the structure of the receptors within, for example, the hypothetical benzene subfamily could facilitate the recognition and discrimination of various substituted derivatives such as toluene, xylene, or phenol. It should be noted that such a model, unlike previous stereochemical models, does not necessarily predict that molecules with similar structures will have similar odors. The activation of distinct receptors with similar structures could elicit different odors, since perceived odor will depend upon higher-order processing of primary sensory information.

Evolution of the Gene Family and the Generation of Diversity

Preliminary evidence from PCR analyses suggests that members of this family of olfactory proteins are conserved in lower vertebrates as well as invertebrates. This gene family presumably expanded over evolutionary time, providing mammals with the ability to recognize an increasing diversity of odorants. Examination of the sequences of the family members cloned from mammals provides some insight into the evolution of this multigene family. Although we have not yet characterized the chromosomal loci encoding these genes, it is likely that at least some member genes will be tandemly arranged in a large cluster as is observed with other large multigene families. A tandem array of this sort provides a template for recombination events, including unequal crossing over and gene conversion, that can lead to expansion and further diversification of the sort apparent among the family members we have cloned (for review see Maeda and Smithies, 1986).

The multigene family encoding the olfactory proteins is large: all of the member genes clearly have a common ancestral origin but have undergone considerable divergence such that individual genes encode proteins that share 40%–80% amino acid identity. Subfamilies are apparent, with groups of genes sharing greater homology among themselves than with members of other subfamilies. Examination of the sequences of even the most divergent subfamilies reveals a pattern in which blocks of conserved residues are interspersed with variable regions. This segmental homology is conceptually similar to the

organization of framework and hypervariable domains within the families of immunoglobulin and T cell receptor variable region sequences (for reviews see Tonegawa, 1983; Hood et al., 1985). This analogy goes beyond structural organization and may extend to the function of these gene families: each family consists of a large number of genes that have diversified over evolutionary time to accommodate the binding of a highly diverse array of ligands. The evolutionary mechanisms responsible for the diversification and maintenance of these large gene families may also be similar. It has been suggested that gene conversion has played a major role in the evolution of immunoglobulin and T cell receptor variable domains (Baltimore, 1981; Egel, 1981; Flanagan et al., 1984). Analysis of the sequences of the putative olfactory receptors reveals at least one instance where a motif from a variable region of one subfamily is found embedded in the otherwise divergent sequence of a second subfamily, suggesting that conversion has occurred. Such a mixing of motifs from one subfamily to another over evolutionary time would provide additional combinatorial possibilities leading to the generation of diversity.

It should be noted, however, that the combinatorial joining of gene segments by DNA rearrangement during development, which is characteristic of immunoglobulin loci (Tonegawa, 1983), is not a feature of the putative odor receptor gene family. We have observed no evidence for DNA rearrangement to generate the diversity of genes we have cloned. We have sequenced the entire coding region along with parts of the 5' and 3' untranslated regions of ten different cDNA clones. The sequences of the coding regions are all different; we have not obtained any evidence for constant regions that would suggest DNA rearrangements of the sort seen in the immune system. These observations indicate that the diverse olfactory proteins are coded by a large number of distinct gene sequences.

Although it is unlikely from our data that DNA rearrangement is responsible for the generation of diversity among the putative odorant receptors, it remains possible that DNA rearrangements may be involved in the regulation of expression of this gene family. If each olfactory neuron expresses only one or a small number of genes, then a transcriptional control mechanism must be operative to choose which of the more than 100 genes within the family will be expressed in a given neuron. Gene conversion from one of multiple silent loci into a single active locus, as observed for the trypanosome variable surface glycoproteins (Van der Ploeg, 1991), provides one attractive model. The gene conversion event could be stochastic, such that a given neuron could randomly express any one of several hundred receptor genes, or regulated (perhaps by positional information), such that a given neuron could express only one or a small number of predetermined receptor types. Alternatively, it is possible that positional information in the olfactory epithelium controls the expression of the family of olfactory receptors by more classical mechanisms that do not involve DNA rearrangement. Whatever mechanisms will regulate the expression of receptor genes within this large multigene family, these

mechanisms must accommodate the requirement that olfactory neurons are regenerated every 30–60 days (Graziadei and Monti Graziadei, 1979), and therefore the expression of the entire repertoire of receptors must be accomplished many times during the life of an organism.

Receptor Diversity and the Central Processing of Olfactory Information

Our results suggest the existence of a large family of distinct odorant receptors. Individual members of this receptor family are likely to be expressed by only a small set of the total number of olfactory neurons. The primary sensory neurons within the olfactory epithelium will therefore exhibit significant diversity at the level of receptor expression. The question then emerges as to whether neurons expressing the same receptors are localized in the olfactory epithelium. Does the olfactory system employ a topographic map to discriminate among the numerous odors? The spatial organization of distinct classes of olfactory sensory neurons, as defined by receptor expression, can now be determined by using the procedures of *in situ* hybridization and immunohistochemistry with probes specific for the individual receptor subtypes. This information should help to distinguish between different models that have been proposed to explain the coding of diverse odorant stimuli (for review see Shepherd, 1985).

In one model, sensory neurons that express a given receptor and respond to a given odorant may be localized within defined positions within the olfactory epithelium. This topographic arrangement would also be reflected in the projection of olfactory sensory axons into discrete regions (glomeruli) within the olfactory bulb. In this scheme, the central coding to permit the discrimination of discrete odorants would depend, in part, on the spatial segregation of different receptor populations. Attempts to discern the topographic localization of specific receptors at the level of the olfactory epithelium has led to conflicting results. In some studies, electrophysiological recordings have revealed differences in olfactory responses to distinct odorants in different regions of the olfactory epithelium (Mackay-Sim et al., 1982; Thommesen and Doving, 1977). However, these experiments have been difficult to interpret since the differences in response across the epithelium are often small and are not observed in all studies (for example, see Sicard, 1985).

A second model argues that sensory neurons expressing distinct odorant receptors are randomly distributed in the epithelium but that neurons responsive to a given odorant project to restricted regions within the olfactory bulb. In this instance, the discrimination of  odors would be a consequence of the position of secondary neurons in the olfactory bulb but would be independent of the site of origin of the afferent signals within the epithelium. Mapping of the topographic projections of olfactory neurons has been performed by extracellular recordings from different regions of the bulb (Thommesen, 1978; Doving et al., 1980) and by 2-deoxyglucose autoradiography to map regional activity after exposure to different odorants (Stewart et al., 1979). These studies suggest that spatially localized groups of bulbar neurons preferentially respond to

different odorants. The existence of specific odorant receptors, randomly distributed through the olfactory epithelium, which converge on a common target within the olfactory bulb, would raise additional questions about the recognition mechanisms used to guide these distinct axonal subsets to their central targets.

Other sensory systems also spatially segregate afferent input from primary sensory neurons. The spatial segregation of information employed by the visual and somatosensory systems, for example, is used to define the location of the stimulus within the external environment as well as to indicate the quality of the stimulus. In contrast, olfactory processing does not extract spatial features of the odorant stimulus. Relieved of the necessity to encode information about the spatial localization of the sensory stimulus, the olfactory system of mammals may use the spatial segregation of sensory input solely to encode the identity of the stimulus itself. The molecular identification of the genes likely to encode a large family of olfactory receptors should provide initial insights into the underlying logic of olfactory processing in the mammalian nervous system.

Experimental Procedures



PCR

RNA was prepared from the olfactory epithelia of Sprague-Dawley rats according to Chirgwin et al. (1979) or using RNAzol B (Cinna/Biotex) and then treated with DNase I (0.1 U per μ g of RNA) (Promega). To obtain cDNA, this RNA was incubated at 0.1 μ g/ μ l with 5 μ M random hexamers (Pharmacia), 1 mM each of dATP, dCTP, dGTP, and TTP, and 2 U/ μ l RNase inhibitor (Promega) in 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2.5 mM MgCl₂, and 0.001% gelatin for 10 min at 22°C, and then for a further 45 min at 37°C following the addition of 20 U/ μ l Moloney murine leukemia virus reverse transcriptase (BRL). After heating at 95°C for 3 min, cDNA prepared from 0.2 μ g of RNA was used in each of a series of PCR experiments containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, 0.001% gelatin, 200 μ M each of dATP, dCTP, dGTP, and TTP, 2.5 U of Taq polymerase (Perkin Elmer Cetus), and 2 μ M of each PCR primer. PCR amplifications were performed according to the following schedule: 96°C for 45 s, 55°C for 4 min (or 45°C for 2 min), and 72°C for 3 min with 6 s extension per cycle for 48 cycles. The primers used for PCR were a series of degenerate oligonucleotides made according to the amino acid sequences found in transmembrane domains 2 and 7 of a variety of different members of the seven transmembrane domain protein superfamily (for example, see O'Dowd et al., 1989b). The regions used correspond to amino acids 60–70 and 286–295 of clone l15 (Figure 4). Each of five different 5' primers was used in PCR reactions with each of six different 3' primers. The 5' primers had the following sequences:

A1: AA(T/C)T(G/A)(G/C)ATI(C/A)TI(G/C)TIAA(T/C)(C/T)TIGCIGTIG-CIGA;
A2: AA(T/C)TA(T/C)TT(T/C)(C/A)TI(G/A)TIAA(T/C)CTIGCI(T/C)TIG-CIGA;
A3: AA(T/C)(T/C)(T/A)TT(T/C)(A/C)TIATI(T/A)CTCTIGCIT(G/C)IG-CIGA;
A4: (C/A)GTTI(C/T)TIATGTG(T/C)AA(C/T)CTI(T/A)(G/C)(C/T)TT(T/C)GCIGA;
A5: ACIGTTA(T/C)ATIACICA(T/C)(C/T)TI(A/T)(C/G)ATIGCIGA.

The 3' primers were:

B1: CTGII(C/T)(G/T)(G/A)TTCATIA(A/T)I(A/C)(C/A)(A/G)TAIA(T/C)IA-(T/C)GG(G/A)TT;
B2: (G/T)(A/G)T(C/G)(G/A)TTIAG(A/G)CA(A/G)CA(A/G)TAIATIATIGG(G/A)TT;
B3: TCIAT(G/A)TT(A/G)AAIGTIGT(A/G)TAIATIATIGG(G/A)TT;
B4: GC(C/T)TTIGT(A/G)AAIATIGC(A/G)TAIAG(G/A)AAIGG(G/A)TT;

B5: AA(A/G)TCIGG(G/A)(C/G)(T/A)ICGI(C/G)A(A/G)TAIAT(C/G)AI-
IGG(G/A)TT;
B6: (G/C)(A/T)I(G/C)(A/T)ICCIAC(A/G)AA(A/G)AA(A/G)TAIAT(A/
G)AAIGG(G/A)TT.

An aliquot of each PCR reaction was analyzed by agarose gel electrophoresis, and bands of interest were amplified further by performing PCR reactions on pipet tip ($\sim 1 \mu\text{l}$) plugs of the agarose gels containing those DNAs. Aliquots of these semipurified PCR products were digested with the restriction enzyme HaeIII or Hinfl, and the digestion products were compared with the undigested DNAs on agarose gels.

Isolation and Analysis of cDNA Clones

cDNA libraries were prepared according to standard procedures (Maniatis et al., 1982; Sambrook et al., 1989) in the cloning vector λ ZAP II (Stratagene) using poly(A)⁺ RNA prepared from Sprague-Dawley rat epithelia (see above) or from an enriched population of olfactory neurons that had been obtained by a "panning" procedure (L. B. and R. A., unpublished data) using an antibody against the H blood group antigen (Chembimed) found on a large percentage of rat olfactory neurons. In initial library screens, 8.5×10^5 independent clones from the olfactory neuron library and 1.8×10^6 clones from the olfactory epithelium library were screened (Maniatis et al., 1982) with a ^{32}P -labeled probe (Prime-it, Stratagene) consisting of a pool of gel-isolated PCR products obtained using primers A4 and B6 (see above) in PCR reactions, using as template olfactory epithelium cDNA, rat liver DNA, or DNA prepared from the two cDNA libraries. In later library screens, a mixture of PCR products obtained from 20 cDNA clones with the A4 and B6 primers was used as probe ("P1" probe). In initial screens, phage clones were analyzed by PCR using primers A4 and B6, and those that showed the appropriate size species were purified. In later screens all positive clones were purified, but only those that could be amplified with the B6 primer and a primer specific for vector sequence were analyzed further. To obtain plasmids from the isolated phage clones, phagemid rescue was performed according to the instructions of the manufacturer of λ ZAP II (Stratagene). DNA sequence analysis was performed on plasmid DNAs using the Sequenase system (United States Biochemical Corp.), initially with the A4 and B6 primers and later with oligonucleotide primers made according to sequences already obtained.

Northern and Southern Blot Analyses

For Northern blots, poly(A)⁺ RNAs from various tissues were prepared as described above or purchased from Clontech. One microgram of each RNA was size fractionated on formaldehyde-agarose gels and blotted onto nylon membranes (Maniatis et al., 1982; Sambrook et al., 1989). For Southern blots, genomic DNA prepared from Sprague-Dawley rat liver was digested with the restriction enzyme EcoRI or HindIII, size fractionated on agarose gels, and blotted onto nylon membranes (Maniatis et al., 1982; Sambrook et al., 1989). The membranes were dried at 80°C and then prehybridized in 0.5 M sodium phosphate buffer (pH 7.3) containing 1% bovine serum albumin and 4% SDS. Hybridization was carried out in the same buffer at 65°C – 70°C for 14–20 hr with DNAs labeled with ^{32}P . For the first Northern blot shown, the "P1" probe (see above under cDNA clone isolation) was used. For the second Northern blot shown, a mix of PCR fragments from seven divergent cDNA clones was used (see Southern blot below). For Southern blots, the region indicated in clone I15 by amino acids 118 through 251 was amplified from a series of divergent cDNA clones using PCR. The primers used for these reactions had the following sequences:

P1: ATGGCITA(T/C)GA(T/C)(C/A)GITA(T/C)GTIGC;
P4: AAIA(G/A)I(G/C)(A/T)IACIA(T/C)(G/C)(A/T)IA(G/A)(A/G)TGI(G/C)-
(A/T)I(C/G)C.

These DNAs (or a DNA encompassing transmembrane domains 2 through 7 for clone F6) were labeled and tested for cross-hybridization at 70°C . Those DNAs that did not show appreciable cross-hybridization were hybridized individually, or as a pool, to Southern blots at 70°C .

Acknowledgments

We would like to thank Ora Pearlstein, Bruce Brender, Tom Livelli, and

Lillian Eoyang for expert technical assistance. We are also grateful to Andrew Chess, Jane Dodd, Tom Jessell, Eric Kandel, and John Ngai for helpful discussions and for critically reading the manuscript, and to George Gaitanaris for helpful discussions during the course of this work. We also wish to thank Phyllis Kisloff and Miriam Gutierrez for expert assistance in the preparation of this manuscript.

This work was supported by the Howard Hughes Medical Institute and by grants from the Frederick R. Adler Education Fund and the National Institutes of Health (PO1-CA 23767) (to R. A.).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact. 

Received March 14, 1991; revised March 19, 1991.

References

- Amoore, J. E. (1982). Odor theory and odor classification. In *Fragrance Chemistry* (New York: Academic Press, Inc.), pp. 27–73.
- Baltimore, D. (1981). Gene conversion: some implications for immunoglobulin genes. *Cell* 24, 592–594.
- Boekhoff, I., Tareilus, E., Strotmann, J., and Breer, H. (1990). Rapid activation of alternative second messenger pathways in olfactory cilia from rats by different odors. *EMBO J.* 9, 2453–2458.
- Bonner, T. I., Buckley, N. J., Young, A. C., and Brann, M. R. (1987). Identification of a family of muscarinic acetylcholine receptor genes. *Science* 237, 527–532.
- Bouvier, M. W., Hausdorff, A., De Blasi, A., O'Dowd, B. F., Kobilka, B. K., Caron, M. G., and Lefkowitz, R. J. (1988). Removal of phosphorylation sites from the β -adrenergic receptor delays the onset of agonist-promoted desensitization. *Nature* 333, 370–373.
- Breer, H., Boekhoff, I., and Tareilus, E. (1990). Rapid kinetics of second messenger formation in olfactory transduction. *Nature* 345, 65–68.
- Bronshtain, A. A., and Minor, A. V. (1977). Regeneration of olfactory flagella and restoration of the electroolfactogram following application of Triton X-100 to the olfactory mucosa of frogs. *Tsitologija* 19, 33–39.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18, 5294–5299.
- Dhallan, R. S., Yau, K.-W., Schrader, K. A., and Reed, R. R. (1990). Primary structure and functional expression of a cyclic nucleotide-activated channel from olfactory neurons. *Nature* 347, 184–187.
- Doving, K. B., Selset, R., and Thommessen, G. (1980). Olfactory sensitivity to bile acids in salmonid fishes. *Acta Physiol. Scand.* 108, 123–131.
- Egel, R. (1981). Intergenic conversion and reiterated genes. *Nature* 290, 191–192.
- Flanagan, J. G., Lefranc, M.-P., and Rabbitts, T. H. (1984). Mechanisms of divergence and convergence of the human immunoglobulin $\alpha 1$ and $\alpha 2$ constant region gene sequences. *Cell* 36, 681–688.
- Graziadei, P. P. C., and Monti Graziadei, G. A. (1979). Neurogenesis and neuron regeneration in the olfactory system of mammals. I. Morphological aspects of differentiation and structural organization of the olfactory sensory neurons. *J. Neurocytol.* 8, 1–18.
- Hamm, H. E., Deretic, D., Arendt, A., Hargrove, P. A., Koenig, B., and Hofmann, K. P. (1988). Site of G protein binding to rhodopsin mapping with synthetic peptides to the α subunit. *Science* 241, 832–835.
- Hood, L., Kronenberg, M., and Hunkapiller, T. (1985). T cell antigen receptors and the immunoglobulin supergene family. *Cell* 40, 225–229.
- Jones, D. T., and Reed, R. R. (1989). $\text{G}_{\alpha\text{o}}$: an olfactory neuron-specific G-protein involved in odorant signal transduction. *Science* 244, 790–795.
- Kobilka, B. K., Kobilka, T. S., Daniel, K., Regan, J. W., Caron, M. G., and Lefkowitz, R. J. (1988). Chimeric α_2 - β_2 -adrenergic receptors: delineation of domains involved in effector coupling and ligand binding specificity. *Science* 240, 1310–1316.
- Lancet, D. (1986). Vertebrate olfactory reception. *Annu. Rev. Neurosci.* 9, 329–355.

- Mackay-Sim, A., Shamar, P., and Moulton, D. G. (1982). Topographic coding of olfactory quality: odorant-specific patterns of epithelial responsiveness in the salamander. *J. Neurophysiol.* **48**, 584–596.
- Maeda, N., and Smithies, O. (1986). The evolution of multigene families: human haptoglobin genes. *Annu. Rev. Genet.* **20**, 81–108.
- Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).
- Moulton, D. G., and Beidler, L. M. (1967). Structure and function in the peripheral olfactory system. *Physiol. Rev.* **47**, 1–52.
- Nakamura, T., and Gold, G. (1987). A cyclic nucleotide-gated conductance in olfactory receptor cilia. *Nature* **325**, 442–444.
- Nathans, J., Thomas, D., and Hogness, D. S. (1986). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science* **232**, 193–202.
- O'Dowd, B. F., Hnatowich, M., Caron, M. G., Lefkowitz, R. J., and Bouvier, M. (1989a). Palmitoylation of the human β_2 -adrenergic receptor. Mutation of CYS³⁴¹ in the carboxyl tail leads to an uncoupled non-palmitoylated form of the receptor. *J. Biol. Chem.* **264**, 7564–7569.
- O'Dowd, B. F., Lefkowitz, R. J., and Caron, M. G. (1989b). Structure of the adrenergic and related receptors. *Annu. Rev. Neurosci.* **12**, 67–83.
- Pace, U., Hanski, E., Salomon, Y., and Lancet, D. (1985). Odorant-sensitive adenylate cyclase may mediate olfactory reception. *Nature* **316**, 255–258.
- Reed, R. R. (1990). How does the nose know? *Cell* **60**, 1–2.
- Rhein, L. D., and Cagan, R. H. (1980). Biochemical studies of olfaction: isolation, characterization, and odorant binding activity of cilia from rainbow trout olfactory rosettes. *Proc. Natl. Acad. Sci. USA* **77**, 4412–4416.
- Rhein, L. D., and Cagan, R. H. (1983). Biochemical studies of olfaction: binding specificity of odorants to a cilia preparation from rainbow trout olfactory rosettes. *J. Neurochem.* **41**, 569–577.
- Rushton, W. A. H. (1955). Foveal photopigments in normal and colour-blind. *J. Physiol.* **129**, 41–42.
- Rushton, W. A. H. (1965). A foveal pigment in the deuteranope. *J. Physiol.* **176**, 24–37.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Second Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).
- Shepherd, G. M. (1985). Are there labeled lines in the olfactory pathway? In *Taste, Olfaction and the Central Nervous System* (New York: The Rockefeller University Press), pp. 307–321.
- Sicard, G. (1985). Olfactory discrimination of structurally related molecules: receptor cell responses to camphoraceous odorants. *Brain Res.* **326**, 203–212.
- Sklar, P. B., Anholt, R. R. H., and Snyder, S. H. (1986). The odorant-sensitive adenylate cyclase of olfactory receptor cells: differential stimulation by distinct classes of odorants. *J. Biol. Chem.* **261**, 15538–15543.
- Stewart, W. B., Kauer, J. S., and Shepherd, G. M. (1979). Functional organization of the rat olfactory bulb analysed by the 2-deoxyglucose method. *J. Comp. Neurol.* **185**, 715–734.
- Strader, C. D., Sigal, I. S., Register, R. B., Candelore, M. R., Rands, E., and Dixon, R. A. F. (1987). Identification of residues required for ligand binding to the β -adrenergic receptor. *Proc. Natl. Acad. Sci. USA* **84**, 4384–4388.
- Strader, C. D., Sigal, I. S., and Dixon, R. A. F. (1989). Structural basis of β -adrenergic receptor function. *FASEB J.* **3**, 1825–1832.
- Thommesen, G. (1978). The spatial distribution of odour induced potentials in the olfactory bulb of char and trout (Salmonidae). *Acta Physiol. Scand.* **102**, 205–217.
- Thommesen, G., and Doving, K. B. (1977). Spatial distribution of the EOG in the rat. A variation with odour stimulation. *Acta Physiol. Scand.* **99**, 270–280.
- Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature* **302**, 575–581.
- Van der Ploeg, L. H. T. (1991). Antigenic variation in African trypanosomes: genetic recombination and transcriptional control of VSG genes. In *Gene Rearrangement* (New York: Oxford University Press), pp. 51–98.
- Wald, G., Brown, P. K., and Smith, P. H. (1955). Iodopsin. *J. Gen. Physiol.* **38**, 623–681.