

Tissue and Zonal-Specific Expression of an Olfactory Receptor Transgene

Pankaj Qasba and Randall R. Reed

The Howard Hughes Medical Institute, Departments of Molecular Biology and Genetics and Neuroscience, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Discrimination of odorants is thought to arise from the selective expression of one of a small number of individual receptors in any single olfactory neuron. Receptor genes are expressed in a small subset of neurons throughout a zonally restricted region of the sensory epithelium. We demonstrate that a 6.7 kb region upstream of the *M4* olfactory receptor coding region was sufficient to direct expression in olfactory epithelium. Moreover, reporter expression recapitulated the zonal restriction and distributed neuronal expression observed for endogenous olfactory receptors. Transgenic lines were obtained that directed

expression in two different receptor zones, one of which was identical to the endogenous *M4* receptor. When the reporter was expressed in the same zone as the endogenous *M4* receptor, the two expression patterns were, in large part, non-overlapping. These results suggest a model in which important regulatory elements are located in close proximity to transcription initiation sites of the olfactory receptor genes and receive information defining zonal patterning via long-range processes.

Key words: transgene; receptors; zones; M4; sensory neurons; olfactory epithelium

The remarkable sensitivity of the mammalian olfactory system to discriminate a vast variety of diverse odors is the result of complex anatomical, cellular, biochemical, and genetic specialization. Humans, for example, can discriminate among 10,000 or so distinct odors, which include resolution of many stereoisomers into distinct perceived odor quality (Amoore, 1970). The underlying mechanism or mechanisms that contribute to the exquisite discriminatory capability of the olfactory system remain to be answered.

Cells within the olfactory epithelium follow a highly ordered developmental program, resulting in a high level of expression of gene products essential for odorant signal transduction. The primary event in signal transduction occurs in the olfactory sensory neurons that line the posterior nasal cavity (Jones et al., 1988; Reed, 1992b). Unlike other neurons, olfactory neurons are replaced continuously by the underlying layer of stem cells (Graziadei and Graziadei, 1979; Monti-Graziadei and Graziadei, 1979). The neuronal precursor cells undergo a complex developmental program, followed by differentiation to mature olfactory neurons.

The assumption that olfactory receptors would display homology to other G-protein-coupled receptors led to the identification of a large multigene family thought to encode odorant receptors in rat (Buck and Axel, 1991). Subsequently, additional members of this family have been identified (Levy et al., 1991; Parmentier et al., 1992; Strotmann et al., 1992, 1994; Raming et al., 1993; Sullivan, 1996). These receptors exhibit, as predicted, a seven transmembrane domain topology characteristic of the G-protein-coupled receptor superfamily. The size of the odorant receptor repertoire in mammals may include as many as 1000 genes (Levy

et al., 1991; Reed, 1992a) that are divided into subfamilies of multiple, highly related receptors (Buck and Axel, 1991). The small number (0.5–1%) of olfactory neurons that express individual odorant receptor genes within the rodent neuroepithelium suggest that each neuron might express only one or a few odorant receptor genes (Ressler et al., 1993; Vassar et al., 1994). Therefore, each cell possesses a unique functional identity or receptor phenotype that is specific to the receptor or receptors it expresses. How does an olfactory neuron express a restricted subset of receptors selectively from the enormous odorant receptor repertoire? This differential expression of individual members of the olfactory receptor family in the sensory neurons represents a complex problem in gene regulation.

In mammals, the olfactory epithelium appears to be organized into distinct topographic regions or zones in which expression of a particular receptor gene appears to be restricted to one of the four zones in the epithelium (Ressler et al., 1993; Vassar, 1993). Within the zone the distribution of neurons expressing a given receptor is random. Chromosomal mapping studies have revealed clusters of odorant receptor genes at a single locus, and numerous such loci have been mapped to different chromosomes (Sullivan, 1996). However, receptors expressed in the same zone map to different loci, and a single locus can contain genes expressed in different zones. Selective receptor expression could depend on using transcription factors or novel mechanisms such as gene rearrangement or gene conversion (Hicks et al., 1977; Davis and Davidson, 1986; Reed, 1992a).

To assess the feasibility of these models, we examined the pattern of odorant receptor gene expression with a transgenic approach, using a single-copy mouse olfactory receptor gene from a subfamily that contained only one member. We constructed transgenic lines, using the receptor 5' flanking DNA to define the regulatory regions that may be present within the putative 6.7 kb promoter region involved in tissue specificity and zonal patterns of receptor expression for this receptor. We demonstrate that expression of a reporter gene driven by the 6.7 kb DNA fragment

Received Sept. 23, 1996; revised Oct. 17, 1997; accepted Oct. 22, 1997.

We are grateful for the contributions of Karen Schrader and to the members of the Reed laboratory for stimulating and supportive discussions.

Correspondence should be addressed to Dr. Randall R. Reed, Room 800 PCTB, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205.

Copyright © 1997 Society for Neuroscience 0270-6474/97/180227-10\$05.00/0

upstream of a receptor coding region imparts several aspects of the complex receptor regulation. Specifically, this putative odorant receptor promoter was sufficient to direct olfactory receptor expression in a tissue-specific, zonal-specific manner, and this expression within the zone was seen only in a subset of neurons.

MATERIALS AND METHODS

Cloning of mouse *M4*

A 550 bp PCR product, representing a heterogeneous mix of mouse genes encoding odorant receptor coding region segments (transmembrane region II and VI; Levy et al., 1991), was used as a probe to screen a mouse genomic library in λ FIX at low stringency (wash conditions, $2\times$ SSC at 55°C). One of the clones isolated, *M4*, was restriction-mapped further and analyzed in detail. The majority of the *M4* coding sequence was contained on a 1.0 kb *Hind*III fragment. An adjacent 5' *Hind*III fragment of 6.7 kb, potentially containing *M4* upstream sequences, was cloned and characterized.

For genomic DNA hybridization analysis, mouse DNA was digested separately with *Hind*III, *Bam*HI, and *Eco*RI, size-fractionated on a 1% agarose gel, and transferred to Genescreen Plus. The filter was UV-cross-linked (UV Stratalinker 2400, Stratagene, La Jolla, CA); prehybridized at 65°C in $6\times$ SSC, $5\times$ Denhardt's, 0.5% SDS, and 100 $\mu\text{g}/\text{ml}$ denatured, fragmented salmon sperm DNA; and hybridized in the same buffer at 65°C with ^{32}P -labeled fragment. Wash conditions were as indicated in the figure legends.

5' rapid amplification of the cDNA ends (RACE)

One microgram of total RNA from mouse olfactory tissue was reverse-transcribed by the Life Technologies (Gaithersburg, MD) first-strand cDNA synthesis kit (Superscript). An aliquot (10 μl) was used for the 5' RACE reaction as described (Frohman et al., 1988). The primers, RACE I and RACE II, and the nested *M4*-specific primers, PQ3 and PQ4, were used for the amplification reaction. Amplified products were resolved on a native polyacrylamide gel, cloned into pBluescript, and sequenced with T7 and T3 primers. The PQ5 and PQ6 primers, based on the sequence of the 305 and 283 bp products from independent RACE reactions, were used in combination with the T7 or T3 primers to map the location of the upstream exon on the 6.7 kb *Hind*III genomic fragment cloned in pBluescript.

Primers used

Following is a list of the primers used: RACE I, 5'-GATGGATCCTG-CAGAAGCTTTTTTTTTTTTTTTTTT; RACE II, 5'-GATGGATC-CTGCAGAAGC; PQ1, 5'-CTCATCAGCCTCAGGAGCCATG (+*M4*); PQ2, 5'-CAGAAGTGACACTTAGATACTT (-*M4*); PQ3, 5'-GAA-GGCTGAGAAACCAAGAG (*M4*, -39 bp from ATG); PQ4, 5'-CTAGTCGAATTCCTGATATGGTGTAGTTCTGAC (*M4*, -7 bp from ATG with *Eco*RI); PQ5, 5'-CAGGATTGTTTCAGAAACCCAGC-CCA (RACE, bottom strand); PQ6, 5'-TCCCACAGGATTGTTTCAG (RACE, top strand); PQ33, 5'-CGCTCGCCACTTCAACATCAACG (β -galactosidase top strand, +2400); PQ36, 5'-GCGACCCGCATTGAC-CCCT (β -galactosidase bottom strand, 2651); PQ37, 5'-GCATGCC-ACATACATGCCAG (junction fragment *E5*, top strand); PQ39, 5'-CCTCTGTATCTCAGGGGCACA (junction fragment *E5*, bottom strand); PQ57, 5'-CTTAAGTGCGGCCGACCATCCAGCGCCAC (β -galactosidase top strand, +1980); PQ59, 5'-GAATTCAGCGG-CGCGCCGGCAACTCTGGCTC (β -galactosidase bottom strand, 2260); PQ74, 5'-GGTTATCTCCAACATACCCAACCTTCG (junction fragment *D5*, top strand); PQ79, 5'-CATAATCTCACTTTGCTG-AAACATCC (junction fragment *D5*, bottom strand); RR318, 5'-GGACTTGCTATTCTGTGCATCTAG (mouse Protamine 1, bottom strand 920); and RR319, 5'-GAAGTGCGGAGCGATACACCGCAT (β -galactosidase top strand, +2700).

Cloning of junction fragment or fragments

The *E5* transgene integration target site was isolated by using genomic DNA from the liver tissue of *E5* animals, followed by digestion with *Bam*HI and size fractionation on a 0.75% agarose gel. DNA fragments (6–12 kb) were excised from the gel, electroeluted, and cloned into λ DASHIII (Stratagene). Approximately 6×10^4 plaques were screened with a ^{32}P -labeled 2 kb (*Bam*HI-*Sst*I) fragment isolated from the β -galactosidase reporter gene. Analysis of one of the positive clones revealed an 11 kb DNA fragment that contained β -galactosidase se-

quences fused to mouse genomic DNA and thus represented the insertion target site and the 3' flanking transgene-genomic DNA junction region.

The *D5* transgene integration site was cloned by an inverse PCR approach. Genomic DNA from liver tissue of *D5* transgenic mice was digested to completion with *Sac*I and ligated overnight at 1 $\mu\text{g}/\text{ml}$ final DNA concentration. Primers PQ33 and PQ36 were used in a PCR reaction, followed by a second round of PCR amplification with PQ57 and PQ59 primers containing *Not*I restriction sites. A unique 1.8 kb PCR product was generated by the ligated reaction. No products were seen in the ligase minus controls or in PCR reactions containing a single primer. The 1.8 kb product was cloned in pBluescript and sequenced. Sequence analysis showed that the 1.8 kb fragment contained 330 bp of apparent mouse genomic DNA sequence immediately adjacent to sequences derived from the reporter. Primer pairs PQ 37/PQ39 and PQ74/PQ79 derived from the *E5* and *D5* junction flanking DNA, respectively, were each used to screen a mouse yeast artificial chromosome (YAC) DNA library (Research Genetics, Huntsville, AL) according to the manufacturer's recommendations. Overlapping YAC clones were identified for each of the transgene insertion sites. All YAC clones were grown routinely in $\text{ura}^- \text{trp}^-$ minimal media. Total genomic DNA from the YAC-containing strains was digested with *Hind*III, size-fractionated on 1% agarose gels, and transferred to nylon membrane.

Chromosome mapping

DNA samples derived from 94 backcrossed animals [(*C57BL/6J* \times *SPRET/Ei*) \times *SPRET/Ei*, BSS panel, Jackson Laboratories, Bar Harbor, ME] were used to map the two transgene insertion sites. Primer pairs PQ37/39 and PQ74/79, derived from the *E5* and *D5* sites of transgene insertion, were used and scored for the presence (BL6) or absence (*SPRET/Ei*) of the expected band. Linkage analysis was performed with Map Manager (Manly and Elliott, 1991; Manly, 1993).

Transgenic DNA construct

The DNA fragment containing the 6.7 kb region upstream of the *M4* translation start site was fused to the *Escherichia coli* β -galactosidase gene/mouse protamine-I 3' untranslated sequence (Peschon et al., 1987) such that the translation start of the reporter ATG was at the same position as the mouse *M4* initiation codon. The 10.3 kb fragment was prepared for injection by restriction digestion with *Xho*I and *Not*I to excise the fragment from the vector backbone and was purified by agarose gel electrophoresis. The transgene was injected into the pronuclei of *B6/C3* embryos at 4–6 ng/ μl , and injected eggs were transplanted into *CD1* mothers (Hogan et al., 1986). Potential founder animals were tested for the presence of the transgene by PCR analysis of tail DNA, using the primer pair RR318/RR319 that generated a 728 bp PCR product. Animals that were positive by PCR were crossed to *C57BL/6J* or *DBA/2J* animals (Jackson Laboratories) to yield F_1 offspring, hemizygous for the transgene locus. Each transgenic line was maintained by sister-brother mating of the transgenic F_1 animals.

Detection of transgene expression and immunocytochemistry

Histochemical methods were adapted and modified from Danciger et al. (1989) and Stefanin et al. (1967). Transgene expression in the mouse olfactory epithelium was examined by staining with the chromogenic substrate X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside). Mice (4–6 weeks old) were anesthetized deeply and fixed by intracardiac perfusion with >50 ml of PBS, followed immediately by >50 ml of fixative [0.5% glutaraldehyde and picric acid-formaldehyde (PAF/Zamboni) (Stefanin et al., 1967) in buffered PBS, or Bouin's fixative (for antibody staining)]. The mouse nose was dissected, post-fixed for 2–4 hr at 4°C , and submerged in 30% sucrose in PBS overnight. The tissue was embedded in OCT medium (Miles, Elkhart, IN) and quick-frozen in liquid nitrogen. Serial coronal sections (10 μm) were cut on a Microm cryostat at -20°C , mounted on Superfrost/Plus slides (Fisher Scientific, Pittsburgh, PA), air-dried, and stored at -80°C until use. On thawing, the slides were washed in PBS-Mg $^{2+}$, fixed for 5 min in the appropriate fixative, and subjected to X-gal or antibody staining. The pattern of expression of *M4* was defined by the following methods.

β -Galactosidase staining. The fixed slides were washed in PBS-Mg $^{2+}$ for 5 min and incubated in X-gal staining solution (Emson et al., 1990) at 37°C for 12–18 hr. After washes in PBS-Mg $^{2+}$, the color reaction was stopped by a final wash in 2% paraformaldehyde in PBS and mounted in Aquamount (Polysciences, Warrington, PA).

Antibody. Rabbit antisera directed against the last 20 amino acids of

M4 were affinity-purified over a specific peptide column and used in the immunohistochemical experiments. Sections were rehydrated and post-fixed in Bouin's solution as described above and immersed in ice-cold methanol for 3 min to ensure permeabilization of cell membranes. After several washes in PBS-Mg²⁺, slides were incubated with 5% normal goat serum for 1 hr, reacted with the affinity-purified antibody (diluted 1:500 with 2% normal goat serum in PBS-Mg²⁺) for 4 hr at room temperature, and incubated with biotinylated goat anti-rabbit IgG at 1:200 for 2 hr (Vector Laboratories, Burlingame, CA), followed by avidin–biotin–peroxidase (ABC) complex at 1:100 for 1 hr (Vector). The slides were washed four times for 5 min in PBS-Mg²⁺ between each incubation. Immunoreactive material was visualized by incubation with diaminobenzidine (DAB, 0.5 mg/ml) and hydrogen peroxide (0.015%), followed by a 10 min wash in H₂O.

Antibody and β -galactosidase double labeling. Sections were perfused and post-fixed in Zamboni and stained for β -galactosidase activity with X-gal (5 hr), followed by antibody staining as described above except that incubation with primary antibody was extended to 12–18 hr at 4°C.

RESULTS

Isolation of a mouse odorant receptor gene

The olfactory receptor gene family displays highest homology in the regions corresponding to transmembrane domains II and VI. Degenerate oligonucleotides based on these regions were used in an reverse transcription-PCR (RT-PCR) reaction with olfactory cDNA to generate a product of 550 bp containing a heterogeneous population mixture of receptor sequences (Levy et al., 1991). A mouse genomic DNA library was screened with this probe at low stringency, and positive clones were identified at a frequency of 0.3%, suggesting the existence of ~600 receptor-related sequences per haploid genome (Buck and Axel, 1991; Reed, 1992a). Among 12 recombinant phages that were analyzed in some detail, the isolate *M4* was chosen for further study on the basis of several criteria. The *M4* phage DNA encoded an open reading frame of 317 amino acids that share homology with other members of the odorant receptor family. The genomic DNA in the *M4* phage extended >6.7 kb upstream and 2.2 kb downstream of the coding region. Finally, Southern blot analysis demonstrated that the *M4* coding region hybridized at high stringency to a single restriction fragment in mouse genomic DNA (Fig. 1). Although many odorant receptors exist in subfamilies of highly homologous genes, the use of *M4*, which appeared to have only a single member in its subfamily, was chosen to simplify the analysis of transcriptional regulation and expression of odorant receptors. Specifically, we wished to compare the expression of the endogenous *M4* gene by *in situ* hybridization and specific antibodies with the expression of reporter constructs carrying potential regulatory sequences.

Endogenous expression pattern of *M4*

The pattern of *M4* expression was examined by RT-PCR in six different tissues, using primers specific for the *M4* coding region. A 1 kb product corresponding to the *M4* RNA was seen in olfactory tissue as well as a faint product in pancreas that was not reproduced in subsequent experiments and may represent low-level genomic DNA contamination of this sample (Fig. 2a). The use of RT-PCR to examine the expression pattern of olfactory receptors was hampered by the characteristic absence of introns within the coding regions. Nonetheless, this experiment demonstrated that *M4* expression was restricted predominantly to olfactory epithelium.

The restricted expression of the *M4* receptor within the olfactory epithelium was assessed by immunohistochemistry, using specific antibody to this receptor. The polyclonal antisera were generated to the last 20 amino acids of the *M4* coding sequence in

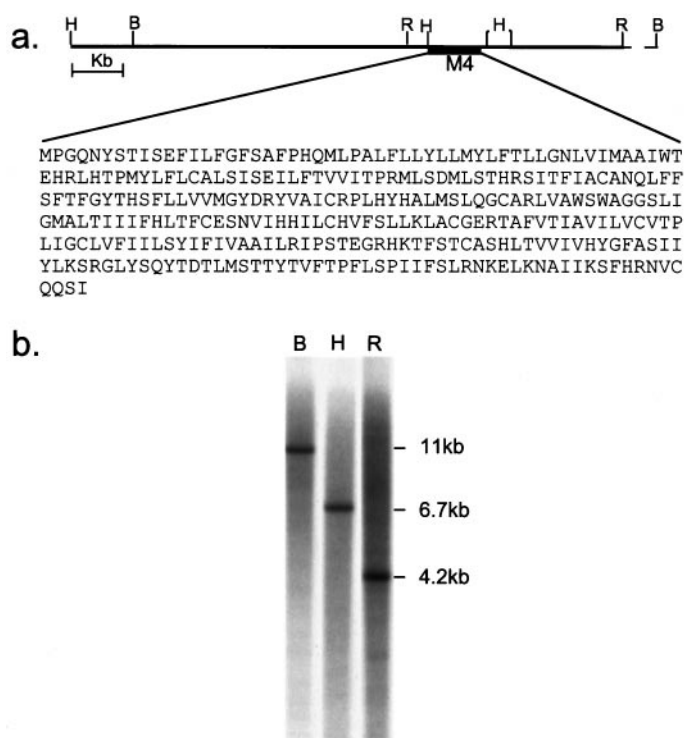


Figure 1. *M4* is a single copy gene. *a*, Partial restriction map of the mouse *M4* gene locus showing the schematic location of the *M4* open reading frame (filled box). Deduced protein sequence shows 317 amino acids, representing the *M4* coding region. *b*, *Bam*HI; *H*, *Hind*III; *R*, *Eco*RI. *b*, Five micrograms of mouse liver DNA were digested with the indicated enzymes in each lane. The Southern blot was probed with the ³²P-labeled *M4* coding sequence. The probe showed a single band hybridizing in each lane. The size shown on the right corresponds to each of the hybridized bands.

a region that displays the greatest divergence among members of this receptor family (Buck and Axel, 1991; Levy et al., 1991). Immunoreactive cells were observed in the epithelial layer at positions characteristic of the mature sensory neurons. The most intense staining was seen in the dendrites, dendritic knobs, and cilia of the receptor cells, thereby supporting the identification of the immunoreactive cells as the sensory neurons (Figure 2b).

The *M4* expression pattern was examined by using serial coronal sections spanning the entire anterior–posterior extent of the olfactory neuroepithelium. Expression of *M4* followed a zonal pattern, and within this zone only a subset of neurons expressed the *M4* protein. This pattern was consistent in the two nasal cavities and between different animals. The restriction of odorant receptor family expression to a single zone within the olfactory epithelium has been characterized extensively in several laboratories (Ressler et al., 1993; Vassar, 1993). The zone of expression of the *M4* receptors correlates precisely with the K4/K7 receptor subfamily (Ressler et al., 1993), using *in situ* hybridization studies. In the nomenclature of Buck, this pattern corresponds to zone-II (Buck, 1992). No antibody staining was seen in the olfactory stem cells or in the supporting sustentacular cells. *In situ* results using the *M4* coding region as a probe (data not shown) paralleled the data obtained with the *M4*-specific antibody. On the basis of the reported density of olfactory neurons within the mouse neuroepithelium (Ressler et al., 1993) and the densities of the *M4*-positive neurons in our studies, we estimated that *M4* is expressed in ~0.5–1% of olfactory neurons within the zone.



Figure 2. Endogenous expression of *M4*. RNA from different tissues was used in a RT-PCR assay with primers PQ1 and PQ2, spanning the *M4* coding region. PCR products (1 kb) are shown resolved on a 1% agarose gel. The grouped tissues represent two independent sets of RT-PCR reactions with brain, olfactory bulb, olfactory neuroepithelium (nose), liver, pancreas, and testis. Sections (10 μ m) cut through the mouse nasal cavity were used in immunohistochemical staining, using anti-*M4* antibody. A restricted distribution of the *M4*-positive cells is shown confined to the posterior loop, indicated by arrowheads (middle panel). A higher magnification of the section is shown in the bottom panel; *M4*-expressing neurons are marked with arrowheads.

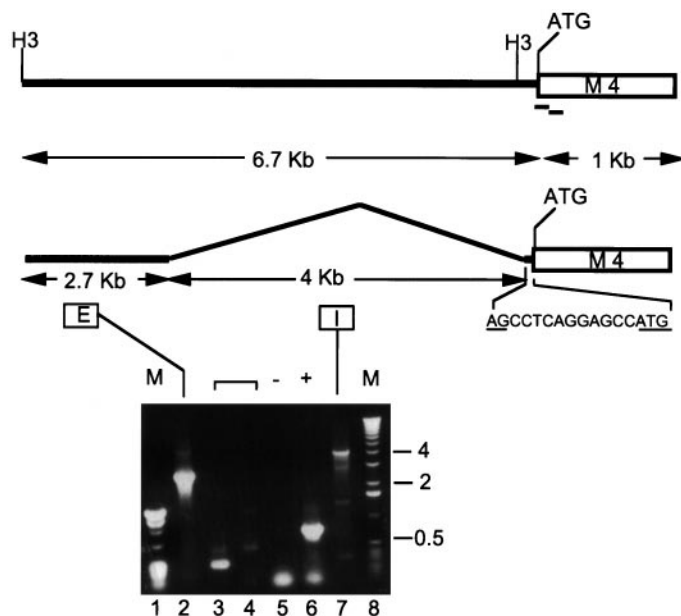


Figure 3. Mapping and sequence analysis of the *M4* transcription start site. A schematic diagram shows the *M4* ATG and the 5' proximal region (6.7 kb). The open box represents the 1 kb *M4* coding region, and the *M4*-specific (PQ3-PQ4 nested) primers are shown by the two bars used in the RACE reaction. Results from the RACE analysis mapped the transcription start site 4 kb 5' to the ATG, resulting in 2.7 kb of untranscribed sequences 5' of the 300 bp exon. The underlined *AG* and *ATG* represent the splice donor site and the translation start site, respectively. The bottom panel shows the PCR products generated from the genomic DNA as the template, 2.7 kb containing the 300 bp exon (*E*), lane 2. A 4 kb PCR product represents the intron (*I*), lane 7. Lanes 3-5 represent the negative controls and lane 6 the positive PCR control (see Materials and Methods). *M*, Marker lanes.

Proximal promoter has a 4 kb intron

A 5' RACE analysis was used to identify sequences upstream of the translation and transcription start sites (Frohman et al., 1988) and to define the intron/exon boundaries in the *M4* 5' untranslated region. A 330 bp product observed in reactions with olfactory epithelial RNA, *M4*-specific and RACE-specific primers represented 5' untranslated sequences of the *M4* message. Three experiments with different RACE primers generated products that were slightly smaller than the original 330 bp fragment, suggesting that the 5' end of the *M4* RNA and the site of transcription initiation had been reached. Oligonucleotide primers were designed to the 5' and 3' ends of the 330 bp RACE and used in a PCR reaction to identify the location of these sequences on the *M4* genomic template. All of the sequences obtained in the RACE experiments resided in a single 300 bp exon located 4 kb 5' to the *M4* initiation codon (Fig. 3). All of the olfactory receptors analyzed to date contain one or multiple additional exons upstream of a single coding-region exon. A canonical 3' splice site was observed 12 bp upstream of the *M4* coding region. The results suggested that the 6.7 kb fragment upstream of the *M4* translation start site region contained the transcription initiation site and possibly the regulatory regions/*cis*-elements that may be involved in *M4* receptor regulation and zonal expression.

To determine whether these upstream sequences in the mouse *M4* odorant gene were sufficient to direct expression in the olfactory neuroepithelium, we generated a reporter construct consisting of 6.7 kb upstream of the *M4* coding region fused to a β -galactosidase-mouse protamine reporter cassette (Fig. 4). This

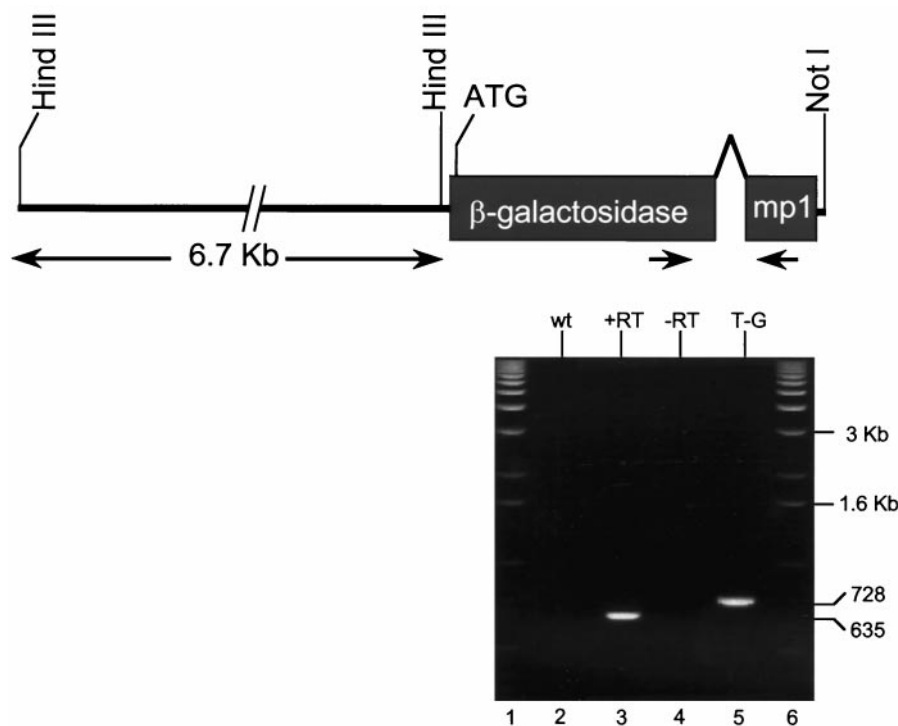


Figure 4. Structure of the *M4* transgene construct. The 6.7 kb DNA 5' to the *M4* ATG was fused to the *E. coli* gene encoding *lacZ* and the 3' end of the mouse protamine 1 structural gene (β -galactosidase, *mp1*, shaded boxes). The *mp1* gene provides a poly(A⁺) addition site and a 93 bp intron. Primers designed across the intron are shown by the arrows. Below, An agarose gel showing a 635 bp PCR product generated from the transgene message, lane 3, and a 728 bp from the genomic transgene (*T-G*) DNA, lane 5. No products are seen in lanes 2 and 4, representing the wild-type mouse DNA and samples with no reverse transcriptase (–RT), respectively. Marker lanes are shown as lanes 1 and 6.

reporter gene construct hopefully would enable a functional analysis of the putative promoter region of *M4* in transgenic animals and would allow identification of functional regulatory regions present within this 6.7 kb fragment. Eight lines of transgenic mice were identified that carried the introduced construct and transmitted the integrated sequences through the germline. The copy number in these transgenic lines varied between 2 and 10 copies integrated at a single site. Although each transgenic line contained copies that appeared to be intact by restriction analysis, additional copies with truncations of the 5' and 3' sequences were observed in most of the lines.

***M4* transgene expression is zonally restricted and, within the zone, expressed only by a subset of neurons**

Expression of the reporter gene was examined initially by RT-PCR in five tissues: liver, brain, spleen, testis, and olfactory epithelium. Primers (RR318 and RR319) were chosen that flanked a 93 bp intron at the 3' end of the reporter cassette and generated a 728 bp product from the genomic DNA and a 635 bp product from messenger RNA to permit discrimination between the transgene reporter expression and genomic contamination (Fig. 4). In three of the lines, RT-PCR analysis revealed transgene expression in olfactory epithelium and in testis, whereas in the other lines no expression was detected in any of the tissues examined. Expression in testis has been reported for other members of the odorant receptor family and may reflect the propensity of this tissue to display "fortuitous" expression of transgenes and endogenous genes, with no apparent function in that tissue. Expression in two of the lines, *D5* and *E5*, was examined in additional tissues, and expression was still restricted to olfactory tissue and testis (Fig. 5).

The three transgenic lines (*D2*, *D5*, and *E5*) were examined for β -galactosidase (*lacZ*) enzyme expression in tissue sections from nose, liver, spleen, and testis by staining with the chromogenic substrate X-gal. No staining was detected in the spleen or liver in

any of the lines. In testis, moderate staining was observed in a subset of cells but was indistinguishable in wild-type and transgenic mice. In each of the three lines in which transgene expression was detected, expression within the olfactory epithelium was restricted zonally and within that zone was observed only in a subset of neuronal cells.

Two lines, *D5* and *E5*, revealed robust expression of β -galactosidase in the neuroepithelium that resembled the endogenous *M4* expression in several respects. The reporter expression was confined to the epithelial layer of the olfactory tissue and displayed a bilaterally symmetric and zonally restricted pattern (Fig. 6 *a,b*). Within the neuroepithelium only a small fraction (~1%) of the cells expressed the reporter. These β -galactosidase-expressing cells displayed X-gal staining in dendritic processes leading to the luminal surface. No X-gal staining was observed in other major cell types of the epithelium, the stem cells and the supporting sustentacular cells (Fig. 6*d*).

The reporter transgenes were expressed in two distinct topographical patterns within the neuroepithelium. The β -galactosidase-positive cells in the *D5* line were restricted to the same zone as the endogenous *M4* receptor (zone II), whereas staining in the *E5* line was confined exclusively to zone IV. These observations suggest that a construct carrying the 6.7 kb 5' segment was sufficient to direct β -galactosidase expression in a tissue-specific and zonal-specific manner in the olfactory neuroepithelium but was insufficient by itself to define the particular zone where the expression will occur.

***M4* and *D5* express in the same zone**

With analysis of the zonal expression pattern of the two transgene lines, it was evident that *D5* X-gal staining was in the same zone as that seen with *M4* endogenous antibody staining. Figure 6, *e* and *f*, shows the colocalization of anti-*M4* peroxidase staining and histochemical X-gal staining in zone II of the olfactory neuroepithelium. Furthermore, serial coronal sections spanning

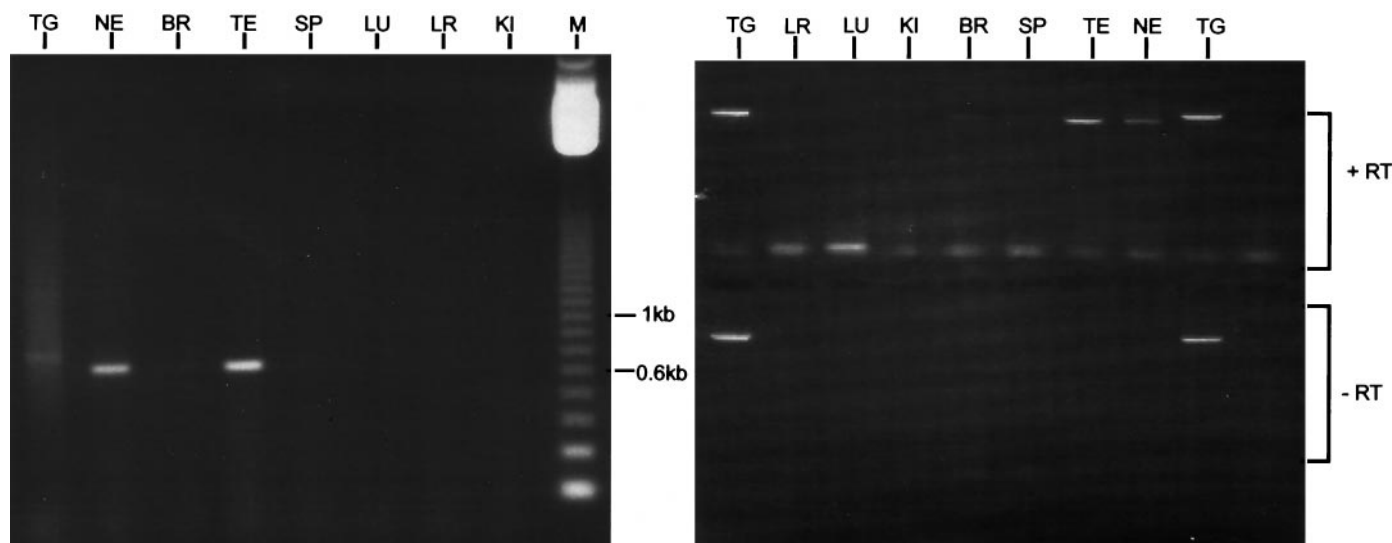


Figure 5. Reporter transgene expression of *E5* and *D5* in different tissues. RT-PCR experiment was performed on the cDNA from seven different tissues, using the primer pair RR318 and RR319. A 728 bp PCR product was generated with the genomic DNA from the transgenic mouse (*TG*), whereas the cDNA from various tissues generated a 635 bp product that is 93 bp smaller. The *right panel* represents the β -galactosidase reporter expression in the *D5* transgenic line. The *top panel* (+RT) shows PCR products with reverse transcriptase, and the *bottom panel* (–RT) is shown without reverse transcriptase. Expression in the *E5* transgenic line is shown on the *left panel* with a 123 bp marker lane represented by *M*. The tissues are *NE*, nose; *BR*, brain; *TE*, testis; *SP*, spleen; *LU*, lung; *LR*, liver; and *KI*, kidney.

the anterior–posterior extents of the olfactory epithelia also revealed that the regions that stained with the anti-*M4* antibody and β -galactosidase activity were always coincident. It was possible to perform double-labeling experiments by incubating the sections first with the X-gal chromogenic reagent to reveal the transgene-expressing cells and subsequently staining them with the *M4*-specific antibody. From 15 complete sections derived from several animals ($n = 3$), a total of 1031 *M4*-labeled and 500 β -galactosidase-positive cells were observed. A total of eight double-positive cells ($M4^+ \beta$ -gal $^+$) were observed on these sections (Fig. 6*g*). This frequency corresponds to what might be expected if expression of the two proteins were independent events, i.e., ~ 0.5 – 1% of the $M4^+$ cells should be β -galactosidase-positive and vice versa.

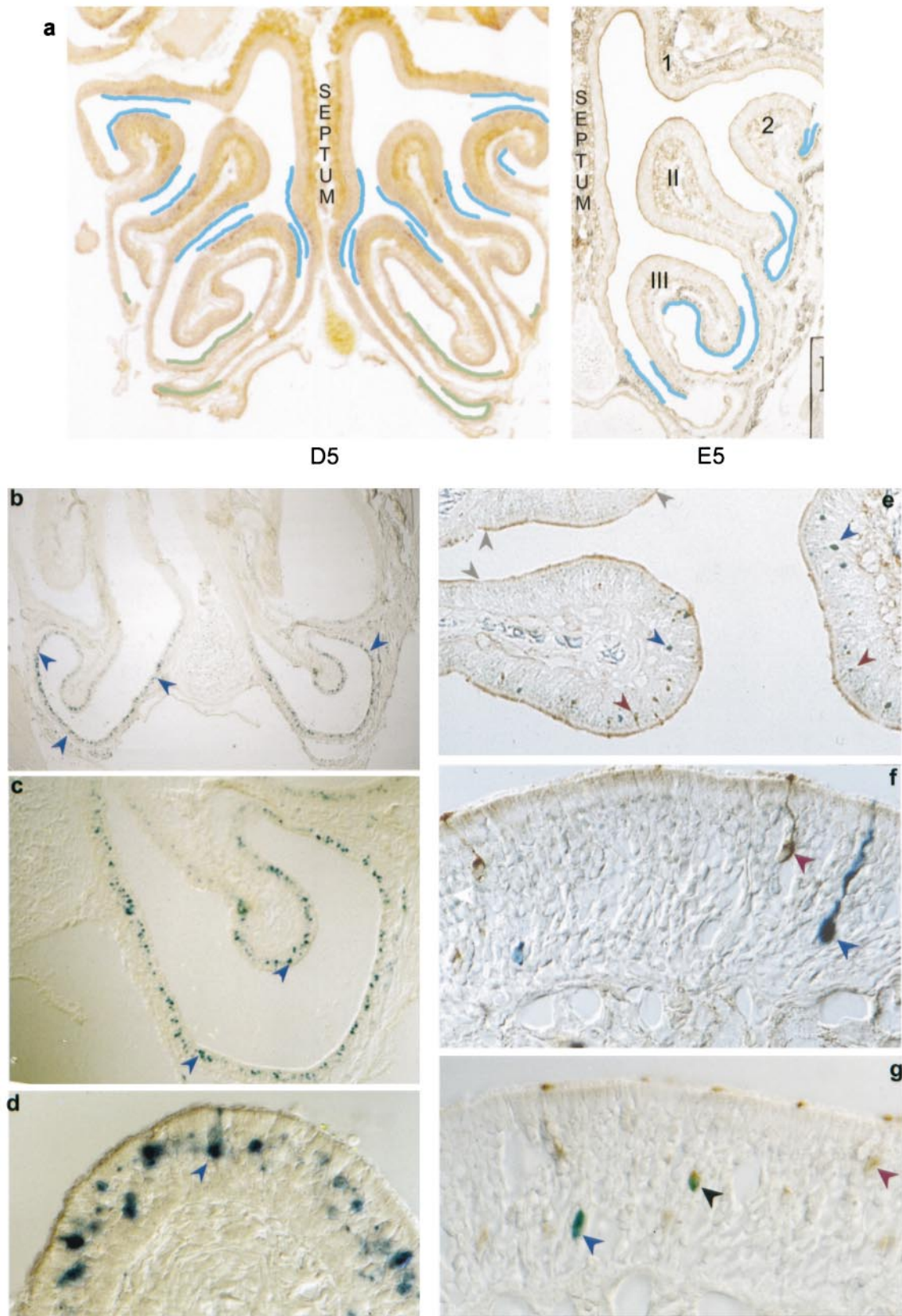
Analysis and chromosomal mapping the transgene

Although the frequency of obtaining lines that expressed the introduced transgene and the observation of expression was similar to or only slightly lower than that observed for other olfactory promoter-driven reporters, the observation of transgene expression in two different zones suggested that the site of integration might be important in achieving the expression pattern. With a large family of the odorant receptors distributed throughout the genome, it was possible that the transgene was in proximity of another receptor or receptors and had acquired the zonal expression pattern appropriate for the flanking receptor. Subsequent

breeding experiments with the transgene progeny and genomic Southern analysis confirmed a single integration site in the *E5* and *D5* lines. To explore the possibility that other receptors present in proximity to the transgene insertion site might influence the pattern of expression, we cloned and examined the sites of genomic insertion for *E5* and *D5*.

The *E5* transgene insertion site was cloned from a size-fractionated genomic library (9–15 kb) as described in Materials and Methods. The transgene–genomic junction fragment in the *D5* line was identified by an inverse PCR approach (Cuyper et al., 1984; Benkel et al., 1995) (data not shown). Genomic DNA from the parental strain as well as the two transgenic lines was subjected to hybridization with the isolated junction fragments. In each case a novel mobility fragment was observed in the appropriate lane that corresponded to the junction between the transgene sequences and the target fragment (Fig. 7). To explore sequences adjacent to the site of transgene insertion, we identified three and four overlapping YACs for *D5* and *E5*, respectively. Total genomic DNA from yeast cultures containing the YACs was hybridized with a 550 bp PCR product representing a complex mixture of odorant receptors (Levy et al., 1991). No receptor-related sequences were identified by using low-stringency hybridization or by PCR amplification of receptors with the use of degenerate oligonucleotides capable of amplifying most members of the odorant receptor family.

Figure 6. Patterns of expression of β -galactosidase in mouse nasal cavity of *D5* and *E5* transgenic mice. *a*, Coronal sections of the mouse nasal cavity are shown with the schematic localization of the *D5* (*left*) and the right nasal cavity of an *E5* transgenic mouse stained with X-gal at a similar sectioning level. The extents of the transgene expression (which coincided with *M4* antibody staining) are indicated in *blue*. The regions in a comparable section that were not expressing OMP by *in situ* hybridization are indicated in *green*. The septum (*SEPTUM*), the positions of ectoturbinates *I* and *2*, and the positions of endoturbinates *II* and *III* are indicated. The dorsal region is displayed at the *top* of the image. Overlays were produced by scanning images and processing them in Photoshop. Coronal sections through the *E5* mouse nasal cavity (*b–d*) were stained with X-gal. The reporter expression is localized to the mature neurons within the neuroepithelial layer. A low magnification of the *E5* coronal section (*a–c*) shows X-gal staining, marked with *blue arrowheads*. The staining is confined to the most lateral and ventral regions of the epithelium on both sides of the septum (zone IV). Higher magnification of the section (*d*) shows X-gal-positive neurons over the central portion of the epithelium (*blue arrowheads*). Double-labeled sections from a *D5* transgene animal were post-fixed in Zamboni and stained for activity with X-gal, followed by *M4* antibody (*e, g*). *Figure legend continues.*



A coronal section at low magnification shows double-stained neuroepithelium with X-gal staining in blue; peroxidase-stained *M4*-positive cells are represented by brown arrowheads (zone II). Regions of the epithelium showing no X-gal or *M4* staining are depicted by gray arrows. A high-magnification micrograph (f) shows *M4*-positive neurons colocalized with neurons expressing β -galactosidase, which is driven by the putative *M4* promoter region. Approximately 1% of the *M4* or β -galactosidase-positive cells showed double staining (black arrowheads) within the same cell with peroxidase and X-gal (g).

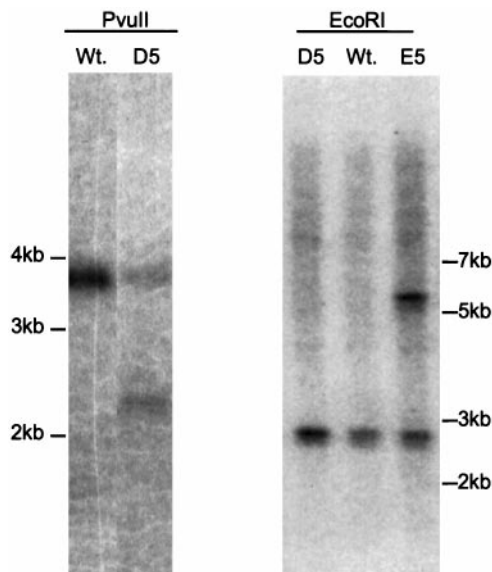


Figure 7. Disruption of the transgene integration sites in *D5* and *E5*. Five micrograms of *C57BL/6J* (*Wt*) and *D5* and *E5* mouse liver DNA were digested with *PvuII* (*left*) or *EcoRI* (*right*), electrophoresed in 0.75% agarose, blotted onto a nylon membrane, and hybridized with the ^{32}P -labeled probe. The primer pair PQ74/PQ79 generated a *D5*-specific junction probe used on the *left*, and PQ 37/PQ39 generated a probe for *E5* junction that was used on the *right*. DNA sizes are indicated on the *sides* of the panels. Both *D5* (*PvuII*-digested) and *E5* (*EcoRI*) lanes show an extra band, representing the transgene insertion-induced disruption.

The chromosomal location of the two integration target sites for *E5* and *D5* was determined by using a *C57Bl6xMus Spretus* backcross panel that previously had been analyzed for a large number of other chromosomal markers. The presence of specific PCR products from the insertion sites that were capable of distinguishing between the two parental chromosomes permitted mapping of the insertion site. The *E5* junction site is located near D6MIT8 at 35 cm on mouse chromosome six (LOD score 8.8). The *D5* insertion site mapped near D10Bir6 at 26 cm on chromosome 10 (LOD score 6.1). The *D5* insertion site is located 18 cm from olfactory receptor *Olf8* recently mapped to chromosome 10 (Sullivan, 1996). To date, no olfactory receptors have been mapped to chromosome six.

DISCUSSION

The restricted/differential expression of odorant receptors within the olfactory epithelium has been hypothesized to play an important role in the ability of this tissue to discriminate among olfactory stimuli. The identification of a receptor family that displays a complex expression pattern consisting of zonal restriction and an apparently stochastic distribution within that zone lends support to the hypothesized role of these receptors in odorant detection. The regulatory mechanisms that underlie these expression patterns remain mainly unknown. However, we have described here the extent of the DNA sequences that are required to recapitulate the zonally restricted and stochastic expression of a reporter within the sensory neurons of the olfactory epithelium.

M4 and its properties

The mouse *M4* receptor, identified initially in a screen of genomic DNA, shares the characteristic expression patterns of the members of the family originally described by Buck and Axel (1991).

The *M4* gene, existing as a subfamily with itself as the only member, is expressed preferentially in the olfactory epithelium. An antibody to the highly divergent region at the C terminus of the protein revealed staining in the dendrite, dendritic knob, and cilia of a small fraction of the sensory neurons within a single zone of the epithelium. The localization of the receptor protein to these structures parallels the observations made for other components of the odorant signaling pathway.

Analysis of the genomic organization of the *M4* gene confirmed the absence of introns within the coding region, consistent with all other known members of this family. The 5' end of the mRNA, identified by RACE-PCR, was encoded by a single exon separated by a 4.0 kb intron from the translation initiation site. All of the sequences present in mRNA could be identified in the cloned genomic fragment, suggesting that gene conversion events in the flanking untranslated sequence were not likely to contribute to the regulation of this receptor.

Transgenic lines tissue expression

The *M4* promoter/ β -galactosidase reporter construct led to detectable expression in three of the eight lines that were generated. This frequency was comparable to that observed for other

promoter/reporter constructs in transgenic mice (Walters et al., 1996). Notably, in each of the lines examined, no expression of the reporter was detected in tissues other than olfactory epithelium and testis. The expression of the β -galactosidase reporter within the olfactory epithelium in all three expressing lines was restricted to the mature olfactory neurons and paralleled in several respects the endogenous pattern of the *M4* receptor. The earliest expression within the neuronal lineage was in the region corresponding to immature neurons and was most intense in cells with characteristic mature neuronal morphology. In both *D5* and *E5* the reporter expression was restricted topographically to zones previously described for members of the odorant receptor family (Ressler et al., 1993), and, like the endogenous receptors within a zone, positive neurons were scattered randomly. The stereotypic distribution of staining in these transgenic lines was constant in all animals and was maintained through the germline. The appearance of the reporter in two different zones in the two lines has important implications for the mechanisms responsible for this patterning. Although we cannot rule out the possibility that zonal expression arises from deletion of negative regulatory elements in truncated copies of the inserted transgene, aspects of the regulation must arise from elements contained within the *M4* promoter region. We conclude from these studies that the 6.7 kb fragment from the *M4* gene is sufficient to direct most aspects of the regulated expression associated with odorant receptors.

The ability of the reporter construct to define a zonally restricted pattern of expression without the specification of the zone suggests a hierarchical mechanism in which distinct processes direct the observed pattern. Perhaps the zonal patterning is predetermined and specified for a large region of the genome by processes analogous to the genomic imprinting and X inactivation described in other systems. The transgene inherits the zonal pattern native to that locus. Endogenous receptors then can be seen as committed to expression in a particular zone by virtue of their location within the genome.

The selection and restriction of receptor expression within a single neuron are poorly understood. The pattern of expression of

endogenous M4 protein and the reporter in the *D5* line suggest that they occur as independent events, with each occurring in ~1% of the neurons within the zone. The anticipated frequency of double-labeled cells then would be 1% of the *M4*-positive cells, approximately the number that was observed in these experiments. One should not conclude from these experiments that multiple receptors are expressed in individual neurons. The absence of functional receptor protein expression from the reporter construct might prevent the feedback mechanisms that could serve to restrict expression to a single functional receptor in each cell.

The examination of the chromosomal distribution of odorant receptors and their zonal expression within the epithelium (Sullivan, 1996) demonstrated that receptors expressed in different zones could cosegregate in these linkage studies. *In situ* hybridization studies and detailed analysis of genomic organization of receptor subfamilies (Sullivan, 1996) indicate that members of a single subfamily are restricted to a single zone and are closely juxtaposed in the genome. Taken together, these data suggest that zonal patterns of expression arise from control elements that dictate some aspects of receptor expression (zones) over long genomic distances. A second aspect of receptor regulation, the expression within a small subset of the neurons within a zone, might be achieved via competition between discrete promoters for active enhancers. Models for regulation of clustered genes with a globin system have been described (Engel and Choi, 1988). The data presented here suggest that small segments of DNA upstream from odorant receptor coding regions are receptive to both kinds of regulation. The observation that transgene expression respected zonal boundaries but could be relocated to a different zone suggests that distant signals are responsible for defining the zone, functioning perhaps via elements located within the proximal promoter.

Models for receptor regulation

The selective expression of odorant receptor genes is likely to arise from a hierarchical series of regulated transcriptional controls. The data presented here in combination with the work of others (Sullivan et al., 1995) suggest that an early event in receptor expression is the establishment of zones within the epithelium. An attractive model is that the zones are specified by an activation process in which regions of the genome are made competent for expression in one of the zones. Alternatively, receptor-encoding regions of the genome could be repressed for expression (Chess et al., 1994) in other than the appropriate zone via a process much like inactivation on the X chromosome. Once the repertoire of receptors that are competent for expression in any particular zone is established, the selection of the individual receptor or small number of receptors to be expressed in a single olfactory neuron is achieved via subsequent competition for *trans*-acting factors or relief from negative regulation mediated by DNA methylation or inhibitory factors. These processes are consistent with the data from Chess et al. (1994), suggesting that receptor selection does not involve the activation of both alleles of a single receptor gene; rather, expression appears to occur stochastically on a small fraction of receptor genes.

In the experiments with a transgenic receptor promoter construct, we have defined sequences that are sufficient for receiving zonal specification and permitting the stochastic selection of transgene expression in the appropriate cell type. These results permit the dissection of essential sequences for each of these

processes and the *trans*-acting factors that result in the remarkable regulation of odorant receptor expression.

REFERENCES

- Amoore JE (1970) Molecular basis of odor. Springfield, IL: Thomas.
- Benkel BF, Perreault J, Gagnon C, Conklin K (1995) A rapid PCR-based test for the endogenous viral element ev3 of chickens. *Anim Genet* 26:189–191.
- Buck LB (1992) The olfactory multigene family. *Curr Opin Genet Dev* 2:467–473.
- Buck L, Axel R (1991) A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* 66:175–187.
- Chess A, Simon I, Cedar H, Axel R (1994) Allelic inactivation regulates olfactory receptor gene expression. *Cell* 78:823–824.
- Choi OR, Engel DJ (1988) Developmental regulation of β -globin gene switching. *Cell* 55:17–26.
- Cuypers HT, Selten G, Quint W, Zijlstra M, Berns A (1984) Murine leukemia virus-induced T-cell lymphomagenesis: integration of proviruses in a distinct chromosomal region. *Cell* 37:141–150.
- Danciger E, Mettling C, Vidal M, Morris R, Margolis F (1989) Olfactory marker protein gene: its structure and olfactory neuron-specific expression in transgenic mice. *Proc Natl Acad Sci USA* 86:8565–8569.
- Davis RL, Davidson N (1986) The memory gene *dunce*⁺ encodes a remarkable set of RNAs with internal heterogeneity. *Mol Cell Biol* 6:1464–1470.
- Emson PC, Shoham S, Feler C, Buss T, Price J, Wilson CJ (1990) The use of a retroviral vector to identify foetal striatal neurons transplanted into the adult striatum. *Exp Brain Res* 79:427–430.
- Frohman MA, Dush MK, Martin GR (1988) Rapid amplification of full-length cDNA from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* 85:8998–9002.
- Graziadei PPC, Graziadei GAM (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.
- Hicks JB, Strathern JN (1977) Interconversion of mating type in *S. cerevisiae* and the Cassette model for gene transfer. *Brookhaven Symp Biol* 29:233–242.
- Hogan BLM, Horsburgh G, Cohen J, Hetherington CM, Fisher G, Lyon MF (1986) Small eyes (Sey): a homozygous lethal mutation on chromosome 2 which affects the differentiation of both lens and nasal placodes in the mouse. *J Embryol* 97:95–110.
- Jones DT, Barbosa E, Reed RR (1988) Expression of G-protein α subunits in rat olfactory epithelium: candidates for olfactory signal transduction. *Cold Spring Harb Symp Quant Biol* 53:349–353.
- Levy NS, Bakalyar HA, Reed RR (1991) Signal transduction in olfactory neurons. *J Steroid Biochem Mol Biol* 39:633–637.
- Manly KF (1993) A Macintosh program for storage and analysis of experimental genetic mapping data. *Mamm Genome* 4:303–313.
- Manly KF, Elliott RW (1991) RI Manager, a microcomputer program for analysis of data from recombinant inbred strains. *Mamm Genome* 1:123–126.
- Monti-Graziadei AG, Graziadei PPC (1979) Studies on neuronal plasticity and regeneration in the olfactory system: morphologic and functional characteristics of the olfactory sensory neuron. In: *Neural growth and differentiation*, pp 373–396. New York: Raven.
- Parmentier M, Libert S, Schiffmann S, Lefort A, Eggericks D, Ledent C, Mollereau C, Gerard C, Perret J, Grootegoed A, Vassart G (1992) Expression of members of the putative olfactory receptor gene family in mammalian germ cells. *Nature* 355:453–455.
- Peschon JJ, Behringer RR, Braister RL, Palmiter RD (1987) Spermatid-specific expression of protamine 1 in transgenic mice. *Proc Natl Acad Sci USA* 84:5316–5319.
- Raming K, Krieger J, Strotmann J, Boekhoff I, Kubick S, Baumstark C, Breer H (1993) Cloning and expression of odorant receptors. *Nature* 361:353–356.
- Reed RR (1992a) The mechanisms of sensitivity and specificity in olfaction. *Cold Spring Harb Symp Quant Biol* 57:501–504.
- Reed RR (1992b) Signaling pathways in odorant detection. *Neuron* 8:205–209.
- Ressler KJ, Sullivan SL, Buck LB (1993) A zonal organization of odorant receptor gene expression in the olfactory epithelium. *Cell* 73:597–609.

- Stefanin M, De Martino C, Zamboni L (1967) Fixation of ejaculated spermatozoa for electron microscopy. *Nature* 216:173-174.
- Strotmann J, Wanner I, Krieger J, Raming K, Breer H (1992) Expression of odorant receptors in spatially restricted subsets of chemosensory neurones. *NeuroReport* 3:1053-1056.
- Strotmann J, Wanner I, Helfrich T, Beck A, Meinken C, Kubick S, Breer H (1994) Olfactory neurones expressing distinct odorant receptor subtypes are spatially segregated in the nasal neuroepithelium. *Cell Tissue Res* 276:429-438.
- Sullivan SL, Bohm S, Ressler KJ, Horowitz LF, Buck LB (1995) Target-independent pattern specification in the olfactory epithelium. *Neuron* 15:779-789.
- Sullivan SL, Adamson MC, Ressler KJ, Kozak CA, Buck LB (1996) The chromosomal distribution of mouse odorant receptor genes. *Proc Natl Acad Sci USA* 93:884-888.
- Vassar R, Ngai J, Axel R (1993) Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. *Cell* 74:309-318.
- Vassar R, Chao SK, Sitcheran R, Nunez JM, Vosshall LB, Axel R (1994) Topographic organization of sensory projections to the olfactory bulb. *Cell* 79 6:981-991.
- Walters E, Grillo M, Tarozzo G, Stein-Izsak C, Corbin J, Bocchiaro C, Margolis FL (1996) Proximal regions of the olfactory marker protein gene promoter direct olfactory neuron-specific expression in transgenic mice. *J Neurosci Res* 43:146-160.