Evolution of Odorant Receptors Expressed in Mammalian Testes

Aron Branscomb,* Jon Seger† and Raymond L. White*

**Department of Human Genetics and Huntsman Cancer Institute and †Department of Biology, University of Utah, Salt Lake City, Utah 84112*

Manuscript received January 10, 2000
Accepted for publication June 15, 2000

ABSTRACT

About 10% of mammalian odorant receptors are transcribed in testes, and odorant-receptor proteins have been detected on mature spermatozoa. Testis-expressed odorant receptors (TORs) are hypothesized to play roles in sperm chemotaxis, but they might also be ordinary nasal odorant receptors (NORs) that are expressed gratuitously in testes. Under the sperm-chemotaxis hypothesis, TORs should be subject to intense sexual selection and therefore should show higher rates of amino acid substitution than NORs, but under the gratuitous-expression hypothesis, TORs are misidentified NORs and therefore should evolve like other NORs. To test these predictions, we estimated synonymous and nonsynonymous divergences of orthologous NOR and TOR coding sequences from rat and mouse. Contrary to both hypotheses, TORs are on average more highly conserved than NORs, especially in certain domains of the OR protein. This pattern suggests that some TORs might perform internal nonolfactory functions in testes; for example, they might participate in the regulation of sperm development. However, the pattern is also consistent with a modified gratuitous-expression model in which NORs with specialized ligand specificities are both more highly conserved than typical NORs and more likely to be expressed in testes.

VERTEBRATE odorant receptors (ORs) were identified by Buck and Axel (1991). They form a large clade within the seven-transmembrane-domain (7TMD) G-protein-coupled receptor (GPCR) superfamily, which includes opsins and a great diversity of neurotransmitter and hormone receptors (Yokoyama and Starmer 1996). Testis-expressed ORs (TORs) were discovered by Parmentier *et al.* (1992) during a study of other GPCRs, and OR expression has been detected subsequently in various nonolfactory tissues of several mammalian species (Abe *et al.* 1993; Vanderhaeghen *et al.* 1993, 1997; Drutel *et al.* 1995; Walensky *et al.* 1995, 1998; Asai *et al.* 1996; Nef and Nef 1997; Dreyer 1998; Raming *et al.* 1998). Testis expression has been characterized by RNase-protection assays (RPA), *in situ* hybridizations, and protein immunohistochemistry (Vanderhaeghen *et al.* 1993, 1997; Walensky *et al.* 1995, 1998; Asai *et al.* 1996). Some TORs are known to be transcribed in the olfactory epithelium as well as in the testis (*e.g.*, Vanderhaeghen *et al.* 1997), but patterns of expression have been characterized directly for only a few OR genes, so most tissue assignments are based on cloning [by reverse transcriptase (RT)-PCR with degenerate primers] from a nasal or a testis cDNA library.

There appear to be many fewer TORs (~50 per species in rodents; Vanderhaeghen *et al.* 1997) than NORs (500–1000 per species; Buck and Axel 1991; Levy *et al.* 1991; Chess *et al.* 1994; Mombaerts 1999a), but TORs occur throughout the odorant-receptor gene family (Parmentier *et al.* 1992; Vanderhaeghen *et al.* 1997; see Figure 1) rather than being clustered in a few clades. Several TORs have been cloned more than once from the same species, and apparent orthologs have been cloned from different species (Vanderhaeghen *et al.* 1997); these independent rediscoveries of individual TORs support the inference (originally based on other lines of evidence) that the number of TORs is not very large, and they imply that patterns of testis expression may remain evolutionarily stable for at least a few tens of millions of years.

Why are odorant receptors expressed in the testis? The scattered phylogenetic distribution of TORs within the OR gene family can be taken to support either of two artifactual explanations: first, that TORs are ordinary NORs transcribed gratuitously in the testis but not performing any function there; and second, that most putative testis "cDNA" clones are amplified from contaminating genomic DNA (R. Axel, personal communication; but see Vanderhaeghen *et al.* 1997). In either case, "TORs" would be misidentified NORs and therefore would be expected to evolve like NORs.

Alternatively, TOR proteins might mediate sperm chemotaxis, as suggested by Parmentier *et al.* (1992), Vanderhaeghen *et al.* (1993, 1997), Walensky *et al.* (1995, 1998), and others. In this case, TORs would determine phenotypes likely to become involved in male-male competition and female choice. For example, males could gain mating advantages by adding "decoy"

Corresponding author: J. Seger, Department of Biology, University of Utah, 257 S. 1400 East, Salt Lake City, UT 84112-0840.
E-mail: seger@bionix.biology.utah.edu

compounds to their seminal fluids, to which other males' TORs (but not their own) were vulnerable. Signals used in mate choice are expected to be evolutionarily unstable owing to the "antagonism" inherent in such interactions; and as expected, sexual signals often evolve rapidly (Andersson 1994; Eberhard 1996; Rice 1996, 1998; Tsaur and Wu 1997; Arnqvist 1998; Metz *et al.* 1998; Partridge and Hurst 1998; Vacquier 1998; Clark *et al.* 1999; Holland and Rice 1999). Thus, on the sperm-chemotaxis hypothesis, TORs might be expected to show high rates of amino acid substitution.

To test these predictions, we compared the evolution of orthologous NOR and TOR genes in rat and mouse. We found, to our surprise, that the amino acid sequences of 10 TORs are more highly conserved, on average, than those of 8 NORs. The greater conservation of TORs is concentrated in the extracellular end of the fourth transmembrane domain (TM4, thought to be involved in ligand binding) and the third intracellular loop (IC3, which interacts with G-proteins). This distinctive pattern of amino acid substitution contradicts straightforward predictions of the gratuitous-expression, nonexpression, and sperm-chemotaxis hypotheses. We consider two alternative models that might explain the relatively stringent conservation of at least some TORs. In the first model, some TORs are recruited to novel internal (developmental or physiological) functions that differ from those performed by canonical nasal odorant receptors. In the second model, some NORs evolve highly focused specificities for odorants of special ecological importance; this tight focus on single ligands causes these specialized NORs to evolve relatively slowly at the amino acid sequence level and to acquire increased levels of expression in the olfactory epithelium; as a side effect, they are at greater than average risk of being transcribed gratuitiously in nonolfactory tissues, especially the testis. We discuss the kinds of evidence needed to test these two models.

A phylogenetic analysis of 160 paralogous OR genes confirms that TOR and NOR lineages interdigitate extensively and suggests that recruitment between nasal and testicular expression patterns may occur in both directions. However, small differences in the assumed ease of recruitment in each direction lead to large differences in the estimated numbers of $N \rightarrow T$ and $T \rightarrow N$ recruitments. In the future, when patterns of nasal and testicular expression have been documented for many orthologs and closely related paralogs in species representing a range of divergence times, it should become possible to resolve histories of expression and of amino acid sequence change with enough precision to say whether sequence conservation precedes or follows expression in the testis. In either case, conserved ORs could be important to the evolution of the odorant-receptor gene family as a whole if they periodically give rise (by duplication or gene conversion) to new NOR lineages of greater average longevity than those derived from typical, less well-conserved NORs.

MATERIALS AND METHODS

Sequence names: We add lowercase species prefixes to names that do not otherwise indicate their source (m, mouse; r, rat; h, human; d, dog; p, pig), and we shorten some names. Thus F6 becomes rF6 and MTPCR09 becomes mT09, but CfOLF1 (from dog, *Canis familiaris*) remains CfOLF1. A convenient feature of this system is that corresponding postfixes can be used to identify orthologs (*e.g.*, mT09r is the rat ortholog of mT09m).

PCR and sequencing: Published NOR or TOR cDNA sequences from rat or mouse were used to design primers that specifically amplify both the original sequence (in the source species) and a presumptive ortholog (in the other species) from genomic DNA (Sprague-Dawley rat or BALB/c mouse; Clontech, Palo Alto, CA). Several partial TOR sequences were extended to nearly full length by inverse PCR. Primer sequences and reaction conditions can be obtained from the first author (A.B.). PCR products were sequenced directly on ABI (Columbia, MD) 373 and 377 automated fluorescent sequencing machines.

The original sequences are rF6 (M64378), rF12 (M64381), rI8 (M64387), rI9 (M64388), mK7 (L14566), mK20 (U28770), mM31 (U28777), mT09 (X89681), mT15 (X89683), mT33 (X89685), rT07 (X89697), rT09 (X89698), and rT18 (X89702). We redetermined all except rI9 as positive controls, and these differ by no more than three nucleotide substitutions from the originals. Several genes from each tissue are represented by nearly full-length coding sequences (TM1 through TM7), and several are represented by shorter sequences. Five genes are represented by published sequences: the apparently orthologous relationships of mOR3/rT44 (M84005, X89706), rT05/mT07 (X89695, X89680), rT19/mT18 (X89703, X89684), and rT38/mT53 (X89705, X89691) were noted independently by us and by Vanderhaeghen *et al.* (1997), and rI7m (AF106007, the mouse ortholog of rI7r, M64386) was obtained intentionally by Krautwurst *et al.* (1998). Also, mM64 (U28781) appears to be an allele or a very closely related paralog of rF12m. The sequences newly described here have been submitted to GenBank under accession nos. AF271033–AF271057. We classify genes known to be expressed both in the testis and the nose as TORs.

Sequence divergence: Synonymous (K_S) and nonsynonymous (K_A) substitutions were estimated by the method of Li (1993) and Pamilo and Bianchi (1993). Protein domain boundaries were defined by alignment with the structural model of Pilpel and Lancet (1999), and domain-specific nonsynonymous divergence estimates and K_A/K_S ratios were analyzed by nested ANOVA as implemented in JMP 3.1 (SAS Institute, Cary, NC). Tissue (nose or testis) and domain (major extracellular, intracellular, and transmembrane segments of the protein) were treated as fixed effects, with sequence identity as a random effect nested within tissues.

Phylogeny: We aligned DNA sequences for mammalian odorant receptors derived from tissue-specific cDNA libraries, or for which a site of expression was determined directly. We reduced sets of alleles, closely related paralogs, and obvious orthologs to a single representative each. Many of the sequences are fragments representing one-third to one-half of a complete coding region; our alignment comprises a 112-codon region of shared homology that begins immediately after the MAYDRYVAIC motif at the boundary between TM3 and IC2, which is frequently used as a binding site for degenerate primers. One hundred and sixty sequences (94 NORs and 66 TORs) are included. Accession numbers and the alignment can be obtained from the second author (J.S.). Phylogenetic relationships were estimated by DNA maximum-likelihood analysis as implemented in DNAML 3.6 (Felsenstein 1981, 1989); we used the fast-search and global-rearrangement op-

tions, with four categories of sites evolving at relative rates of 1, 8, 16, and 24; sites were assigned to categories on the basis of a parsimony reconstruction of nucleotide substitutions on a preliminary three-category maximum-likelihood (ML) tree. The categories' template and other details can be obtained from J.S. The tree shown in Figure 1 is the best of five found with different random input orders of the sequences. Histories of recruitment between nose and testis were estimated by MacClade 3.07 (Maddison and Maddison 1992). Histories of amino-acid substitution for clades of interest were estimated by MacClade, by PROTPARS 3.57 (Felsenstein 1989), and from the ancestral DNA sequences estimated by DNAML.

RESULTS

Divergence between orthologs: Synonymous substitution rates vary widely among hundreds of genes that have been sequenced in rat and mouse (Wolfe and Sharp 1993; Makalowski and Boguski 1998); this variation is thought to be caused mainly by regional variation in mutation rates (Wolfe *et al.* 1989; Li 1997; McVean and Hurst 1997). The distribution of K_S for the 18 ortholog pairs (Table 1) is fully consistent with the overall distribution for rat and mouse orthologs; the mean and variance fall well within the ranges expected for a sample of this size, and the combined synonymous divergence for NORs ($K_S = 0.20$) is very close to that for TORs ($K_S = 0.19$). However, the nonsynonymous divergence for NORs ($K_A = 0.040$) is twice as large as that for TORs ($K_A = 0.020$). This difference falls just short of formal significance by *t*-tests on the 18 gene-specific K_A values, raw amino acid differences, and K_A/K_S ratios, owing to the large variance within each tissue and consequently broad overlap between

TABLE 1

Divergences of orthologous odorant-receptor genes in rat and mouse

Orthologs			Raw differences					Estimated substitutions				
Rat	Mouse	cdn	d1	d2	d3	aa	aa/cd	K_S	(SD)	K_A	(SD)	K_A/K_S
						NORs						
mK20r	mK20m	86	9	1	9	9	0.105	0.12	(0.042)	0.048	(0.0164)	0.39
rF12r	rF12m	265	17	16	46	32	0.121	*0.23	(0.039)	0.066	(0.0112)	0.28
rI8r	rI8m	267	25	17	59	40	0.150	*0.32	(0.046)	0.078	(0.0120)	0.25
rI9r	rI9m	262	12	8	33	19	0.073	0.18	(0.033)	0.035	(0.0080)	0.20
mK7r	mK7m	133	4	2	17	6	0.045	0.17	(0.045)	0.024	(0.0092)	0.14
mM31r	mM31m	89	5	1	19	7	0.079	*0.32	(0.082)	0.037	(0.0142)	0.12
rF6r	rF6m	263	4	6	34	8	0.030	0.17	(0.030)	0.017	(0.0058)	0.10
rI7r	rI7m	270	6	1	24	5	0.019	0.13	(0.026)	0.009	(0.0041)	0.07
All sequences		1635	82	52	241	126	0.077	0.20	(0.014)	0.040	(0.0034)	0.20
$K_S \leq 0.2$		1014	35	18	117	47	0.046	0.16	(0.015)	0.023	(0.0033)	0.15
						TORs						
rT19r	mT18m	112	12	5	16	19	0.170	*0.24	(0.074)	0.075	(0.0177)	0.32
rT07r	rT07m	126	4	2	8	5	0.040	0.08	(0.028)	0.019	(0.0086)	0.24
mT09r	mT09m	268	19	10	50	24	0.092	*0.27	(0.041)	0.050	(0.0096)	0.18
rT09r	rT09m	261	9	4	37	9	0.034	0.19	(0.033)	0.020	(0.0058)	0.11
rT44r	mOR3m	157	3	3	19	5	0.032	0.17	(0.042)	0.018	(0.0076)	0.10
mT15r	mT15m	133	1		17	2	0.015	0.13	(0.036)	0.010	(0.0060)	0.08
rT05r	mT07m	157	2	1	20	3	0.019	0.16	(0.038)	0.009	(0.0054)	0.06
mT33r	mT33m	264	6	2	38	4	0.015	0.20	(0.034)	0.009	(0.0041)	0.04
rT38r	mT53m	157	3		21	2	0.013	0.20	(0.048)	0.008	(0.0045)	0.04
rT18r	rT18m	263	3		31		0.000	0.18	(0.034)	0.002	(0.0018)	0.01
All sequences		1898	62	27	257	73	0.038	0.19	(0.012)	0.020	(0.0023)	0.11
$K_S \leq 0.2$		1518	31	12	191	30	0.020	0.17	(0.013)	0.011	(0.0019)	0.07

The length of each aligned sequence pair is given in codons (cdn), with differences at first, second, and third positions (d1, d2, d3), and the number (aa) and proportion (aa/cd) of amino acid differences. Estimated synonymous substitutions per synonymous site (K_S) and nonsynonymous substitutions per nonsynonymous site (K_A) are given with their standard deviations. K_S is highly variable ($s_{Ks} = 0.064$), and K_S and K_A are correlated over these 18 gene pairs ($r = 0.57$), as they are for other pairs of putative orthologs in rat and mouse (Wolfe and Sharp 1993; Makalowski and Boguski 1998). *, the five K_S values >0.2 [slightly less than one standard deviation above the mean for hundreds of rat-mouse orthologs (Makalowski and Boguski 1998)]; most analyses are performed both with and without these homolog pairs, which are presumably the ones at greatest risk of being paralogous. The ratio K_A/K_S estimates the probability of fixation for a nonsynonymous (amino acid changing) mutation, relative to the probability for a synonymous mutation, and therefore indicates the apparent efficiency with which selection resists amino acid substitutions (smaller values of K_A/K_S indicating more stringent selection). The 10 TOR sequences combined show a K_A/K_S ratio 55% as large as for the 8 NORs (47% as large for genes with $K_S \leq 0.2$). The overall K_A for TORs is 50% as large as for NORs (48% as large for genes with $K_S \leq 0.2$), and TORs show 49% as many amino acid differences per codon (43% as many for genes with $K_S \leq 0.2$).

the two distributions. (It is not legitimate to combine the genes within each tissue/species combination and then test the overall difference between the resulting aggregated rat-mouse NOR and TOR divergences against its formal standard error, because of the rate heterogeneity among genes.) However, several other features of the pattern indicate that many TORs belong to a population of genes different from that represented by most odorant receptors.

First, some TORs are highly conserved. Three of the 10 in Table 1 have K_A/K_S ratios <0.05, and 1 has identical amino acid sequences in rat and mouse, with a K_A/K_S <0.01 despite a typical level of synonymous divergence. The most highly conserved NOR shows a K_A/K_S of 0.07 and 5 amino acid differences; a typical full-length NOR ($K_A/K_S \approx 0.20$) would show ~21 amino acid differences. For comparison, a sample of 14 other (non-OR) GPCR rat-mouse ortholog pairs from the GPCR database shows a mean (combined) K_A/K_S of 0.13 (data not shown), and a sample of 470 rat-mouse ortholog pairs of all kinds shows a mean K_A/K_S of 0.19, with 23% of the individual values <0.05 and 6% <0.01 (Makalowski and Boguski 1998).

Second, the average relative difference between NORs and TORs increases when the samples are culled to remove ortholog pairs with synonymous divergences (K_S) >0.2 (slightly above the mean K_S for all rat-mouse comparisons). The resulting samples, restricted to include only pairs with $K_S \leq 0.2$, should be relatively unlikely to include putative ortholog pairs that are, in fact, closely related paralogs. If the overall nonsynonymous difference between NORs and TORs had been caused by a greater number of misidentified paralogs in the NOR sample, then the exclusion of pairs with $K_S > 0.2$ should have reduced the difference between the samples rather than increased it.

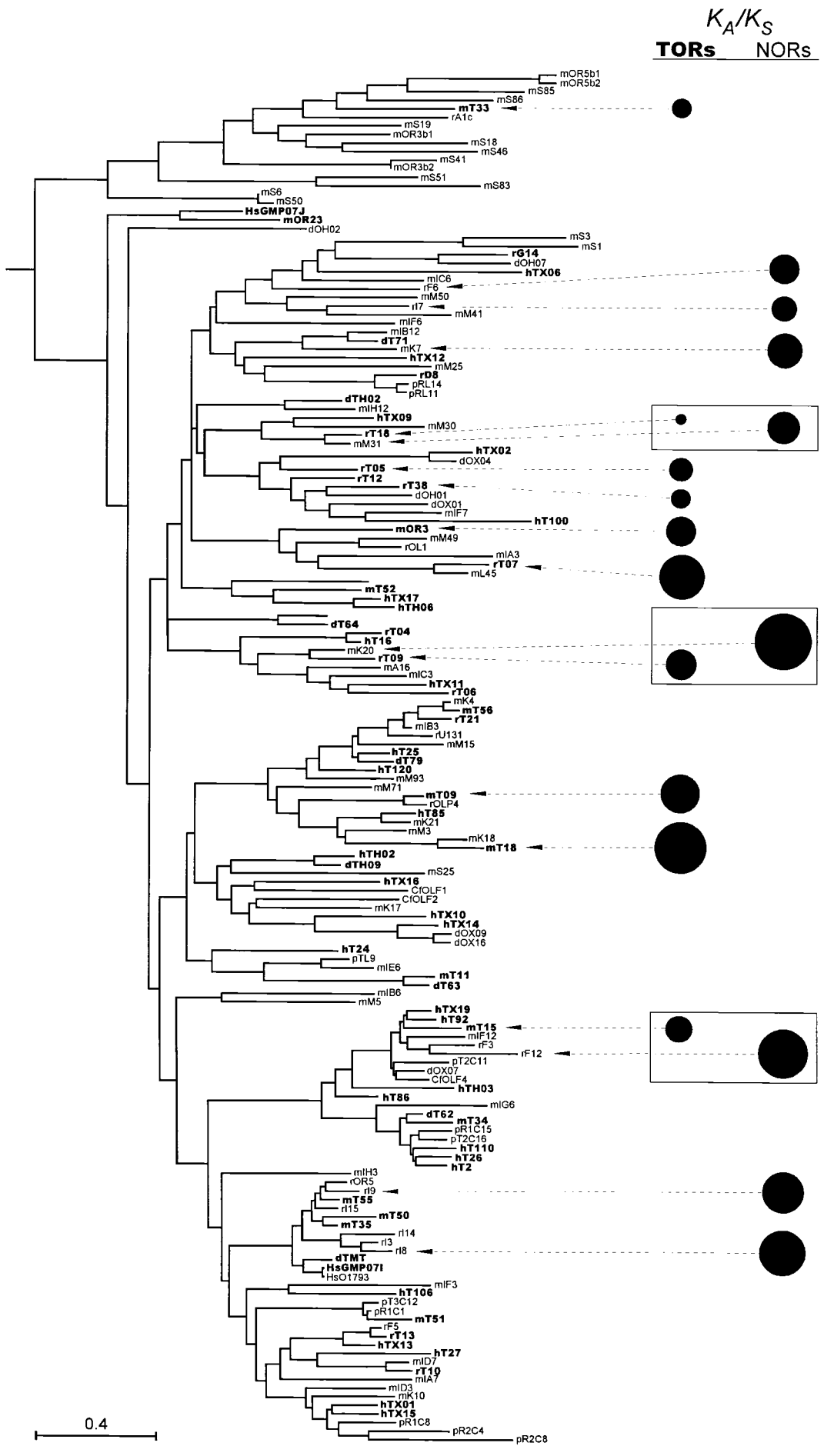
Third, each of three TOR ortholog pairs is significantly more strongly conserved than a closely related NOR pair (Figures 1 and 2). These comparisons are of particular interest because, in principle, some odorant-receptor subfamilies might be inherently less tolerant of amino acid substitutions than others; if our 10 orthologous TOR pairs happened to be sampled largely from such subfamilies, then their slower evolution might have nothing to do with expression in the testis (L. B. Buck, personal communication). On this hypothesis, there should be a large phylogenetic component to the variation in K_A/K_S ratios. In particular, closely related NOR/TOR pairs should show similar ratios. This specific prediction is not satisfied by the three such cases in our data set (Figure 2), and the more general prediction is not supported by the overall distribution of K_A/K_S ratios on the phylogeny (Figure 1). In principle, the high average K_A/K_S ratios of NORs might be an artifact if some are pseudogenes, but this seems unlikely because all show ratios significantly <1.0, and all show several strictly conserved amino acids shared by other odorant receptors (data not shown).

Fourth, amino acid substitutions are distributed heterogeneously among functional domains of the OR protein, both in NORs and TORs (Table 2), but the pattern differs significantly between the two tissues (Table 3). For example, the fourth transmembrane domain (TM4) includes amino acid positions that vary extensively among paralogous members of the OR family; this variation is concentrated in the extracellular end of the domain and along the face of the α-helix that is inferred to orient inward toward TM5 and to participate in ligand binding (Pilpel and Lancet 1999). We find a similar pattern in the substitutions that have accumulated between rat and mouse orthologs (Table 2). Overall, TM4 shows a higher rate of amino acid substitution than any other domain, but the rate in TORs is far lower than the rate in NORs; in the restricted sample ($K_S \leq 0.2$), only 6 amino acid substitutions occur in the TM4 domains of eight TORs, but 12 occur in five NORs; all 6 of the TOR substitutions are located near the cytoplasmic end of TM4, while the 12 NOR substitutions are distributed roughly uniformly along the length of TM4 (data not shown). TM1 and especially IC3 also appear to be more strongly conserved in TORs than in NORs (Table 2).

In summary, some of the TORs in our sample appear to have experienced stronger or more constant purifying selection than typical NORs, especially in certain functional domains of the protein. This distinctive history of selection appears to be inconsistent with the sperm-chemotaxis hypothesis and with at least the simplest version of the gratuitous-expression hypothesis.

Phylogenetic distribution of testis expression: NOR and TOR lineages interdigitate extensively in the odorant receptor phylogeny (Figures 1–3), suggesting that recruitments between tissues have occurred on many occasions and in both directions. This pattern is somewhat surprising. If TORs were functionally distinct from NORs then they might be expected to be evolutionarily distinct as well; they might be expected to derive from one or a few recruitments and thus to form one or a

Figure 1.—Phylogenetic relationships of mammalian odorant-receptor genes as estimated by maximum-likelihood analysis. TORs are indicated in boldface type and NORs in regular type. K_A/K_S ratios from Table 1 are represented by the areas of solid circles connected to individual sequences by dashed arrows. Three closely related NOR-TOR pairs (Figure 2) are boxed. The tree is rooted arbitrarily at its midpoint. Almost all internal branches are of lengths significantly greater than zero, and bootstrap analyses of neighbor-joining trees derived from protein distances suggest that many of the relatively distinct clades that appear here are probably real (data not shown). However, many features of this tree are undoubtedly incorrect, so it should be viewed only as a rough guide to the probable history of the odorant-receptor gene family.


K_A/K_S
TORs
NORs
mOR5b1
mOR5b2
mS85
mS86
mT33
rA1c
mS19
mOR3b1
mS18
mS46
mS41
mOR3b2
mS51
mS83
mS6
mS50
HsGMP07J
mOR23
dOH02
mS3
mS1
rG14
dOH07
hTX06
mIC6
rF6
mM50
rI7
mM41
mIF6
mIB12
dT71
mK7
hTX12
mM25
rD8
pRL14
pRL11
dTH02
mIH12
hTX09
mM30
rT16
mM31
hTX02
dOX04
rT05
rT12
rT38
dOH01
dOX01
mIF7
hT100
mOR3
mM49
rOL1
mIA3
rT07
mL45
mT52
hTX17
hTH06
dT64
rT04
hT16
mK20
rT09
mA16
mIC3
hTX11
rT06
mK4
mT56
rT21
mIB3
rU131
mM15
hT25
dT79
hT120
mM93
mM71
mT09
rOLP4
hT85
mK21
mM3
mK18
mT18
hTH02
dTH09
mS25
hTX16
CfOLF1
CfOLF2
mK17
hTX10
hTX14
dOX09
dOX16
hT24
pTL9
mIE6
mT11
dT63
mIB6
mM5
hTX19
hT92
mT15
mIF12
rF3
rF12
pT2C11
dOX07
CfOLF4
hTH03
hT86
mIG6
dT62
mT34
pR1C15
pT2C16
hT110
hT26
hT2
mIH3
rOR5
rI9
mT55
rI15
mT50
mT35
rI14
rI3
rI8
dTMT
HsGMP07I
HsO1793
mIF3
hT106
pT3C12
pR1C1
mT51
rF5
rT13
hTX13
hT27
mID7
rT19
mIA7
mID3
mK10
hTX01
hTX15
pR1C8
pR2C4
pR2C8
0.4

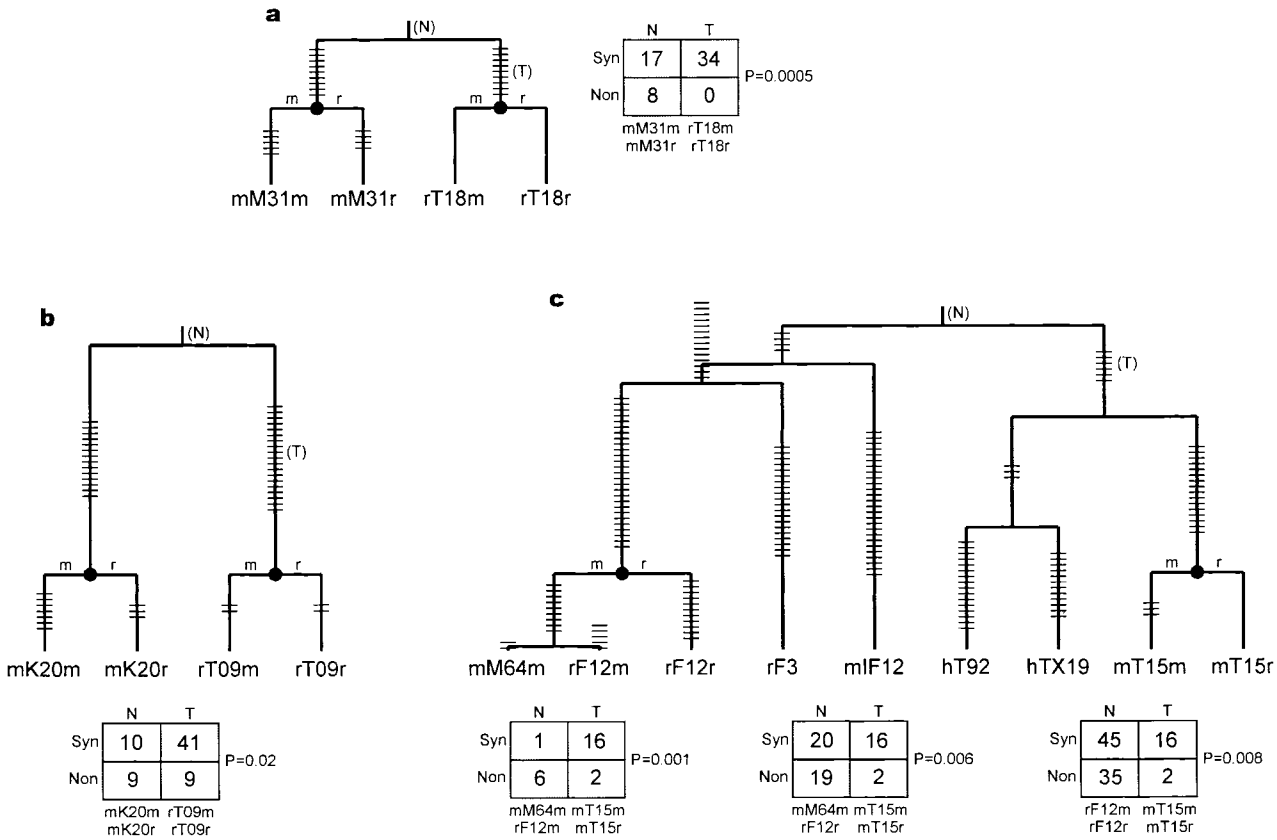


FIGURE 2.—Amino acid substitutions in the recent histories of closely related NOR and TOR ortholog pairs. Branch lengths are roughly proportional to K_S values, except that the splits between orthologs (solid circles) are forced to occur at the same depth (representing the time of the last common ancestor of rats and mice), and all tips are forced to occur at the same distance from the root. At each speciation, the rat branch (r) goes right and the mouse branch (m) goes left. In all three cases illustrated here, the ancestral sequence is inferred to be an NOR (N); thus testicular expression (T) arises on the right-hand branch following the gene duplication represented by the first split. Ticks represent amino acid substitutions (inferred by PROTPARS) in the 112-codon region used to construct Figure 1; in each case, many related sequences not shown here were included in the analysis. The 2 × 2 contingency tables show synonymous (Syn) and nonsynonymous (Non) nucleotide substitutions for the nasal (N) and testicular (T) orthologs in each pairwise comparison. Synonymous and nonsynonymous substitutions were determined by inspection of the aligned orthologous sequences and were unambiguous in every case; the tabulations use all nucleotide positions available for each orthologous pair. Two-tailed significance levels (P) were estimated by Fisher's exact test (SOKAL and ROHLF 1995); similar values (all formally significant) were obtained by other procedures (*e.g.*, G-tests) and for sequences restricted to the 112-codon region. (a) The NOR ortholog pair in this comparison (mM31) shows above-average synonymous divergence (K_S = 0.32) and modestly strong conservation (seven amino acid differences in 89 codons; K_A/K_S = 0.12). The TOR pair (rT18) shows typical synonymous divergence (K_S = 0.18) and extremely strong amino acid conservation (no amino acid differences in 263 codons; K_A/K_S = 0.01). The branches leading to mM31m and mM31r carry a total of nine ticks even though the sequences differ at only seven amino acid positions, because at two positions, different amino acid substitutions are inferred to have occurred on each branch. The number of nonsynonymous substitutions (8) is also greater than the number of amino acid differences because nonsynonymous differences occur at both the first and second nucleotide positions in one codon. (b) This NOR pair (mK20) shows very low synonymous divergence (K_S = 0.12), a typical level of amino acid divergence (nine differences in 86 codons; K_A = 0.048), and thus a high K_A/K_S ratio (0.39). The TOR pair (rT09) is typical for TORs (K_S = 0.19, K_A = 0.02, K_A/K_S = 0.10). (c) This NOR (rF12) and TOR (mT15) are not each others' closest relatives among the sequences in Figure 1, so additional sequences (not orthologous to either one) are included in this tree. rF12 is represented here by two sequences from mouse (the full-length rF12m from this study and the shorter mM64m from SULLIVAN *et al.* 1996). When paired with the rat ortholog rF12r, both show significant excess nonsynonymous substitution relative to the TOR pair (mT15). Remarkably, when paired with each other these two alleles (or very closely related paralogs) also show a significant excess of nonsynonymous (6) to synonymous (1) substitution relative to the ratio in mT15 (2–16).

few clades adjacent to or nested within the larger family of NORs. However, if TORs are simply misidentified NORs, then they should be scattered randomly through the family.

In fact, the observed distribution appears to be nearly random. When the history of expression is estimated by parsimony analysis (MADDISON and MADDISON 1992), NOR lineages derive from TORs at least 2 times (and potentially as many as 18 times), while TOR lineages derive from NORs at least 32 times (and potentially as many as 48). Similar numbers are obtained from trees on which nasal and testis tissue assignments have been scrambled randomly. A relative excess of NOR → TOR derivations is expected under the random-assignment

TABLE 2

Amino acid, nonsynonymous, and synonymous substitutions in orthologous rodent odorant receptors, by segment

	NORs ($N = 8$)						TORs ($N = 10$)					
Segment	N_{AA}	DIF	P_{DIF}	K_A	K_S	K_A/K_S	N_{AA}	DIF	P_{DIF}	K_A	K_S	K_A/K_S
					All sequences							
TM1[a]	126	14	0.11	0.061	0.18	0.34	104	6	0.06	0.027	0.27	0.10
IC1	24	1	0.04	0.016	0.31	0.05	20		0.00	0.000	0.27	0.00
TM2	115	2	0.02	0.009	0.18	0.05	92	3	0.03	0.016	0.19	0.09
EC1	94	10	0.11	0.064	0.34	0.19	72	5	0.07	0.033	0.07	0.45
TM3	119	9	0.08	0.040	0.22	0.19	96	1	0.01	0.008	0.26	0.03
IC2	76	4	0.05	0.022	0.16	0.14	104		0.00	0.004	0.24	0.02
TM4[a]	208	36	0.17	0.092	0.18	0.51	267	18	0.07	0.036	0.19	0.20
EC2	263	12	0.05	0.021	0.21	0.10	330	14	0.04	0.022	0.11	0.21
TM5	175	11	0.06	0.030	0.18	0.17	220	12	0.06	0.027	0.22	0.12
IC3[a]	170	19	0.11	0.062	0.19	0.32	254	9	0.04	0.019	0.22	0.09
TM6	120	1	0.01	0.004	0.11	0.04	180	2	0.01	0.008	0.11	0.07
EC3	48	3	0.06	0.036	0.22	0.16	63	2	0.03	0.025	0.17	0.14
TM7	98	4	0.04	0.020	0.31	0.06	110	1	0.01	0.008	0.34	0.03
All	1636	126	0.08	0.040	0.20	0.20	1912	73	0.04	0.020	0.19	0.11
	NORs ($N = 5$)						TORs ($N = 8$)					
Segment	N_{AA}	DIF	P_{DIF}	K_A	K_S	K_A/K_S	N_{AA}	DIF	P_{DIF}	K_A	K_S	K_A/K_S
					Restricted sample ($K_S \leq 0.2$)							
TM1[a]	74	7	0.09	0.045	0.14	0.32	78		0.00	0.000	0.28	0.00
IC1	14		0.00	0.000	0.25	0.00	15		0.00	0.000	0.20	0.00
TM2	69		0.00	0.000	0.12	0.00	69	1	0.01	0.008	0.26	0.03
EC1	58	3	0.05	0.025	0.24	0.10	54	2	0.04	0.017	0.06	0.29
TM3	71	1	0.01	0.006	0.19	0.03	72	1	0.01	0.011	0.25	0.04
IC2	45		0.00	0.000	0.13	0.00	80		0.00	0.006	0.24	0.03
TM4[a]	127	12	0.09	0.049	0.13	0.39	213	6	0.03	0.014	0.14	0.10
EC2	164	4	0.02	0.010	0.15	0.07	264	6	0.02	0.014	0.09	0.15
TM5	109	4	0.04	0.018	0.19	0.10	176	6	0.03	0.018	0.19	0.10
IC3[a]	113	11	0.10	0.054	0.16	0.34	207	3	0.01	0.009	0.22	0.04
TM6	80		0.00	0.000	0.08	0.00	160	2	0.01	0.006	0.10	0.06
EC3	32	2	0.06	0.037	0.20	0.19	55	2	0.04	0.028	0.17	0.16
TM7	58	3	0.05	0.028	0.23	0.12	90	1	0.01	0.005	0.28	0.02
All	1014	47	0.05	0.023	0.16	0.15	1533	30	0.02	0.011	0.17	0.07

Rows summarize amino acid and nucleotide substitutions within individual structural domains of the odorant-receptor protein, for all of the ortholog pairs sampled from a given tissue. Transmembrane (TM*x*), intracellular (IC*x*), and extracellular (EC*x*) domains are listed as they occur topologically in the primary structure of the protein; the poorly conserved amino- and carboxyl-terminal tails are omitted. For each domain, the columns give total numbers of codons compared (N_{AA}), raw numbers of amino acid differences (DIF), proportion of amino acid differences (P_{DIF}), and estimated rates of nonsynonymous and synonymous substitution (K_A, K_S, and their ratio K_A/K_S). Data for NORs and TORs are summarized separately in the left-hand and right-hand sides, respectively. The top summarizes data for all 18 genes (8 NORs and 10 TORs). The bottom excludes 3 NORs and 2 TORs for which $K_S > 0.2$; these genes are presumably the ones at greatest risk of being represented by mouse and rat sequences that are paralogous rather than orthologous (see Table 1).

[a] Relative rates of amino acid substitution are strikingly higher for NORs than for TORs in these domains, but not in any other domains with large numbers of sampled codons. The first five and last three domains are represented by fewer genes (and hence relatively fewer codons) than the block from IC2 through IC3, because some genes in the data set are represented by partial sequences (Table 1).

null model because there are many more NORs than TORs on the tree; nasal expression therefore tends to reconstruct as ancestral under the default assumption that recruitments in either direction are equally likely *a priori*.

However, there is no reason to expect recruitments in both directions to occur with equal ease. If the penalty for NOR → TOR changes is increased from 1 evolutionary step to 1.2 steps, then the most parsimonious expression history includes just 18 NOR → TOR recruitments and the number of TOR → NOR recruitments increases to 34, despite the numerical preponderance of NORs. An equivalent (20%) penalty bias in the other direction gives 48 unambiguous NOR → TOR recruitments and

TABLE 3

ANOVA significance tests of effects on domain-specific $K_A/\bar{K}_S$ values

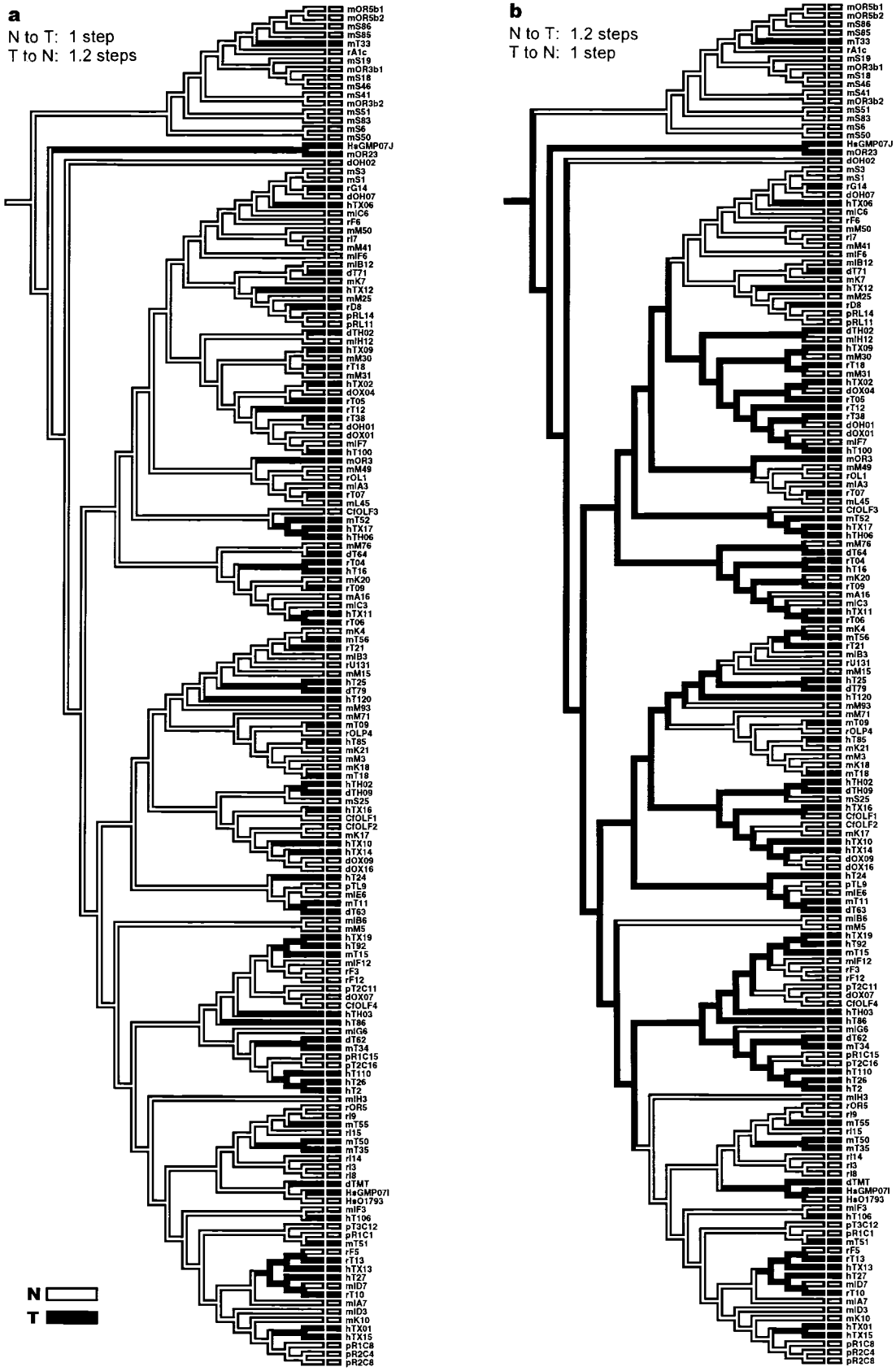
Source	SS	MS	DF	*F*	*P*
		All sequences			
Tissue	0.248	0.248	1	3.61	0.072
Gene[tissue]{random}	1.31	0.082	16	2.99	0.0005
Domain	0.899	0.112	8	4.10	0.0003
Domain*Tissue	0.529	0.066	8	2.41	0.021
		Restricted sample ($K_S \leq 0.2$)			
Tissue	0.397	0.397	1	6.17	0.027
Gene[tissue]{random}	0.835	0.076	11	2.94	0.0034
Domain	0.519	0.065	8	2.52	0.019
Domain*Tissue	0.605	0.076	8	2.93	0.0076

K_A was estimated separately for each domain of each pair of orthologous OR sequences and then normalized by the overall estimate of K_S for that gene; these normalized nonsynonymous substitution rates were used as the dependent (response) variable. The independent variables were tissue (nose or testis), domain, and gene identity (which was nested within tissues and treated as a random effect). The very short and relatively poorly represented IC1 domain was omitted. The analysis was performed both with and without the five genes showing K_S values >0.2. Not surprisingly, genes and domains show strong and significant main effects in both analyses. The most interesting result is the significant domain*tissue interaction, which also appears in both analyses. This indicates that the NOR and TOR samples show distinct patterns of differences (in levels of amino acid conservation) among domains of the OR protein (see Table 2). Owing to the significance of this interaction effect, the main effect of tissue has no simple interpretation and therefore should not be taken at face value.

2 unambiguous TOR → NOR recruitments (Figure 3). In the absence of detailed information on patterns of expression for orthologs and for entire subfamilies of paralogs in several species, we doubt that robust estimates of these relative recruitment probabilities can be made.

It is surprising (on any hypothesis) to find the known TORs distributed as randomly on the OR phylogeny as they are, because most of them were cloned in one laboratory during what amounts to one study (Vanderhaeghen *et al.* 1997). NORs tend to show phylogenetic biases among studies (see Figure 3 legend) for reasons easily attributed to primer sequences and other aspects of cloning strategies that vary among studies and presumably select for different sequence subfamilies. As a consequence, distinctive NOR subfamilies are still being discovered nearly a decade after Buck and Axel (1991) identified the odorant-receptor family. These empirical ascertainment biases would be expected to apply to TORs as much as to an equivalent collection of NORs.

Figure 3.—Inferred histories of nasal and testicular expression within the odorant-receptor gene family. Branches are colored by MacClade (Maddison and Maddison 1992) to indicate TOR (solid) and NOR (open) lineages, as reconstructed under two contrasting models for recruitment between nasal and testicular expression. (a) In the first model, recruitments from the testis to the nose cost 1.2 evolutionary steps relative to a cost of 1 step for recruitments from the nose to the testis; in other words, $T \rightarrow N$ recruitments are assumed to be inherently ~20% less likely to occur (per opportunity) than $N \rightarrow T$ recruitments. The most parsimonious character-state reconstruction (on this assumption) is one in which there are 48 recruitments from the nose to the testis, but just 2 from the testis to the nose. (There are no ambiguous nodes under either model.) (b) In the second model, the relative costs are reversed to favor recruitment from the testis to the nose. On this assumption, the most parsimonious history is one with 18 recruitments from the nose to the testis and 34 from the testis to the nose, and testis expression reconstructs as ancestral for the odorant-receptor family as a whole. Taken literally, this latter inference seems highly implausible, but the lineages represented as "testis-expressed" can be interpreted more generally as "relatively well conserved" (for example, under the focused-olfaction model), in which case a bias toward relatively well-conserved ancestral lineages makes obvious sense. The visual impression that TORs are distributed randomly (uniformly) over the phylogeny is supported by comparisons of average evolutionary distances between sequences on the maximum-likelihood tree (Figure 1). For all 160 sequences, the average sum of intervening branch lengths is 1.62 estimated substitutions per nucleotide position. For 66 TORs, the average pairwise distance is 1.48 (91% of the expected value), indicating that TORs are only weakly clustered within the OR family (at least as currently represented by sequences in the database). By contrast, 10 rat NORs identified by Buck and Axel (1991) show an average distance of 1.17 (72% of expectation), indicating much stronger clustering. Other collections of NORs derived from cDNA libraries by PCR with degenerate primers are less strongly clustered, but often more so than the TORs. For example, 20 mouse NORs from Sullivan *et al.* (1996) show an average distance of 1.43 (88%), and 11 pig NORs from Matarazzo *et al.* (1998) show an average distance of 1.33 (82%). However, 17 mouse NORs from Krautwurst *et al.* (1998) are as uniformly distributed as TORs, with an average distance of 1.51 (93%), and 13 mouse NORs selected for their functional responses to a panel of chemically related odorants (Malnic *et al.* 1999) show an average distance of 2.00 (123%).


a
N to T: 1 step
T to N: 1.2 steps
mOR5b1
mOR5b2
mS86
mS85
mT33
rA1c
mS19
mOR3b1
mS18
mS46
mS41
mOR3b2
mS51
mS83
mS6
mS50
HsGMP07J
mOR23
dOH02
mS3
mS1
rG14
dOH07
hTX06
mIC6
rF6
mM50
rI7
mM41
mIF6
mIB12
dT71
mK7
hTX12
mM25
rD8
pRL14
pRL11
dTH02
mIH12
hTX09
mM30
rT18
mM31
hTX02
dOX04
rT05
rT12
rT38
dOH01
dOX01
mIF7
hT100
mOR3
mM49
rOL1
mIA3
rT07
mL45
CfOLF3
mT52
hTX17
hTH06
mM76
dT64
rT04
hT16
mK20
rT09
mA16
mIC3
hTX11
rT06
mK4
mT56
rT21
mIB3
rU131
mM15
hT25
dT79
hT120
mM93
mM71
mT09
rOLP4
hT85
mK21
mM3
mK18
mT18
hTH02
dTH09
mS25
hTX16
CfOLF1
CfOLF2
mK17
hTX10
hTX14
dOX09
dOX16
hT24
pTL9
mIE6
mT11
dT63
mIB6
mM5
hTX19
hT92
mT15
mIF12
rF3
rF12
pT2C11
dOX07
CfOLF4
hTH03
hT86
mIG6
dT62
mT34
pR1C15
pT2C16
hT110
hT26
hT2
mIH3
rOR5
rI9
mT55
rI15
mT50
mT35
rI14
rI3
rI8
dTMT
HsGMP07I
HsO1793
mIF3
hT106
pT3C12
pR1C1
mT51
rF5
rT13
hTX13
hT27
mID7
rT10
mIA7
mID3
mK10
hTX01
hTX15
pR1C8
pR2C4
pR2C8
N
T
b
N to T: 1.2 steps
T to N: 1 step
mOR5b1
mOR5b2
mS86
mS85
mT33
rA1c
mS19
mOR3b1
mS18
mS46
mS41
mOR3b2
mS51
mS83
mS6
mS50
HsGMP07J
mOR23
dOH02
mS3
mS1
rG14
dOH07
hTX06
mIC6
rF6
mM50
rI7
mM41
mIF6
mIB12
dT71
mK7
hTX12
mM25
rD8
pRL14
pRL11
dTH02
mIH12
hTX09
mM30
rT18
mM31
hTX02
dOX04
rT05
rT12
rT38
dOH01
dOX01
mIF7
hT100
mOR3
mM49
rOL1
mIA3
rT07
mL45
CfOLF3
mT52
hTX17
hTH06
mM76
dT64
rT04
hT16
mK20
rT09
mA16
mIC3
hTX11
rT06
mK4
mT56
rT21
mIB3
rU131
mM15
hT25
dT79
hT120
mM93
mM71
mT09
rOLP4
hT85
mK21
mM3
mK18
mT18
hTH02
dTH09
mS25
hTX16
CfOLF1
CfOLF2
mK17
hTX10
hTX14
dOX09
dOX16
hT24
pTL9
mIE6
mT11
dT63
mIB6
mM5
hTX19
hT92
mT15
mIF12
rF3
rF12
pT2C11
dOX07
CfOLF4
hTH03
hT86
mIG6
dT62
mT34
pR1C15
pT2C16
hT110
hT26
hT2
mIH3
rOR5
rI9
mT55
rI15
mT50
mT35
rI14
rI3
rI8
dTMT
HsGMP07I
HsO1793
mIF3
hT106
pT3C12
pR1C1
mT51
rF5
rT13
hTX13
hT27
mID7
rT10
mIA7
mID3
mK10
hTX01
hTX15
pR1C8
pR2C4
pR2C8

The apparently random distribution of TORs therefore hints at a potentially excessive evenness. Such evenness might be expected if TORs frequently give rise to NORs.

DISCUSSION

Testis-expressed odorant receptors appear to represent a population of genes that differs in significant respects from the population of canonical nasal odorant receptors. Typical TORs appear to evolve more slowly than typical NORs, especially in certain domains of the OR protein, and some are very well conserved between rat and mouse. These differences contradict predictions derived from both the sperm-chemotaxis and the gratuitous-expression hypotheses. The most obvious alternative explanation is that at least some TORs perform nonolfactory internal functions and that selection for precise or efficient performance of such functions conserves the amino acid sequences of these TORs. Compared to many kinds of genes, typical NORs evolve fairly rapidly. This could reflect changing olfactory environments, or weak selective constraints associated with the broadly overlapping ligand specificities of most NORs (see below), or both.

The hypothesis that nasal odorant receptors might occasionally be recruited to participate in various developmental processes (*e.g.*, sperm maturation) seems plausible given that NORs mediate axonal pathfinding by olfactory sensory neurons (Mombaerts *et al.* 1996; Wang *et al.* 1998; Mombaerts 1999a,b; Ebrahimi and Chess 2000). Odorant receptors constitute ~1% of mammalian genes and undoubtedly possess diverse functional properties, so perhaps it should be expected that some members of this huge family would acquire novel functions in nonolfactory tissues (Drutel *et al.* 1995; Nef and Nef 1997; Dreyer 1998). Even within a given tissue, different ORs could play different roles. For example, TORs expressed postmeiotically in spermatids may represent a small subset of those expressed in the testis as a whole (Walensky *et al.* 1998), so our sample of 10 TORs from whole-testis cDNA libraries might contain few if any genes expressed in spermatids. Perhaps we failed to sample (or to sample adequately) a small class of TORs that mediate sperm chemotaxis as originally proposed (Parmentier *et al.* 1992) and that evolve rapidly owing to their involvement in sexual selection.

But might the testis expression of conserved TORs be a consequence of their conservation, rather than their conservation being a consequence of nonolfactory functions? We considered several models that reverse the chain of causation in this way. Most rest on questionable assumptions, but one seems plausible. It begins with the observation that some odorants are likely to be ecologically critical, in the sense that an individual's fitness will depend strongly on detecting these odorants at low concentrations and distinguishing them accurately from other, chemically similar odorants. In response to selection for "focused" olfactory functions, some NORs would become specialized for detection of certain odorants, to the exclusion of others. Such specialist NORs would tend to evolve relatively slowly after they became optimized for detection of single odorants. Such NORs might also acquire relatively high levels of expression in the olfactory epithelium so as to lower the animal's threshold for detection of the critical odorants, either by increasing receptor concentrations within individual sensory neurons or by increasing the numbers of neurons that select these NORs for expression. As an incidental consequence of the properties that lead to high expression in the olfactory epithelium, some such odorant-specialist NORs might come to be expressed gratuitously in testes (see below). Odorant receptors cloned from testis cDNA libraries therefore would tend to show greater than average levels of amino acid sequence conservation.

This model rests on testable assumptions. For example, it assumes that odorant receptors vary in their degrees of ligand specificity. Typically only one odorant receptor is expressed in a given olfactory sensory neuron, and neurons expressing a given receptor are distributed more or less randomly within one of several zones in the olfactory epithelium (Nef *et al.* 1992; Vassar *et al.* 1993; Ressler *et al.* 1993; Chess *et al.* 1994; Malnic *et al.* 1999). It has long been believed that typical odorant receptors must respond to at least a few different odorants, because mammalian olfactory systems are able to discriminate tens of thousands of odorants using less than 1000 odorant receptors. Thus the olfactory "code" must be combinatorial, with chemically similar ligands being distinguished through patterns of differential response by sensory neurons that respond at least weakly to a number of related ligands. This model has received support from studies that characterize the responses of individual receptors, neurons, and olfactory-bulb glomeruli (Krautwurst *et al.* 1998; Duchamp-Viret *et al.* 1999; Malnic *et al.* 1999; Rubin and Katz 1999; Touhara *et al.* 1999; reviewed by Buck 2000). In one experiment, isolated mouse sensory neurons (expressing single odorant receptors that were subsequently identified) responded with varying strengths to several different chemically related odorants. Each odorant produced a different pattern of responses among the members of a set of neurons, each of which responded to at least a few of the odorants in the set (Malnic *et al.* 1999). The fact that individual odorant receptors may discriminate more or less sharply among closely related odorants implies that their global breadths of response probably vary as well and have been subject to adjustment by natural selection. The focused-olfaction model predicts that highly conserved receptors should tend to have narrowly tuned patterns of response; this prediction could be tested by means of techniques like those used by Krautwurst *et al.* (1998), Malnic *et al.* (1999), and Rubin and Katz (1999).

The assumption that different NORs might be expressed at different levels in the olfactory epithelium also seems plausible, and some evidence supports it. Widely varying numbers of olfactory neurons are labeled by *in situ* hybridizations with probes made from different odorant-receptor genes (*e.g.*, Ressler *et al.* 1993, 1994). The interpretation of these experiments is complicated by the fact that hybridization probes may identify several closely related paralogous genes, but there is no reason to assume that all odorant receptors should be expressed by equal numbers of sensory neurons, and no evidence points toward such a conclusion. The focused-olfaction model predicts that highly conserved receptors should often be expressed abundantly, and this prediction could be tested through approaches that minimize the detection of more than one related sequence.

The focused-olfaction model predicts that odorants recognized by highly conserved receptors will be ones of special ecological importance; this implies that such odorants signal either great danger or great opportunity. The identities of such odorants, once determined, should be consistent with this interpretation. For example, animals exposed to such odorants might show signs of unusual agitation or interest in the source of the odor. These predicted associations could also be pursued in the opposite direction, from an ecologically derived understanding of significant odorants, toward an analysis of the expression, tuning, and evolutionary conservation of their receptors. Related species with similar ecologies should tend to share more of these "most significant" odorants than equally related species with dissimilar ecologies; thus, under the focused-olfaction model (but not under the internal-function model), ecologically similar species should share greater numbers of highly conserved TORs than ecologically dissimilar species, all else being equal.

Male and female ecologies often differ, especially with respect to reproductive strategies. The focused-olfaction model therefore suggests that some specialist odorant receptors might be expressed mainly or exclusively in the members of one sex. Indeed, a potential mechanistic explanation for the gratuitous expression of some conserved ORs in testes might be that they are specialized to detect odorants of particular importance to males and that their expression tends to increase in response to androgens (appropriately in the olfactory epithelium, gratuitously in the testes). Sexually dimorphic expression of conserved odorant receptors would support the focused-olfaction model, but dimorphic expression is not a strong or necessary prediction of the model in its general form. Many components of the RNA polymerase II transcription complex appear to be expressed at high levels during the early haploid phases of meiosis, giving rise to "a permissive environment for transcription initiation" (Schmidt 1996). If ORs that are relatively highly expressed in the olfactory epithelium tend to be especially liable to gratuitous expression under such conditions, then a correlation between evolutionary conservation and testicular expression could occur under the focused-olfaction model, even in the absence of sexually dimorphic expression in the olfactory epithelium.

Four TORs in our data set (rT19/mT18, rT09, rT05/mT07, and rT38/mT53) were tested for expression in the olfactory epithelia of mice and/or rats by Vanderhaeghen *et al.* (1997). Two show obvious signals (rT09 and rT05/mT07, with intermediate to modestly high conservation, respectively) and two do not (rT19/mT18 and rT38/mT53, with very low and rather high conservation). These RPA experiments used mRNA pools derived from both male and female olfactory epithelia (P. Vanderhaeghen, personal communication), so they imply that not all well-conserved TORs are expressed at high levels in the nose. However, they do not provide evidence for or against sexually dimorphic expression.

Distinctive predictions can be derived also from the internal-function model. For example, mice carrying targeted disruptions of conserved TORs (or carrying constructs that overexpress such genes) might show reproductive deficits either in traditional phenotypic screens or in tests of reproductive success in competitive seminatural social settings, which can reveal subtle functional differences that would otherwise be difficult to detect (Potts *et al.* 1994). Cetaceans (especially toothed whales) have greatly reduced olfactory systems (Oelschläger 1989), so under either model their odorant-receptor gene families should be reduced. However, the internal-function model predicts that cetaceans will retain a set of conserved TORs equivalent to those of their closest relatives (hippopotamuses and ruminants; Nikaido *et al.* 1999), while the focused-olfaction model predicts that cetacean TORs will be as severely reduced as their NORs. This latter prediction is supported weakly by a sample of 17 cetacean OR sequences cloned by PCR with degenerate primers from genomic DNA of the striped dolphin, *Stenella coeruleoalba* (Freitag *et al.* 1998). Every one of the 17 sequences contains at least one frameshift mutation or premature stop codon, consistent with the hypothesis that cetacean ORs have lacked function for millions of years and therefore have not been subject to selection. However, these sequences were cloned from the genome at large, not from a testis cDNA library, so this result does not rule out the possibility that a few ORs [not sampled in the study by Freitag *et al.* (1998)] remain active in cetacean testes and subject to selection for some function there.

If typical NORs respond to several different odorants, and if typical odorants stimulate sensory neurons expressing several different NORs, then most NORs are somewhat functionally redundant. How can such a large set of individually nonessential genes be maintained? As subfamilies of NORs diversify, selection on their individual member genes should often become rather weak. Disabling mutations of weakly selected ORs might there-

fore be carried to fixation at relatively high rates, converting formerly functional NORs to pseudogenes. NOR "deaths" would need to be replenished from some source. To the degree that successful new NOR lineages tend to derive from relatively well-conserved sequences, and to the degree that the birth-and-death process turns over quickly, most functional NORs might come to descend (if not immediately, then at modest removes) from a relatively small number of highly conserved (perhaps often testis-expressed) ancestors. In this model, conserved OR lineages (whatever the reason for their conservation) would form the "trunks" of most subfamily gene trees and would thereby contribute strongly to the shape and diversity of the entire odorant-receptor family.

We thank Fred Adler, Richard Axel, Linda Buck, Joshua Cherry, Patrice Corneli, Joe Dickinson, Maria Ganfornina, Barbara Graves, Glenn Herrick, Stephen Heinemann, Laurie Issel-Tarver, Kevin Johnson, David Mason, Dustin Penn, Wayne Potts, John Roth, Victoria Rowntree, Diego Sanchez, Pierre Vanderhaeghen, Lisa Vawter, Loren Walensky, David Witherspoon, and two anonymous reviewers for helpful discussions and for comments on the manuscript. We thank Joe Felsenstein for the prerelease version of DNAML 3.6, and we thank Doron Lancet and Yitzhak Pilpel for a preprint of Pilpel and Lancet (1999). This study was supported in part by the Center for Biopolymers, the Dee Endowment, the Howard Hughes Medical Institute, and the Huntsman Cancer Institute.

LITERATURE CITED

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

Andersson, M., 1994 *Sexual Selection.* Princeton University Press, Princeton, NJ.

Arnqvist, G., 1998 Comparative evidence for the evolution of genitalia by sexual selection. Nature **393:** 784–786.

Asai, H., H. Kasai, Y. Matsuda, N. Yamazaki, F. Nagawa *et al.*, 1996 Genomic structure and transcription of a murine odorant receptor gene: differential initiation of transcription in the olfactory and testicular cells. Biochem. Biophys. Res. Commun. **221:** 240–247.

Buck, L., 2000 The molecular architecture of odor and pheromone sensing in mammals. Cell **100:** 611–618.

Buck, L., and R. Axel, 1991 A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell **65:** 175–187.

Chess, A., I. Simon, H. Cedar and R. Axel, 1994 Allelic inactivation regulates olfactory receptor gene expression. Cell **78:** 823–834.

Clark, A. G., D. J. Begun and T. Prout, 1999 Female × male interactions in *Drosophila* sperm competition. Science **283:** 217–220.

Dreyer, W. J., 1998 The area code hypothesis revisited: olfactory receptors and other related transmembrane receptors may function as the last digits in a cell surface code for assembling embryos. Proc. Natl. Acad. Sci. USA **95:** 9072–9077.

Drutel, G., J. M. Arrang, J. Diaz, C. Wisnewsky, K. Schwartz *et al.*, 1995 Cloning of OL1, a putative olfactory receptor and its expression in the developing rat heart. Recept. Channels **3:** 33–40.

Duchamp-Viret, P., M. A. Chaput and A. Duchamp, 1999 Odor response properties of rat olfactory receptor neurons. Science **284:** 2171–2174.

Eberhard, W. G., 1996 *Female Control: Sexual Selection by Cryptic Female Choice.* Princeton University Press, Princeton, NJ.

Ebrahimi, A. W., and A. Chess, 2000 Olfactory neurons are interdependent in maintaining axonal projections. Curr. Biol. **10:** 219–222.

Felsenstein, J., 1981 Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17:** 368–376.

Felsenstein, J., 1989 *PHYLIP*—phylogeny inference package (version 3.2). Cladistics **5:** 164–166.

Freitag, J., G. Ludwig, I. Andreini, P. Rössler and H. Breer, 1998 Olfactory receptors in aquatic and terrestrial vertebrates. J. Comp. Physiol. A **183:** 635–650.

Holland, B., and W. R. Rice, 1999 Experimental removal of sexual selection reverses intersexual antagonistic coevolution and removes a reproductive load. Proc. Natl. Acad. Sci. USA **96:** 5083–5088.

Krautwurst, D., K. W. Yau and R. R. Reed, 1998 Identification of ligands for olfactory receptors by functional expression of a receptor library. Cell **95:** 917–926.

Levy, N. S., H. A. Bakalyar and R. R. Reed, 1991 Signal transduction in olfactory neurons. J. Ster. Biochem. Mol. Biol. **39:** 633–637.

Li, W.-H., 1993 Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J. Mol. Evol. **36:** 96–99.

Li, W.-H., 1997 *Molecular Evolution.* Sinauer, Sunderland, MA.

Maddison, W. P., and D. R. Maddison, 1992 *MacClade: Analysis of Phylogeny and Character Evolution, Version 3.* Sinauer, Sunderland, MA.

Makalowski, W., and M. S. Boguski, 1998 Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. Proc. Natl. Acad. Sci USA **95:** 9407–9412.

Malnic, B., J. Hirono, T. Sato and L. B. Buck, 1999 Combinatorial receptor codes for odors. Cell **96:** 713–723.

Matarazzo, V., A. Tirard, M. Renucci, A. Belaich and J. L. Clement, 1998 Isolation of putative olfactory receptor sequences from pig nasal epithelium. Neurosci. Lett. **249:** 87–90.

McVean, G. T., and L. D. Hurst, 1997 Evidence for a selectively favourable reduction in the mutation rate of the X chromosome. Nature **386:** 388–392.

Metz, E. C., R. Robles-Sikisaka and V. D. Vacquier, 1998 Nonsynonymous substitution in abalone sperm fertilization genes exceeds substitution in introns and mitochondrial DNA. Proc. Natl. Acad. Sci. USA **95:** 10676–10681.

Mombaerts, P., 1999a Molecular biology of odorant receptors in vertebrates. Annu. Rev. Neurosci. **22:** 487–509.

Mombaerts, P., 1999b Seven-transmembrane proteins as odorant and chemosensory receptors. Science **286:** 707–711.

Mombaerts, P., F. Wang, C. Dulac, S. K. Chao, A. Nemes *et al.*, 1996 Visualizing an olfactory sensory map. Cell **87:** 675–686.

Nef, S., and P. Nef, 1997 Olfaction: transient expression of a putative odorant receptor in the avian notochord. Proc. Natl. Acad. Sci. USA **94:** 4766–4771.

Nef, P., I. Hermans-Borgmeyer, H. Artières-Pin, L. Beasley, V. E. Dionne *et al.*, 1992 Olfaction in birds: differential embryonic expression of nine putative odorant receptor genes in the avian olfactory system. Proc. Natl. Acad. Sci. USA **89:** 8948–8952.

Nikaido, M., A. P. Rooney and N. Okada, 1999 Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. Proc. Natl. Acad. Sci. USA **96:** 10261–10266.

Oelschläger, H. A., 1989 Early development of the olfactory and terminalis systems in baleen whales. Brain Behav. Evol. **34:** 171–183.

Pamilo, P., and N. O. Bianchi, 1993 Evolution of the Zfx and Zfy genes: rates and interdependence between the genes. Mol. Biol. Evol. **10:** 271–281.

Parmentier, M., F. Libert, S. Schurmans, S. Schiffmann, A. Lefort *et al.*, 1992 Expression of members of the putative olfactory receptor gene family in mammalian germ cells. Nature **355:** 453–455.

Partridge, L., and L. D. Hurst, 1998 Sex and conflict. Science **281:** 2003–2008.

Pilpel, Y., and D. Lancet, 1999 The variable and conserved interfaces of modeled olfactory receptor proteins. Protein Sci. **8:** 969–977.

Potts, W. K., C. Manning and E. K. Wakeland, 1994 The role of infectious disease, inbreeding and mating preferences in main-

taining MHC genetic diversity: an experimental test. Philos. Trans. R. Soc. Lond. Ser. B **346:** 369–378.

Raming, K., S. Konzelmann and H. Breer, 1998 Identification of a novel G-protein coupled receptor expressed in distinct brain regions and a defined olfactory zone. Recept. Channels **6:** 141–151.

Ressler, K. J., S. L. Sullivan and L. B. Buck, 1993 A zonal organization of odorant receptor gene expression in the olfactory epithelium. Cell **73:** 597–609.

Ressler, K. J., S. L. Sullivan and L. B. Buck, 1994 Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell **79:** 1245–1255.

Rice, W. R., 1996 Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. Nature **381:** 232–234.

Rice, W. R., 1998 Male fitness increases when females are eliminated from gene pool: implications for the Y chromosome. Proc. Natl. Acad. Sci. USA **95:** 6217–6221.

Rubin, B. D., and L. C. Katz, 1999 Optical imaging of odorant representation in the mammalian olfactory bulb. Neuron **23:** 499–511.

Schmidt, E. E., 1996 Transcriptional promiscuity in testes. Curr. Biol. **6:** 768–769.

Sokal, R. R., and J. F. Rohlf, 1995 *Biometry: The Principles and Practice of Statistics in Biological Research*, Third Edition. Freeman, New York.

Sullivan, S. L., M. C. Adamson, K. J. Ressler, C. A. Kozak and L. B. Buck, 1996 The chromosomal distribution of mouse odorant receptor genes. Proc. Natl. Acad. Sci. USA **93:** 884–888.

Touhara, K., S. Sengoku, K. Inaki, A. Tsuboi, J. Hirono *et al.*, 1999 Functional identification and reconstitution of an odorant receptor in single olfactory neurons. Proc. Natl. Acad. Sci. USA **96:** 4040–4045.

Tsaur, S.-C., and C. I. Wu, 1997 Positive selection and the molecular evolution of a gene of male reproduction, Acp 26Aa of *Drosophila*. Mol. Biol. Evol. **14:** 544–549.

Vacquier, V. D., 1998 Evolution of gamete recognition proteins. Science **281:** 1995–1998.

Vanderhaeghen, P., S. Schurmans, G. Vassart and M. Parmentier, 1993 Olfactory receptors are displayed on dog mature sperm cells. J. Cell Biol. **123:** 1441–1452.

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

Vassar, R., J. Ngai and R. Axel, 1993 Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell **74:** 309–318.

Walensky, L. D., J. Roskams, R. J. Lefkowitz, S. H. Snyder and G. V. Ronnett, 1995 Odorant receptors and desensitization proteins colocalize in mammalian sperm. Mol. Med. **1:** 130–141.

Walensky, L. D., M. Ruat, R. E. Bakin, S. Blackshaw, G. V. Ronnett *et al.*, 1998 Two novel odorant receptor families expressed in spermatids undergo 5′-splicing. J. Biol. Chem. **273:** 9378–9387.

Wang, F., A. Nemes, M. Mendelsohn and R. Axel, 1998 Odorant receptors govern the formation of a precise topographic map. Cell **93:** 47–60.

Wolfe, K. H., and P. M. Sharp, 1993 Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. J. Mol. Evol. **37:** 441–456.

Wolfe, K. H., P. M. Sharp and W.-H. Li, 1989 Mutation rates differ among regions of the mammalian genome. Nature **337:** 283–285.

Yokoyama, S., and W. T. Starmer, 1996 Evolution of G-protein coupled receptor superfamily, pp. 93–119 in *Human Genome Evolution*, edited by M. S. Jackson, G. Dover and T. Strachan. BIOS Scientific, Oxford.

Communicating editor: S. Yokoyama