Genomics **61,** 24–36 (1999)
Article ID geno.1999.5900, available online at http://www.idealibrary.com on IDEAL®

Primate Evolution of an Olfactory Receptor Cluster: Diversification by Gene Conversion and Recent Emergence of Pseudogenes

Dror Sharon,*[,1] Gustavo Glusman,* Yitzhak Pilpel,* Miriam Khen,* Frank Gruetzner,† Thomas Haaf,† and Doron Lancet*[,2]

*Department of Molecular Genetics and the Crown Human Genome Center, The Weizmann Institute of Science, Rehovot 76100, Israel; and †Max-Planck-Institute of Molecular Genetics, Ihnestrasse 73, 14195 Berlin, Germany

Received February 11, 1999; accepted June 10, 1999

The olfactory receptor (OR) subgenome harbors the largest known gene family in mammals, disposed in clusters on numerous chromosomes. We have carried out a comparative evolutionary analysis of the best characterized genomic OR gene cluster, on human chromosome 17p13. Fifteen orthologs from chimpanzee (localized to chromosome 19p15), as well as key OR counterparts from other primates, have been identified and sequenced. Comparison among orthologs and paralogs revealed a multiplicity of gene conversion events, which occurred exclusively within OR subfamilies. These appear to lead to segment shuffling in the odorant binding site, an evolutionary process reminiscent of somatic combinatorial diversification in the immune system. We also demonstrate that the functional mammalian OR repertoire has undergone a rapid decline in the past 10 million years: while for the common ancestor of all great apes an intact OR cluster is inferred, in present-day humans and great apes the cluster includes nearly 40% pseudogenes. © 1999 Academic Press

INTRODUCTION

Olfactory receptor (OR) proteins are G-protein-coupled receptors (Buck and Axel, 1991; Lancet and Pace, 1987; Reed, 1990), expressed mainly in the olfactory neuroepithelium, but also found in other tissues (Drutel *et al.,* 1995; Walensky *et al.,* 1998), including mammalian germ cells (Parmentier *et al.,* 1992). Their major role is the activation of a signal transduction pathway leading to odorant recognition and discrimination (Buck and Axel, 1991; Lancet and Ben-Arie, 1993; Ngai *et al.,* 1993). Each neuronal cell may express only one OR gene (Lancet, 1991), and even just one allele at a given locus (Chess *et al.,* 1994). This expression pattern is believed to provide the molecular basis of odor discrimination.

OR proteins were first cloned from rat (Buck and Axel, 1991) and later shown to be present in the genome of a wide variety of species (Barth *et al.,* 1997; Ben-Arie *et al.,* 1994; Freitag *et al.,* 1995; Issel-Tarver and Rine, 1997; Nef *et al.,* 1996; Sullivan *et al.,* 1996). The total number of OR genes in the mammalian "olfactory subgenome" is estimated to be 300–1000 (Buck and Axel, 1991; Lancet *et al.,* 1993a; Ressler *et al.,* 1994), constituting a very large gene superfamily, divided into more than 20 families (Lancet and Ben-Arie, 1993). Analyses of OR genomic sequences provided evidence for a simple gene structure, which includes, in addition to the first exon, the ~1-kb long intronless coding region, a 5- to 6-kb upstream intron, a short 5′ untranslated exon, and a putative control region (Asai *et al.,* 1996; Glusman *et al.,* 1996; Walensky *et al.,* 1998).

OR genes have been shown to be disposed in clusters on several chromosomes in human (Ben-Arie *et al.,* 1994; Rouquier *et al.,* 1998b) and other species (Barth *et al.,* 1997; Issel-Tarver and Rine, 1997; Sullivan *et al.,* 1996), as is the case for other multigene families (Heim and Meyer, 1992; Higgs *et al.,* 1989). The gene clusters have likely arisen by a process of repeated duplication of individual genes (Glusman *et al.,* 1996) as well as of entire clusters (Lancet and Ben-Arie, 1993; Trask *et al.,* 1998). The cluster organization may be related to the regulation of gene expression, which leads to the unique patterns of OR cellular expression (Chess *et al.,* 1994).

In the human genome, OR clusters have been identified on numerous chromosomes (Buettner *et al.,* 1998; Olsen *et al.,* 1993; Rouquier *et al.,* 1998b), and more than 100 OR genes have been partially or fully sequenced so far. One of the OR clusters, encompassing an ~450-kb region on human chromosome 17p13.3, has been studied in our laboratory in considerable de-

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession Nos. AF101730 to AF101778 and AF087915 to AF087930.

[1] Current address: Ocular Molecular Genetics Laboratory, Massachusetts Eye and Ear Infirmary, Harvard Medical School, 243 Charles Street, Boston, MA 02114.
[2] To whom correspondence should addressed.

0888-7543/99 $30.00
Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

tail (Ben-Arie *et al.,* 1994; Glusman *et al.,* 1996), including physical mapping and DNA sequencing (Glusman *et al.* 1996, Glusman, submitted). The 16 OR coding regions in this cluster belong to seven OR subfamilies, potentially generated by a complex multistep duplication mechanism. This well-characterized cluster constitutes an ideal target for comparative studies of OR evolution.

One of the most important questions in the field of olfaction is how OR gene diversity has been generated. In contrast to the case of immune genes, no evidence exists for somatic DNA rearrangements or somatic mutations in the coding regions of olfactory genes, and OR diversity therefore appears to be wholly germline-encoded. Single point mutations in newly duplicated genes are a potential diversification pathway, but it may not suffice to allow organisms to cope with a constantly changing chemosensory environment. In some other multigene families, gene conversion among paralogs has been proposed to be a pathway for diversification (Slightom *et al.,* 1985; Wines *et al.,* 1991), a process believed to be enhanced by gene clustering. We report here a comparison of the OR genes in the human chromosome 17 cluster to their orthologous coding regions in chimpanzee, as well as in other primates. The data lend considerable support for the role of gene conversion in the evolution of olfactory receptor diversity.

In the human genome, OR clusters have been reported to harbor a very high proportion of pseudogenes (Ben-Arie *et al.,* 1994; Buettner *et al.,* 1998; Rouquier *et al.,* 1998b), potentially reflecting a recent diminution of the olfactory repertoire. Indeed, in the currently studied cluster at least 40% of the coding regions represent pseudogenes. However, relatively little information has been available on the timing of this gene inactivation process. The present comparative sequence analysis reveals that rapid genome dynamics applies also to pseudogene generation and that the entire human chromosome 17 OR cluster may have been functionally intact at the dawn of great ape evolution.

MATERIALS AND METHODS

Fluorescence in situ hybridization (FISH). Chromosomes were prepared from human (HSA) peripheral blood lymphocytes, from EBV-transformed lymphoblastoid cells of chimpanzee (PTR), gorilla (GGO), orangutan (PPY), and *Presbytis cristata,* and from fibroblast cultures of *Callithrix geoffrey.* For FISH (Ward *et al.,* 1995), chromosome preparations were treated with 100 μg/ml RNase A in 2× SSC at 37°C for 60 min and with 0.01% pepsin in 10 mM HCl at 37°C for 10 min and then dehydrated in an ethanol series (70, 85, and 100%). Slides were denatured at 80°C in 70% formamide, 2× SSC, pH 7.0, and again dehydrated in an alcohol series. PAC DNA was labeled by standard nick translation with biotin-16–dUTP (Boehringer Mannheim). Biotinylated PAC DNA (10 ng/μl) was coprecipitated with 100 ng/μl human cot-1 competitor DNA (Gibco) and 500 ng/μl salmon sperm carrier DNA and redissolved in 50% formamide, 10% dextran sulfate, 2× SSC. After 10 min of denaturation at 70°C, 30 μl of hybridization mixture was applied to each slide and sealed under a coverslip. Slides were left to hybridize in a moist chamber at 37°C for 1 to 3 days. Slides were washed 3 × 5 min in 50% formamide, 2× SSC at 42°C followed by a 5-min wash in 0.1× SSC at 65°C. Hybridized probes were detected by fluorescein isothiocyanate (FITC)-conjugated avidin (Vector). Chromosomes and cell nuclei were counterstained with 1 μg/ml 4,6-diamidino-2-phenylindole (DAPI) in 2× SSC for 5 min. The slides were mounted in 90% glycerol, 100 mM Tris–HCl, pH 8.0, and 2.3% DABCO. Images were taken with a Zeiss epifluorescence microscope equipped with a thermoelectronically cooled charge-coupled device camera (Photometrics CH250), which was controlled by an Apple Macintosh computer. Oncor imaging software was used to capture grayscale images and to superimpose the images onto a color image. Oncor imaging software was also used to invert the DAPI image into a G-banded metaphase for identification of the chromosomes.

The PACs used are LLNLP704E02527Q3 (P123), LLNLP704M22845Q3 (P129), LLNLP704C10910Q3 (P008), LLNLP704O19796Q3 (P107), and LLNLP704P041058Q3 (P110) from human PAC library 704 (Ioannou *et al.,* 1994).

DNA isolation. Genomic DNA from chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), and macaque (*Macaca fascicularis*) individuals was isolated from whole blood (kindly provided by Dr. Yigal Horvitz from the Israeli Safari Zoo) using the Genomix DNA preparation kit (Talent srl, Trieste Italy). Genomic DNA from chimpanzee, bonobo (*Pan paniscus*), and gorilla individuals was kindly provided by Dr. Keneth K. Kidd (Yale University), and genomic DNA from orangutan (*Pongo pygmaeus*) individuals was kindly provided by Cedric C. Muir (Simon Fraser University).

PCR and primers. Primers for PCR amplification and for sequencing were designed to amplify the full open-reading frame of the 16 human OR genes, based on the available sequences (Glusman, in preparation) as follows: OR17-2, 5′-GATACTGTTGTTTTTCCATTA-3′ and 5′-TTATGTACGATGCCATTTCA-3′ (1100 bp, 55°C); OR17-4, 5′-TTTGGTAAACATTTGCTGGT-3′ and 5′-TCCACTTTAATGCTGTCTTT-3′ (1016 bp, 55°C); OR17-7, 5′-TGATATTCCTCTCCCCTTTC-3′ and 5′-ATATACCGTAACTGGTGACTA-3′ (1013 bp, 55°C); OR17-23, 5′-ATTTGTTGGTGTTAATGTTGC-3′ and 5′-CACTTCAATGCCCAAAATTAC-3′ (999 bp, 55°C); OR17-24, 5′-CTAACACTGCTCAGTCCACT-3′ and 5′-CGTCCAGCAGAGATGTCCAG-3′ (935 bp, 60°C); OR17-25, 5′-CTGGACATCTCTGCTGGACG-3′ and 5′-GCCCATACCTAGTCCTTCAG-3′ (823 bp, 55°C); OR17-30, 5′-GTTGGTGTTAATGTTACAGAA-3′ and 5′-CCAGTCTCCACTTCAATCT-3′ (1007 bp, 55°C); OR17-31, 5′-GTTGGTGTTAATGTTACAGAG-3′ and 5′-CCGGTCTCCACTTCAATCC-3′ (1006 bp, 55°C); OR17-40, 5′-TTGGGAGGAGTGATTCTTTTC-3′ and 5′-ATTTTTTTCCTAGTTTCTGGCT-3′ (1076 bp, 55°C); OR17-93, 5′-TTTTATTCACGGTTTTGAGG-3′ and 5′-AAGACTCCCAAGTATTACTGTC-3′ (1051 bp, 55°C); OR17-201, 5′-TGAACTGATACCTCCCCTGC-3′ and 5′-AAAAGGCAGGAAAGGAGGTCA-3′ (1017 bp, 55°C); OR17-208, 5′-CACCTTGGGTTTCTCAGCAG-3′ and 5′-GAATAGCATGCTCACCAATC-3′ (1064 bp, 60°C); OR17-209, 5′-CCCCCATATTGGATGTCAAG-3′ and 5′-GGTCCCAAAGAGTCAAATAG-3′ (1117 bp, 55°C); OR17-210, 5′-GAGTCACAGAATGATGAAGAAG-3′ and 5′-AAGCCACTGATTTAGAGTGA-3′ (1039 bp, 55°C); OR17-228, 5′-TGTCTTCATGGGAAATTTTCCT-3′ and 5′-TGAACTGATACCTCCCCTGC-3′ (1054 bp, 60°C). Primers for PCR amplification of noncoding regions located on the human chromosome 17 OR cluster were as follows: NC2, 5′-CCCCTATTCTGTCGCTCAAG-3′ and 5′-AGAACGGACACTGGGTAAGG-3′ (1075 bp, 55°C); NC4, 5′-TTCGATTATTAACAGGTTCTG-3′ and 5′-TTTAGCAACTCTCCTTCTGC-3′ (785 bp, 55°C); NC5, 5′-AAAGACATGGGCTAAATCAC-3′ and 5′-CTATATAAGGAGGTATCCTC-3′ (779 bp, 55°C).

The PCR mixture contained a total volume of 25 μl containing 50 mM KCl, 10 mM Tris–HCl, pH 8.3, 1.5 mM $MgCl_2$, a 200 mM concentration of each deoxyribonucleoside triphosphate (dNTP), 0.1 mM concentration of each primer, 1 unit of *Taq* DNA polymerase (Boehringer Mannheim, Germany), and 50 ng of genomic DNA. PCR products were electrophorated in a 1% agarose gel to view their size and then purified using the High Pure PCR Product Purification Kit (Boehringer Mannheim).

Cloning of PCR products. When more than one type of PCR product was obtained (namely two closely related paralogs or two alleles of the same gene), the PCR products were cloned using the TOPO TA Cloning kit (Invitrogen, Carlsbad, CA) or the CloneAMP pAMP1 system for rapid cloning kit (Gibco BRL). Plasmid DNA for sequencing was purified with the Wizard Plus SV Miniprep kit (Promega).

Sequencing. Sequencing reactions were performed on PCR products or clones in both directions with dye terminators (Dye terminator cycle sequencing kit; Perkin–Elmer) on an ABI 373 or ABI 377 automated sequencer. After base calling with ABI Analysis Software (version 3.0), the analyzed data were edited using Sequencher 3.0 (GeneCodes Corp., Ann Arbor, MI).

Sequence analysis. The new sequences described in this paper have been deposited with GenBank under Accession Nos. AF101730 through AF101778. Additional OR sequences used in this study are as follows: OR17-1 (AF087915), OR17-2 (AF087916), OR17-4 (AF087917), OR17-7 (AF087918), OR17-23 (AF087919), OR17-24 (AF087920), OR17-25 (AF087921), OR17-30 (AF087922), OR17-31 (AF087923), OR17-40 (AF087924), OR17-93 (AF087925), OR17-201 (AF087926), OR17-208 (AF087927), OR17-209 (AF087928), OR17-210 (AF087929), and OR17-228 (AF087930).

Nucleotide sequence alignments were performed using the ClustalW program with standard parameters, and phylogeny was assessed using the neighbor-joining analysis (Higgins *et al.,* 1996). Family assignments and nomenclature are derived from a scheme for olfactory receptor gene classification and nomenclature (Glusman, in preparation).

The detection of gene conversion events was performed in two ways: (1) A pretty box analysis of the relevant sequences was drawn using the MacBoxShade 1.0.8 computer program (http://www.netaxs.com/~jayfar/mops.html), and a site by site search for shared nucleotides between paralogous sequences was performed. (2) The nucleotide alignment was searched for recombination breakpoints using the TOPAL package (McGuire and Wright, 1998). The window sliding graph of the difference in the sum of squares (Dss) was analyzed, and the regions having a high Dss level were further analyzed by eye.

The identification of human–chimpanzee orthologous pairs was accomplished using the HoverGen database (Duret *et al.,* 1994). The following orthologous pairs were used to compute a typical value for interortholog distances for coding regions: α-globin (HS661), β-globin (HUMHBB222), insulin (HSINSU), relaxin 1 (HSREL2), relaxin 2 (A06925), interleukin 8 (HSAL8RB4), *N*-formyl-1 (HUMFMLP), C5aR (HSC5AR), low-affinity *N*-formyl (HSLIPA4R), *N*-formyl2 (HUMFMLPY), C-myc (HUMMYCPOA), interferon inducible (HSINFIN3), complement component C4A (HUMMHC4AB), intercellular adhesion molecule 1 (HSICAM1), von Willebrand factor 23 (HUMVWFAA), complement receptor type 1- CR1 (HUMCR1SF41), MHC-class I B(2) microglobulin (HSMGLO), non-Anon-B hepatitis-associated (HUMHCAMAO8), and amiloride-sensitive Na^+ channel (HSU38165). In those cases where intronic sequences were available from both human and chimpanzee, the data were used to obtain an estimate of noncoding nucleotide divergence. The statistical significance of the differences between the group averages was determined by a *t*-test analysis. Nucleotide divergence values were computed using the GeneAssist program (Shpaer, 1997).

RESULTS

The Orthologous OR Cluster in Great Apes

A minimal set of five PAC clones covering the human cluster were used as fluorescence *in situ* hybridization probes. These PACs were hybridized to both human and great ape metaphase chromosome spreads. In human, each probe showed specific hybridization to chromosome 17p13 (Fig. 1A). In the chimpanzee, each probe was localized to the subtelomeric short arm of chromosome 19 (Fig. 1B). The hybridized chimpanzee chromosomal band (PTR 19p15) is homologous to the studied human chromosomal band (HSA 17p13) (Paris Conference, 1975). An apomorphic reciprocal translocation between ancestral chromosomes homologous to human chromosomes 5 and 17 has occurred in the gorilla (Stanyon *et al.,* 1992). Thus, sequences homologous to human chromosomes 17 are diverged on gorilla chromosomes 4 and 19. Interestingly, each PAC was mapped on gorilla chromosomal band 19q23–q24 (Fig. 1C), the homolog of HSA 17p13, in close proximity to the evolutionary chromosome breakpoint. In orangutan, all five PACs were localized in the subtelomeric region 19pter (Fig. 1D). Orangutan chromosome 19 is the counterpart of HSA 17 (Jauch *et al.,* 1992).

Under routine conditions of high stringency, no hybridization signals were visible on other chromosomes. This supports the conclusion that in all great apes, similar to humans, this group of OR genes is disposed in a contiguous genomic cluster, which is located in the syntenic chromosomal region. This cluster conservation may actually be more general. Two PACs, namely 107 and 129, were also hybridized on chromosomes of *Pr. cristata* and *C. geoffrey.* In both Old and New World monkeys, specific hybridization signals were detected exclusively on the homologs of HSA 17 (data not shown).

The OR Orthologous Gene Pairs

As a first step in the characterization of the syntenic genomic cluster in chimpanzee, we sequenced this simian OR orthologs by generating gene-specific primers for PCR amplification, based on the knowledge of the full-length sequences of all 16 OR genes from the human chromosome 17 cluster (Glusman, submitted). PCR amplification was performed at high stringency on chimpanzee genomic DNA, with 16 primer pairs corresponding to the DNA sequences immediately flanking the human open reading frames. These sequences harbor considerable variability, thus allowing most of the OR open reading frames to be uniquely amplified. Using these primer pairs, 16 chimpanzee coding regions have been amplified and sequenced. The relatively low degree of interspecific divergence (mean of 2.1 $\pm$ 0.1%), as well as the fact that a single PCR amplification product was obtained for almost all of the OR genes, is evidence that we have likely identified unique orthologs of the human OR gene cluster members. This, together with the results of the *in situ* hybridization experiments, suggests, although not demonstrated by direct mapping data, that the chimpanzee cluster organization may be the same as that in human (Fig. 2).

The dendrogram presented in Fig. 3 shows two exceptions to the straightforward orthology relationships. One case is the human paralogous pair hOR17-30 and hOR17-31 (99.0% identity), for which

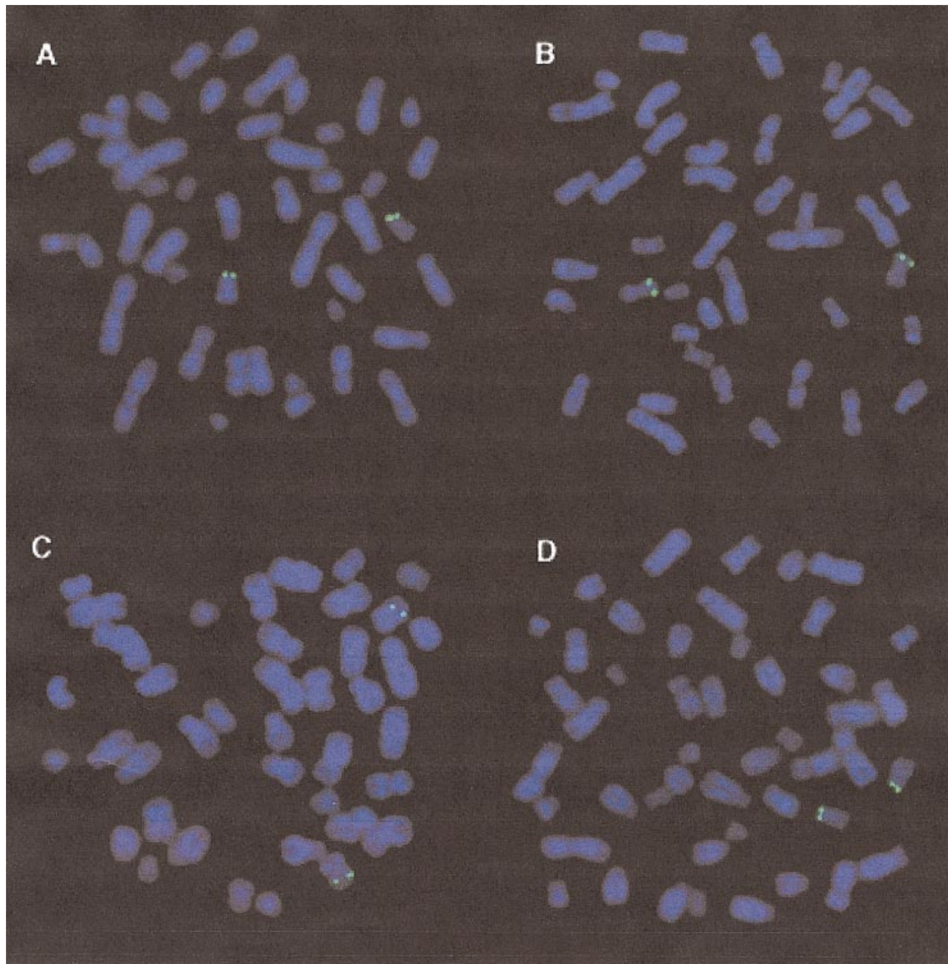

FIG. 1. Fluorescence *in situ* hybridization on primate chromosomes using human PAC 110. The biotinylated probe is detected by FITC–avidin (green fluorescence). Chromosomes are counterstained with DAPI (blue). (**A**) human, (**B**) chimpanzee, (**C**) gorilla, and (**D**) orangutan metaphase spreads.

the human–chimpanzee orthology relationships are uncertain (Fig. 3, upper arrow). In a second case (hOR17-93), the primer pair amplified two different products. Subcloning and sequencing led to the identification of a new OR gene, OR-923, common to human, chimpanzee, and gorilla, that belongs to the same subfamily (1E) as OR-93 and OR-2 (Fig. 3, lower arrow). At present it is not clear whether OR-923 is located in the flanking region of the currently studied OR cluster or elsewhere in the human genome.

Recent Pseudogene Formation

Of the 16 human ORs identified by genomic sequencing, 6 are fixed in the human population as pseudogenes, while one (OR17-31) is polymorphic, having one intact allele and one interrupted allele (Sharon, in preparation). In chimpanzee, four pseudogenes were found to bear mutations that disrupt the open reading frame (Fig. 4). Two of the pseudogenes (OR-23 and OR-25) were rather ancient, with the deleterious mutations shared by humans, chimpanzees, and gorillas. In two other cases (OR-208 and OR-24), both human and chimpanzee have pseudogenes, but these have been generated by independent events in the two species (Figs. 2 and 4). A fifth pseudogene (OR-210) is human-specific, having a premature stop codon, while in chimpanzee it has an intact open reading frame. For the sixth human pseudogene, which has a very high GC content and could not be amplified from any primate genomic DNA, no comparative data are available.

Additional partial analyses of the orthologous OR genes in gorilla, orangutan, and macaque provided crucial timing information on the process of primate OR pseudogene formation. We found that the two pseudogenes (OR-23 and OR-25) shared by the African apes—chimpanzee and gorilla—and human are intact in orangutan and macaque. This led to the conclusion that they were intact at the time of orangutan–African ape divergence (~9 million years ago (Easteal and Herbert, 1997)).

None of the other five primate pseudogenes analyzed is shared by all African apes. These include two of the OR genes (hOR17-7 and OR17-209), which are intact in human and chimpanzee, but are pseudogenes in gorilla, due to one base deletions. In both cases, the gorilla pseudogenes are accompanied by an intact vari-

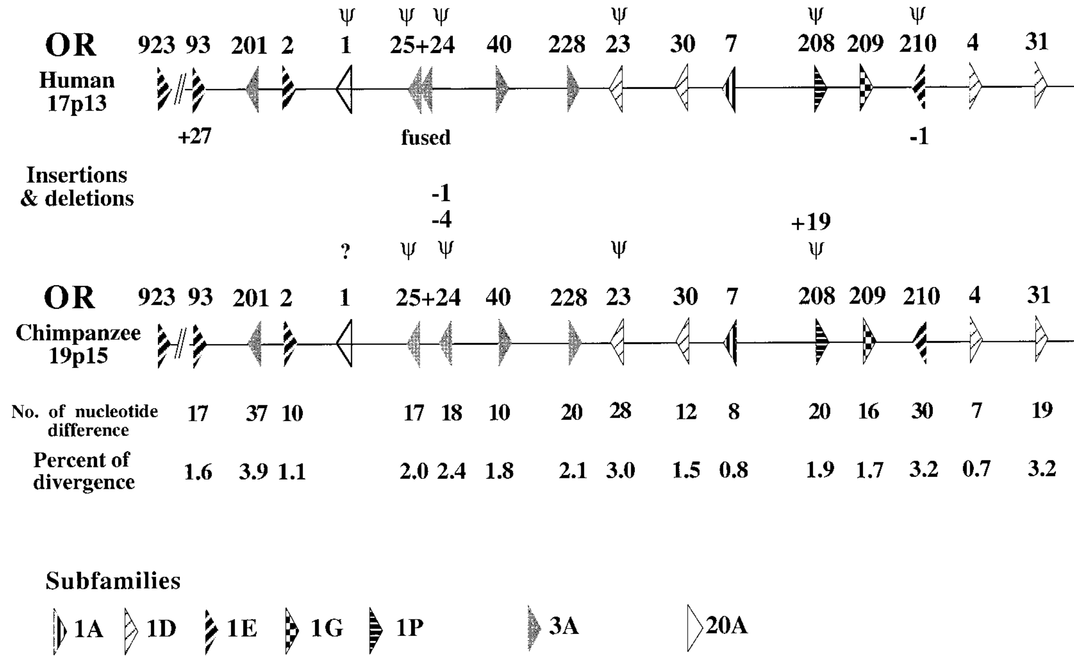

FIG. 2. The human OR cluster on 17p13.3 and the syntenic cluster on chimpanzee chromosome 19p15. The direction of the coding regions was determined only in the human gene cluster (Glusman, in preparation) and is speculated to remain unchanged in the chimpanzee genomic clusters. Pseudogenes are marked by ψ, and the mutations affecting the length of the open reading frames are shown between the two cluster lines. Percentage divergence is simply the percentage of nucleotide differences. The grouping into subfamilies was performed according to Glusman (in preparation). The location of OR-923 in the human genome was not determined, and thus its location is indicated by a broken line.

ant, a potential case of heterozygosity with one of the alleles being a pseudogene. It may thus be inferred that OR-7, OR-24, OR-208, OR-209, and OR-210 were all functional at the time of the orangutan–African ape divergence. In other words, it is likely that as recently as ~9 million years ago all 15 OR genes in the genomic cluster analyzed by us were intact (Fig. 4).

A Human-Specific Deleterious Gene Fusion

A previous analysis of a region within the human OR cluster on chromosome 17 (Glusman *et al.,* 1996) revealed that two of the OR pseudogenes, hOR17-24 and OR17-25, were fused by nonhomologous recombination. Here we used PCR to amplify across the recombination point. The results indicated that this gene pair is fused in all 80 unrelated human individuals tested (data not shown), and thus this dual gene inactivation event seems to be fixed in the human population. We have obtained sequence evidence that these genes are not fused in other primates. This is indicated by our ability to sequence OR-25 in chimpanzee and orangutan and to obtain 138 bp of the region that was lost in hOR17-25 due to the fusion event. The chimpanzee orthologs (cOR19-24 and cOR19-25) still have small deletions that render them pseudogenes. One of these, a 2-bp deletion in OR-25, was found in the human, chimpanzee, bonobo, and gorilla orthologs, but not in the orangutan ortholog, and thus seems to have occurred before the orangutan–African ape divergence, much earlier than the human-specific hOR17-24/OR17-25 fusion event.

Punctuated Evolutionary Changes

Ten of the 16 OR genes in the cluster have intact open reading frames in both human and chimpanzee. Nine have been inherited from the common ancestor only with point mutations, and one (hOR17-93) has a 27-bp in-frame internal duplication (Ben-Arie *et al.,* 1994), which is shown here to be human-specific.

Aiming to study further the differences between intact ORs and pseudogenes, we compared the nucleotide and amino acid sequences of OR orthologous pairs with other intact genes and with noncoding regions (see Materials and Methods for details). While the seemingly intact genes showed an average of 1.9 ± 0.7% single nucleotide difference between respective orthologs, a somewhat higher average nucleotide substitution rate (2.5 ± 0.5%) was seen for the pseudogenes ($P = 0.09$). The values for the intact OR open reading frames were statistically different from those seen for other genes (1.2 ± 0.6%; $P = 0.005$). On the other hand, the degree of single nucleotide substitution for the OR pseudogenes is rather similar to that found in noncoding regions along the cluster and elsewhere in the primate genome (2.5 ± 1.1%; $P = 0.89$). At the protein level, a clearer picture emerged. Intact OR genes show a rate of change significantly higher than that for open reading frames from other genes (3.6 ± 1.6% vs 1.8 ±

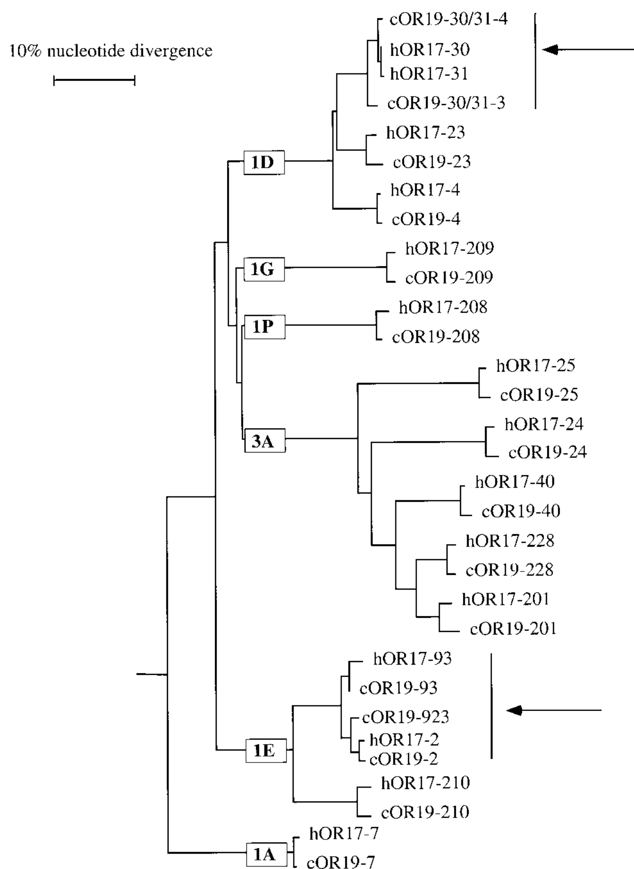

FIG. 3. A phylogenetic dendrogram representing the human and chimpanzee OR orthologs (names beginning with h and c, respectively). The tree was rooted using a fish OR sequence (ICTORDA, Accession No. L09217). Arrows mark cases of ambiguous orthology relationships. Subfamily affiliation is shown within boxes.

1.6%; $P = 0.009$), but significantly lower than for OR pseudogenes ($5.7 \pm 1\%$; $P = 0.017$). Thus, intact OR genes, but not pseudogenes, appear to be subject to a selective pressure, although lower than seen in other gene coding regions.

An interesting comparison may be made among OR genes and pseudogenes in terms of the distribution of the human–chimpanzee differences along the coded protein, summed over all relevant genes. For intact genes, but not for pseudogenes, two of the transmembranal helices (TM2 and TM3) were found to be practically devoid of mutations, i.e., to be highly conserved between the two species. Unexpectedly, normal genes showed a relatively large number of amino acid differences in the second intracellular loop (between TM3 and TM4), containing the conserved MAYDRYVAIC motif and known to be highly conserved in GPCRs (Baldwin, 1994).

Gene Conversion Events

Aiming to identify those sequences that most likely underwent gene conversion events, we carried out a systematic comparison of all available sequences from each subfamily. Candidate regions were those that harbor new mutations shared between members of a paralogous pair. Three of the gene subfamilies within the currently analyzed OR cluster are represented by three or more members per subfamily. In all three cases, there is evidence for gene conversion events, which occurred among subfamily members, and not across subfamily boundaries (Table 1 and Fig. 5A). In family 3A, 11 events may be discerned, most of them sharing a common donor OR coding region, OR-228. Two of the events involve an independent conversion of a specific segment from OR-228 to OR-40, in both human and chimpanzee. Repeated gene conversion, in which at least two instances of genetic exchange could be identified, involves a rather long (610 bp) transfer, with OR-228 as a donor and OR-201 as an acceptor. The first conversion event appears to have occurred prior to the separation of Old World monkeys from apes. This is evidenced by the presence of the same configuration in the chimpanzee, gorilla, and macaque orthologs of OR-201. In a much more recent event, the same region underwent human-specific gene conversion, as evidenced by the 99.5% nucleotide identity shared by the first 610 bases of the hOR17-201 and OR17-228 coding regions (Fig. 5B). In contrast, the ~400 nucleotides at the 3′ end of hOR17-201 shows only 87.1% nucleotide identity to the nearest subfamily member, hOR17-24, and thus seems to represent the original state of hOR17-201 prior to the occurrence of the gene conversion events.

In another subfamily, 1E, four gene conversion events appear to have occurred. In two of these, short segments were donated by as yet unidentified OR genes, while in one case, a chimpanzee specific short gene conversion event could be identified between cOR19-93 and cOR19-210. In the fourth case, OR17-2 served as a donor of a central segment that spans most of the coding region (at least 857 bp), and contains an older short converted segment. The acceptor is OR-923 (Fig. 5A), the newly identified OR gene.

In the case of family 1D, gene conversion events could be identified among most subfamily members (Fig 5A). A special case in this subfamily consists of the two highly similar OR paralogs, hOR17-30 and OR17-31, which share 99.0% nucleotide identity. This suggests a very recent human-specific duplication, since the observed similarity is even higher than that encountered among human–chimpanzee orthologs in the cluster (97.9 ± 0.6). Interestingly, when a pair of oligonucleotides designed to amplify both hOR17-30 and OR17-31 were used to PCR-amplify the expected single chimpanzee ortholog, two highly similar genes could be detected (99.4%). However, since the two human coding regions were found to be much more similar to each other than to any of the chimpanzee homologs (Fig. 3), no obvious orthology relationships could be established. The simplest hypothesis would be, therefore, that a duplication occurred prior to the human–chimpanzee divergence and that the two duplicated genes were then homogenized in human, through a gene conversion event.

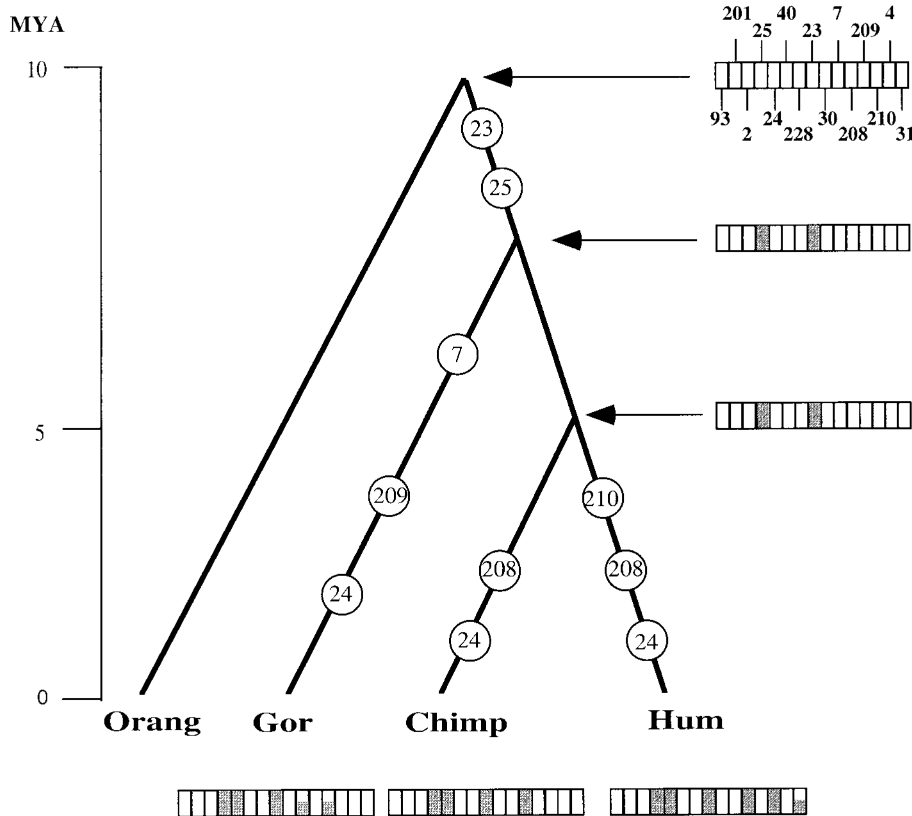

FIG. 4. A primate evolutionary tree representing the inferred formation of different OR pseudogenes. The identifying numbers of genes turned to pseudogenes are marked in circles along the evolutionary branches. The reconstructed configuration of pseudogenes in the OR cluster is shown on the right and bottom striped bars, where pseudogenes within the cluster are shown in gray (half bar is a polymorphism). Estimated divergence times: human–chimpanzee, 5 million years ago (MYA); hominoid–gorilla, 8 MYA; orangutans–African apes, 9 MYA. Hum, human; Chimp, chimpanzee; Gor, gorilla; Orang, orangutan.

To study this question further, we sequenced the gorilla and macaque orthologs. The results showed that each species contains a pair of paralogous genes, which are highly similar (nucleotide identity ranging from 99.2 to 97.7), confirming the notion of an even earlier duplication, followed by recent homogenizing gene conversion events in each of the species.

A summary of all the significant gene conversion traces seen in primate OR genes (Fig. 6) indicates that they belong to two classes: long range (more than 50% of the coding region length) and short range (considerably smaller segments). There is a clear indication that the C-terminal third of the OR coding region has a much lower propensity of taking part in gene conversion. A considerable fraction of these events (43%, Fig. 6) have their beginning and/or end fall within the region of the OR molecule, encompassing the hypervariable helices 3, 4, and 5 and the second extracellular loop, suggesting events that could generate new binding site specificities.

DISCUSSION

While displaying a dynamic process of internal evolutionary change, the orthologous simian OR cluster appears to have been kept intact in its overall layout and have remained as a single genomic copy. Except in cases of paralogs with very high mutual similarity (e.g., OR17-30 and OR17-31), pairwise orthology relationships have been established for the OR genes within the primate clusters. Such local chromosomal conservation was shown, among others, in related human, dog, and mouse OR clusters (Carver *et al.,* 1998; Issel-Tarver and Rine, 1997) and within the globin gene cluster (Slightom *et al.,* 1985). In closely related species such as human and chimpanzee, detailed synteny relationships may extend to much larger genomic segments, as exemplified by the syntenic region on human chromosome 4p (Crouau-Roy *et al.,* 1996). Alternative scenarios could have been frequent interchromosomal duplications as was found in a block of three OR genes (Brand-Arpon *et al.,* 1999; Trask *et al.,* 1998) or where a simian cluster is duplicated or split into two or more segments on different chromosomes, as has been shown for other gene-rich clusters on human chromosomes X and 16 (Eichler *et al.,* 1996). Our results plausibly suggest that most or all the orthologs for genes in the human chromosome 17 OR cluster are included in a chimpanzee syntenic chromosomal cluster on chromosome 19.

Whereas the entire OR cluster seems to have remained undisturbed, at least two of its OR genes have

TABLE 1

Putative Gene Conversion Events among OR Genes

Acceptor gene[a]	Donor gene[a]	Sub family	Minimal region	Maximal region	Length of transfer	Timing of transfer[b]
HumanOR17-40	HumanOR17-228	3A	391–600	337–607	210–271	<5MYA
cOR19-40	cOR19-228	3A	391–530	337–607	140–271	<5MYA
OR-40	Unknown	3A	58–76	58–76	19	>5MYA
OR-40	Unknown	3A	123–127	Unknown	5	>5MYA
HumanOR17-201	HumanOR17-228	3A	1–610	1–610	610	<5MYA
OR-201	OR-228	3A	1–610	1–610	610	>30MYA
cOR19-201	cOR19-228	3A	153–154	126–177	2–52	<5MYA
gOR19-201	gOR19-228	3A	7–7	1–23	1–23	<8MYA
gOR19-201	gOR19-228	3A	525–530	506–561	6–56	<8MYA
OR-201	Unknown	3A	82–87	Unknown	6	>30MYA
mOR19-228	mOR19-201	3A	23–100	8–123	78–116	<30MYA
OR-2	Unknown	1E	219–246	135–291	28–157	>8MYA
OR-93	Unknown	1E	313–330	292–330	18–39	>5MYA
cOR19-93	cOR19-210	1E	260–265	226–281	6–56	<5MYA
OR-923	OR-2	1E	1–857	1–904	857–904	>8MYA
OR-4	OR-23	1D	1–358	1–358	358	>8MYA
OR-23	OR-4	1D	119–119	98–160	1–63	<8MYA
cOR19-23	OR-30	1D	97–98	44–130	2–87	<5MYA
mOR19-23	mOR19-31	1D	363–417	358–423	55–66	<30MYA
mOR19-23	mOR19-31	1D	513–556	496–574	44–79	<30MYA
OR17-30	OR17-31	1D	1–939	1–939	939	<30MYA

[a] The acceptor or donor gene is indicated by the species name (c, chimpanzee; g, gorilla, and m, macaque), chromosome number, and gene number. The species and chromosome number are indicated only in species-unique events.

[b] Estimated divergence times: human–chimpanzee, 5 MYA; hominoid–gorilla, 8 MYA; Old World monkeys–primates, 30 MYA.

been reported to have close paralogs on other human chromosomes. These include two paralogs of hOR17-23 on human chromosome 11 (Buettner *et al.,* 1998) and two paralogs of hOR17-2 on chromosomes 5 and 13 (Rouquier *et al.,* 1998b). Such paralogs were probably generated by a rather recent interchromosomal duplication. The gene OR-923 reported here may be another such case, if future analyses locate it in a cluster different from that of human chromosome 17.

A central aspect of the evolution of the olfactory subgenome is the level and type of selection acting on OR genes. Since a typical "general" odorant is likely to be sensed by several receptors with different affinities (Lancet *et al.,* 1993b), only a low level of selection is expected against the loss of single OR genes. Still, some selection does appear to operate, as indicated by the discernibly lower rate of variation among intact OR gene orthologs, than among pseudogenic orthologs.

Two of the most interspecifically conserved OR genes are the intact coding regions of OR-2 and OR-4. These coding regions have been found to be intact in all primate species studied. Interestingly, hOR17-2 was found to be expressed in the human male germline. The OR17-2 gene product therefore potentially interacts with specific sperm-related chemical signals (Parmentier *et al.,* 1992), which may explain its conservation among species. The higher degree of conservation in these OR genes could also be ascribed to an involvement with highly specific olfactory cues, typified by behaviorally important odorants such as pheromones.

The availability of an entire group of full-length OR coding region sequences permitted us to conduct a significant analysis of pseudogene incidence. Six of the 16 genes (38%) on the human chromosome 17 OR cluster are found to be pseudogenes. The other 10 genes carry no clear signs of sequence deterioration within the coding regions. Some of them, however, could still be functionally defective. The lower rate of differences for the putative intact genes provides, nevertheless, an independent indication that these genes may be functional. For OR17-2, -40, -93, -201, and -228, there is, in addition, corroborative experimental evidence showing that these are transcribed in the olfactory epithelium (Ben-Arie *et al.,* 1994; Crowe *et al.,* 1996; A. Sosinsky and D. Lancet, unpublished data) and in sperm (Parmentier *et al.,* 1992).

In the case of the OR clusters on human chromosome 11, a much higher overall incidence of pseudogenes was observed (nearly 80%), and some of the OR clusters were shown to contain only pseudogenes (Buettner *et al.,* 1998). It is reasonable to speculate that these gene groups may have been inactivated as entire clusters, perhaps because of a defective expression control mechanism. In contrast, it is likely that for the cluster studied here such shared control elements are still intact.

The high incidence of OR pseudogenes in human

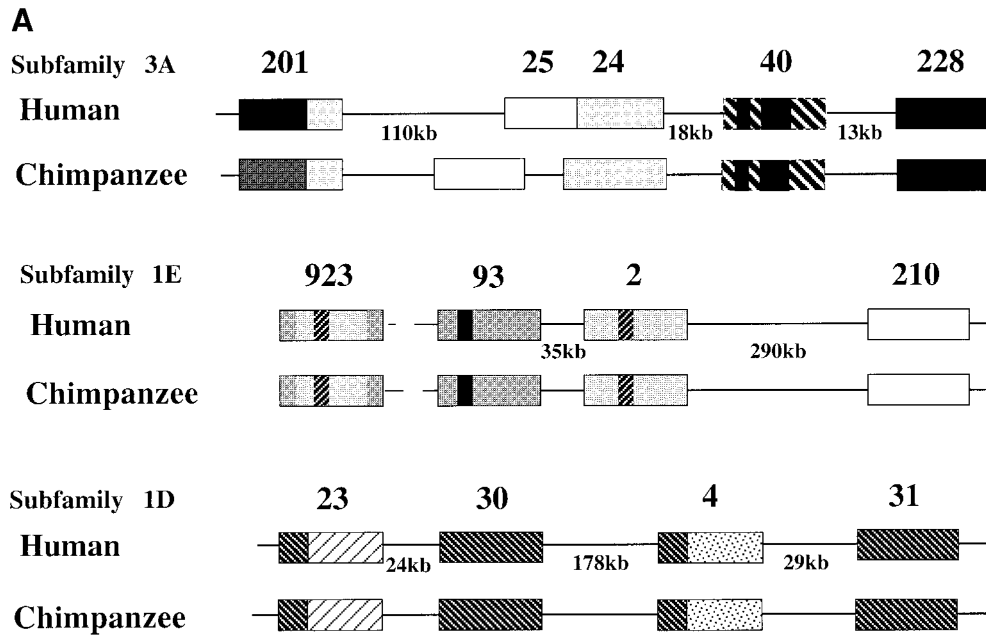

FIG. 5. (**A**) Gene conversion events between clustered OR genes within subfamilies 3A, 1E, and 1D. Physical distances in kilobases separating the OR coding regions along the cluster are indicated, as derived from our unpublished sequencing (Glusman, in preparation). Pattern identity indicates a high level of sequence identity. (**B**) Repeated gene conversion event between OR genes. (**I**) A nucleotide identity plot between humanOR17-201 and humanOR17-228. (**II**) A nucleotide alignment encompassing the gene conversion region. (**III**) Phylogeny analysis for the two parts of humanOR17-201. Trees were rooted using humanOR17-31. Phylogeny was assessed using the parsimony analysis of the PAUP (Phylogenetic Analysis Using Parsimony) program available from D. L. Swofford.

compared to other species such as mouse (<2% by our unpublished GenBank search) and zebrafish (about 10% (Barth *et al.,* 1997)) suggests a time-dependent evolutionary process of OR gene loss. Overall, it is possible that more than half of all human OR genes are inactive (Rouquier *et al.,* 1998b). A better understanding of this phenomenon requires information on the timing of the process that has turned genes into pseudogenes. We infer that all OR genes on the cluster we have studied were intact before the divergence of orangutans from African apes, estimated to be ~9 million years ago (Easteal and Herbert, 1997). This drastic change in the number of functional genes over a short period of evolution is noteworthy and is consistent with a trend indicated in a recent study of a single human OR pseudogene (Rouquier *et al.,* 1998a). This change probably stems from the considerable reduction of the selection level acting on OR genes during primate evolution, in accordance with the decreasing olfactory dependence of this “micro-osmatic” order. Interestingly, a large array of class II olfactory receptor genes in the dolphin were found to be nonfunctional pseudogenes (Freitag *et al.,* 1998). Class II receptors are speculated to recognize volatile odorants and thus might be turned into pseudogenes during evolution, in aquatic mammals, in a mechanism similar to the evolution of OR genes in primates. Based on the above-mentioned results, we assume that gradual loss of functional OR genes is not a human-specific phenomenon and might take place whenever there are genes encoding receptors with a negligible contribution to chemosensory faculties.

It is possible that the main selective force in the evolution of OR genes is that which favors the conservation of the overall size of the functional repertoire. If this size decreased below a specific threshold, a deficit could ensue in the animal’s overall olfactory faculties, including a potential decrement in the average affinity toward general odorants (Lancet *et al.,* 1993b).

We describe here a large incidence of gene conversion events in the OR cluster. In all cases but five, the conversion donor was identified within the same OR gene cluster. Similarly, intrachromosomal gene conversion was previously reported to be much more frequent than interchromosomal gene conversion (Liao *et al.,* 1997). The availability of the sequence of the monkey orthologs of hOR17-228 and OR17-201 shed light on an interesting phenomenon: two gene conversion events that occurred in the exact same position. This may indicate that, beyond straightforward sequence similarity, there are additional signals that regulate gene conversion among OR genes, perhaps recombination “hot spots” (Amor, 1988). A similar repeated gene conversion phenomenon was found in other GPCR genes, the color vision receptor genes, in which intron 4 of the green and red receptor genes was converted twice, with the last event probably being human-specific (Zhou and Li, 1996).

Gene conversion between OR genes had been previously postulated (Ben-Arie *et al.,* 1994; Buck and Axel, 1991), but we provide here the first direct evidence that this mechanism is widespread among OR genes. Gene conversion may be important in generating the diversity of OR genes (see below). Notably, in three cases,

B

I

% Difference

bp in coding region

II

```
                                610
                                 |
hOR17-228 AACTCAATGAGCTGCTGCTCTTTGCTGTGGGTTTCATCATGGCAGGCACACCTTTGGTTCTCATCATCACTG
hOR17-201 ------------------------TA-CA-CAGC-T-------T-TGG----C-----CT------GTGTGT
cOR19-228 ----------------------------------------------------------------------CT
cOR19-201 ------------------------TA-CA-CAGC-G-------T-TGG-G--C-----CT------G-GTGT
```

III

1 610 948

a b

hOR17-228a

hOR17-201a

cOR19-228a

cOR19-201a

hOR17-228b

cOR19-228b

cOR19-201b

hOR17-201b

FIG. 5—*Continued*

OR-30/31, OR-201, and OR-923, the observed gene conversion events appear to involve large segments of genes, resulting in the replacement of most of the open reading frame by a paralogous sequence. Such "homogenization" events would actually result in an opposite effect, namely, the decreased diversity of the OR repertoire, because they generate two largely identical paralogs (Liao *et al.,* 1997; Schlotterer and Tautz, 1994). A similar homogenization process could also be important in pseudogene "rescue," whereby a pseudogene may be corrected by an intact gene, as was found in other gene families (Benevolenskaya *et al.,* 1997).

In contrast to the homogenizing conversion events, we report here 10 other cases of gene conversion events, involving smaller segments of the coding regions. Such instances of short gene conversion were found in other gene families as well (Wines *et al.,* 1991). These events may be essential in OR gene diversification. OR proteins are known to exist in several hundred variants, whose diversity is germline-coded. This situation is different from that encountered in immune receptors, whereby diversity is generated via somatic DNA rearrangements and mutations, as well as by combinatorial association at the protein level (Lieber, 1996).

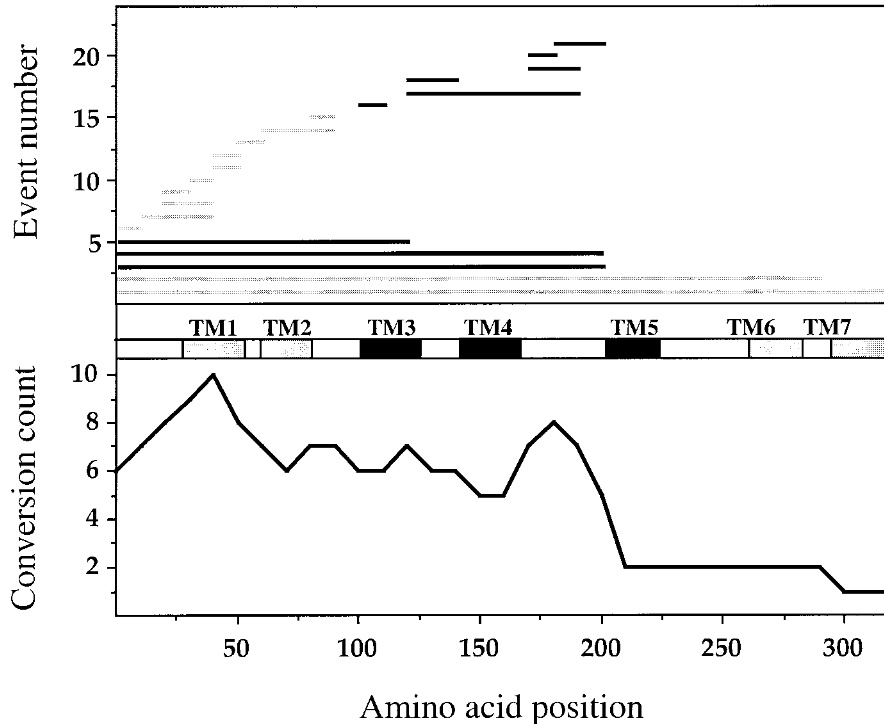

FIG. 6. A schematic representation of the location of gene conversion events in the chromosome 17 OR cluster. (**Top**) Bars indicate the boundaries of the inferred gene conversion. Darker color indicates conversion bars whose beginning and/or end is located within the region that includes the hypervariable transmembrane (TM) helices 3, 4, and 5. (**Bottom**) The cumulative count of gene conversion beginnings and ends as a function of the amino acid position.

A straightforward scenario would anticipate that OR diversity arises by gene duplication (Barth *et al.,* 1997; Ben-Arie *et al.,* 1994; Issel-Tarver and Rine, 1997), followed by the gradual accumulation of point mutations in the duplicated variant. The rather frequent occurrence of gene conversion reported here suggests that an additional germline mechanism for the generation of diversity may be at work in the olfactory receptor repertoire, similar to that found in other multigene families (Rodakis *et al.,* 1984; Wines *et al.,* 1991). It may thus be argued that many of the OR coding regions actually comprise mosaics of protein segments generated by repeated gene conversion events. If true, this would considerably accelerate the evolution of new OR variants and lead to a combinatorial process that brings together segments from several OR gene paralogs, to generate new odorant binding site configurations. That nearly a half of all the observed gene conversion events result in combinatorial joining within the putative OR binding site encompassing the hypervariable helices 3, 4, and 5 provides credence to this hypothesis.

Obviously, there is a subtle interplay between gene conversion, which may increase diversity, and pseudogene formation, which tends to decrease the multiplicity of the functional repertoire. Our results do not address the relative strength of these two effects during primate evolution. More extended comparative studies that might encompass complete OR gene repertoires in several species would likely shed light on this intriguing question.

ACKNOWLEDGMENTS

This research was supported by grants to D.L. from the U.S. National Institutes of Health (DC00305), a Wolfson Research Award of the Israel Academy of Science, the Israel Ministry of Science (in collaboration with the German BMBF and an Infrastructure grant), the German–Israeli Foundation for Scientific Research and Development, and the Krupp Foundation. The fluorescence *in situ* hybridization work was supported by Research Grant Ha1374/5-1 to T.H. from the Deutsche Forschungsgemeinschaft. We thank Dr. Yigal Horvitz for his kind gifts of blood, Dr. Cedric Campbell Muir and Dr. Keneth Kidd for the DNA samples, Dr. Edna Ben-Asher for the PAC clones used for FISH, Elena Robinson and Hilla German-Shashoua for technical assistance, and Moshe Eisenberg, Martin Kupiec, and Avraham Levi for fruitful discussions. D. L. holds the Ralph and Lois Silver chair in Nerogenomics.

REFERENCES

Amor, M., Parker, K. L., Globerman, H., New, M. I., and White, P. C. (1988). Mutation in the CYP21B gene (Ile-172—-Asn) causes steroid 21-hydroxylase deficiency. *Proc. Natl. Acad. Sci. USA* **85:** 1600–1604.

Asai, H., Kasai, H., Matsuda, Y., Yamazaki, N., Nagawa, F., Sakano, H., and Tsuboi, A. (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.

Baldwin, J. M. (1994). Structure and function of receptors coupled to G proteins. *Curr. Opin. Cell. Biol.* **6:** 180–190.

Barth, A. L., Dugas, J. C., and Ngai, J. (1997). Noncoordinate expression of odorant receptor genes tightly linked in the zebrafish genome. *Neuron* **19:** 359–369.

Ben-Arie, N., Lancet, D., Taylor, C., Khen, M., Walker, N., Ledbetter, D. H., Carrozzo, R., Patel, K., Sheer, D., Lehrach, H., and North, M. A. (1994). Olfactory receptor gene cluster on human chromosome 17: Possible duplication of an ancestral receptor repertoire. *Hum. Mol. Genet.* **3:** 229–235.

Benevolenskaya, E. V., Kogan, G. L., Tulin, A. V., Philipp, D., and Gvozdev, V. A. (1997). Segmented gene conversion as a mechanism of correction of 18S rRNA pseudogene located outside of rDNA cluster in *D. melanogaster. J. Mol. Evol.* **44:** 646–651.

Brand-Arpon, V., Rouquier, S., Massa, H., de Jong, P. J., Ferraz, C., Ioannou, P. A., Demaille, J. G., Trask, B. J., and Giorgi, D. (1999). A genomic region encompassing a cluster of olfactory receptor genes and a myosin light chain kinase (MYLK) gene is duplicated on human chromosome regions 3q13–q21 and 3p13. *Genomics* **56:** 98–110.

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

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

Carver, E. A., Issel-Tarver, L., Rine, J., Olsen, A. S., and Stubbs, L. (1998). Location of mouse and human genes corresponding to conserved canine olfactory receptor gene subfamilies. *Mamm. Genome* **9:** 349–354.

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

Crouau-Roy, B., Service, S., Slatkin, M., and Freimer, N. (1996). A fine-scale comparison of the human and chimpanzee genomes: Linkage, linkage disequilibrium and sequence analysis. *Hum. Mol. Genet.* **5:** 1131–1137.

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

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

Duret, L., Mouchiroud, D., and Gouy, M. (1994). HOVERGEN: A database of homologous vertebrate genes. *Nucleic Acids Res.* **22:** 2360–2365.

Easteal, S., and Herbert, G. (1997). Molecular evidence from the nuclear genome for the time frame of human evolution. *J. Mol. Evol.* **44:** S121–132.

Eichler, E. E., Lu, F., Shen, Y., Antonacci, R., Jurecic, V., Doggett, N. A., Moyzis, R. K., Baldini, A., Gibbs, R. A., and Nelson, D. L. (1996). Duplication of a gene-rich cluster between 16p11.1 and Xq28: A novel pericentromeric-directed mechanism for paralogous genome evolution. *Hum. Mol. Genet.* **5:** 899–912.

Freitag, J., Krieger, J., Strotmann, J., and Breer, H. (1995). Two classes of olfactory receptors in *Xenopus laevis. Neuron* **15:** 1383–1392.

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

Glusman, G., Clifton, S., Roe, R., and Lancet, D. (1996). Sequence analysis in the olfactory receptor gene cluster on human chromosome 17: Recombinatorial events affecting receptor diversity. *Genomics* **37:** 147–160.

Heim, M. H., and Meyer, U. A. (1992). Evolution of a highly polymorphic human cytochrome P450 gene cluster: CYP2D6. *Genomics* **14:** 49–58.

Higgins, D. G., Thompson, J. D., and Gibson, T. J. (1996). Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* **266**: 383–402.

Higgs, D. R., Vickers, M. A., Wilkie, A. O., Pretorius, I. M., Jarman, A. P., and Weatherall, D. J. (1989). A review of the molecular genetics of the human alpha-globin gene cluster. *Blood* **73:** 1081–1104.

Ioannou, P. A., Amemiya, C. T., Garnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A., and de Jong, P. J. (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat. Genet.* **6:** 84–89.

Issel-Tarver, L., and Rine, J. (1997). The evolution of mammalian olfactory receptor genes. *Genetics* **145:** 185–195.

Jauch, A., Wienberg, J., Stanyon, R., Arnold, N., Tofanelli, S., Ishida, T., and Cremer, T. (1992). Reconstruction of genomic rearrangements in great apes and gibbons by chromosome painting. *Proc. Natl. Acad. Sci. USA* **89:** 8611–8615.

Lancet, D. (1991). Olfaction. The strong scent of success. *Nature* **351**(6324): 275–276.

Lancet, D., and Ben-Arie, N. (1993). Olfactory receptors. *Curr. Biol.* **3:** 668–674.

Lancet, D., Gross-Isseroff, R., Margalit, T., Seidmann, E., and Ben-Arie, N. (1993a). Olfaction: From signal transduction and termination to human genome mapping. *Chem. Senses* **18:** 217–225.

Lancet, D., and Pace, U. (1987). The molecular basis of odor recognition. *Trends Biochem. Sci.* **12:** 63–66.

Lancet, D., Sadovsky, E., and Seidemann, E. (1993b). Probability model for molecular recognition in biological receptor repertoires: Significance to the olfactory system. *Proc. Natl. Acad. Sci. USA* **90:** 3715–3719.

Liao, D., Pavelitz, T., Kidd, J. R., Kidd, K. K., and Weiner, A. M. (1997). Concerted evolution of the tandemly repeated genes encoding human U2 snRNA (the RNU2 locus) involves rapid intrachromosomal homogenization and rare interchromosomal gene conversion. *Embo J.* **16:** 588–598.

Lieber, M. (1996). Immunoglobulin diversity: Rearranging by cutting and repairing. *Curr. Biol.* **6:** 134–136.

McGuire, G., and Wright, F. (1998). TOPAL: Recombination detection in DNA and protein sequences. *Bioinformatics* **14:** 219–220.

Nef, S., Allaman, I., Fiumelli, H., De Castro, E., and Nef, P. (1996). Olfaction in birds: Differential embryonic expression of nine putative odorant receptor genes in the avian olfactory system. *Mech. Dev.* **55:** 65–77.

Ngai, J., Chess, A., Dowling, M. M., Necles, N., Macagno, E. R., and Axel, R. (1993). Coding of olfactory information: Topography of odorant receptor expression in the catfish olfactory epithelium. *Cell* **72:** 667–680.

Olsen, A. S., Reed, R. R., Brandriff, B. F., Gordon, L. A., Hoffman, S. M. G., Schrader, K., Tsujimoto, S., Carrano, A. V., and Mohrenweiser, H. (1993). Assembly and analysis of a cosmid contig map of the olfactory receptor gene family on chromosome 19. *Am. J. Hum. Gen.* **51**(Suppl.): A243 [Abstract 957]

Paris Conference (1971, supplement 1975). Standardization in human cytogenetics. *Cytogenet. Cell. Genet.* **15:** 203–238.

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

Reed, R. R. (1990). How does the nose know? *Cell* **60:** 1–2.

Ressler, K. J., Sullivan, S. L., and Buck, L. B. (1994). A molecular dissection of spatial patterning in the olfactory system. *Curr. Opin. Neurobiol.* **4:** 588–596.

Rodakis, G. C., Lecanidou, R., and Eickbush, T. H. (1984). Diversity in a chorion multigene family created by tandem duplications and a putative gene-conversion event. *J. Mol. Evol.* **20:** 265–273.

Rouquier, S., Friedman, C., Delettre, C., van den Engh, G., Blancher, A., Crouau-Roy, B., Trask, B. J., and Giorgi, D. (1998a). A gene recently inactivated in human defines a new olfactory receptor family in mammals. Hum. Mol. Genet. **7:** 1337–1345.

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

Schlotterer, C., and Tautz, D. (1994). Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. *Curr. Biol.* **4:** 777–783.

Shpaer, E. G. (1997). GeneAssist. Smith-Waterman and other database similarity searches and identification of motifs. *Methods Mol. Biol.* **70:** 173–187.

Slightom, J. L., Chang, L. Y., Koop, B. F., and Goodman, M. (1985). Chimpanzee fetal G gamma and A gamma globin gene nucleotide sequences provide further evidence of gene conversions in hominine evolution. *Mol. Biol. Evol.* **2:** 370–389.

Stanyon, R., Wienberg, J., Romagno, D., Bigoni, F., Jauch, A., and Cremer, T. (1992). Molecular and classical cytogenetic analyses demonstrate an apomorphic reciprocal chromosomal translocation in *Gorilla gorilla. Am. J. Phys. Anthropol.* **88:** 245–250.

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

Trask, B. J., Friedman, C., Martin-Gallardo, A., Rowen, L., Akinbami, C., Blankenship, J., Collins, C., Giorgi, D., Iadonato, S., Johnson, F., Kuo, W. L., Massa, H., Morrish, T., Naylor, S., Nguyen, O. T., Rouquier, S., Smith, T., Wong, D. J., Youngblom, J., and van den Engh, G. (1998). Members of the olfactory receptor gene family are contained in large blocks of DNA duplicated polymorphically near the ends of human chromosomes. *Hum. Mol. Genet.* **7:** 13–26.

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

Ward, D. C., Boyle, A., and Haaf, T. (1995). Fluorescence in situ hybridization techniques. Metaphase chromosomes, interphase nuclei, and extended chromatin fibers. *In* “Human Chromosomes: Principles and Techniques” (R. S. Verma and A. Babu, Eds.), pp. 184–192, MacGraw-Hill, New York.

Wines, D. R., Brady, J. M., Southard, E. M., and MacDonald, R. J. (1991). Evolution of the rat kallikrein gene family: Gene conversion leads to functional diversity. *J. Mol. Evol.* **32:** 476–492.

Zhou, Y. H., and Li, W. H. (1996). Gene conversion and natural selection in the evolution of X-linked color vision genes in higher primates. *Mol. Biol. Evol.* **13:** 780–783.