Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 346 (2005) 23–28
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Evolutionary changes of the number of olfactory receptor genes in the human and mouse lineages

Yoshihito Niimura*, Masatoshi Nei

Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University, 328 Mueller Laboratory, University Park, PA 16802, USA

Received 1 July 2004; received in revised form 17 September 2004; accepted 28 September 2004
Available online 31 December 2004
Received by T. Gojobori

Abstract

The numbers of functional olfactory receptor (OR) genes are quite variable among mammalian species. Previously we have reported that humans have 388 functional OR genes and 414 pseudogenes, while mice have 1037 functional genes and 354 pseudogenes. These observations suggest either that humans lost many functional OR genes after the human–mouse divergence (HMD) or that mice gained many functional genes. To distinguish between these two hypotheses, we devised a new method of inferring the number of functional OR genes in the most recent common ancestor (MRCA) of humans and mice. An application of this method suggested that the MRCA had ~750 functional OR genes and that mice acquired ~350 new OR genes after the HMD whereas ~430 OR genes in the MRCA have become pseudogenes or eliminated in the human lineage. Therefore, the two evolutionary hypotheses mentioned above are not mutually exclusive and both are nearly equally responsible for the difference in the number of OR genes between humans and mice.

Keywords: Gene expansion; Gene gain; Gene loss; Most recent common ancestor; Pseudogene

1. Introduction

Olfactory receptor (OR) genes constitute the largest known multigene family in mammalian genomes (Niimura and Nei, 2005 and the references therein). Identification of all the OR genes from the human and mouse genomes (Niimura and Nei, 2003, 2005) revealed that mice have ~2.7 times as many functional OR genes as humans and have a slightly smaller number of pseudogenes than humans. The larger number of functional OR genes and the smaller fraction of pseudogenes in mice (25%) than in humans (52%) has been explained by the assumption that many OR genes were lost in the human lineage after the divergence of the two species. Rouquier et al. (2000) randomly sequenced OR genes from 10 primate species and mice and found that the fraction of pseudogenes increased in the order rodents, Old World monkeys, apes, and humans. From this observation and others, they hypothesized that the reduction of the sense of smell could correlate with the loss of functional OR genes. In this study, however, they incorrectly reported the fraction of pseudogenes in mice and humans to be zero and ~70%, respectively. Gilad et al. (2003) compared 50 human OR genes with their putative orthologs in other primates and showed that humans have accumulated dysfunctional mutations in OR coding regions much faster than other primates. This study supports the idea of human-specific loss of OR genes, although the sample size was limited. Recently, Gilad et al. (2004) proposed that the loss of OR

Abbreviations: OR, olfactory receptor; HMD, human–mouse divergence; MRCA, most recent common ancestor; MY, million years; MYA, million years ago.

* Corresponding author. Present address: Department of Bioinformatics, Medical Research Institute, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan. Tel.: +81 3 5803 5840; fax: +81 3 5803 0247.
E-mail address: niimura@bioinfo.tmd.ac.jp (Y. Niimura).

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.09.027

genes coincided with the acquisition of full trichromatic vision in primates.

However, the larger number of functional OR genes in mice than in humans can be explained by another evolutionary scenario: the mouse lineage gained more new functional genes by repeated gene duplication. In fact, this explanation seems quite likely, because the total number of functional and nonfunctional OR genes is larger in mice than in humans. However, to study these possible evolutionary changes of OR genes, it is necessary to estimate the number of OR genes in the most recent common ancestor (MRCA) between humans and mice. If we know this number, we would be able to tell which of the above two hypotheses is correct. The purpose of this paper is to study this problem. For this purpose, we invented a new statistical method to estimate the number of genes in the MRCA and inferred the evolutionary changes of OR gene families in the human and mouse lineages.

2. Materials and methods

2.1. Data

We used 388 functional and 414 nonfunctional OR genes from humans and 1037 functional and 354 nonfunctional OR genes from mice that were identified in our previous works (Niimura and Nei, 2003, 2005).

2.2. Number of genes that have been generated in the human and mouse lineages

Fig. 1A shows a schematic phylogenetic tree for 10 mouse and 4 human functional genes. In this tree, gene duplication occurred four times (filled circles) in the mouse lineage after the human–mouse divergence (HMD), and two times (open circles) in the human lineage. (Here we assume that one gene duplication generates one new gene.) The increments of the number of genes by duplication can easily be obtained from the diagrams in Fig. 1B and C (bottom). The tree in Fig. 1B (top) was obtained by excluding all the human genes in Fig. 1A. The y-axis in the bottom diagram of this figure shows the increase of the number of mouse genes by gene duplication. (Note that here we do not consider pseudogenes or the genes eliminated from the genome. We are considering only the genes of which descendents are functional. Therefore, the numbers of genes in the mouse and human lineages are different even before the HMD as shown in Fig. 2A.) Therefore, the number of genes generated in the mouse lineage after the HMD can be obtained by subtracting the number of genes at the time of the HMD from the present number. In the example of Fig. 1B, this number is 10−6=4. Similarly, the number of genes generated in the human lineage after the HMD is equal to 2 (=4−2) from Fig. 1C.

The actual numbers of increase of OR genes were computed by constructing a linearized tree for all the mouse and human functional OR genes. In the preceding paper, we have constructed a neighbor-joining (NJ) tree for human and mouse OR genes (Fig. 3 in Niimura and Nei, 2005), showing that these genes are clearly separated into class I and class II genes. A linearized tree was constructed from this NJ tree by using the algorithm developed by Takezaki et al. (1995). Outgroup sequences are necessary to estimate branch lengths in the linearized tree. We therefore used all the class I genes as the outgroup for estimating the branch lengths of class II genes and used all the class II genes as the outgroup for estimating the branch lengths of class I genes. In the case that an interior branch length was negative, it was assumed to be zero. Using the linearized tree, the

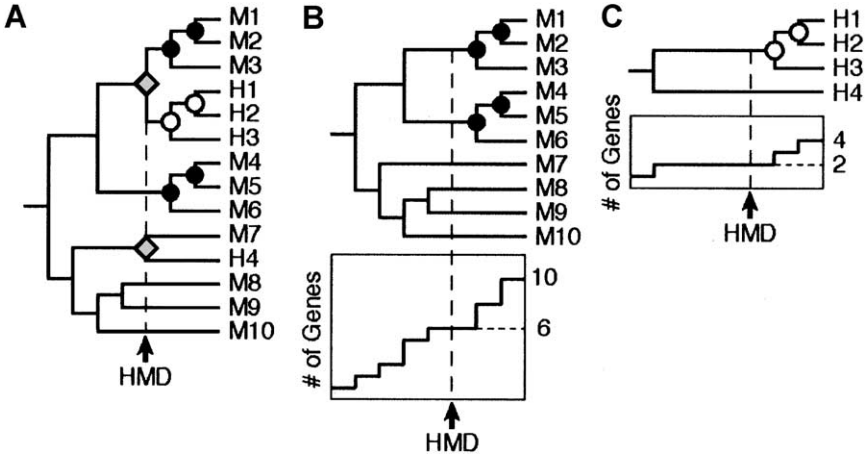

Fig. 1. Increase of the number of OR genes in the mouse and human lineages. (A) Schematic phylogenetic tree containing 10 mouse genes (M1, M2,. . ., M10) and 4 human genes (H1, H2,. . ., H4). A diamond shows the divergence time between mice and humans. Filled and open circles indicate gene duplication events that occurred after the HMD in the mouse and human lineages, respectively. (B) (Top) Schematic phylogenetic tree obtained by excluding all the human genes in (A). (Bottom) Diagram indicating the increase of genes in the mouse lineage (see Section 2.2). (C) (Top) Phylogenetic tree obtained by excluding all the mouse genes in (A). (Bottom) Diagram indicating the increase of genes in the human lineage.

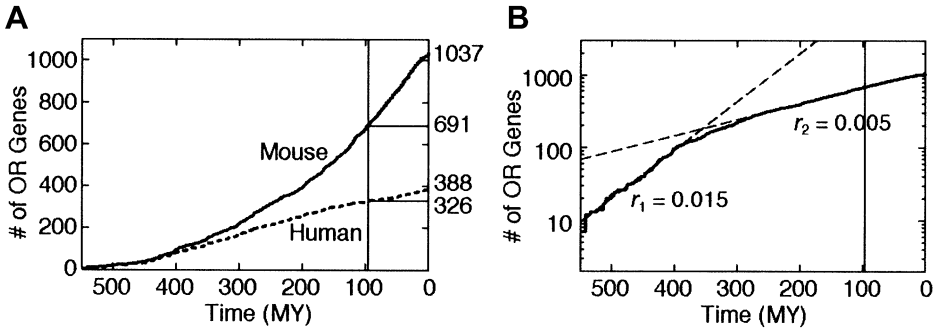

Fig. 2. (A) Diagram indicating the increase of OR genes in the mouse (solid line) and the human (dashed line) lineages. A vertical line indicates the divergence time between mice and humans (96 MY; Nei et al., 2001). (B) Log plot indicating the increase of OR genes in the mouse lineage. The numbers of OR genes for the periods of 550 to 400 MYA and of 300 MYA to the present time were separately fitted to x_0e^{rt} by the least square method. Two dashed lines indicate regression lines for these time periods. A vertical line indicates the divergence time between mice and humans.

numbers of genes that have been generated in the mouse and human lineages were estimated by the method mentioned above (Fig. 1).

3. Results

3.1. Increase of the number of OR genes in the mouse and human lineages

We first studied the numbers of OR genes that have been newly generated by gene duplication in the mouse and human lineages after divergence of the two species. Fig. 2A shows a diagram indicating the increases of OR genes in the human and mouse lineages (see Section 2.2). The time scale in Fig. 2A was obtained by using a molecular time estimate of about 96 million years ago (MYA) (Nei et al., 2001) for the human and mouse divergence as the calibration point. In the preceding paper (Niimura and Nei, 2005), we have identified 205 MRCA genes that generated orthologous human and mouse OR genes. We computed the branch length between the MRCA and the present-day sequences in terms of the number of amino acid substitutions per site and obtained the average branch length (b) for all MRCA genes, which was b=0.073 (Supplementary data, Fig. S1). This value corresponds to the human–mouse divergence time of 96 million years (MY), and therefore the time scale for the diagram in Fig. 2A could be obtained. Note that the divergence time (96 MYA) was used only for obtaining the time scale in Fig. 2A and B and therefore does not affect the estimate of the number of OR genes in the MRCA.

The numbers of OR genes counted at the time of the HMD (shown by the vertical line in Fig. 2A) were 691 for the mouse lineage and 326 for the human lineage. Therefore, the number of new OR genes generated after the HMD is estimated to be 346 (=1037−691) for the mouse and 62 (=388−326) for the human. These results indicate that the increment of the number of OR genes is much larger in the mouse lineage than in the human. These are very rough estimates, because the constancy of evolutionary rate is not validated. However, these estimates appear to be reasonably accurate, because similar estimates are obtained by the following independent method.

In the preceding paper (Niimura and Nei, 2005), we have identified 205 orthologous gene pairs that contained 306 mouse genes and 241 human genes. This means that 205 MRCA genes have increased to 306 functional genes in the mouse lineage and to 241 functional genes in the human. It is therefore possible to compute the number of MRCA genes (N_M) that generated 1037 mouse functional genes and the number of MRCA genes (N_H) that generated 388 human functional genes, if we assume that all genes multiplied at the same rate. In the mouse lineage, we obtain the equation N_M×(306/205)=1037. Therefore, N_M is estimated to be 695. Similarly, N_H can be estimated by N_H×(241/205)=388. It becomes 330. These estimates are very close to those obtained above.

Using the diagram in Fig. 2A, it is possible to estimate the rate of increase of the number of OR genes. If the rate of increase is constant over time, the number of genes at time t is given by $x(t)=x_0e^{rt}$, where r is the rate of increase of the number of genes per gene per MY and x_0 is the number of genes at t=0 (Nei, 1969). We estimated r using mouse functional OR genes. Fig. 2B indicates that r changed considerably during the time period of 300–400 MYA, but it was nearly constant before and after the time period. Therefore, we calculated r for the time period from 550 to 400 MYA (r_1) and for the period from 300 MYA to the present time (r_2) separately. We then obtained r_1=0.015 and r_2=0.005.

3.2. Loss of OR genes in the mouse and human lineages

Let us now estimate the number of MRCA genes that have become pseudogenes in the human and mouse lineages. In the preceding paper, we assigned each OR pseudogene to one of the phylogenetic clades by constructing a tree composed of one pseudogene and 1425 human and mouse functional OR genes (Fig. 4 in Niimura and Nei, 2005). In this process, we identified three different cases of clustering of a pseudogene with its closest functional gene

or genes. The upper part of Fig. 3A shows these three cases (types I, II, and III trees) for a human pseudogene (Hψ). In the case of type I trees, Hψ is closest to human functional gene or genes. This type of tree is expected to occur if the MRCA gene generates both functional genes and pseudogenes in the human lineage (Fig. 3B). In our data analysis, we found 193 pseudogenes showing type I trees.

In the case of type II trees, Hψ is closest to mouse functional gene or genes, and this represents the case where the MRCA gene remained functional in the mouse lineage but became inactivated in the human lineage (Fig. 3C). In practice, the same MRCA gene may generate more than one pseudogene. Therefore, when all human pseudogenes for type II trees are considered, the number of MRCA genes involved (one in Fig. 3C) may be smaller than the number of pseudogenes generated (three in Fig. 3C). The number of MRCA genes involved can be estimated by counting the number of different mouse functional gene or genes that are closest to the human pseudogenes showing type II trees. In our data analysis, we found 158 pseudogenes for type II trees and 130 MRCA genes involved.

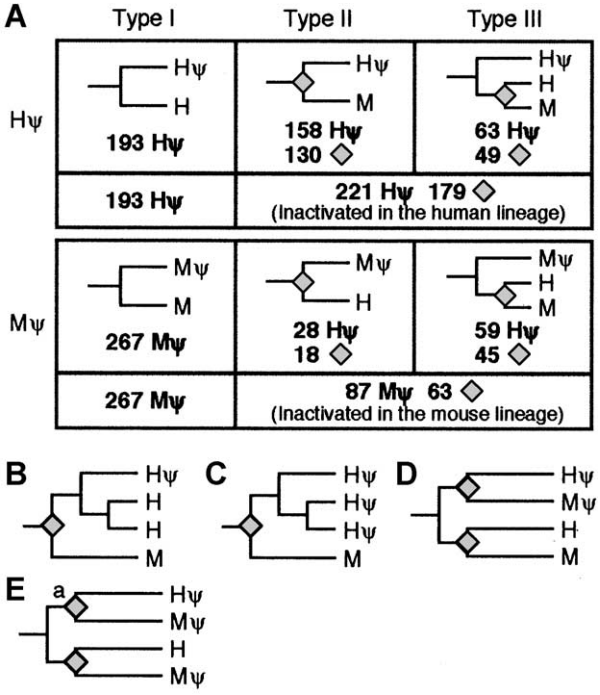

Fig. 3. Loss of OR genes in the mouse and human lineages. (A) Numbers of types I, II, and III trees for human (Hψ) and mouse (Mψ) pseudogenes. For types II and III trees, the estimated numbers of MRCA genes involved are also shown (indicated by diamonds). The estimated number of MRCA genes that became inactivated in the human and mouse lineages was obtained by adding the numbers of MRCA genes involved for types II and III trees. The estimated number of pseudogenes that have been generated from these MRCA genes in each lineage was obtained by adding the numbers of pseudogenes for types II and III. (B) Phylogenetic tree indicating that the MRCA gene (shown by a diamond) has generated both functional genes (H) and pseudogenes (Hψ) in the human lineage. This is a type I tree, because Hψ is closest to human functional gene or genes rather than mouse functional genes. (C) Phylogenetic tree indicating that the MRCA gene remained functional in the mouse lineage (M) but became inactivated in the human lineage (Hψ). In these cases, Hψ is closest to a mouse functional gene (type II trees). Note that each of the phylogenetic trees classified into types I, II, and III contained only one pseudogene. (D) Phylogenetic tree indicating that one MRCA gene became pseudogenes in both human and mouse lineages (Hψ and Mψ) and another MRCA gene remained functional in both lineages (H and M). The tree for Hψ is classified into type III, because Hψ is closest to both human and mouse functional genes. (E) Phylogenetic tree indicating that one MRCA gene became pseudogenes in both human and mouse lineages (Hψ and Mψ) and another MRCA gene remained functional in the human lineage (H) but became a pseudogene in the mouse lineage. In this case, although the MRCA gene (a) became inactivated in the human lineage, the phylogenetic tree for Hψ is classified into type I, because the closest functional gene to Hψ is from humans.

Type III trees, in which Hψ is closest to both human and mouse functional genes, are expected to occur when two MRCA genes were involved and one of them generated pseudogenes in both human and mouse lineages (Fig. 3D). The number of MRCA genes that have generated human pseudogenes showing type III trees can be estimated by counting the number of different pairs of functional human and mouse genes involved (Supplementary data, Fig. S2A) under the assumption that the possibility of the involvement of three MRCA genes is negligible (Supplementary data, Fig. S2B). We found 63 human pseudogenes showing type III trees and estimated the number of MRCA genes that have generated these pseudogenes to be 49. Here we also assume that each of the 49 MRCA genes was functional and became pseudogenes in both human and mouse lineages independently, because the probability that a pseudogene survives for nearly 100 MY without being decayed into an undetectable DNA sequence or being deleted from the genome appears to be small. However, actually, it is possible that the MRCA gene was a pseudogene (see Discussion).

We can now compute the total number of inactivated MRCA genes and that of pseudogenes generated from the MRCA genes in the human lineage by considering all cases of types II and III trees. They became 179 and 221, respectively (Fig. 3A). We have done similar computations for the mouse lineage and obtained 63 inactivated MRCA genes and 87 pseudogenes generated from these MRCA genes (lower part of Fig. 3A).

Previously, we identified 193 human pseudogenes showing type I trees. Theoretically, a type I tree can be generated even when two MRCA genes are inactivated in the mouse lineage, as shown in Fig. 3E. Therefore, the numbers of inactivated MRCA genes may be underestimated. To infer how frequently such events occurred, we constructed several phylogenetic trees for all functional genes and pseudogenes longer than 250 codons. (There are 312 human and 199 mouse pseudogenes >250 codons long.) Inspection of these trees suggested that the probability of occurrence of Fig. 3E trees is quite low. Similarly, probability that a type II tree is generated owing to the inactivation of two MRCA genes in the human lineage appeared to be very small.

4. Discussion

In this paper, we have made the following observations. (i) The mouse lineage acquired ~350 new OR genes after the HMD, whereas the human lineage acquired much smaller number of genes (~60). (ii) The human lineage has lost ~180 functional OR genes in the MRCA as pseudogenes, whereas the mouse lineage has lost much smaller number of genes (~60). These observations indicate that both gene expansion in the mouse lineage and gene loss in the human lineage are responsible for the difference in the number of OR genes between the two species. In this paper, we clearly showed that OR genes have substantially expanded in the mouse lineage by comparing the number of functional genes in mice with that in the MRCA. This is in contrast to the previous report that 80% of mouse genes have one-to-one orthologous genes in humans and mouse-specific gene expansion compared with human genes is rare (Mouse Genome Sequencing Consortium, 2002).

Mouse-specific gene expansion can be explained by the hypothesis that rodents require a broader range of OR genes to cope with a wide variety of environments in which they live. By contrast, the reduction of functional OR genes in the human lineage is probably due to the reduced importance of olfaction, as suggested by Rouquier et al. (2000) and Gilad et al. (2003, 2004). However, several studies suggested that the relationship between the number of functional OR genes in the genome and the ability of olfaction is not straightforward. For example, Laska et al. (2000, 2004) reported that the sense of smell in primates is better than rats for some types of odors, although it was reported that rats have ~2000 OR genes (Rat Genome Sequencing Project Consortium, 2004). Shepherd (2004) pointed out that higher brain mechanisms in humans may compensate for the smaller repertoire of OR genes. Nevertheless, the number of functional OR genes in the genome is expected to reflect some aspect of the ability of olfaction, such as an ability to distinguish between two similar odor molecules.

The evolutionary changes of the numbers of OR genes in the human and mouse lineages can be described in some more detail (Fig. 4), although the numbers obtained are rough estimates. We estimated that 1037 functional OR genes in mice originated from 691 OR genes in the MRCA (Fig. 2A). This estimate appears to be quite reliable, because another independent method gave nearly the same number. We also estimated that 63 functional genes in the MRCA have become pseudogenes in the mouse lineage (Fig. 3A). These results suggest that the number of functional OR genes in the MRCA was 754 (Fig. 4A). By contrast, we estimated that 388 functional OR genes in humans have been derived from 326 OR genes in the MRCA (Fig. 2A). These numbers suggest that 428 (=754−326) MRCA genes have been inactivated in the human lineage after the HMD (Fig. 4B). We estimated that 179 functional MRCA genes have become pseudogenes in the human lineage (Fig. 3A).

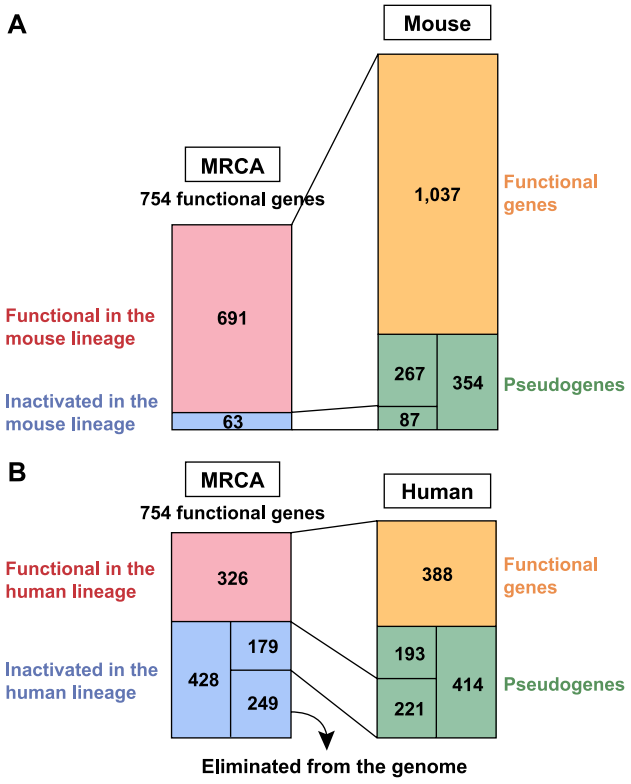

Fig. 4. Evolutionary changes of the numbers of OR genes in the mouse lineage (A) and in the human lineage (B). (A) We estimated that 691 genes in the MRCA (red) have become 1037 functional genes (orange) and 267 pseudogenes for type I trees out of 354 pseudogenes (green) in the mouse lineage. We also estimated that 63 genes in the MRCA (blue) have generated 87 mouse pseudogenes showing types II and III trees. From these numbers, the number of functional genes in the MRCA is estimated to be 754. (B) We estimated that 326 MRCA (red) genes have generated 388 functional genes and 193 pseudogenes for type I trees out of 414 pseudogenes (green) in the human lineage. The other 428 genes (blue) out of 754 genes in the MRCA were inactivated in the human lineage. Out of the 428 MRCA genes, 179 genes have become 221 pseudogenes showing types II and III trees. The other 249 genes appear to have been eliminated from the genome.

Therefore, 249 (=428−179) genes appear to have been eliminated from the genome or become undetectable by homology search due to accumulation of mutations. Here we did not consider the genes eliminated from the genome in the mouse lineage. However, this number should be relatively small, because the estimated number of MRCA genes that have become pseudogenes in the mouse lineage is small (63). These results indicates that the number of functional OR genes in the MRCA was close to the average between those in humans and mice.

As we mentioned earlier, we assumed that the MRCA genes that generated pseudogenes showing type III trees were functional, but it is possible that they were actually pseudogenes. Therefore, the number of functional OR genes in the MRCA may be overestimated. However, this possibility does not essentially change our conclusion, because the estimated number of functional genes in the MRCA is 709 (=691+18) even if all the MRCA genes

involved in type III trees are assumed to have been pseudogenes. Recently Zhang et al. (2004) identified ~5000 processed pseudogenes from the mouse genome and estimated that ~40% of them were generated before the HMD. Ancestral pseudogenes that have emerged before the HMD can be identified by finding processed pseudogenes that are located in syntenic regions between the two species. However, it is quite difficult to identify ancestral pseudogenes for OR pseudogenes, because most of them are located within large OR genomic clusters and therefore are not processed pseudogenes.

As shown in Fig. 3A, the number of human pseudogenes showing type I trees is nearly the same as the total number of human pseudogenes showing types II and III trees. Therefore, almost half of the human pseudogenes are not responsible for the loss of functional OR genes in the human lineage. This observation is explained by the presence of H* pseudogenes, which appear to have spread over the human genome by duplication after pseudogenization (Niimura and Nei, 2005). This indicates that a comparison of the number of pseudogenes between different species should be done with caution. A large number of pseudogenes may not mean that many functional genes have been lost but reflect frequent duplication of a particular type of pseudogenes.

One may criticize our method for counting the number of MRCA genes (Fig. 1), because the constancy of the evolutionary rate was not validated. However, our method is expected to give reasonably good estimates of the number of genes even if the evolutionary rate is not the same for all phylogenetic clades for the following reason. Suppose that one phylogenetic clade evolves faster than the average. In this case, a branch length within the clade is overestimated, and thus the divergence time for the genes in the clade is estimated to be too ancient. Therefore, even if gene duplication occurred after the HMD, the gene duplication may be inferred to have occurred before the HMD. For this reason, fast-evolving clades make the number of genes generated after the HMD underestimated. By contrast, slow-evolving clades make the number of genes generated after the HMD overestimated. Under the assumption that the rate of gene expansion is nearly the same for fast- and slow-evolving clades, the number of cases of overestimation of divergence time and the number of cases of underestimation is nearly the same. Therefore, the estimated number of genes that have been generated after the HMD is expected to be close to the actual number.

We estimated the rate of increase of the number of OR genes to be 0.005–0.015 per gene per MY for mouse functional genes. Interestingly, these estimates are similar to Lynch and Conery's (2003) estimates of genome-wide rate of gene duplication in humans, i.e., 0.009 per gene per MY. However, this estimate was obtained by considering only recently duplicated genes including those that may disappear relatively soon. In fact, they estimated that the average half-life of duplicated genes among eukaryotes is ~4 MY. By contrast, our estimates were obtained by considering only functional genes. Therefore, it may mean that a substantial number of duplicated OR genes have stayed functional in the genome in the mouse lineage for a relatively long evolutionary time.

Acknowledgements

This work was supported by National Institutes of Health Grant GM20293 (to M.N.).

Appendix A

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2004.09.027.

References

Gilad, Y., Man, O., Pääbo, S., Lancet, D., 2003. Human specific loss of olfactory receptor genes. Proc. Natl. Acad. Sci. U. S. A. 100, 3324–3327.

Gilad, Y., Wiebe, V., Przeworski, M., Lancet, D., Pääbo, S., 2004. Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. PLoS Biol. 2, 120–125.

Laska, M., Seibt, A., Weber, A., 2000. Microsmatic primates revisited: olfactory sensitivity in the squirrel monkey. Chem. Senses 25, 47–53.

Laska, M., Wieser, A., Rivas Bautista, R.M., Hernandez Salazar, L.T., 2004. Olfactory sensitivity for carboxylic acids in spider monkeys and pigtail macaques. Chem. Senses 29, 101–109.

Lynch, M., Conery, J.S., 2003. The evolutionary demography of duplicate genes. J. Struct. Funct. Genomics 3, 35–44.

Mouse Genome Sequencing Consortium, 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520–562.

Nei, M., 1969. Gene duplication and nucleotide substitution in evolution. Nature 221, 40–42.

Nei, M., Xu, P., Glazko, G., 2001. Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. Proc. Natl. Acad. Sci. U. S. A. 98, 2497–2502.

Niimura, Y., Nei, M., 2003. Evolution of olfactory receptor genes in the human genome. Proc. Natl. Acad. Sci. U. S. A. 100, 12235–12240.

Niimura, Y., Nei, M., 2005. Comparative evolutionary analysis of olfactory receptor gene clusters between humans and mice. Gene 346, 13–21.

Rat Genome Sequencing Project Consortium, 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428, 493–521.

Rouquier, S., Blancher, A., Giorgi, D., 2000. The olfactory receptor gene repertoire in primates and mouse: evidence for reduction of the functional fraction in primates. Proc. Natl. Acad. Sci. U. S. A. 97, 2870–2874.

Shepherd, G.M., 2004. The human sense of smell: are we better than we think? PLoS Biol. 2, 572–575.

Takezaki, N., Rzhetsky, A., Nei, M., 1995. Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. 12, 823–833.

Zhang, Z., Carriero, N., Gerstein, M., 2004. Comparative analysis of processed pseudogenes in the mouse and human genomes. Trends Genet. 20, 62–67.