Evolution of the rodent eosinophil-associated RNase gene family by rapid gene sorting and positive selection

Jianzhi Zhang, Kimberly D. Dyer, and Helene F. Rosenberg*

Laboratory of Host Defenses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892

Communicated by Masatoshi Nei, Pennsylvania State University, University Park, PA, February 17, 2000 (received for review November 16, 1999)

The mammalian RNase A superfamily comprises a diverse array of ribonucleolytic proteins that have a variety of biochemical activities and physiological functions. Two rapidly evolving RNases of higher primates are of particular interest as they are major secretory proteins of eosinophilic leukocytes and have been found to possess anti-pathogen activities *in vitro*. To understand how these RNases acquired this function during evolution and to develop animal models for the study of their functions *in vivo*, it is necessary to investigate these genes in many species. Here, we report the sequences of 38 functional genes and 23 pseudogenes of the eosinophil-associated RNase (EAR) family from 5 rodent species. Our phylogenetic analysis of these genes showed a clear pattern of evolution by a rapid birth-and-death process and gene sorting, a process characterized by rapid gene duplication and deactivation occurring differentially among lineages. This process ultimately generates distinct or only partially overlapping inventories of the genes, even in closely related species. Positive Darwinian selection also contributed to the diversification of these EAR genes. The striking similarity between the evolutionary patterns of the EAR genes and those of the major histocompatibility complex, immunoglobulin, and T cell receptor genes stands in strong support of the hypothesis that host-defense and generation of diversity are among the primary physiological function of the rodent EARs. The discovery of a large number of divergent EARs suggests the intriguing possibility that these proteins have been specifically tailored to fight against distinct rodent pathogens.

How gene families evolve and new gene functions originate are among the central issues in evolutionary biology. The mammalian RNase A superfamily is an excellent system for studying this problem because the many members of this superfamily each have evolved unique molecular structures and profiles of biochemical activities and physiological functions in a relatively short period of evolutionary time. At the same time, each RNase A superfamily member maintains the characteristic six to eight cysteines, a catalytic lysine, and two catalytic histidines (1, 2). We have been interested in a pair of primate RNase genes that emerged from a gene duplication that occurred about 30 million years (MY) ago in the evolutionary lineage of Old World monkeys and hominoids (3–8). The protein products of these two genes, eosinophil-derived neurotoxin (EDN) and eosinophil cationic protein (ECP), are found in the large specific granules of eosinophilic leukocytes. After gene duplication, ECP acquired a cell membrane-disruptive function that is likely to be responsible for its *in vitro* activity against bacteria and parasites (7, 9–11). EDN, on the other hand, was found to reduce the infectivity of RSV (respiratory syncytial virus) and HIV *in vitro* (12, 13), an activity that is RNase-dependent, at least in the case of RSV (12). Sequence analysis showed that the ECP and EDN genes are among the fastest evolving genes of primates (5), and the evolution of ECP appears to be driven by directional positive selection (8).

To study the molecular details of the origins of the anti-pathogen activities in ECP and EDN, and to develop animal models for the study of their functions *in vivo*, it is important to explore these genes in many species. It will also be interesting to investigate whether this lineage of the RNase A superfamily has evolved new functions in other species as well, and to determine what these functions might be. Through purification and sequencing of protein products from eosinophils of the house mouse (*Mus musculus*), Larson *et al.* (14) successfully cloned two eosinophil-associated RNase (EAR) genes. These two genes appear to have been generated by gene duplication after the separation of rodents and primates. Using conserved regions of the mouse EAR genes as primers, six additional genes were amplified from the mouse and eight from the rat (*Rattus norvegicus*) (15, 16), and a phylogenetic analysis showed that the mouse and rat genes form separate clusters in the tree (16). To understand how such distinct gene clusters evolved and to examine whether this pattern is unique to *Mus* and *Rattus* (i.e., subfamily Murinae of family Muridae), we report here the isolation of EAR genes from species of the subfamilies Gerbillinae (gerbils) and Cricetinae (hamsters), as well as from three additional species of the *Mus* genus. Unexpectedly, we found evidence suggesting massive gene deactivation in addition to extensive gene duplication. Thus, in contrast to our earlier thoughts—that rodent EAR genes emerged through recent, independent gene family expansions alone (16)—our results here suggest instead that rapid gene birth-and-death and gene sorting coupled with positive selection are more likely to have been the mechanisms used in generating the diversity among the EAR genes.

Materials and Methods

Isolation of EAR Genes. Genomic DNAs of the Mongolian gerbil *Meriones unguiculatus* and Chinese hamster *Cricetulus griseus* were isolated from cells of the American Type Culture Collection cell lines CCL-100 and CRL-9618, respectively, and genomic DNAs of the ricefield mouse *Mus caroli*, spiny mouse *Mus saxicola*, and shrew mouse *Mus pahari* were a gift from Anthony Furano (National Institutes of Health, Bethesda, MD). Each DNA sample was derived from an individual organism. The RNase genes were amplified by PCR with the primers derived from the published mEAR1 sequence of the house mouse (14) as follows: 5′-ATG GGT CCG AAG CTG CTT GAG TCC-3′ and 5′-CTA AAA TGT CCC ATC CAA GTG AAC-3′. The PCR were performed as described in ref. 16, and the multiple products present in a single band of ≈450 bp were identified by

Abbreviations: EAR, eosinophil-associated RNase; ECP, eosinophil cationic protein; EDN, eosinophil-derived neurotoxin; MY, million years; RSV, respiratory syncytial virus.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF238385–AF238445.

*To whom reprint requests should be addressed. E-mail: h2k@nih.gov.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.080071397. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.080071397

Table 1. Numbers of functional genes and pseudogenes of the EAR family in different rodent species

Species	No. of genes	
	Functional	Pseudo
House mouse (*Mus musculus*)	8	1
Ricefield mouse (*Mus caroli*)	17	2
Spiny mouse (*Mus saxicola*)	7	7
Shrew mouse (*Mus pahari*)	4	13
Rat (*Rattus norvegicus*)	8	0
Gerbil (*Meriones unguiculatus*)	6	0
Hamster (*Cricetulus griseus*)	4	1

dideoxy-sequencing of individual plasmids on both directions after subcloned into the pCR II TA cloning vector (Invitrogen).

Sequence Analysis. DNA sequences of functional genes were aligned based on the translated protein sequences by using CLUSTAL V (17) with visual adjustments. Pseudogenes were aligned similarly based on their DNA sequences. The primer-encoded regions, as well as those sites that contain gaps in the alignment, were not used in any evolutionary analysis (complete-deletion option). Gene trees were reconstructed by using the Neighbor–Joining method (18) with Kimura's two-parameter distance (19), and 1,000 bootstrap replications were conducted to evaluate the reliability of the trees. Rates of synonymous and nonsynonymous nucleotide substitutions were estimated by the method of Zhang *et al.* (8), which was modified from the method of Nei and Gojobori (20) to account for unequal rates of transitions and transversions. The software MEGA (21) and a prerelease version of MEGA2 (S. Kumar, personal communication) were used for the above analyses. The hypothesis of molecular clock was examined by the two-cluster test of Takezaki *et al.* (22). Gene conversion among paralogous genes of a species was tested by using Sawyer's method (23), implemented in the program of Drouin *et al.* (24). Bonferroni correction was used when multiple tests were performed. Isoelectric points (pI) of mature proteins were computed by the Wisconsin GCG program available on-line at National Institutes of Health.

Results

EAR Functional Genes. A total of 17, 7, 4, 6, and 4 putative functional genes of the EAR family were obtained from the ricefield mouse *Mus caroli* (abbreviated as rfm), spiny mouse *Mus saxicola* (spm), shrew mouse *Mus pahari* (shm), gerbil *Meriones unguiculatus* (ger), and hamster *Cricetulus griseus* (ham), respectively (Table 1). Presence of multiple EAR genes in these species is not surprising, as both genomic Southern analyses and molecular sequencing revealed multiple EAR genes in the house mouse and rat (14–16). The multiple sequences of each species reported here have pairwise distances of at least three amino acids, which essentially rules out the possibility that PCR errors have generated these sequence differences (16). In addition, almost every sequence reported here was isolated at least twice. An important note: although our phylogenetic analysis showed close relationship of the newly identified EAR genes to the mEAR1 and mEAR2 genes of the house mouse, which are both expressed in eosinophils, we have no information on the expression patterns of the genes identified in this study. As such, our use of "EAR" in the gene name is an indication of classification by sequence similarity only.

An alignment of the amino acid sequences encoded by these genes and the house mouse mEAR1 gene (14) is shown in Fig. 1*A*. Except for one sequence (rfmEAR5), all show features typical of the RNase A superfamily. Each sequence has an amino-terminal signal peptide of 27 amino acids. The mature proteins are 125–131 aa long, with eight characteristic cysteines as well as a catalytic lysine and two catalytic histidines at

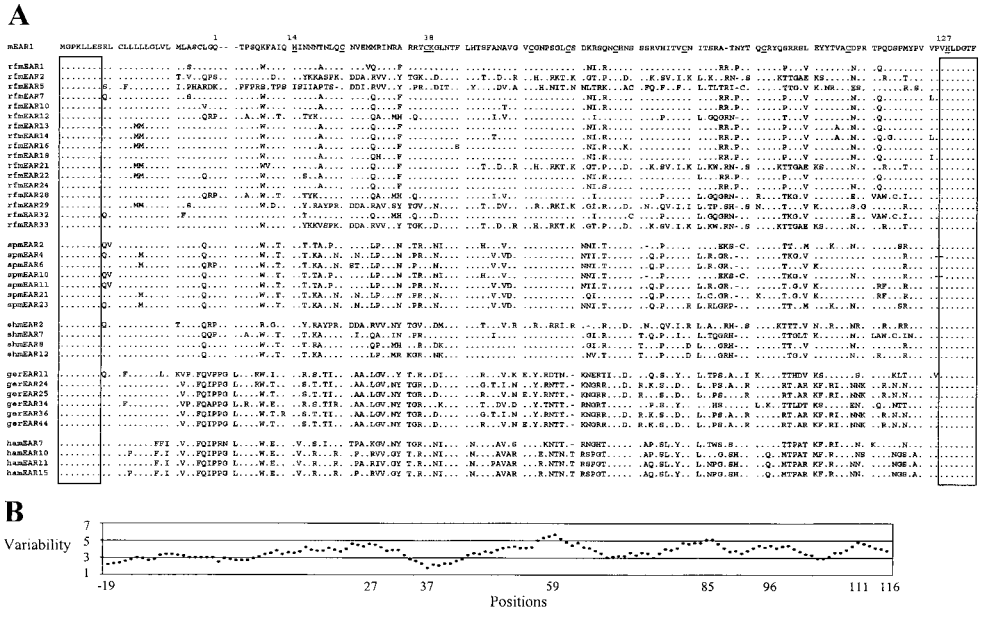

Fig. 1. Rodent EAR sequences. (*A*) Alignment of the newly sequenced EARs with mEAR1 of the house mouse. The eight conserved cysteines as well as the three catalytic sites at positions 14, 38, and 127 are underlined. Position 1 of mEAR1 is known to be the initial amino acid of the mature protein as determined by amino-terminal sequencing (14). The first eight and last seven amino acids are primer-encoded regions and are boxed. Abbreviations for the species: rfm, ricefield mouse; spm, spiny mouse; shm, shrew mouse; ger, gerbil; ham, hamster. (*B*) Sliding-window analysis of sequence variability. The number of different amino acids observed at each site is counted, and the average number over a window size of 10 sites is shown. Some peaks and valleys in the plot are marked by the starting positions of the windows. The average variability over all sites is 3.58. All available rodent EAR sequences except rfm5 are used in this analysis with the primer-encoded regions not included.

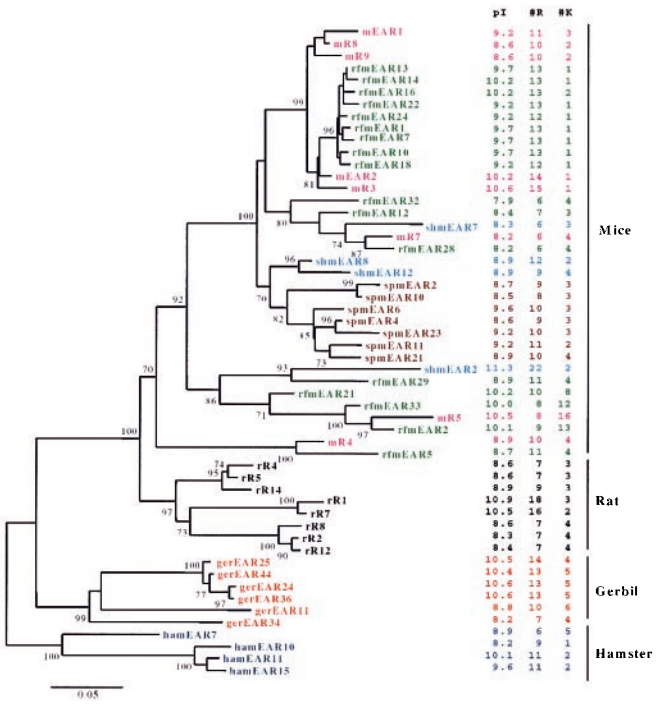

Fig. 2. Phylogenetic tree of functional EAR genes of rodents. After the removal of gaps, a total of 390 nucleotide sites are used in tree-making. Bootstrap percentages that are equal to or higher than 70 are shown on interior branches. Following convention (14–16), house mouse genes are designated either by mEAR or mR, whereas rat genes are designated by rR. The pI values, as well as the numbers of arginine and lysine residues in the mature proteins are shown to the right of the tree. Genes from different species are depicted with different colors.

appropriate positions. The enormous sequence variation among the EARs is evident from the alignment, with multiple regions of high variability (Fig. 1*B*). The rfmEAR5 gene is different from all others in that there is a 1-nt deletion at the third position of the 21st codon and a 2-nt deletion 73 nucleotides thereafter, causing a frame-shift of 25 codons and absence of the first catalytic histidine (Fig. 1). Although it is unclear whether this gene is expressed, the fact that there are no premature stop codons and that the eight cysteines, the second catalytic histidine, and the catalytic lysine are all well conserved suggest that the gene may in fact be subject to purifying selection. The protein product of this gene, if it exists, probably has no RNase activity due to the missing catalytic histidine, but a formal examination of the protein will be necessary to prove this point.

When an evolutionary tree was reconstructed by using the newly sequenced genes, house mouse and rat EAR genes (14–16), and other members of the RNase A superfamily from the mouse and human (1, 2, 7, 14–16), the rodent EAR genes form a statistically supported group with human EDN and ECP as its closest relatives (data not shown). This suggests that the newly obtained sequences belong to the EAR family of the RNase A superfamily. Fig. 2 shows a Neighbor–Joining tree including all rodent EAR genes. The genes from the rat, gerbil, and hamster form three separate clusters with significant (>95%) bootstrap support. Previously, the house mouse and rat genes were found to form separate clusters (16). Our results here indicate that this evolutionary pattern is not limited to *Mus* and *Rattus* and the subfamily Murinae, in which they reside, but rather extends to the subfamilies Gerbillinae (gerbils) and Cricetinae (hamsters) of the family Muridae.

Genes from the four species of the *Mus* genus form another cluster in the tree. Within this cluster, the seven EAR genes of the spiny mouse group together with 82% bootstrap support. Genes from the other three species intermingle to some extent, and most major groups comprise genes from two to three species. It is believed that the house mouse and ricefield mouse diverged about 4–6 MY ago, and the trifurcation of the spiny mouse, shrew mouse, and the house mouse/ricefield mouse pair occurred about 8–10 MY ago (25). A comparison between the house mouse and ricefield mouse shows that although several genes may share orthologous relationships (e.g., mR4 with rfmEAR5, mR5 with rfmEAR2, and mR7 with rfmEAR28; see Fig. 2 for notations of the house mouse genes), others have been duplicated since speciation. Of particular interest is a group of nine ricefield mouse genes and five house mouse genes that form a statistically supported group (the top 14 genes in Fig. 2). Although relationships among these genes cannot be resolved because of relatively low levels of sequence divergence, these genes are likely to be products of very recent duplications.

EAR Pseudogenes. It is apparent from the tree in Fig. 2 that numerous EAR gene duplications took place before the divergence of the four *Mus* species, but some genes of one species lack orthologous counterparts in other species. This suggests that functional genes of one species may have become pseudogenes in others. Our study has revealed EAR pseudogenes in all *Mus* species examined. Specifically, we found 2, 7, and 13 pseudogenes in the ricefield mouse, spiny mouse, and shrew mouse, respectively (Table 1). Previous studies in the house mouse also identified a pseudogene (15). In addition, one pseudogene was found in the hamster, whereas none were found in the gerbil.

A Neighbor–Joining tree of 54 functional genes and 23 pseudogenes from the EAR family of seven rodent species is shown in Fig. 3. The single pseudogene from the house mouse was not included in this analysis because it contains a large deletion. Because of the insertions and deletions in the pseudogenes, the number of nucleotide sites shared by all sequences is limited. This reduces the reliability of the estimated tree somewhat, as is evident from the bootstrap values. Nevertheless, sequences from the hamster, gerbil, rat, and the *Mus* genus still form separate clusters with moderate bootstrap support. The single pseudogene of the hamster (hamEAR30p, p stands for *p*seudogene) appears to have branched off earlier than its functional genes. For the four *Mus* species, the branching patterns of their functional genes are somewhat different from those in Fig. 2. In such cases, the phylogenetic relationships reconstructed in the tree of Fig. 2 are probably more reliable because the tree was based on a larger number of nucleotide sites. Interestingly, for the three *Mus* species examined in this study, the one that has fewest functional EAR genes (shrew mouse, 4) has the largest number (13) of pseudogenes, and vice versa. This makes the sum of the numbers of functional genes and pseudogenes nearly constant for the three *Mus* species examined here (Table 1). The total number of EAR genes is relatively small in the house mouse, which may be the result of incomplete reporting of pseudogenes in previous studies or a genuine distinction between this species and the other three *Mus* species in their EAR families.

The presence of a large number of pseudogenes helps to answer several questions in EAR evolution. For instance, functional genes of the spiny mouse form a closely related cluster in the tree of Fig. 2. This gives an impression that there was only one EAR gene in the spiny mouse lineage until the common ancestor of the cluster, which is in contrast to the fact that many genes seem to have been duplicated before the divergence of the spiny mouse and other *Mus* species. Now it is clear that many ancient EAR genes became pseudogenes in the spiny mouse. This can be seen in Fig. 3, where spiny mouse pseudogenes are found in a number of places outside the functional gene cluster.

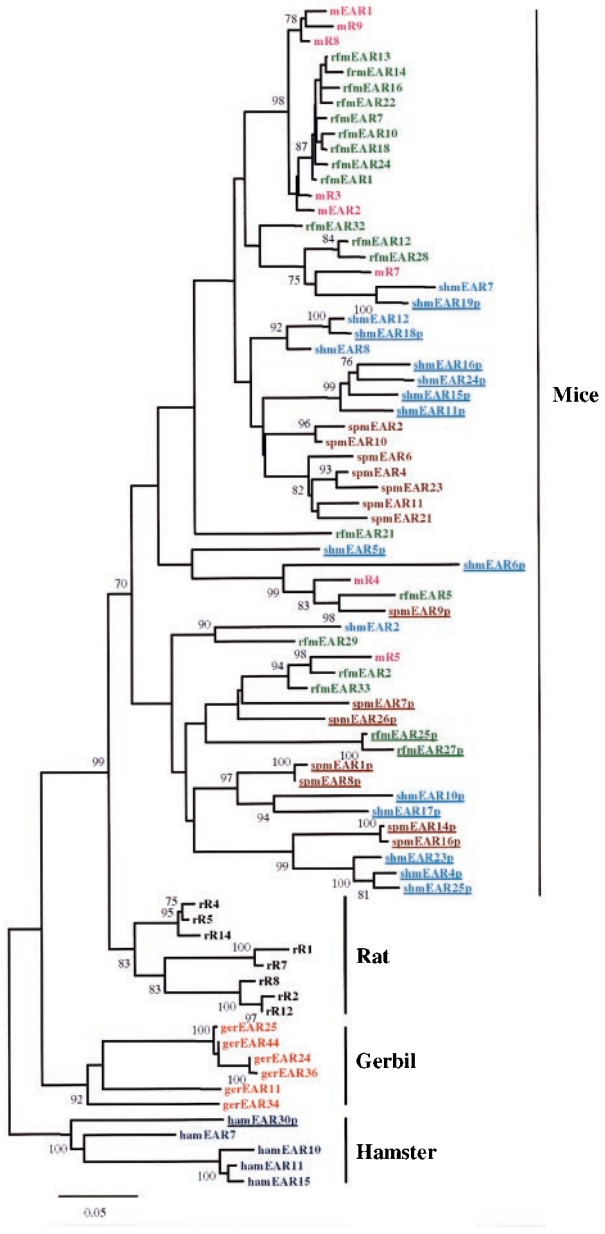

Fig. 3. Phylogenetic tree of functional genes and pseudogenes of the rodent EAR family. A total of 262 nucleotide sites are used in tree-making. Bootstrap percentages that are equal to or higher than 70 are shown on interior branches. Genes from different species are depicted with different colors. Pseudogenes bear a ''p'' at the end of their names and are underlined.

Similar patterns are found in other *Mus* species as well. It is apparent that when the four *Mus* species diverged, there were already a number of EAR genes in the genomes of their respective ancestors. Further gene duplication and gene deactivation ultimately resulted in what we observed here—that closely related species possess distinctive or only partially overlapping inventories of members of the EAR family.

It is likely that the distinct EAR gene clusters in the gerbil, hamster, rat, and other *Mus* species were produced in a fashion similar to that observed in the spiny mouse. That is, there were multiple genes in the common ancestor of these species, but a process of rapid gene deactivation in most gene lineages, coupled with a high rate of gene duplication, resulted in separate gene clusters in different species. Because of the relatively long time of divergence among the gerbil, hamster, rat, and *Mus* species (>40 MY; ref. 26), most pseudogenes in non-*Mus* species may have changed so much that our primers based on small regions of the house mouse functional genes can no longer detect them. This may be the primary reason why almost no EAR pseudogenes were identified outside the *Mus* genus.

Functional Divergence of EAR Genes Under Positive Selection. Do the many rodent EAR genes have different functions and specificity? Although there is no direct experimental evidence supporting this hypothesis at present, our sequence analysis suggests that the answer is most likely yes. The primate ECP and EDN genes are known to have distinct biochemical properties and functions, some of which are related to their distinct isoelectric points (pI). For example, the human ECP has 12 more arginines than human EDN, and its pI (11.4) is subsequently much higher than that (8.9) of EDN (5). The high arginine content of ECP makes it very positively charged, which presumably is critical for the tight contact of ECP to the negatively charged bacterial cell membrane, which in turn is likely to be important for its membrane-disruptive activity and anti-bacterial function (8, 9, 27). Because evolution of the new functions in primate EARs appear to be related to changes in pI and content of basic amino acids, we examined these properties of the rodent EARs (Fig. 2). Rodent EARs are found to have a wide range of pIs, from 7.9 to 11.3. For individual species, except the spiny mouse (which has a pI range of 1.1, probably because its EARs are all closely related), all other species have wide ranges (1.9–3.0) of pI among their EARs. For instance, the pI of shmEAR2 is 11.3, whereas that of shmEAR7 is 8.3. Similar to the primate ECPs and EDNs, the number of arginine residues in the rodent EARs varies considerably. For example, shmEAR2 has 22 arginines, whereas shmEAR7 has only 6. It is apparent that the amino acid substitutions toward arginine in certain rodent EARs and in primate ECP (8) occurred independently of one another. It will be interesting to examine whether these parallel changes in arginine content have resulted in a corresponding parallel evolution of the cell membrane-disruptive activity in the rodent EARs, enabling them to kill bacteria. Furthermore, rodent EARs show variation in the number of lysines as well. For example, whereas most EARs have no more than 5 lysines, 3 closely related EARs, rfmEAR33, rfmEAR2, and mR5, have 12–16 lysines, and their pIs are also high (10.0–10.5) (Fig. 2). It is unclear, however, whether lysines and arginines are interchangeable in the determination of EAR functions. In another arginine-rich protein of mammals, protamine P1, arginines cannot be replaced by lysines (28, 29) because the polyarginine motif can activate casein kinase II, whereas polylysine cannot (28). Regardless, diversification in pI and the number of basic amino acids in paralogous EARs suggests that distinct biochemical activities or functions have evolved.

What then is the driving force underlying the functional diversification of the EARs? Here, we tested the role of positive Darwinian selection by computing the synonymous (d_S) and nonsynonymous (d_N) nucleotide substitutions per site in comparisons of paralogous EAR genes of individual species (Fig. 4). For the rat, gerbil, and hamster, although the d_N/d_S ratios are smaller than 1 in most cases, they are much higher than the ratio (0.19) in an average rodent gene (30, 31). For the four species of the *Mus* genus, most comparisons show higher d_N than d_S, suggesting operation of positive selection in these genes. In particular, for the eight house mouse genes, the average pairwise d_N (0.181 ± 0.020) is significantly greater than the average pairwise d_S (0.102 ± 0.017) ($P < 0.01$, one-tail t test). Similarly, for the 17 ricefield mouse genes, the average d_N (0.166 ± 0.018) is also significantly greater than the average d_S (0.115 ± 0.019) ($P < 0.05$). It can also be shown by a phylogeny-based method (8) that the nonsynonymous substitution rate is significantly

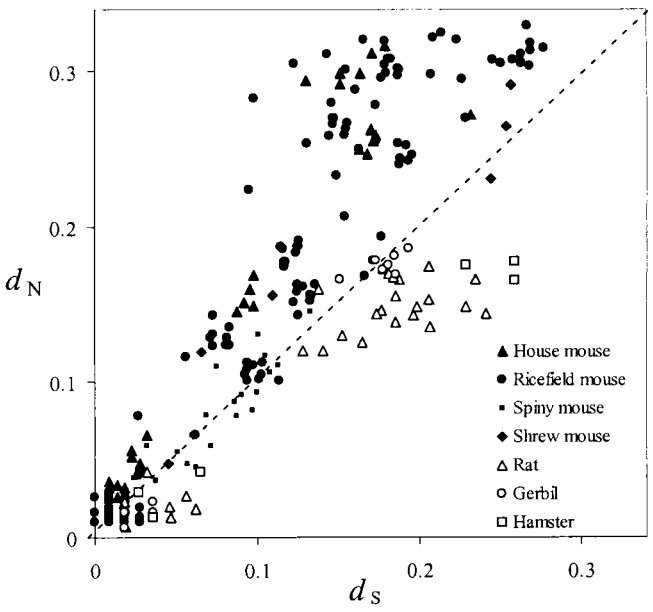

Fig. 4. Pairwise synonymous (d_S) and nonsynonymous (d_N) nucleotide distances between paralogous EAR genes belonging to individual species. The average transition/transversion ratio is approximately 0.9, computed by the Kimura's method (19).

higher than the synonymous rate in the branches linking the eight house mouse genes ($P < 0.01$). These results strongly suggest that positive selection has been operating on the EAR genes in the *Mus* species. A test of molecular clock based on the tree of Fig. 2 also indicates that the EAR genes of the *Mus* genus evolve significantly faster than those of the other rodents ($P < 0.01$), consistent with the occurrence of positive selection in *Mus* species. Because positive selection occurs only when there is difference in fitness (function), our results strongly suggest that the paralogous EAR genes have evolved distinct functions or specificity.

Discussion

In this study, we identified and sequenced 38 potentially functional genes and 23 pseudogenes of the EAR family from 5 rodent species. Our analyses suggest that extensive gene duplication and frequent gene deactivation occurred during the evolution of the rodent EARs, which resulted in distinct or only partially overlapping gene inventories in these closely related species. Based on the separate gene clusters of the house mouse EARs and rat EARs in their phylogenetic tree, Singhania *et al.* (16) suggested that the EAR family expanded independently in the two species, implying a single common ancestral gene in genera *Mus* and *Rattus*. This now seems unlikely when the distinct EAR gene clusters from the gerbil and hamster are considered, because one then needs to invoke at least four independent gene family expansions to explain the findings presented here. Furthermore, the observation that seven pseudogenes of the spiny mouse diverged before the separation of the seven functional genes indicates that independent gene family expansion is an oversimplified explanation. Rather, "gene sorting" followed by gene duplication is likely to be the mechanism used. Here, gene sorting is defined as the process leading to differential retention of ancestral genes or gene lineages in different species. The gene sorting observed among the rodent EARs can be regarded as a special case of what has been described as the birth-and-death process of gene family evolution (32), where "birth" implies gene duplication and "death" refers to gene deactivation. Although birth-and-death process occurs universally, in most known gene families it proceeds at a very low rate. Thus, straightforward identification of orthologous genes even among different mammalian orders is usually possible (32). Among the rodent EARs, the rates of gene birth and gene death are unusually high, such that orthologous gene pairs cannot be found even between the mouse and rat or, in some cases, between species within the *Mus* genus.

It is also worth noting that this study highlights the importance of sequencing and analyzing pseudogenes to understand the evolution of gene families. As noted above, it was the sequence analysis of several EAR pseudogenes that prompted the consideration of evolution by gene sorting as opposed to independent gene family expansion. Despite the prevalence of pseudogenes in higher vertebrates (e.g., 20% of all genes on the human chromosome 22 are pseudogenes; ref. 33), they are usually considered as the "dead ends" of evolution and are therefore often neglected inappropriately by many biologists.

Gene conversion has often been invoked to explain evolution of gene families, particularly when genes from different species form distinct clusters in the tree. Although it is difficult to rule out completely the possibility of gene conversion in the present case, five lines of evidence suggest that gene conversions, if they happened at all, are infrequent and local, and therefore, are not a major force in the origin and diversification of the EAR genes. First, the nucleotide differences among the paralogous genes from each species are considerable (over 10%, see Fig. 2), unlike those typical cases of gene conversion where the nucleotide differences are no more than 1–3% (34, 35). Second, the EAR genes have diverged in pI and presumably in function as well, suggesting that they have been diversified, not homogenized. Third, intermingling of EAR genes from three *Mus* species indicates that the rate of gene conversion is quite low. Fourth, gene conversion cannot be statistically proven in any of the seven species examined. Finally, if functional diversity in EARs is advantageous, as suggested by the occurrence of positive selection, alleles homogenized by gene conversion are unlikely to be fixed in the population (34, 36).

The chromosomal organization of the rodent EAR genes is also unknown, although human EAR genes (ECP, EDN, and the EDN pseudogene) are closely linked in chromosome 14 (5). The information on the chromosomal organization is important because gene duplication by unequal crossing-over is particularly easy when the paralogous genes are linked (34), as in the case of Ig genes. As the mouse genome sequencing project proceeds, the genomic organization and complete inventory of the mouse EAR family is expected to be revealed soon. It is also important to ask whether sequence polymorphism at the EAR loci is high, as it is in some of the MHC loci (37) and whether there is variation in the number of EAR genes among individuals, as in the case of Igκ chain variable region genes of humans (38). In the present study, however, we only examined one individual per species, and these interesting questions remain to be addressed in the future.

Rapid birth-and-death and gene sorting are not unique to the rodent EAR family. Before this, rapid birth-and-death process had been described in only a few major gene families, all of which are associated with the generation of diversity in support of immunity and host defense: the major-histocompatibility-complex (MHC), Ig, and T cell receptor (TCR) gene families (32, 36, 38–41). Two specific examples—Cadavid *et al.* (39) were unable to find orthologous genes of the MHC class I loci among different genera of New World monkeys, but instead found that genes from these organisms clustered in a tree in a genus-specific fashion. Similarly, Sitnikova and Su (40) determined that the Ig heavy chain variable region genes of the cattle and sheep form separate clusters in the tree, despite the recent divergence of the two species (20 MY ago; ref. 26). In each case, the ancestral

species appear to have had more than one gene, so the simple explanation—independent gene family expansions—does not hold. Similarly, positive Darwinian selection, another feature contributing to the divergence of EARs, is also a hallmark of vertebrate host-defense genes (37, 42–44) and their counterpart genes in pathogens (45, 46), although a few other genes including those involved in reproduction (47–51) and color vision (52) are also known to be under strong positive selection.

Most important, it is interesting to consider the implications that this unusual evolutionary pattern, one that is shared by MHC, Ig, TCR, and now EAR gene families, has on our understanding of the function of the latter. As part of our larger interest in the role of eosinophils, eosinophil RNases, and host defense, we have recently begun to explore the associations linking eosinophils with respiratory viruses, specifically the paramyxovirus pathogen RSV. We have shown that the human eosinophil RNase, EDN, reduces the infectivity of isolated virions of RSV in a dose-dependent, RNase-dependent fashion, and that this anti-viral activity also involves a specific, saturable interaction between EDN and its target (12, 53) (H.F.R. and J. B. Domachowske, unpublished data). Based on our earlier evolutionary and functional studies, we have hypothesized that there are unusual evolutionary pressures on the EAR lineages related to the acquisition of specialized anti-viral activity, and that each RNase has diverged to interact specifically with individual targets relating to host defense against viral pathogens (54). Our findings here—that the evolution of the rodent EARs parallels that observed in gene families whose function is based directly on their ability to generate diversity—stands in strongest support to date of this hypothesis. It is intriguing to consider the possibility that the function of the EARs is likewise based on the generation of diversity. It is possible that species-specific EARs are in the process of being tailored to fight against distinct, species-specific pathogens. In this respect, the multiple hypervariable regions (Fig. 1*B*) may be of particular importance in determining the specificity of each EAR, a hypothesis that can ultimately be tested experimentally. Our recent experiments with the murine paramyxovirus pathogen, pneumonia virus of mice (PVM), represent a step in this direction (55).

We thank Dr. Anthony Furano for the genomic DNAs of three *Mus* species, Dr. Sudhir Kumar for a test version of MEGA2, and Drs. Masatoshi Nei and Jaap Beintema for their insightful comments and suggestions.

1. D'Alessio, G. & Riordan, J. F. (1997) *Ribonucleases, Structures and Functions* (Academic, San Diego).
2. Beintema, J. J. & Kleineidam, R. G. (1998) *Cell. Mol. Life Sci.* **54,** 825–832.
3. Hamann, K. J., Ten, R. M., Loegering, D. A., Jenkins, R. B., Heise, M. T., Schad, C. R., Pease, L. R., Gleich, G. J. & Barker, R. L. (1990) *Genomics* **7,** 535–546.
4. Rosenberg, H. F. & Dyer, K. D. (1995) *J. Biol. Chem.* **270,** 21539–21544.
5. Rosenberg, H. F., Dyer, K. D., Tiffany, H. L. & Gonzalez, M. (1995) *Nat. Genet.* **10,** 219–223.
6. Rosenberg, H. F. & Dyer, K. D. (1997) *Nucleic Acids Res.* **25,** 3532–3536.
7. Rosenberg, H. F. (1998) *Cell. Mol. Life Sci.* **54,** 795–803.
8. Zhang, J., Rosenberg, H. F. & Nei, M. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 3708–3713.
9. Young, J. D. E., Peterson, C. G. B., Venge, P. & Cohn, Z. A. (1986) *Nature (London)* **321,** 613–616.
10. Waters, L. S., Taverne, J., Tai, P. C., Spry, C. J., Targett, G. A. & Playfair, J. H. (1987) *Infect. Immun.* **55,** 877–881.
11. Lehrer, R. I., Szklarek, D., Barton, A., Ganz, T., Hamann, K. J. & Gleich, G. J. (1989) *J. Immunol.* **142,** 4428–4434.
12. Domachowske, J. B., Dyer, K. D., Bonville, C. A. & Rosenberg, H. F. (1998) *J. Infect. Dis.* **177,** 1458–1464.
13. Lee-Huang, S., Huang, P. L., Sun, Y., Huang, P. L., Kung, H. F., Blithe, D. L. & Chen, H. I. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 2678–2681.
14. Larson K. A., Olson, E. V., Madden, B. J., Gleich, G. J., Lee, N. A. & Lee, J. J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 12370–12375.
15. Batten, D., Dyer, K. D., Domachowske, J. B. & Rosenberg, H. F. (1997) *Nucleic Acids Res.* **25,** 4235–4239.
16. Singhania, N. A., Dyer, K. D., Zhang, J., Deming, M. S., Bonville, C. A., Domachowske, J. B. & Rosenberg, H. F. (1999) *J. Mol. Evol.* **49,** 721–728.
17. Higgins, D. G., Bleasby, A. J. & Fuchs, R. (1992) *Comput. Appl. Biosci.* **8,** 189–191.
18. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
19. Kimura, M. (1980) *J. Mol. Evol.* **16,** 111–120.
20. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
21. Kumar, S., Tamura, K. & Nei, M. (1993) MEGA, *Molecular Evolutionary Genetics Analysis* (The Pennsylvania State University, University Park, PA), Version 1.02.
22. Takezaki, N., Rzhetsky, A. & Nei, M. (1995) *Mol. Biol. Evol.* **12,** 823–833.
23. Sawyer, S. (1989) *Mol. Biol. Evol.* **6,** 525–538.
24. Drouin, G., Prat, F., Ell, M. & Clarke, G. D. P. (1999) *Mol. Biol. Evol.* **16,** 1369–1390.
25. Sing, N., Barbour, K. W. & Berger, F. G. (1998) *Mol. Biol. Evol.* **15,** 312–325.
26. Kumar, S. & Hedges, S. B. (1998) *Nature (London)* **392,** 917–920.
27. Rosenberg, H. F., Ackerman, S. J. & Tenen, D. G. (1989) *J. Exp. Med.* **170,** 163–176.
28. Ohtsuki, K., Nishikawa, Y., Saito, H., Munakata, H. & Kato, T. (1996) *FEBS Lett.* **8,** 115–120.
29. Rooney, A. P., Zhang, J. & Nei, M. (2000) *Mol. Biol. Evol.* **17,** 278–283.
30. Makalowski, W. & Boguski, M. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 9407–9412.
31. Zhang, J. (2000) *J. Mol. Evol.* **50,** 56–68.
32. Nei, M. & Hughes, A. L. (1992) in *11th Histocompatibility Workshop and Conference*, eds. Tsuji, K., Aizawa, M. & Sasazuki, T. (Oxford Univ. Press, Oxford), Vol. 2, pp. 27–38.
33. Dunham, I., Shimizu, N., Roe, B. A., Chissoe, S., Hunt, A. R., Collins, J. E., Bruskiewich, R., Beare, D. M., Clamp, M., Smink, L. J., *et al.* (1999) *Nature (London)* **402,** 489–495.
34. Li, W. H. (1997) *Molecular Evolution* (Sinauer, Sunderland, MA).
35. Swanson, W. J. & Vacquier, V. D. (1998) *Science* **281,** 710–712.
36. Nei, M., Gu, X. & Sitnikova, T. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 7799–7806.
37. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335,** 167–170.
38. Sitnikova, T. & Nei, M. (1998) *Mol. Biol. Evol.* **15,** 50–60.
39. Cadavid, L. F., Shufflebotham, C., Ruiz, F. J., Yeager, M., Hughes, A. L. & Watkins, D. I. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 14536–14541.
40. Sitnikova, T. & Su, C. (1998) *Mol. Biol. Evol.* **15,** 617–625.
41. Su, C., Jakobsen, I., Gu, X. & Nei, M. (1999) *Immunogenetics* **50,** 301–308.
42. Tanaka, T. & Nei, M. (1989) *Mol. Biol. Evol.* **6,** 447–459.
43. Hughes, A. L. & Yeager, M. (1997) *BioEssays* **19,** 777–786.
44. Duda, T. F., Jr, & Palumbi, S. R. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 6820–6823.
45. Boenhoeffer, S., Holmes, E. C. & Nowak, M. A. (1995) *Nature (London)* **376,** 125.
46. Hughes, M. K. & Hughes, A. L. (1995) *Mol. Biochem. Parasitol.* **71,** 99–113.
47. Lee, Y.-H. & Vacquier, V. D. (1992) *Biol. Bull. (Woods Hole, Mass.)* **182,** 97–104.
48. Swanson, W. J. & Vacquier, V. D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 4957–4961.
49. Metz, E. C. & Palumbi, S. R. (1996) *Mol. Biol. Evol.* **13,** 397–406.
50. Tsaur, S.-C. & Wu, C.-I (1997) *Mol. Biol. Evol.* **14,** 544–549.
51. Rooney, A. P. & Zhang, J. (1999) *Mol. Biol. Evol.* **16,** 706–710.
52. Yokoyama, R. & Yokoyama, S. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9315–9318.
53. Domachowske, J. B., Bonville, C. A., Dyer, K. D. & Rosenberg, H. F. (1998) *Nucleic Acids Res.* **26,** 5327–5332.
54. Rosenberg, H. F. & Domachowske, J. B. (1999) *Immunol. Res.* **20,** 261–274.
55. Domachowske, J. B., Bonville, C. A., Dyer, K. D. & Rosenberg, H. F. (2000) *Cell. Immunol.*, in press.