CONCERTED EVOLUTION OF PARALOGOUS Oas1 GENES IN RODENTIA AND CETARTIODACTYLA

Perelygin A.A.*, Zharkikh A.A.2, Brinton M.A.1

1 Biology Department, Georgia State University, Atlanta, GA 30302, USA; 2 Bioinformatics Department, Myriad Genetics, Inc., Salt Lake City, UT 84108, USA

*Corresponding author: e-mail: aperelygin@gsu.edu

Key words: 2'-5' oligoadenylate synthetase, concerted evolution, gene conversion

SUMMARY

Motivation: The mouse 2'-5' oligoadenylate synthetase-1B (Oas1b) gene confers resistance to disease induced by flaviviruses, including West Nile virus (Perelygin et al., 2002). Oas1b is a member of the Oas gene family, which is located on mouse chromosome 5. The molecular evolution of paralogous mammalian 2'-5' oligoadenylate synthetase-1 genes has not been previously described.

Results: A total of fourteen new mRNA sequences of paralogous Oas1 genes were determined in mice, rats, cattle and pigs. These sequences as well as seven Oas1 gene sequences from GenBank were mapped to genomic regions and also used to build a phylogenetic tree. The majority of the eight mouse Oas1 genes clustered with their rat orthologs. However, the differences between paralogous rodent Oas1c and Oas1d genes in each species were smaller than the distances between the corresponding orthologs suggesting a concerted evolution of these genes. A new method was developed to compare the distribution of substitutions along the Oas1 nucleotide sequences in rodent or even-toed ungulate evolutionary lineages and to quantify these differences as probabilities. The distributions for pairs of sequences were then compared using the non-parametric Kolmogorov-Smirnov test and the results suggested that the homogenization of paralogous 2'-5' oligoadenylate synthetase-1 genes was due to gene conversion.

INTRODUCTION

2'-5' oligoadenylate synthetases are important components of an interferon-mediated antiviral pathway, but are also involved in other cellular processes such as apoptosis, cell growth and differentiation, gene regulation, DNA replication and RNA splicing. The sequences of most members of the mouse 2'-5' oligoadenylate synthetase (Oas) gene family and all members of the human OAS gene family have previously been reported (Justesen et al., 2000; Perelygin et al., 2002). Four OAS genes (hOAS1, hOAS2, hOAS3 and hOASL1) are located on human chromosome 12. The murine Oas family includes eight small mOas1 genes (mOas1a through mOas1h), a medium sized mOas2 gene and a large mOas3 gene, as well as two Oas-like genes, mOasl1 and mOasl2, which contain two tandemly repeated ubiquitin-like sequences at their 3'-ends. Because only few additional mammalian 2'-5' oligoadenylate synthetase genes were previously reported, it was not known whether the gene families in other mammals were more like the one in mice or the one in humans. Two rat rOas1 genes located on rat chromosome 12 were previously detected. Also, single OAS1 gene sequences from both pig and cattle were previously submitted to GenBank. In this study, we analyzed the 2'-5' oligoadenylate synthetase gene families of cattle, pig, mouse and rat and obtained evidence that supports concerted
evolution of paralogous 2'-5' oligoadenylate synthetase-1 genes within the orders Rodentia and Cetartiodactyla.

METHODS

GenBank was searched for 2'-5' oligoadenylate synthetase genes and pseudogenes using the Blast program (Altschul et al., 1997). The Dotmap program was used to visually identify additional exons and pseudogenes. Partial rat, cattle and pig mRNA sequences predicted by the bl2seq program were used to design gene-specific primers, which were then utilized to amplify full-length sequences from commercial cDNAs. Multiple alignments of the full-length sequences were performed with the Align program. Manual editing was used to synchronize the positions of gaps between pair-wise alignments. The phylogenetic tree was constructed using the njtree program as described previously (Perelygin et al., 2005). The confidence of each node was estimated using a bootstrap algorithm (Zharkikh, Li, 1995).

RESULTS AND DISCUSSION

Eight new rat 2'-5' oligoadenylate synthetase genes and two pseudogenes were predicted in the NW 047376 supercontig. Predicted sequences were used to amplify six full-length rOas1 cDNAs. Maps of mouse and rat Oas1 clusters were compared (Fig. 1) to suggest the orthologous relationships between individual genes. These relationships were validated by comparison of Oas1 tissue expression patterns in mice and rats.

A new pig pOAS1Y gene and two new cattle genes, cOAS1Y and cOAS1Z, were also identified. Exact orthologous relationships between cattle and pig OAS1 genes could not be established due to the lack of a pig OAS1 map. New and previously reported sequences from even-toed ungulates and rodents were used to build a phylogenetic tree (Fig. 2). Two parts of the tree are in discordance with the map of the rodent Oas gene cluster. First, the rodent Oas1c and Oas1d genes are separated from each other on the map by six other genes and show a good orthological correspondence between mouse and rat. However, comparison of their sequences showed that the differences between paralogous rodent Oas1c and Oas1d genes in each species are smaller than those between the corresponding orthologs. Second, the rOas1g gene is more similar to rOas1i than to mOas1g. Therefore, in two cases, the similarities between rodent Oas1 orthologous cDNA sequences are lower than those between Oas1 paralogous sequences within each species. The observed deviation from the principle of divergent evolution, which assumes that more recent gene duplication produces more similar sequences, supports concerted
evolution of paralogous Oas1 genes in rodents presumably due to gene conversion. Even in cases where the tree topology was correct, the effect of gene conversion resulted in non-uniform occurrence of convergent substitutions. There were 151 and 159 convergent substitutions between paralogous genes in mice and rats, respectively, whereas only 111 convergent substitutions were observed between non-orthologous rodent Oas1 genes. Comparison of the substitution distributions along the nucleotide sequence in different evolutionary lineages can also unambiguously demonstrate the presence of gene conversion. A new approach was developed to quantify the differences between distributions of substitutions. We assumed that the substitution rate at any particular nucleotide position does not change among the genes and species being considered. However, different nucleotide positions may have different substitution rates. For each pair of genes, the cumulative distribution of differences can be built by assigning to each position \( i \) the number of differences \( N_i \) located to the left of this position normalized by the total number of differences \( N \) observed between the two genes, i.e., \( F_i = N_i / N \). The distributions for two pairs of sequences are then compared by a non-parametric Kolmogorov-Smirnov (KS) test, which estimates the maximum difference between the two distributions. The distributions are expected to be similar when there are no evolutionary events or restrictions differentially affecting the occurrence of sequence changes in different lineages. For the cOAS1Y/cOAS1Z pair of genes, the KS test was not significant (\( P = 0.512 \)). The cOAS1X/cOAS1Z pair showed a higher substitution rate (\( P = 0.069 \)) in the proximal half of the gene compared to the distal half. For the cOAS1X/cOAS1Y pair, the result of the KS test was highly significant (\( P = 4.9 \times 10^{-9} \)). This pair has a low substitution rate in the first half of the sequence and a significantly higher rate in the second half. For the cOAS1X/cOAS1Z and cOAS1X/cOAS1Y pairs, the point at which the substitution rates switch is located at approximately the same position, namely position 500 of the cOAS1X coding sequence (difference 0.47). Therefore, cOAS1X is more similar to cOAS1Y before position 500, but cOAS1X is more similar to cOAS1Z after this position. The simplest explanation of this observation is that cOAS1X originated by conversion between the cOAS1Z and cOAS1Y genes and then diverged. A similar analysis was performed for the pig genes, pOAS1X and pOAS1Y. The distribution of differences between these two genes reached 0.47 at position 940 (\( P = 1.4 \times 10^{-9} \)), which could be due to a recent transfer of a large portion from one gene to another. The presence of interlocus transfers between duplicated copies of genes violates the basic principle of phylogenetic inference, i.e. the principle of divergence of related sequences, suggesting that gene conversion is the major mechanism by which homogenization occurs in paralogous 2'-5' oligoadenylate synthetase-1 genes.

Figure 2. A phylogenetic tree of mouse (m), rat (r), cattle (c) and pig (p) Oas1 genes.
ACKNOWLEDGEMENTS

This work was supported by Public Health Service grants AI045135 from the Institute of Allergy and Infectious Diseases, National Institutes of Health and CI000216 from the National Center for Infectious Diseases, Centers for Disease Control and Prevention.

REFERENCES


