Genome Analysis

Duplication, degeneration and subfunctionalization of the nested synapsin–Timp genes in Fugu

Wei-Ping Yu1, Sydney Brenner1,2 and Byrappa Venkatesh1

1Institute of Molecular and Cell Biology, 30 Medical Drive, Singapore 117609 Singapore
2Salk Institute, 10010 North Torrey Pines Road, La Jolla, San Diego, CA 92037-1099, USA

The genes encoding synapsin (Syn) and the tissue inhibitor of metalloproteinase (Timp) exhibit a nested organization that is conserved in fruit fly and vertebrates. Analysis of the human and Fugu genomes show that the evolution of Syn–Timp gene families is characterized by duplications, secondary loss and the partitioning of ancestral functions. Of particular interest are two duplicate Syn–Timp loci in Fugu that have accumulated complementary degenerate mutations such that each Syn duplicate produces one of the two transcripts generated from the single ancestral gene, and one of the Timp genes is lost.

The fruit fly genome contains a single alternatively spliced synapsin (Syn) gene [1]. Interestingly, the sole fruit-fly gene encoding the tissue inhibitor of metalloproteinase (Timp) is nested within the intron of Syn gene in reverse orientation [2]. Vertebrates contain several genes encoding Syn and Timp, some nested and some independent, indicating that the ancestral Syn–Timp locus has undergone duplications during vertebrate evolution. Thus, these gene families constitute an interesting and informative example to understand the history of gene duplications and the fate of duplicate genes during vertebrate evolution.

The human genome contains three synapsin genes, SYN1, SYN2 and SYN3, each of which include a TIMP gene [2], as well as a fourth TIMP gene (TIMP2) that is not associated with a SYN gene (http://genome.ucsc.edu). We searched the draft Fugu genome sequence (http://www.fugu-sg.org) for members of Syn and Timp families and identified a pair of nested Syn–Timp genes, and a single Syn and two Timp genes that are independent. We confirmed their identity (orthology) based on their genomic organization and phylogenetic analysis of amino acids.

![Fig. 1. Phylogenetic analysis of Syn (a) and Timp (b) sequences. Protein sequences were aligned using ClustalX and the phylogenetic trees were generated by the Neighbor-Joining method based on the maximum likelihood of distances. Numbers at the nodes are bootstrap values for 1000 replicate analyses. Fugu Syn genes were identified by BLAST searching the draft Fugu genome sequence (http://www.fugu-sg.org) with human synapsins. cDNA for Fugu Syn and Timp genes were cloned and sequenced to confirm the exon–intron organization. Syn and Timp sequences for other species were retrieved from the NCBI database (http://www.ncbi.nlm.nih.gov).](http://tigs.trends.com/0168-9525/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved. doi:10.1016/S0168-9525(03)00048-9)
acid sequences (Figs 1,2), and propose a model for the evolution of vertebrate Syn and Timp genes (Fig. 3). According to this model, the invertebrate Syn–Timp locus was duplicated at least three times before the divergence of the tetrapod and teleost lineages. The independent Timp2 could be the result of either degeneration of the flanking Syn gene following the locus duplication or partial duplication of the locus involving only the Timp sequence. Subsequently, the Syn2–Timp4 and Timp2 loci were again duplicated in the fish lineage, followed by secondary loss of two Timp genes.

The conserved synteny between the human SYN2 and
TIMP2 loci with their orthologous loci in the Fugu (Fig. 2) indicate that the duplicate Fugu Syn2A and Syn2, and Timp2A and Timp2B genes arose as a result of segmental duplications. Comparative studies of the teleost and mammalian genomes have indicated that the teleost lineage might have undergone a whole-genome duplication after it diverged from the mammalian lineage \[5,4\]. The duplications of Syn2 and Timp2 loci observed in the Fugu could be part of the whole-genome duplication proposed to have occurred in the teleost lineage.

A comparison of the complement of Syn and Timp genes in the Fugu and humans indicates that Fugu contains an additional Syn gene (Syn2B) that could be redundant. However, a careful analysis of the duplicate Fugu Syn2 gene sequences shows that the function of the single human SYN2 gene has been partitioned between the duplicate Fugu genes. The human SYN2 gene comprises 13 exons, which are alternatively spliced to generate isoforms 2a and 2b \[5\] (Fig. 4). The Fugu Syn2A comprises 13 exons, similar to the human SYN2a splice variant, whereas Syn2B contains only 11 exons analogous to those used by human SYN2b splice variant (Fig. 4a). Indeed, the C-terminal 22 residues that are unique to the human SYN2b isoform are highly conserved in Fugu Syn2B (Fig. 4b).

Thus, the two isoforms generated by alternative splicing of the single human gene are encoded separately by the two duplicate Fugu genes. This could have been accomplished by degenerative mutations experienced by both the Fugu genes. The Fugu Syn2A has lost the potential to make the alternate isoform similar to human SYN2b owing to a point mutation in the 5' end of 11th intron \((T\rightarrow A)\) that has changed a Cys codon to a stop codon (Fig. 4b) and the loss of the alternate polyA signal within the 11th intron. In addition, Fugu Syn2B has lost the potential to generate transcripts analogous to the human SYN2a splice variant because of the degeneration of exons 12 and 13. This might have been preceded and prompted by the loss of splice donor site of intron 11 of Syn2B (GT → GC; Fig. 4b). In another duplicated teleost gene, mitf (mitfa and mitfb), it appears that complementary degeneration of alternative 5' exons has resulted in the duplicates separately encoding two of the splice variants of the ancestral gene. Furthermore, the duplicate pair together recapitulate the expression pattern of the ancestral gene \[6,7\].

The classical model for the fate of duplicate genes predicts that one member of the duplicate pair is generally ‘lost’ through accumulation of degenerative mutations, or occasionally preserved owing to beneficial mutations that confer a novel function or an improvement of an existing function. The recently proposed third alternative, termed subfunctionalization, predicts that both members accumulate degenerative mutations such that the function of the ancestral gene is partitioned between the duplicates \[8,9\]. In the present example, we see all the three processes acting on duplicated Syn–Timp loci to shape the evolution of a gene family in vertebrates. Of particular interest is the

---

**Fig. 4.** Comparisons of the genomic organizations and sequences of the human SYN2 gene, and Fugu Syn2A and Syn2B genes. (a) Genomic organization of human SYN2 gene, and Fugu Syn2A and Syn2B genes. Exons are shown as rectangles. Only the last exon of isoform SYN2b is shown. Fragments of Fugu Syn2A gene sequence were found on scaffolds #3205 (27.9 kb) and #1288 (68.8 kb) (http://www.fugu-sg.org, and the sequence was completed by joining the two scaffolds by PCR. Gaps within the scaffold were filled by PCR. The Syn2B gene is located on scaffold #1050 (78.3 kb). (b) Sequences around the 11th exon of human SYN2b owing to a point mutation in the 5' end of 11th intron \((T\rightarrow A)\) that has changed a Cys codon to a stop codon (Fig. 4b) and the loss of the alternate polyA signal within the 11th intron. In addition, Fugu Syn2B has lost the potential to generate transcripts analogous to the human SYN2a splice variant because of the degeneration of exons 12 and 13. This might have been preceded and prompted by the loss of splice donor site of intron 11 of Syn2B (GT → GC; Fig. 4b). In another duplicated teleost gene, mitf (mitfa and mitfb), it appears that complementary degeneration of alternative 5' exons has resulted in the duplicates separately encoding two of the splice variants of the ancestral gene. Furthermore, the duplicate pair together recapitulate the expression pattern of the ancestral gene \[6,7\].
process of subfunctionalization acting on the duplicated Syn2–Timp4 loci in the Fugu lineage such that the complementary degenerate mutations have completely eliminated one of the extra genes (Timp associated with Syn2B), and partitioned the function of the two Syn2 genes. Contrary to the predictions of the classical model for the fate of duplicate genes, vertebrates, particularly teleosts, seem to have retained a large proportion of gene duplicates. It is estimated that teleosts such as zebrafish have retained at least 20% of the duplicate gene pairs following the proposed whole-genome duplication in the teleost lineage [10]. The present example, together with the instances of subfunctionalization of expression domains [8] and splice variants [6,7] described in zebrafish and other teleosts seem to suggest that degeneration and complementation may be a much more common mechanism than previously thought in retaining duplicate genes.

References
2 Pohar, N. et al. (1999) Invertebrate tissue inhibitor of metalloproteinase: structure and nested gene organization within the synapsin locus is conserved from Drosophila to human. Genomics 57, 293–296
5 Porton, B. et al. (1999) Cloning of cDNA encoding human synapsins IIa and IIb. DNA Seq. 10, 49–54