Phylogenetic analysis of dinoflagellate luciferase genes from seven species: a possible role for conserved nucleotides in the circadian regulation of protein synthesis

Liu L and Hastings JW
Dept of Molecular & Cellular Biology, Harvard University, Cambridge, MA 02138 USA

The structures and sequences of luciferase genes from five bioluminescent species, *Alexandrium affin, Alexandrium tamarensis, Pyrocystis noctiluca, Pyrocystis fusiformis*, and *Protoceratium reticulatum* have been determined and compared with two previously reported ones from *Lingulodinium polyedrum* and *Pyrocystis lunula*. All have a structure that is unique to dinoflagellate luciferases: three homologous catalytic domains preceded by an N-terminal region of unknown function. Both pairwise comparison and phylogenetic inference indicate that the similarity of the corresponding individual domains between species is greater than that of the different domains within the polypeptide. Trees constructed from each of the three individual domains are congruent with the tree of the full-length coding sequence, suggesting coevolution of three domains. Luciferase and ribosomal DNA trees both indicate that the *L. polyedrum* luciferase represents the most primitive form. Synonymous but not nonsynonymous substitution rates of the central region of the intramolecularly conserved domains vary considerably among the lineages, being greatly reduced from the flanking regions in *L. polyedrum*, mildly in *Protoceratium* and very little if any in *Alexandrium* and *Pyrocystis* species. The lineage-specific constraints of synonymous substitution in the central region of the domains correlate inversely with the content of GC3 (~87% for *Lingulodinium, 84% for Protoceratium*, and only ~65% for *Pyrocystis*). The difference in GC3 content can be accounted for by the biased usage toward C-ending codons at the degenerate sites.