

Genome-wide analysis reveals conserved promoter regions between single exon gene and multiple exon gene orthologs

Mbaye Tine¹, Heiner Kuhl², Peter Teske³, and Richard Reinhardt¹

¹Max Planck Institute for Plant Breeding Research Department Plant-Microbe Interactions

²Leibniz Institute of Freshwater Ecology and Inland Fisheries

³University of Johannesburg Faculty of Science

August 20, 2020

Abstract

Several studies have attempted to understand the origin and evolution of single exon genes (SEGs) in eukaryotic organisms including fishes, but few have examined the functional and evolutionary relationships between SEG and multiple exon gene (MEG) orthologs, in particular the conservation of promoter regions. Given that SEGs originate via the reverse transcription of mRNA from a “parental” MEG, such comparisons may enable identifying evolutionarily-related SEG/MEG orthologs, which might fulfill equivalent physiological functions. Here, the relationship of SEG proportion with MEG count, gene density, intron count and chromosome size was assessed for the genome of sea bass, *Dicentrarchus labrax*. Then, SEGs with an MEG parent were identified, and promoter sequences of SEG/MEG orthologs compared, to identify highly conserved functional motifs. The results revealed a total proportion of 1585 (8.3%) SEGs evenly distributed in the sea bass genome, which was correlated with MEG count but not with gene density. These results suggest that SEGs are continuously and independently generated after species divergence over evolutionary time, as is evident from the significant proportion of SEGs with an MEG parent. Functional annotation showed that the majority of SEGs are functional, as is evident from their expression in RNA-seq data used to support homology-based genome annotation. Differences in 5'UTR and 3'UTR lengths between SEG/MEG orthologs observed in this study may contribute to gene expression divergence between them, and therefore lead to the emergence of new SEG functions. The comparison of nonsynonymous to synonymous changes (K_a/K_s) between SEG/MEG parents showed that 74 of them are under positive selection ($K_a/K_s > 1$; $P = 0.0447$). An additional fifteen of SEGs with a MEG parent have a common promoter, which implies that they are under the influence of common regulatory networks and may be involved in equivalent functions.

Hosted file

Manuscript.doc available at <https://authorea.com/users/352547/articles/476745-genome-wide-analysis-reveals-conserved-promoter-regions-between-single-exon-gene-and-multiple-exon-gene-orthologs>



