Measures of Phylogenetic Support for Phylogenomic and the Whole Genomes of Two Lungfish Restate Lungfish and Origin of Land Vertebrates

Authors : Yunfeng Shan, Xiaoliang Wang, Youjun Zhou

Abstract : Whole-genome data from two lungfish species, along with other species, present a valuable opportunity to reassess the longstanding debate regarding the evolutionary relationships among tetrapods, lungfishes, and coelacanths. However, the use of bootstrap support has become outdated for large-scale phylogenomic data. Without robust phylogenetic support, the phylogenetic trees become meaningless. Therefore, it is necessary to re-evaluate the phylogenies of tetrapods, lungfishes, and coelacanths using novel measures of phylogenetic support specifically designed for phylogenomic data, as the previous phylogenies were based on 100% bootstrap support. Our findings consistently provide strong evidence favoring lungfish as the closest living relative of tetrapods. This conclusion is based on high gene support confidence with confidence intervals exceeding 95%, high internode certainty, and high gene concordance factor. The evidence stems from two datasets containing recently deciphered whole genomes of two lungfish species, as well as five previous datasets derived from lungfish transcriptomes. These results yield fresh insights into the three hypotheses regarding the phylogenies of tetrapods, lungfishes, and coelacanths. Importantly, these hypotheses are not mere conjectures but are substantiated by a significant number of genes. Analyzing real biological data further demonstrates that the inclusion of additional taxa diminishes the number of orthologues and leads to more diverse tree topologies. Consequently, gene trees and species trees may not be identical even when whole-genome sequencing data is utilized. However, it is worth noting that many gene trees can accurately reflect the species tree if an appropriate number of taxa, typically ranging from six to ten, are sampled. Therefore, it is crucial to carefully select the number of taxa and an appropriate outgroup while excluding fast-evolving taxa as outgroups to mitigate the adverse effects of long-branch attraction (LBA) and achieve an accurate reconstruction of the species tree. This is particularly important as more whole-genome sequencing data becomes available.

Keywords : gene support confidence (GSC), origin of land vertebrates, coelacanth, two whole genomes of lungfishes, confidence intervals
Conference Title : ICBCBBE 2024 : International Conference on Bioinformatics, Computational Biology and Biomedical Engineering
Conference Location : New York, United States
Conference Dates : January 29-30, 2024