



Taxon-Tree Viewer: a visualization tool for exploring distribution of Pin-II PIs across species

Nikhilesh K. Yadav

Information Systems Area, Indian Institute of Management, (IIM), Indore, India

Abstract

Taxonomy is essential for the classification of sequences to help researchers to understand and communicate biological information. PIN2DB a dedicated biological database of Potato inhibitor type-II protease inhibitors (Pin-II type PI) family, covers functional annotations of each available Pin-II type PI sequence. As captured in the database, Pin-II PIs are widely distributed across species with multiple levels of hierarchies and complex relationships. Interactive Visualization can aid in exploring and discerning meaning from such complex data. To meet this requirement, we have developed “Taxon- TreeViewer” (<https://pinir.ncl.res.in/Taxon-TreeView/>) an interactive visualization tool to explore the diversity of Pin-II PI sequences in PIN2DB. It develops an interactive tree which is laid out as a node-link diagram that displays the hierarchical relationships across Pin-II PI taxonomy. Each node label displays name of taxon, and its PIs count. Users can expand or collapse any node by just clicking it. The visualization also provides the users with the facility to orient the tree either horizontally or vertically. To improve the legibility of the visualization the Taxon-TreeViewer provides the facility to select the root node taxon of the tree. Taxon-Tree Viewer is implemented using Java Script and makes use of D3s hierarchical layouts. It is derived from Mike Bostock’s Collapsible Tree example. In future we plan to display distribution of more sequence features across species for each Taxon node. We are also working on a generic open source variant of Taxon- TreeViewer which could be easily implemented for any biological database with taxonomy classification.

Keywords: bioinformatics; data visualization; hierarchical data; protease inhibitor; taxonomy