An Analysis System for Integrating High-Throughput Transcript Abundance Data with Metabolic Pathways in Green Algae

Authors : Han-Qin Zheng, Yi-Fan Chiang-Hsieh, Chia-Hung Chien, Wen-Chi Chang

Abstract : As the most important non-vascular plants, algae have many research applications, including high species diversity, biofuel sources, adsorption of heavy metals and, following processing, health supplements. With the increasing availability of next-generation sequencing (NGS) data for algae genomes and transcriptomes, an integrated resource for retrieving gene expression data and metabolic pathway is essential for functional analysis and systems biology in algae. However, gene expression profiles and biological pathways are displayed separately in current resources, and making it impossible to search current databases directly to identify the cellular response mechanisms. Therefore, this work develops a novel AlgaePath database to retrieve gene expression profiles efficiently under various conditions in numerous metabolic pathways. AlgaePath, a web-based database, integrates gene information, biological pathways, and next-generation sequencing (NGS) datasets in Chlamydomonasreinhardtii and Neodesmus sp. UTEX 2219-4. Users can identify gene expression profiles and pathway information by using five query pages (i.e. Gene Search, Pathway Search, Differentially Expressed Genes (DEGs) Search, Gene Group Analysis, and Co-Expression Analysis). The gene expression data of 45 and 4 samples can be obtained directly on pathway maps in C. reinhardtii and Neodesmus sp. UTEX 2219-4, respectively. Genes that are differentially expressed between two conditions can be identified in Folds Search. Furthermore, the Gene Group Analysis of AlgaePath includes pathway enrichment analysis, and can easily compare the gene expression profiles of functionally related genes in a map. Finally, Co-Expression Analysis provides co-expressed transcripts of a target gene. The analysis results provide a valuable reference for designing further experiments and elucidating critical mechanisms from high-throughput data. More than an effective interface to clarify the transcript response mechanisms in different metabolic pathways under various conditions, AlgaePath is also a data mining system to identify critical mechanisms based on high-throughput sequencing.
Keywords : next-generation sequencing (NGS), algae, transcriptome, metabolic pathway, co-expression
Conference Title : ICBCSB 2014 : International Conference on Bioinformatics and Computational Systems Biology
Conference Location : Stockholm, Sweden
Conference Dates : July 14-15, 2014