Topic: **Sequence Analysis**

PRELIMINAR STUDY AND MINING OF CHITINASES IN ENVIRONMENTAL METAGENOME

A Fróes[1], DC Buriticá[1], A Davila[2]

[1]*Programa de Pós-Graduação em Biologia Computacional e Sistemas (IOC)*
[2]*Institute Oswaldo Cruz*

Chitin, a homopolymer of β-1,4-linked N-acetyl-D-glucosamine residues, is the most abundant renewable resource in the environment after cellulose. It is widely distributed in nature as a structural component of crustacean, fungi, protozoa and insects skeletons. Chitinases (EC 3.2.1.14) are glycosyl hydrolases which catalyze the degradation of chitin and are present in a wide range of organisms such as bacteria, fungi, insects, plants, and animals. Based on sequence amino-acid similarities chitinases can be divided in two different families (18 and 19) of glycosyl hydrolases and do not share similarities in amino acid sequence or three-dimensional structure. Family 18 encompasses chitinases found in bacteria, fungi, viruses, and animals, and plant chitinases. Family 19 includes chitinases of plants and found in some *Streptomyces* and other Actinobacteria strains. Those enzymes are being studied aiming the development of new compounds for fungi biocontrol and drugs like topic lotions for treating dermatophytosis and treatment of fungal infections, since chitin present in cell wall confers rigidity and protection to fungi against human immunity. The present work aims to analyze conserved domains of chitinases and verify if they are in agreement with their function and evolution in different organisms. A search of protein sequences of chitinases of families 18 and 19 was performed on the environmental metagenome database at NCBI then we provided some preliminary annotation for those results. The search consisted of 3 datasets: Bacteria set (31 sequences), Plants set (28 sequences) and Fungi set (30 sequences). For each dataset multiple alignments were generated with MAFFT then an HMM model created with HMMER. Each model was calibrated (hmmcalibrate) and used to query for similarity against the environmental metagenome database (Local DataBase). The sequences with best hits of similarities were considered. Those were parsed by an in-house perl script which created a table in MS Excel. Finally, those sequences were run analyzed by BLAST against SWISS-PROT and Refseq DB. The preliminary results indicate that those sequences are probably chitinases. Sequences that (i) showed higher similarity with BLAST, along with (ii) the best HMMER hits found in metagenome database and (iii) those used to build the hmm model were used for a phylogenetic tree reconstruction using neighbor joining algorithm. These preliminary results create a basis to begin a more comprehensive search for chitinases, and then do some functional annotation of them, which later will allow detailed studies of function and evolution of chitinases involved in metagenomic studies. Supported by: CNPq, FIOCRUZ