Target Discovery of enzymes involved in degradation of lignocellulosic biomass from Sugar Cane Soil Metagenome

Alvarez, T. M1; Paixão, D. A. A1; Franco Cairo, J. P1; Ruller, R1; Buchli, F1; Ribeiro, D.A1; **Paes Leme, A.F2; Prade, R. A1,3; Squina, F.1**

1- Laboratório Nacional de Ciência e Tecnologia do Bioetanol (CTBE) e 2- Laboratório Nacional de Biociências (LNBio) do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM), Campinas, Brasil, 3- Department of Microbiology & Molecular Genetics, Oklahoma State University, Stillwater 74078 OK United States.

Metagenomics is a new approach that aims to overcome the barriers of microorganism culturability and get full access to microbial biodiversity from environmental samples through molecular biology's tools. The metagenomic DNA represents an enormous genetic and biological pool for mining of new genes, entire pathways and compounds with wide applications in biotechnology. The objective of our work is the generation of a toolkit of lignocellulolytic enzymes from soil metagenome, which could be used as players for the development of strategies for second generation ethanol production. We have focused on soil DNA derived from sugarcane field after harvest, where straws were left on the ground, an environment where it is expected that microbial community is enriched by microorganisms involved in lignocellulosic biomass degradation. Sugarcane Bagasse-Degrading-Soil (SBDS) metagenome was massively-parallel-454-Roche-sequenced. Comparison of 241.630 Mbp obtained by sequencing of SBDS metagenomic DNA revealed over 500,000 unique ORFs with 225,417 significant hits with proteins deposited in GenBank. The majority of ORFs were from bacteria (89.7%) and archaea (1.34%) and only 0.39% were from eukaryotic-origin. We identified a full repertoire of genes with significant match to glycosil hydrolases catalytic domain and carbohydrate-binding modules. A subset of 12 putative carbohydrate-active genes is being target for functional analysis. Soil metagenomic´s libraries into pUC19 were also constructed, which are been screened through enzymatic assays using chromogenic substrates. Additionally, aiming at the creation of a comprehensive catalog of proteins, with activities needed to completely degrade natural lignocellulosic substrates, we have been conducting microbial community enrichment experiments on SBDS under semi-anaerobic and forced-aerobic conditions using plant biomass (bagasse) as sole carbon source. Total proteins from the culture filtrate were collected and shotgun analyzed by Orbitrap/MS/MS to identify extracellular proteins. Mass spectral data were analyzed in MASCOT using a peptide database created from the predicted ORFs derived from the metagenome sequence.

Supported by Fapesp