# Comparative Transcriptome Analysis and Genome Assembly of *Fusarium oxysporum* f. sp. *cubense*

M.A. Dita[1], R. Herai[2], C. Waalwijk[3], M. Yamagishi[4], P. Giachetto[4], G. Ferreira[6], M. Souza[6] and G.H.J. Kema[3]

1Bioversity International, Turrialba, Costa Rica; 2State University of Campinas, Brazil; [3]Plant Research International B.V., PO Box 16, 6700 AA Wageningen, The Netherlands; 4Embrapa Informática Agropecuária, Av. André Tosello, 209 - Barão Geraldo, Caixa Postal 6041- 13083-886 - Campinas, São Paulo; Brazil; 5Embrapa Amazônia Ocidental, Rodovia AM-010, Km 29, Zona Rural - CEP 69010-970 Caixa Postal 319 – Manaus, Amazonas, Brazil; 6Embrapa Agroenergy, Parque Estação Biológica s/n Brasïlia, DF – Brazil, CEP 70770-901

**Keywords**: Panama disease, banana, genomic and transcriptomic analyses

*Fusarium oxysporum* f. sp. *cubense* (Foc), the causal agent of Fusarium wilt of banana, is a highly destructive and genetically diverse pathogen. Despite its economic importance, genomic information of Foc is poor and no transcriptomic analyses have been reported so far. By using 454 sequencing technology, we generated >2.5 million expressed sequenced tags (ESTs) from four Foc strains representing four vegetative compatibility groups (VCGs) and races that infect banana: race 1 (R1, VCG unknown but different from the others here described), race 2 (R2, VCG 0124), subtropical race 4 (SR4, VCG 0120) and tropical race (TR4, VCG 01213). The ESTs were obtained from libraries prepared from mRNA extracted from three physiological states (mycelia, conidia and germinated conidia), which were pooled at a 2:2:1 ratio. Most genes are represented in all libraries, but *in silico* comparative analyses identified a set of unique ESTs for each race (689 for R1, 974 for R2, 296 for SR4 and 555 for TR4), which constitute excellent candidates for diagnostics development, future plant-pathogen interaction studies and functional analyses. In subsequent analyses, a 40x sequencing-coverage (Illumina single reads) of TR4 genomic DNA was assembled in a *de novo* based methodology, resulting in a 29-kb N50 (463 contigs). Preliminary analyses show a high colinearity of EST and genomic data that significantly contributes to the quality of the assembly. Applications of these data will be further discussed.