

Analysis of Genomic Data of *Pyricularia Oryzae* Species Causing Rice Blast

A.C. Samuel^{1*}, L.S. Nawarathna¹ and H.A.C.K. Ariyaratne²

¹*Department of Statistics and Computer Science*

²*Department of Botany*

Faculty of Science, University of Peradeniya, Peradeniya 20400, Sri Lanka

**s15448@sci.pdn.ac.lk*

Blast, caused by the fungus *Pyricularia oryzae* (*P. oryzae*), is the most destructive disease of rice prevalent globally. The rice quantity lost annually to this disease, which is over 30% of its annual production, could sustain approximately 60 million people. Fungicides and other means for disease control are ultimately futile due to the development of fungicidal resistance and adaptation among pathogen populations. Computational methods have been increasingly applied to design accurate and affordable biological experiments. This analysis identified regions of the genome that are discriminative for the selected fungus and hence can be used for pathogen diagnosis. Comparative analysis was performed with four other fungal pathogens that cause the most economically harmful rice diseases in Sri Lanka, namely; *Rhizoctonia solani*, *Cochiobolus miyabeanus*, *Ustilaginoidia virens* (*U. virens*) and *Sarocladium oryzae* (*S. oryzae*). This study used bioinformatics techniques such as pairwise alignments to enumerate mutations. The study was restricted to the 1st chromosome of the genome due to limited processing storage. Markov-models and the Needleman-Wunsch algorithm were employed extensively. Based on the Kruskal-Wallis test, there was no difference in the distributions of proportions of each of the four nucleotides within the genome of *P. oryzae* at 5% significance level. However, the regions from the 300,001bp to 400,000bp; 1,400,001bp to 1,500,000bp and 4,100,001bp to 4,200,000bp contained high proportions of adenine and thymine (AT rich) and low proportions of guanine and cytosine (GC poor) relative to the rest of the genomic divisions. Pairwise alignments between the 1st chromosome of *P. oryzae* and randomly selected regions of *U. virens* and *S. oryzae*, identified distinctive regions in *P. oryzae* from the 720,001bp to 730,000bp pointing to unique genomic sections in comparison to other rice pathogenic fungi. Further analysis of these *P. oryzae* genomic regions will enable developing molecular tools to distinguish and diagnose the pathogen.

Keywords: Markov models, Needleman-Wunsch algorithm, Pairwise alignment, *Pyricularia oryzae*, Rice blast