

README.Rmd

Justyn and Katelyn

12/3/2021

Gene Duplications and the Evolution of Phytochrome in Early Evolving Land Plants

Abstract: 450 million years ago land was dominated by small nonvascular plants that were exposed to sunlight. 300 million years ago, the evolution of vascular plants led to the origin of the first forest with complex canopies and deeply shaded habitats. Around 100 million years ago, complex canopy shade was a major factor in early land plant environments. The complex canopies formed by forest collected most of the available light, leaving only red and far-red wavelengths reaching the understory plants. Plants had to evolve complex phytochrome responses to the differences in light and dark environments. The two proteins that evolved were called Pfr (far red) absorbing light at 730nm and Pr (red) absorbing light at 670nm. The ratio between the presence of red and far-red proteins is crucial for plants survival in different light environments. Phytochrome responses to shade (low R/FR) include (1) inhibition of germination and (2) initiation of shade avoidance responses (less roots produced, early flowering). Phytochrome A (associated with far red light) and Phytochrome B (associated with red light) are the two phytochromes that will be of interest in this study. Phytochrome A and Phytochrome B have been found extensively in seed plants and ferns. Going back further in the plant evolutionary timeline, there is a knowledge gap regarding phytochromes A and B presence in primitive nonvascular plant species. The group of interest is Bryophyta, which has little to no insight into how Phytochrome A and Phytochrome B have evolved. Understanding the prevalence of PHY A and PHY B in the non-vascular clade of Bryophyta is important for gaining insight into the transition periods that cemented the phytochromes from early nonvascular plants to the vascular land plants we see today. Sarah Mathews and Fay Wei Li, as well as GenBank will be the source of the data sets used. GenBank will be used to get data for the organismal clade sequences. The alignments collected from GenBank will then be used in BLAST to find the specific genes of interest. The alignments will then be read, explored, and assessed. Gene trees will be generated in MAFT and inferred in R, APE, and IQ-TREE. Further analysis will be run in Datamonkey to test for positive selection among different branches of the moss phylogenetic tree.