

Exploring diversity in the *Oryza* genus utilizing molecular markers analyzed with distance based and model based statistics

Laura Kelly Vaughan, Ph.D.
Department of Biochemistry, Texas A&M University

Abstract:

Modern molecular biology and biochemical techniques such as microsatellite technology offer powerful tools for the investigation of genetic diversity. Microsatellites are short repeated sequences of five to eight nucleotides in length that can be found in both coding and non-coding regions of the genome. The number of repeats, which become variable, can be used to distinguish between individuals that are closely related. With over 90% of the diversity for a given genus accounted for by wild relatives of the cultivated species, the investigation of the levels of diversity between wild relatives and crops becomes important for the identification of important traits such as disease resistance or increased yield.

My graduate work has focused on a yet unexplored source of genetic variation popularly known as red rice. Red rice is a weedy relative of cultivated rice that has traditionally been considered a useless weed and, in the US, been classified as *Oryza sativa* ssp. *indica*. This assumption was tested by analyzing red rice samples, *Oryza sativa* ssp. *indica*, *Oryza sativa* ssp. *japonica* cultivars, and a collection of wild species *Oryza rufipogon* and *Oryza nivara* with a set of microsatellite markers distributed across all 12 chromosomes. The genetic distance matrix was calculated and then subjected to hierarchical cluster analysis using the average linking method to define related groups and a multidimensional scaling analysis was performed to give a two-dimensional representation of the data. To ensure confidence in the analysis we also used model based relationship statistical analysis to assign samples to groups.

The results of both types of statistical analysis complement each other and clearly demonstrate that the traditional classification of red rice is inadequate. While a number of red rice samples are related to *Oryza sativa* ssp. *japonica* or ssp. *indica*, a portion of the samples collected from the southern United States are very closely related to an accession of the noxious weed, *Oryza rufipogon*. Red rice samples from three states were identified that are identical for all 18 markers with the *Oryza rufipogon* accession IRGC 105491. Currently all accessions of *Oryza rufipogon* are listed as noxious weeds by the USDA and can only be grown in highly controlled quarantine conditions. This suggests that red rice ecotypes may be a useful source of genetic variation for rice improvement as has been suggested for *Oryza rufipogon* IRGC 105491.

During further investigation of this relationship, the apparent misclassification of many accessions of the *Oryza* samples in the United States Small Grains National Seed Bank was discovered. Statistical analysis of the molecular marker data clearly groups some members of both *Oryza rufipogon* and *Oryza nivara* with the cultivated species *Oryza sativa* ssp. *indica*. Other, independently verified *Oryza rufipogon* samples are clearly distinct from both *Oryza nivara* and *Oryza sativa*. While this may help clear up the classification of red rice, it presents many new problems in regard to the national seed bank.

Utilization of the opportunities represented by the diverse germplasm stored in the seed repositories depends on both accurate classification of the accessions and the quantification of the actual diversity of those samples. This research illustrates the importance of applying new techniques to ensure that these importance resources are used to their fullest extent. The use of statistical analysis provided the means to present the molecular marker data in a manner that can be readily utilized. This work also represents an excellent example of the integration of theoretical and applied research in the fields of Biochemistry, Genetics, and Weed Science.