

November 9, 2015

Dr. Marion Röder
Academic Editor
PeerJ

Dear Dr. Röder,

First of all we would like to apologise for the delay in getting the revised version of the manuscript to you. Reviewers suggested an increase in sample sizes, which required further laboratory work and reanalyses. We hope we have addressed the reviewers comments which have clearly improved the quality of the manuscript. The specific comments raised by the reviewers are address in the following paragraphs.

The main concern raised by the reviewers focused on sample size per population.

Natural populations of *O. glumaepatula* in Costa Rica are extremely rare. Populations in Medio Queso River (MQ) are currently under strict protection given that the area is threatened by the construction of a new road. Several environmental agencies are trying to prevent the demise of this population and thus destructive sampling is strictly controlled. Acquiring permits to increase our *O. glumaepatula* sample size at this time would very likely fail or require considerable time. This is the main reason why we used plants growing in the greenhouses at Universidad de Costa Rica for our analyses. Nevertheless, the reviewers main concerns were based on the ability to accurately estimate allele frequencies to determine differences and the likelihood of gene flow among sites and species. Given the lack of genetic differences between closely adjacent populations within MQ river, we grouped populations that were separated by less than 500 meters into three sub-populations within the MQ river. They are almost equidistant from each other along the river, increasing sample size per sub-population to approximately 15-20 individuals, which is comparable to other genetic diversity studies in this species. These sample sizes should provide accurate allele frequency estimates. Given the lack of genetic structure among populations, a Wahlund effect is unlikely to influence our conclusions. The three resulting sub-populations are positioned along the river which still allows us to test the effect of river current on genetic diversity.

We did increase sample sizes for Guanacaste (N=15) and commercial *O. sativa* individuals (N=19). The larger sample sizes did not significantly change our

conclusions, however, estimates are likely to be more accurate. We redid all of our analysis using these new samples.

Reviewers were concerned that our conclusions about introgression may have been overstated and that in order to confirm the claim we would require different or additional markers.

We agree with the reviewers that our results provide only indirect evidence of possible gene flow between *O. sativa* and *O. glumaepatula*, and that further analyses should be conducted to confirm our claims about introgression. However, the use of additional markers is beyond the time and financial scopes of the present project and thus could not be included in this version of the manuscript. Therefore we emphasised genetic diversity analyses, and included additional descriptive analyses such as neighbour-joining trees. We have also tempered our conclusions about introgression and clarified that our results provide only indirect evidence of gene flow and admixture among both rice species.

The length and wording of the manuscript needed revision.

The manuscript has undergone significant rewriting and editing. We have reduced the length of the manuscript considerably. For example, given that Geneland and Structure analyses reached the same conclusions, we eliminated Geneland which contributed only to the length of the manuscript. We checked redundancy throughout the manuscript and clarified the methods section. The manuscript has been checked by two native speakers and by the Scribendi (www.scribendi.com) profesional copyediting service, to improve its readability.

Reviewer #1 asked if our sample of commercial rice truly represented current and historical variation of commercial rice varieties in the vicinity of our *Oryza glumaepatula* populations.

Our commercial rice samples are “CR1821” and “CR5272” which were released by the Ministry of Agriculture in the early 1970s and have been used for more than 30 years. *O. glumaepatula* samples from GU and MQ were collected a few years back before the newly released commercial varieties CFX18 and Palmar18 were widely planted in the country. We are confident that our *O. sativa* samples represent the historic variation in commercial rice in the vicinity of *O. glumaepatula* and the variability at the time of collection. We have included this reasoning in the manuscript.

We hope we have addressed the major concerns of the reviewers. We uploaded a version with Track-changes, however given that the manuscript underwent significant changes this version is somewhat difficult to read. We also include a “clean” version of the new manuscript.

Please do not hesitate to contact us if you have any further questions.

Sincerely yours,

Eric J. Fuchs on behalf of all authors.