**Response to Panel and Reviewers’ Comments of Previous Submission**

We are pleased that the panel has characterized the investigatory team and the scientific environment as “*excellent*”, “*highly experienced and qualified to conduct the proposed projec*t”, our proposal as a “*well thought out*” proposed study “*as regards the recruitment of the AA patients*”, with “*novel concepts, hypotheses”* and “*novel bioinformatics algorithms*”. Moreover we would like to thank them for their very thoughtful and constructive suggestions and criticisms, which we address below:

**Stratification and pre-selection of Genomes.** Major concern was raised by the panel for the need for further stratification of our dataset. R#1 suggested that “*one issue is how the variants (mutations or SNPs) will be segregated or ruled out of analysis, which are not due to race but influenced by other factors such as age, sex..* ”. R#2 noted that the absence of a planned enrollment report made it “*difficult to ascertain the actual ethnic distribution of clear cell and non-clear cell renal cell carcinoma specimens*”, while R#3 pointed that our project *“failed to stratify patients into men, women and children, since mechanisms may vary according to these parameters”*. We have now improved our proposal accordingly in three separate directions. (A) We have included a preliminary analysis accounting for the effect of different factors such as age, sex and smoking from data already available in TCGA, including XXX new genomes from the PCAWG consortium. (B) We have added an additional genotyping analysis, in order to genetically identify/confirm the patient’s ethnicity, as well as improve the comparison between our tested groups. (C) Finally, using TCGA data, we intend to apply a stratification population model to account for the effect of several biological, genetic and clinical factors in our validation dataset. [[need to distinguish aim 1, aim 3 & aim 4]]

**Biological and Clinical implications.[[STL to edit]]** An important issue that was raise by the panel, was the absence of extended discussion concerning the mechanisms that might lead to the development and progression of ccRCC and pRCC. R#1 noted that “*there is a gap between finding the coding or non-coding mutations and clinical features of the patients*” and that “*It will be interesting to know if similar parameters have already been used on the available WGS from the TCGA database (not aiming disparity) to link the coding/non-coding regions to pRCC or ccRCC “*. R#3 added that “*the current knowledge of genetic mechanisms and biological factors associated with renal cell carcinoma in the general population or even on Caucasian Americans should have been provided*”, adding, “especially since the focus is on the genome, and not on functional aspects of the renal cell carcinoma”. Finally, we feel that we have not fully clarify the need for examining both somatic and germline mutations across the genome, as R#3 noted that “*It is not clear why the team proposes to investigate germ line mutations, since the tumor samples do not normally originate from the germ line; at the very least this should be justified*.”[[why say?]] In this direction, we have now addressed these remarks in four ways. (A) We have provided a more detailed description on the current knowledge of the clinical and biological pathology of Kidney cancer. [[where?]] (B) We have added an extra analysis that is focused especially on the study of mutations that may disturb the main metabolic and genetic pathways currently considered to play an important roles in Kidney cancer such as the XXXX pathway. (C) Apart from the prioritization of SNV based on the funseq and LARVA algorithms, we have decided to include two extra analysis including the study of loss of function mutations using ALOFT, as well as examine the prevalence of deleterious mutations using the EXAC database. Finally, (D) we have provided a more detailed justification concerning our decision to study both somatic and germline mutations with regards to mutational burdening in Kidney cancer.

**Improving Deliveries for Specific Aims.** Some reviewers in the panel have expressed their concern with the degree of independence between the aims of our proposal and the significance of their respective deliveries. Specifically, R#3 noted that “*The success of Aim 3 is dependent on Aim 2, and Aim 4 is dependent on Aim 3*”. [[reword: inter-dependent but separate deliverables]] We have now increased the depth of our analysis concerning somatic and germline mutations between different ethnic groups by including a) a loss of function analysis, b) the study of mutational signatures, c) the study of regulatory networks burdened by somatic and germilne mutations, but more importantly, we d) plan to compare the mutational burden between different ethnic groups with respect to the EXAC database. Our preliminary results in individuals from TCGA and PCAWG consortiums show that genes which are considered significant for the onset and progression of kidney cancer, tend to be bear a much higher burden of deleterious mutations in African American individuals.

**Providing Additional Preliminary Results.** Some concerns were also raised concerning the lack of preliminary results that support our hypothesis. More specifically, R#3 has suggested that “*the proposed hypotheses are not supported by provided preliminary data*”. Apart from previous studies that suggest a higher risk of pRCC cancer for African Americans, we have now included an additional figure of preliminary results showing that gene regions previously suggested as important for the development and progression of Kidney cancer are more heavily burdened with deleterious mutations in specific ethnic groups compared to others. Moreover SIGNATURES Shantao to add here [[more specifics]]

**Providing additional Controls.** Concern was raised by the panel for the lack of extra control datasets for the validation of the identified somatic and germline mutation. R#3 suggested that “*the experimental design lacked controls; it is not enough to use TCGA controls because these would not account for experimental variation*”. We have now included in our proposal the validation of additional randomized regions [[[???]] to act as a third level of control in addition to the regions obtained by the different ethnic groups.

**Other additions and changes:**