RESPONSE LETTER

Reviewer #1 (Josh's review)

-- Ref 1.0 Clarifying terminology–--

Reviewer Comment	It is reassuring that the authors used published definitions. In that case, please insert the appropriate references to the literature where the terms are first introduced if that hasn't been done already.
Author Response	We thank the reviewer for the suggestion. The updated manuscript cites references when terms from the literature are introduced.
Excerpt From Revised Manuscript	

Deleted: In the

Deleted: , we have already used

Deleted: at appropriate position

-- Ref 1.1 – Background model--

Reviewer Comment	The authors have addressed the issue of including additional covariates as best as I can envision at this point.
Author Response	We thank the reviewer for recognizing the robustness of our updated background model.
Excerpt From Revised Manuscript	

-- Ref 1.3 –Response variable--

Reviewer Comment	The authors have now clarified their response variable more clearly.
Author Response	We thank the reviewer for confirming our edits to be clear.
Excerpt From Revised Manuscript	

-- Ref 1.4 – Choice of random effects model --

Reviewer Comment	The authors have documented the prior literature and appropriateness of the application of their chosen random effects model.
Author Response	The reviewer's earlier suggestion was very helpful in this regard. We thank him for mentioning this earlier.
Excerpt From Revised Manuscript	

Deleted: Reviewer's

-- Ref 1.5 – Related to LoF spectrum –

Reviewer Comment	The authors have addressed this point by making clarifying changes to the text and Figure 3.
Author Response	We thank the reviewer for going through the updated text and confirming this.
Excerpt From Revised Manuscript	

-- Ref 1.6 – Analysis of samples w/o driving mutations –

Reviewer Comment	Could the authors coordinate with Nuria's group that did the patient specific analysis for this work? They correlate the increase in putative passengers in those samples that lack known drivers. Nuria's group attempted to expand the set of drivers by assessing the expected number of passengers in each sample, calculating an "excess" and then re-analyzing whether variants found in those samples might be reconsidered as drivers. If the authors haven't done so already, I'd suggest they use this new list of samples lacking known drivers for this correlative analysis.
Author Response	As suggested, our updated results are based on the updated version of driver mutation list curated by Nuria's group.
Excerpt From Revised Manuscript	

-- Ref 1.7 – Overlap of results w/ driver group for coding genes --

Reviewer Comment	Overlap of results w/ driver group for coding genes: Did the authors restrict the analysis to coding regions as suggested by Gaddy? I can't tell from the response.
Author Response	In the updated manuscript, we perform the analysis separately for coding and noncoding regions. However, we note that despite being interesting we downplay this result as we restrict our analysis only for PCAWG samples without known driver mutations.
Excerpt From Revised Manuscript	

Deleted: Yes

Deleted: did this

Deleted: both

Deleted: driver elements as suggested

-- Ref 1.8 – Definition of impact assessment for SVs --

Reviewer Comment	The authors point us to the supplement, section 4.2, for their description of how they assessed the impact of SVs. I found this section very confusing and needing a rewrite for clarity.
Author Response	We thank the steering committee for the constructive comments. We updated the paragraphs that describe the methodology behind computing SVIS scores. We have added clear descriptions of features that were used in the random forest algorithm. We also clarified the motivation for selection of the 1000 Genomes SVs in training.
Excerpt From Revised Manuscript	

-- Ref 1.9 – Definition of impact assessment for SVs --

Reviewer Comment	The method is a machine-learning (random forest) based approach that takes in a set of features and predicts the impact of an SV. However, I could not find a clear description of the prediction labels anywhere in the section that would help me understand the gold standard their method is trying to predict. What is the overall target of the prediction? Is it the presence/absence of an SV at all? This needs to be explicitly stated somewhere.
Author Response	We updated the Section 4.2 to clarify the description of the prediction labels. In summary, we have 3 classes of SVs that we deemed most useful to score the SVs. These are somatic SVs, germline SVs, and the polymorphic SVs from the 1000 Genomes Project. Somatic SVs are the SVs that have the most potential for being impactful. There are, however, potentially many passenger SVs that have low impact. Our hypothesis is that these SVs must look similar to the other two classes of SVs (germline and 1KG SVs) when compared with respect to their features. Germline SVs are SVs that most likely have low impact but they may contain an enrichment of SVs that increase susceptibility to cancer. These SVS, therefore are expected to have an average impact score. Finally, the 1KG SVs are expected to be the lowest impact SVs.
Excerpt From Revised Manuscript	

-- Ref 1.10 – Definition of impact assessment for SVs --

Reviewer Comment	Also, the set of features is not listed but instead a windowing approach is defined to account for different SV length sizes. The authors should tabulate the features used for the model.
Author Response	Thanks to this comment, now we have added the list of features for each SV.
Excerpt From Revised Manuscript	

-- Ref 1.11 – Definition of impact assessment for SVs --

Reviewer Comment	Also, it's not clear to me how the 1000 genomes data is being used for this and the rationale on why any of it should be included in the training since one might expect germline events to be generated under very different processes distinct from somatic events?
Author Response	As we explained above, the 1000 Genomes SVs are used to represent the class of SVs that have the lowest impact. This class represents the examples of SVs with neutral effect on the genome, similar to negatives in the classification problems.
Excerpt From Revised Manuscript	

-- Ref 1.8 – Definition of impact assessment for SVs --

Reviewer Comment	Finally, it's not clear what the SV impact score (SVIS) reflects. It is some probability that an SV would be seen in a particular window given the features of that window? If so, how would that correspond to an impact? Would genomic regions with little impact have **more** SV potential since their alterations have relatively smaller functional effect?
Author Response	We believe the description of the windowing procedure was not clear. We now updated the description and we also describe here briefly. Given an SV, we divide it into windows of 10 base pairs and we compute histone modification, conservation signals, and fraction of gene annotations within each 10 bp window. Then, we compute the average and maximum of histone and conservation signals over all windows within the SV. These maximum and average values are used as the feature set for the SV. Second, we also explain here what SVIS reflects. The random forest training learns to discriminate between the 3 classes of SVs (Somatic, Germline, 1000 Genomes) that are used in training. Although the somatic SVs and 1000 Genomes SVs arise from different processes, we hypothesize that the SV impact depends only on the functional elements that the SV affects. We hypothesize further that the somatic SVs that have low impact must resemble the germline or the 1000 genomes SVs. The somatic SVs that do not resemble germline or 1000 Genomes SVs most likely affect functional elements that do not get altered by low impact SVs. Thus, they are assigned high impact. This is why we use the somatic SV class probability that the random forest algorithm computes as the SVIS score.
Excerpt From Revised Manuscript	

Reviewer #2 (Peter's comments)

-- Ref 2.0 Original & random functional impact distribution–--

Reviewer Comment	I would like to see what Figure 1a looks like for totally neutral simulated mutations – it is difficult to know how much the three peaks are explicable by the background mutational process.
Author Response	In the past, we have done this analysis and found significant differences. However, members of the steering committee pointed out these differences could be attributed to imprecision in the background model. With the updated background model, the overall differences are not significant for the majority of cancer cohorts.
Excerpt From Revised Manuscript	

-- Ref 2.1 – BLUP prediction on somatic SNVs in normal tissues--

Reviewer Comment	In the additive model, the question remains as to whether the small excess predictive signal derived from the putative passengers in true cancer samples relates to unmodelled factors influencing mutation distribution or to selection on non-coding mutations. One potential way to assess this would be to apply the BLUP predictor to sets of genome-wide somatic mutations in normal tissues (rather than cancers). The two datasets that I am aware of for this are our one whole genome from normal skin and the 45 genomes in liver, small and large bowel organoids from Ruben van Boxtel (published in Nature, 2016). If selection drives the signal in the additive model, then these should cluster with the simulated samples; if it's unaccounted mutational processes then they should cluster with the cancer samples (albeit the mutational landscape of the normal samples is less rich than the cancers).
Author Response	We thank the reviewer for giving this suggestion. This suggestion is intriguing so we tried follow-up on this. However, as the reviewer pointed out, the number of samples/mutations in both studies are very low to perform the BLUP analysis.
Excerpt From Revised Manuscript	

-- Ref 2.2 – Influence of aneuploidy on additive variance--

Reviewer Comment	The finding of the lower predictive capacity when samples with CNAs are removed is an interesting one – what does arm-level or whole chromosome aneuploidy do to the predictive model? In other words, 1q is commonly gained across many many tumours – and correspondingly has a higher mutation burden overall compared to other regions. Would this lead to apparent discriminative power in the additive model for variants on 1q? One could test this by looking at the genomic distribution of BLUP estimates of the coefficients – at individual-gene level, these should be somewhat correlated, but decay rapidly to be minimal at the chromosome arm-level
Author Response	We thank the reviewer for this suggestion. The current model doesn't explore the influence of aneuploidy on predictability. <we performed this analysis and found XYZ>
Excerpt From Revised Manuscript	

Reviewer #3 (Gaddy's review)

-- Ref 3.0 use predicted impact score–--

Reviewer Comment	Replace impact score with "predicted impact score" throughout the text.
Author Response	We thank the reviewer for suggesting this change. We have updated the text to reflect this change. However, we note that some places in the text we have intentionally omitted the "predicted" key word as it makes the sentence confusing and difficult to read.
Excerpt From Revised Manuscript	

-- Ref 3.1 reorganizing text–--

Reviewer Comment	Reorder sections that you start with mutational processes and clonal vs. sub-clonal.
Author Response	We appreciate this comment by reviewer. However, we feel that the current ordering of sections in the manuscript is suitable for this submission. Considering the timeline of PCAWG submission, we think it will be impractical to reorder sections at this point. However, we will be certainly open to do necessary changes after the peer review process.
Excerpt From Revised Manuscript	

-- Ref 3.2 Clarify what is considered drivers–--

Reviewer Comment	Clarify what you consider a driver: (i) Drivers discovered by the Driver paper, (ii) Events in known cancer genes (e.g. any event in NF1), (iii) Events called as drivers in the Panorama paper.
Author Response	For the majority of our analysis we have used driver events as defined in Nuria's paper. For additive variance analysis, we have done additional analyses on driver elements discovered by the driver paper as well. We have taken care to make this explicit in the new text.
Excerpt From Revised Manuscript	

Deleted: analysis

-- Ref 3.3 Additive variance beyond TERT promoter–--

Reviewer Comment	See how much of the missing variance is explained beyond the TERT promoter mutations.
Author Response	The additive variance measure currently reported in the paper already excludes every mutation in a driver element, including in the TERT promoter.
Excerpt From Revised Manuscript	

Deleted: ,

Deleted: is explained beyond

Deleted: mutations

-- Ref 3.4 Survival and signature–--

Reviewer Comment	Are survival differences account for different signatures and subtypes of disease?
Author Response	The reviewer makes a good point. We have updated the text to point out this caveat in the survival analysis section.
Excerpt From Revised Manuscript	"Finally, we note the potential role of unmeasured patient clinical characteristics or tumor molecular subtypes in influencing these correlations."

Deleted: Reviewer

Deleted: partially

-- Ref 3.2 – TADs and partial SV depletion--

Reviewer Comment	Regarding depletion of partial SVs, can this be related to TADs in both cases?
Author Response	Depletion of partial SVs is observed across different genomic elements. Thus, it's very unlikely that partial SVs in all such regions are related to TADs.
Excerpt From Revised Manuscript	

-- Ref 3.3 – LoF spectrum--

Reviewer Comment	Regarding LoF spectrum, can it be fully explained by the prevalence of signatures in each cohort? If yes, you may want to write a shorter section saying that "STOP codons are distributed as expected by the mutations signatures.
Author Response	In the majority of cases, the LoF spectrum by cohort can be explained by the prevalence of signatures in that cohort. However, in a few cohorts such as Colorectal Adenocarcinoma and Melanoma, we observe differences in the observed and expected LoF distribution. We report this finding in the supplement.
Excerpt From Revised Manuscript	

Deleted: Although, in

Deleted: signature.

Deleted: section

-- Ref 3.4 – Impact score distribution --

Reviewer Comment	Are putative passengers only non-coding? Fig 1a is non-coding
Author Response	Putative passengers are both coding and non-coding. Since the majority of putative passengers are non-coding, we highlight the impact score distribution for just non-coding.
Excerpt From Revised Manuscript	

-- Ref 3.5 – Signature and early vs late mutations --

Reviewer Comment	Is there a difference in signatures between early and late mutations? If yes, can it explain the differences in fraction of impact categories?
Author Response	We thank the reviewer for this helpful comment. We note that the PCAWG-11 group compared signature profiles for early and late subclonal mutations and concluded that "mutational processes act at a rather constant rate during tumor progression". We clarify this point in the updated text.
Excerpt From Revised Manuscript	"We note that different signatures between and early and late subclone mutations have limited contribution to the observed variations[18]."

Deleted: PCAWG11

Deleted: profile

Deleted: subclone