

# The cumulative impact of passenger mutations on cancer development

Corresponding Author: Professor Ilias Georgakopoulos-Soares

**This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.**

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In this interesting, timely study, Nayak et al. study the distributions of (presumed) passenger mutations in driver genes, suggest epistasis with driver mutations, and link them to gene expression/splicing effects as well as with clinical outcomes. Taken together, their findings, if better supported (see below for necessary changes) would support the hypothesis of a continuum of driver potential (or, existence of mini-drivers), which is an important conceptual advance in understanding cancer evolution.

Major issues.

1- The validity of their central claim that (apparent) passengers sometimes function as drivers strongly depends on what mutations are defined as passengers. Here, this is done by elimination of a set of known driver mutations; this is fine in principle however in practice is fraught with difficulties. It appears from their text that they took a set of driver mutations identified (presumably, by recurrence) in the PCAWG study (line 95): this is a fairly small set of 2,700 tumors where only the highest selected and/or highest mutable mutations will be identified as drivers. It is expected to be a high-specificity, low-sensitivity set of driver mutations. Then it follows that exclusion of this set to get to passengers generates a low-specificity set of passengers.

Many of these mutations in their passenger set could, in fact, be recognized drivers in later studies, casting doubt on their principal finding. The core analyses must be repeated with a more comprehensive set of driver mutations (thus, higher-specificity set of passenger mutations). One option would be to check the cancerhotspots.org set, which is derived from ~25k tumor samples (DOI: 10.1158/2159-8290.CD-17-0321 published in 2018).

2- The presence of driver mutations is known to be correlated with the overall, total mutation burden (it increases sub-linearly with the TMB but it does increase). Therefore, all the trends they observe (increase of passenger effects etc) have the potential to be downstream of the mutation burden differences in driver-less samples.

I do see that in Methods line 465-469 the calculations seem to incorporate TMB (as „N“), which does attempt to control for the TMB in the analysis of genomic „subcompartments“ (UTRs, introns, exons...). Firstly, we'd need to see empirical evidence this control was effective e.g. by plotting trends in low, med, high TMB. Secondly, it was not clear to me that this control was performed for other analyses; the one that caught my eye was a reported enrichment of mutations nearby the TSS and the TES (lines 147-149): this is an important result, if correct, and thus should be stringently controlled for TMB. One way is to quantitate this putative enrichment near TSS (or near TES) relative to the gene-flanking regions that are not near TSS/TES, eg 2kb away from the gene, which should be similarly affected by TMB and/or regional changes in mutation distribution. Do the other analyses e.g. survival rigorously control for TMB? Ideally, mutational signatures/spectra would also be explored as a potential confounder, although this is less critical than TMB itself.

Other important issues.

3- The analysis where gene expression drops with multiplicity of mutations was unclear in two ways. First, are they indeed

checking cases with many (sometimes 20+) mutations in the same gene \*in the same tumor\*? It is quite rare to have multiple somatic mutations per gene, and these cases are likely confined only to long genes and/or hypermutated tumors. If they indeed count 20+ per sample, we need to explore how this relates to mutation burden and spectra and gene length. Second, is the decreased expression level simply because of a higher chance of getting a nonsense mutation or a frameshift indel thereby triggering NMD?

4- The definition of drivers (lines 445-448) seems to include flanking regions over-eagerly „examples without any identified driver mutations in the same gene, including promoter upstream regions (-3,000 base-pairs from the transcription start site) and downstream regions (3,000 base-pairs downstream from the transcription end site).“ 200 nt upstream of TSS would be the main promoter region, and up to 1.5k extended promoter (I have never seen 3k used). Downstream of TES, I am not aware of any mention of driver mutations there and I can't envision a mechanism how these mutations would act. I am not sure at all there is a rationale to excluding TES-downstream mutations.

5- CADD score is a little dated and not highly performant by today's standards. I understand the appeal in using the same score for coding and nc mutations, however at least for coding there are much better methods available (AlphaMissense, ESM, REVEL, Varsity etc).

In contrast, for splicing mutations Pangolin indeed is the state-of-the-art, so kudos for that. The Fig 4C plot contains a somewhat surprising result, which is that the predicted splicing effect of drivers (most of which, presumably, are selected via direct changes on aa sequence) seems to be similar to that of deleterious passengers. Can they comment on this, are there known driver hotspots with strong splicing effects, which could be a novel mechanism therefor?

6- For survival analysis, sample sizes are quite limited; these 710 samples are, I infer, the TCGA WGS subset within their PCAWG set. This makes one wonder why did they limit the analysis to the (fairly small) PCAWG, when there is more WGS available. For one, there is the Hartwig cohort; second, TCGA have released WGS mutation calls for several thousand samples via Genomic Data Commons. Having more samples would also stabilize their – currently quite messy and not showing clear trends – per-cancer-type plot in Fig 2E (note: this needs confidence intervals and/or a statistical test for every cell). I am not sure I would exactly call this „consistent enrichment“ (lines 149-150) as it is currently.

7- CNA seem to be fully ignored in the analyses, although they acknowledge at one point (for central nervous system cancer) they might explain some of the trends in their data. I think this is not a dealbreaker point because, if anything, ignoring CNA probably biases their results conservatively (as the CNA could substitute for some of the mini-drivers). That said I would have liked to see some quantitation of how big a role they think CNA would play. I understand the mutual-exclusivity analysis with CNA is complicated due to the co-occurrences CNA\*mutation in the same gene, as two-hit events, however e.g. the associations with survival or similar are more easily checkable.

8- I don't quite see the point in the Fig. 2D analysis of mutation rates of different segments within a gene. I do see the interest in studying TSS proximal mutation rates (upstream: promoter; downstream: 5' UTR and sometimes intragenic enhancers) however the relevance of the rest of this analysis is less obvious. The up-and-down jumps are probably resulting from sparse data, although the „tube“ around the loess fit suggests otherwise; maybe larger bins would be more appropriate given the dataset size, and/or absolute-size bins anchored on the TSS or TES. Consider revising or dropping.

9- The study introducing the SEISMIC method ( <https://doi.org/10.1038/s41467-022-34746-z> ) is conceptually related to their work, yet is not cited. Consider discussing.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

The manuscript "The cumulative impact of passenger mutations on cancer development" reanalyses a large set of somatic mutations from public cancer genome datasets. It critically revises the notion of a passenger mutation as a neutral mutation without significant impact on cancer biology. The manuscript describes a number of statistical analysis, of which some present compelling evidence that there are weak driver mutations among the passengers, which, especially in absence of drivers, can contribute to explaining the genetic causes of cancer. In other cases it is not entirely clear if the observed effects are a consequence of statistical artifacts that need to be corrected for. While the manuscript addresses a very important question and discusses several interesting aspects, the reported results are difficult to interpret, as it is often unclear what and how many datapoints exactly are used (most figures do not report the group sizes, respectively, the n). Also, the nullhypothesis are often unclear or a baseline model is missing.

Before a publication a major revision is required.

Major comments:

1. It is not always clear what statistical hypotheses you are testing exactly. In line 107-110 you test differences on gene level for samples in which a passenger mutation co-occurs with a driver mutation against samples in which passenger mutations do not co-occur with driver mutations. How many observations do you have per group? Why do you compare frequencies instead of events?

2. As mutations in genes are rare events, it is likely that the majority of all mutated cancer genes have exactly one mutation. Therefore, it is not surprising that you rarely observe a co-occurrence between driver and passenger mutations. To understand the relevance of your observation, you have to compute a baseline model in which you test how often you observe a second passenger mutation in a gene and sample in which you have already observed one passenger, against passengers that remain solo. Compute and report on this baseline model.
3. The same holds true for the analysis in figure 2B). The method as described in lines 443-469 indicate that you do not account for the fact that the driver mutation should be counted as a mutation event as well. Instead, it is simply used to define two different groups. Under the null hypothesis that driver and passenger are the same you basically introduce an offset of one between the groups and with the consequence that the driver group is less often mutated. Compute a second model in which you replace "Passenger Count" by "Mutation Count" and reevaluate your analysis.
4. Figure 2 C): Report the number of observations for each column. The reader cannot directly understand if some of the columns are supported by millions of observations, while others may only be backed by a single one. UTR5 mutation are, for instance, much rarer than intronic mutations.
5. Figure 2 D) As discussed above, by not including the driver mutation itself you strongly reduce the mutation frequency. Create a second version of the plot in which you display the "Mean mutation density" on the y-axis instead of the "Mean passenger density". Do you still observe a significant difference between the groups.
6. Figure 3 A)-D) indicate that there may be a subset of passenger mutations that qualifies as weak deleterious drivers mutations (blue curve above 0) in 3 A). This is not informative about the bulk of passenger mutations, but rather hints that the definition of driver mutation you are relying on is not perfect. Your argument that passenger mutations in general have a cumulative effect is therefore not strongly supported by these figures. Separating passenger mutations into real passengers and weak drivers would be a more logical conclusion.
7. You may want to explicitly state that a considerable portion of driver mutations are actually gain-of-function and thus cannot be characterized by the CADD score.
8. From the method section it is unclear if you used scaled or unscaled C-scores as input when you computed the z-score of the CADD score. Furthermore, you do not mention in the caption of Figure 3 that you do not directly show the CADD scores, but their z-scores. This is only indicated in the Method section lines 496-497. Clarify this to avoid confusion.
9. Figure 3 E) in lines 193-195 you mention effect sizes, but cite the Man-Whitney U test, which usually reports significance instead. Clarify were you compute p-values and were you report statistical effect size such as Cohens d. Optimally, report both. As your analysis is overpowered due to pooling mutations across different cohorts, the statistical effect size may be more informative than p-values. Also consider to correct for multiple testing when the number of tests.
10. Figure 4A) The text and the axis label indicate that TF binding data is shown in this panel, while the caption in line 251 refers to „Splicing effects of passenger mutations“. Also the headline of figure 4 only mentions splicing changes. Correct this.
11. Figure 4B while the Method section states that you have computed adjusted p-values, the y-axis label states only p-value, and the caption does not provide additional information. Be precise to avoid confusion. Consider to use the term q-value as abbreviation for adjusted p-value.
12. Regarding your observation that survival times correlate with passenger mutations in driver genes: Survival strongly correlates with overall mutation load in many cancer types. You need to convince the reader that your observation is not a consequence of that effect, as higher baseline mutation rate results in more mutations in all genes, including passenger mutations in driver genes. Explain how you correct for the background mutation rate in this analysis. Furthermore, report the numbers at risk.
13. Figure 6: This analysis may, as well, simply reflect the higher mutation number in cancers with poor outcome. Test this, for instance, by comparing the genome-wide number of mutations in the three ARID2 groups. If passengers=0 has lower average counts than the other two groups, you need to correct for this factor.
14. Your discussion raises a number of very valid points, but it lacks an analysis of how the driver mutation annotation you are using is actually defined. It should be very clear to everyone in the field that the biological impact of mutations is continuum ranging from neutral mutations, over mutations with weak impact to mutations with strong impact. Putting somewhere on this continuum a threshold and calling everything with stronger impact a driver mutation, is a semantic simplification. Some of your analysis reflect this fact, while others may require some additional corrections. Your discussion is hinting at these aspects (e.g. "challenge the canonical driver-passenger dichotomy" line 371-372), but by still using the term passenger mutation for everything that is not a driver, it is unclear if you address relevance to all passenger mutations or if you consider a set of "weak driver mutations" among them. Clarify your position, or at least discuss likelihood of these two interpretations.
15. The definition of what mutation is a driver mutation and what mutation is a passenger mutation is most central to your manuscript. In the method section this definition is given as: „Driver mutations for each patients were obtained from the ICGC and TCGA datasets“ (line 438). This description is not sufficient to reproduce your analysis. Specify in the method section how you objectively define a driver mutation (or cite the exact publication you are relying on), and where exactly you obtained the mutation list from you are using (e.g. name of supplementary datafile, command you used in a data portal, etc.).

#### Minor comments:

- Figures panels are sometimes referenced in lower case and sometimes in upper case. A consistent labeling would be preferable.

(Remarks on code availability)

The manuscript needs to be edited to a sufficient quality level before I would consider to revise the code.

(Remarks to the Author)

Summary

Somatic mutations play critical roles in cancer initiation, progression, and metastasis. Most research has concentrated on driver mutations that are believed to underlie the cancer phenotype, overlooking the more numerous passenger mutations which are generally considered biologically neutral. Nayak et al. provide a comprehensive pan-cancer analysis of passenger mutations from whole-genome sequencing data from 2,263 primary tumors across 31 cancer types, highlighting the cumulative importance of passenger mutations in the absence of bona fide driver mutations. They demonstrate that passenger mutations are significantly enriched in cancer genes that lack canonical driver mutations where they are associated with higher predicted pathogenicity. These mutations exert their impact through changes in transcription factor binding, splicing, and gene expression and their presence correlates with poorer clinical outcomes. This work challenges the traditional binary classification of mutations into driver and passenger categories, proposing instead a continuum model that better captures the complexity of the impact of somatic mutations in cancer.

Strengths

1. The manuscript addresses an understudied aspect of cancer genomics by systematically characterizing the cumulative effect of passenger mutations in cancer genes.
2. The results provide valuable insights into mutation density distribution, functional pathogenicity, gene regulatory disruptions, and prognostic implications.
3. The findings emphasize the necessity of considering passenger mutations in cancer models, contributing to our understanding of cancer evolution and prognosis.

Major concerns

1. Can the authors please clarify the exact driver genes and mutations that are studied here? A representative example of a cancer gene with and without driver mutation in Figure 1 would greatly improve clarity of the approach.
2. Is it generally true that genes without known driver mutations have higher mutation rates, or is this effect driven by few distinct (and potentially very interesting) genes? One would assume that the cumulative passenger effect specifically affects tumor suppressor genes. Can the authors explore this by stratifying their analyses by known oncogenes and tumor suppressors? As they already show in Figure 6A, these gene classes exhibit different properties.
3. Is the increased passenger effect potentially confounded by high-level copy number alterations of cancer genes, which may not carry driver mutations but can accumulate more passenger mutations because there are more copies in the cell?
4. Could the authors please comment in the discussion on negative selection of genes with a driver mutation, where additional passenger mutations might decrease fitness? (contrasting positive selection of passenger mutations).
5. Also please discuss the possibility that the catalog of "driver" mutations in PCAWG is incomplete. This may specifically be true for splice site mutations in PTEN and ARID genes, where the present study shows sometimes profound changes in estimated effect of passenger mutations (which might, in fact, be missed drivers).
6. Is the transcription factor analysis mainly driven by oncogenes, tumor suppressor genes, or a few select cancer genes, which would be very interesting to see? Which of the 107 significantly disrupted TFs are expressed in the respective cancer types?
7. Cancer-Specific Survival Analyses: Survival outcomes vary strongly across cancer types. The current aggregated survival analysis potentially masks critical cancer-type-specific differences. Survival analyses stratified by individual cancer types would yield more precise, clinically relevant insights.
8. Related to survival analyses in Figure 5, do patients with only passenger mutations in EGFR have high-level amplifications that could contribute to the observed effect? Similarly, do the patients with TERT passenger mutations have CNVs, SVs, or activating promoter mutations?
9. Methodological Clarification and Comparison: Kumar et al. (2020) previously conducted a comprehensive analysis of passenger mutations using the same PCAWG datasets. Can the authors please discuss how their methodology and findings provide additional insights, particularly regarding the molecular characterization and modeling of passenger mutations?
10. Please provide additional information (such as the list of tumor-specific driver genes, Cox regression analysis etc) in supplementary tables. It looks like these may eventually be shared in the GitHub repository but are currently not available.
11. Please ensure all files posted on GitHub comply with controlled and open access data rules.

Minor concerns

1. Multiple Driver Mutations Clarification: Could the presence of multiple driver mutations within a single patient influence the passenger mutation rate results? Clarifying this would improve the interpretation of mutation density analyses.
2. Figures:
  - a. Please clearly explain the meaning of symbols (e.g., green triangles in Figure 2a).
  - b. Address apparent discrepancies (such as absent boxes in breast cancer) despite statistical significance in Figure 2A.
  - c. Figure 4A shows TF scores, but the legend describes a splicing analysis.
  - d. Some panels appear to be swapped in Figure 3.
  - e. Please clarify the meaning of the "ncRNA" group for protein-coding cancer driver genes in Figure 3E.
3. Please explain why passenger mutations preferentially create new splice sites while driver mutations predominantly disrupt existing splice sites, potentially offering new mechanistic insight.
4. Please clarify whether the summarized CADD score is normalized for gene length.

(Remarks on code availability)

Some directories mentioned in the GitHub instructions are not available (e.g. results). Due to time constraints, we did not attempt to run the code ourselves.

Reviewer #4

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have thoroughly addressed my remarks. The analysis has been improved by testing against a more comprehensive set of hotspots, and additionally by incorporating additional genome sequences/cohorts, thus results are more compelling with the extra data introduced in revision. The presentation has been improved by omitting some less informative parts, and literature context is now better covered. Overall this study is an important contribution towards understanding the prevalence of mini-drivers and mechanisms by which they act, the analyses are quite solid and presentation is good.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

My comments have been sufficiently addressed.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

In the present manuscript, Nayak and colleagues have greatly improved their evidence for impact of passenger mutations in cancer, specifically through addition of the POG570 and TCGA datasets. This is now a well-thought out and presented story.

I just have a few minor comments:

1. In the Borzoi figure, the predicted track should not be labeled "GTEx" (as GTEx is observed data)
2. In the same figure, it is hard to see whether the predicted sequences line up with the mutation position. Could you please highlight the bases in the sequence logo track that are mutated, maybe with a box that extends through all tracks?
3. In Supp figure 6C,D: "Expression of disrupted TFs": The TFs that are not disrupted themselves, but rather their motifs upstream of cancer genes? (the title of the figure is confusing)

(Remarks on code availability)

**Open Access** This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

## REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In this interesting, timely study, Nayak et al. study the distributions of (presumed) passenger mutations in driver genes, suggest epistasis with driver mutations, and link them to gene expression/splicing effects as well as with clinical outcomes. Taken together, their findings, if better supported (see below for necessary changes) would support the hypothesis of a continuum of driver potential (or, existence of mini-drivers), which is an important conceptual advance in understanding cancer evolution.

Major issues.

1- The validity of their central claim that (apparent) passengers sometimes function as drivers strongly depends on what mutations are defined as passengers. Here, this is done by elimination of a set of known driver mutations; this is fine in principle however in practice is fraught with difficulties. It appears from their text that they took a set of driver mutations identified (presumably, by recurrence) in the PCAWG study (line 95): this is a fairly small set of 2,700 tumors where only the highest selected and/or highest mutable mutations will be identified as drivers. It is expected to be a high-specificity, low-sensitivity set of driver mutations. Then it follows that exclusion of this set to get to passengers generates a low-specificity set of passengers.

Many of these mutations in their passenger set could, in fact, be recognized drivers in later studies, casting doubt on their principal finding. The core analyses must be repeated with a more comprehensive set of driver mutations (thus, higher-specificity set of passenger mutations). One option would be to check the cancerhotspots.org set, which is derived from ~25k tumor samples (DOI: 10.1158/2159-8290.CD-17-0321 published in 2018).

We thank the reviewer for this comment. To address this concern, we repeated the core analyses using a more comprehensive and independent driver mutation set derived from cancerhotspots.org<sup>1</sup> (v2), which is based on recurrence patterns across approximately 25,000 tumor samples, as suggested by the reviewer. To further assess the robustness of our findings, this hotspot annotation was applied to an independent patient cohort, POG570<sup>2</sup> as well as the TCGA dataset. The hotspot mutations were used to define driver events and all remaining mutations classified as passengers. We reproduced the key findings from our original analysis, including the differential abundance of passenger mutations in the presence versus absence of driver events within cancer genes, as well as the association between passenger mutation burden and overall survival.

In both POG570 and TCGA cohorts, passenger mutation density was found to be elevated in samples harboring hotspot mutations within the same gene compared to those without hotspots (Mann-Whitney U test,  $p$ -value < 0.0001; **Figure R1-A** and **Figure R2-A**) and passenger mutations were more frequently observed than expected by chance (Chi-square test of

homogeneity,  $p$ -value $<10^{-6}$ ; **Figure R1-B** and **Figure R2-B**). When stratifying by cancer subtype in the TCGA cohort, the results remained consistent (Mann-Whitney U tests; **Figure R2-C**). We also find that patients with only passenger mutations in a cancer gene had significantly lower median survival times than patients without any mutations in the gene (Mann-Whitney U test,  $p$ -value  $< 0.05$ ; **Figure R1-C**). Further analysis using the Kaplan-Meier estimator revealed that patients with passenger mutations in the absence of driver mutations also demonstrated significantly worse survival outcomes (log-rank test,  $p_1=0.022$ ; **Figure R1-D** and  $p_1<10^{-6}$ ; **Figure R2-D**), with the outcome worsening with increase in passenger burden in cancer genes (log-rank test,  $p_1<10^{-6}$ ,  $p_2<10^{-6}$ ; **Figure R2-E**).

These results demonstrate that our conclusions are robust to both the choice of driver mutation set and the patient cohort, and are not driven by potential limitations in the sensitivity of the original PCAWG driver annotations.

We have included panels R1-A and R2-A in Figure 1A in the manuscript and added figures R1 and R2 as supplementary figures. We have also added the following text in the manuscript:

#### Results Section:

“””

We also reproduced our key findings in the TCGA dataset (dbGAP study phs000178.v11.p8) and the POG570 cohort using an independent set of cancer hotspot mutations (v2) for driver annotation (**Supplementary Figure 2-3**), ensuring the robustness of our results.

...

To further assess robustness, we replicated the analysis in the TCGA dataset and the POG570 cohort using an independent cancer hotspot-based driver annotation. The results were highly concordant, consistently showing an enrichment in the mean adjusted passenger density in the absence of driver (or hotspot) mutations within a cancer gene (Mann-Whitney U test,  $p$ -value $<0.0001$ ; **Figure 2A**).

“””

#### Methods section:

"To ensure the robustness of our findings, we reproduced key analyses in the entire TCGA WGS cohort (dbGaP study phs000178.v11.p8) and the POG570 cohort. In both datasets, driver mutations were annotated using an independent set of cancer hotspot mutations (v2)<sup>36</sup> following a procedure similar to that used in the primary analysis."

#### References:

1. Chang, M. T., Bhattarai, T. S., Schram, A. M., Bielski, C. M., Donoghue, M. T., Jonsson, P., ... & Taylor, B. S. (2018). Accelerating discovery of functional mutant alleles in cancer. *Cancer discovery*, 8(2), 174-183.
2. Pleasance, E., Titmuss, E., Williamson, L. et al. Pan-cancer analysis of advanced patient tumors reveals interactions between therapy and genomic landscapes. *Nat Cancer* 1, 452-468 (2020). <https://doi.org/10.1038/s43018-020-0050-6>

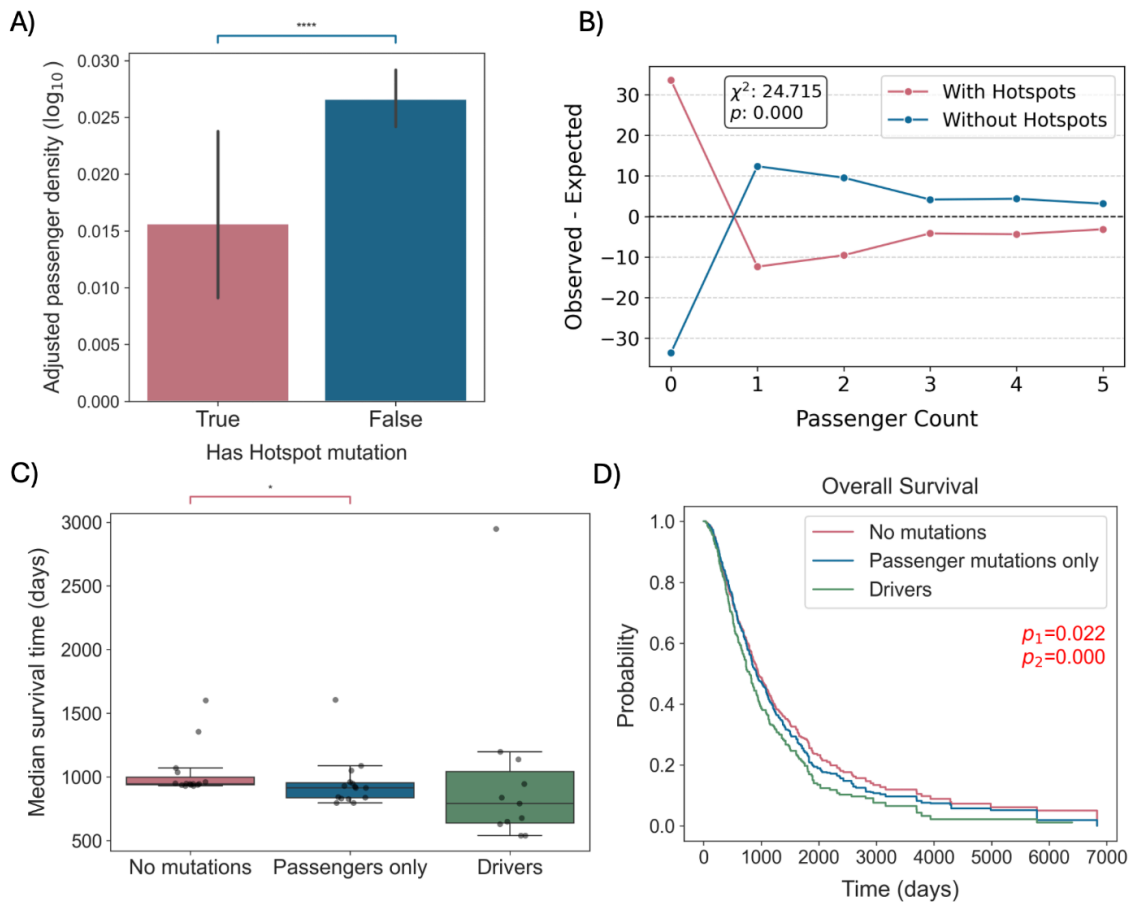


**Figure R1: Results of core analyses reproduced in the POG570 dataset using cancer hotspots data for driver annotation. (A)** Mean adjusted passenger mutation density in samples with and without hotspot mutation in a cancer genes. **(B)** Difference in observed and expected number of samples with varying passenger mutation counts in a cancer gene. **(C)** Distribution of median survival time (in days) of patients. Each data point corresponds to the median survival for a particular cancer gene. **(D)** Kaplan-Meier Survival plot for overall survival (OS) by mutational status of a cancer gene. Statistical significance in panels (A) and (B) is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

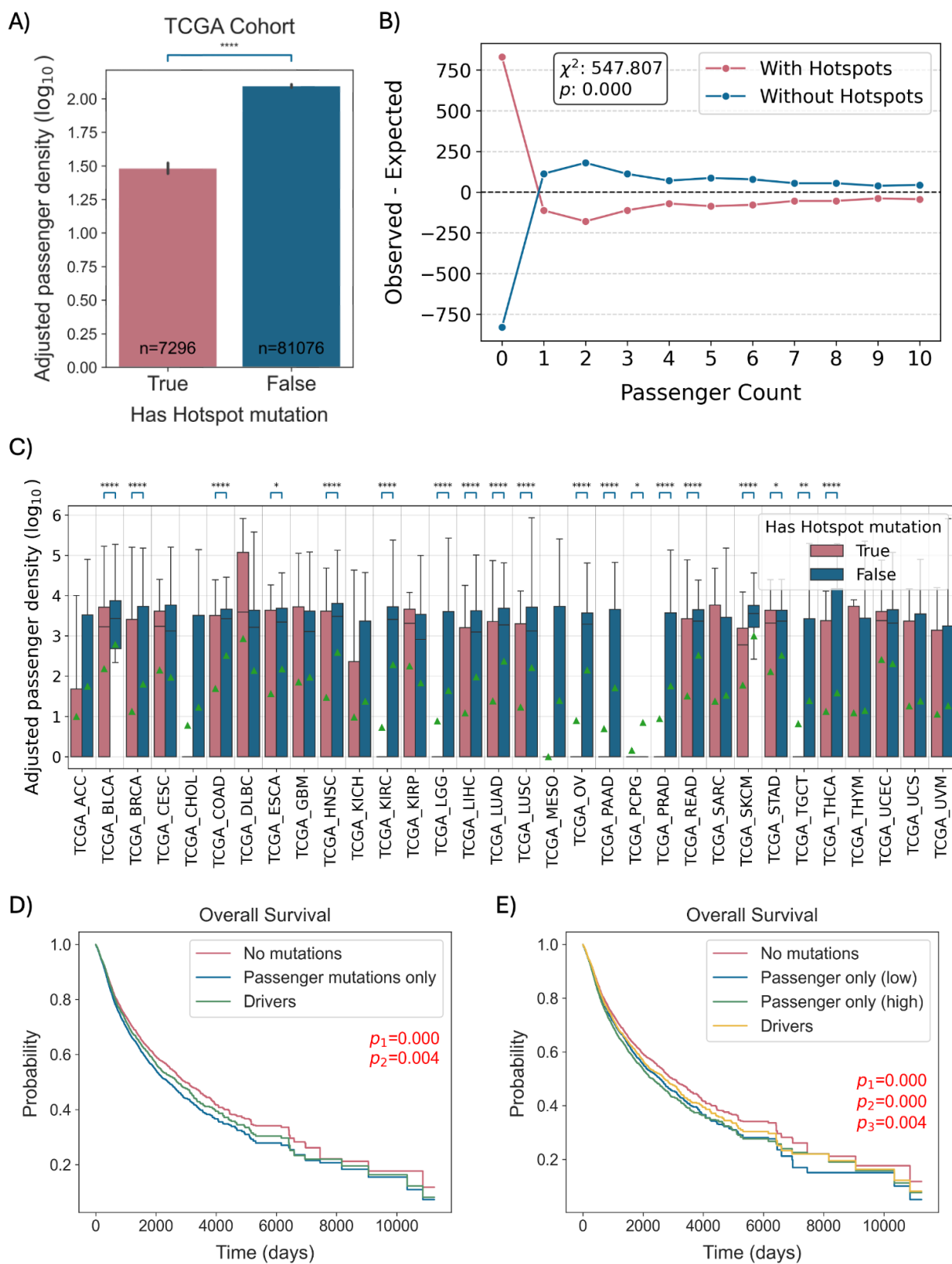


**Figure R2: Results of core analyses reproduced in the TCGA dataset using cancer hotspots data for driver annotation.** (A) Mean adjusted passenger mutation density in samples with and without hotspot mutation in cancer genes. (B) Difference in observed and expected number of samples with varying passenger mutation counts in a cancer gene. (C) Distribution of differences in mean passenger density in samples without and with driver mutations for each cancer gene, stratified by cancer subtype cohort. (D) Kaplan-Meier Survival plot for overall survival (OS) by mutational status of a cancer gene. (E) Kaplan-Meier Survival plot for overall survival (OS) stratified by passenger mutation burden in cancer genes. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

Additionally, we also note that we have observed a separate phenomenon, in which a gene can be disrupted not by one mutation, but by the progressive accumulation of multiple mutations, which in combination lead to its disruption. When occurring individually, these mutations have limited effects on the expression of the cancer gene, while in combination they disrupt its expression, underscoring the importance of holistic models that incorporate the full mutational context of a cancer gene.

2- The presence of driver mutations is known to be correlated with the overall, total mutation burden (it increases sub-linearly with the TMB but it does increase). Therefore, all the trends they observe (increase of passenger effects etc) have the potential to be downstream of the mutation burden differences in driver-less samples.

I do see that in Methods line 465-469 the calculations seem to incorporate TMB (as „N“), which does attempt to control for the TMB in the analysis of genomic „subcompartments“ (UTRs, introns, exons...). Firstly, we'd need to see empirical evidence this control was effective e.g. by plotting trends in low, med, high TMB. Secondly, it was not clear to me that this control was performed for other analyses; the one that caught my eye was a reported enrichment of mutations nearby the TSS and the TES (lines 147-149): this is an important result, if correct, and thus should be stringently controlled for TMB. One way is to quantitate this putative enrichment near TSS (or near TES) relative to the gene-flanking regions that are not near TSS/TES, eg 2kb away from the gene, which should be similarly affected by TMB and/or regional changes in mutation distribution. Do the other analyses e.g. survival rigorously control for TMB? Ideally, mutational signatures/spectra would also be explored as a potential confounder, although this is less critical than TMB itself.

We thank the reviewer for this comment. Analyses involving mutation densities were controlled for tumor mutational burden (TMB) throughout the manuscript. This includes overall passenger density, density by genic subcompartments, density of splicing mutations, and mutation density near TSS/TES. The methods section has been updated to clarify this.

\*\*\*\*

Mutational density was calculated separately for each group, adjusting for the gene length ( $l$ ), genome-wide mutation rates ( $N$ ), and copy number alteration burden ( $C$ ).

\*\*\*\*

To further provide evidence, for the analyses of passenger mutation density across genomic subcompartments, we have stratified the samples by TMB: low TMB ( $\leq 5$  mutations/Mb), medium TMB ( $> 5$  and  $< 10$  mutations/Mb), and high TMB ( $\geq 10$  mutations/Mb), demonstrating that the control was effective (**Figure R3-A**). These results have been included in supplementary material as **Supplementary Figure 4A**, along with the following text in the Results section:

"This effect remained consistent upon further stratification of samples by tumor mutational burden (**Supplementary Figure 4A**)."

Although the original TSS/TES analysis was adjusted for genome-wide TMB, we did not perform a stratified analysis for this result because the corresponding plot was removed from the manuscript in response to another comment by the reviewer (Other issues - 8).

To control for TMB in the survival analysis, we include TMB as a covariate in the Cox proportional hazard models (**Figure 6D**). Specifically, the model incorporated the total number of driver mutations across the genome, total mutational burden, and CNA burden, in addition to age, sex, event occurrence, and time to event.

Further, we have now stratified the median survival analysis by low ( $< 5$  mutations/Mb) and high TMB ( $\geq 5$  mutations/Mb) to demonstrate that our results are consistent (**Figure R3-B**). These results have been included in supplementary material as **Supplementary Figure 4B**, along with the following text in the Results section:

"This effect remained consistent when patients were stratified by tumor mutation burden (**Supplementary Figure 4B**)."

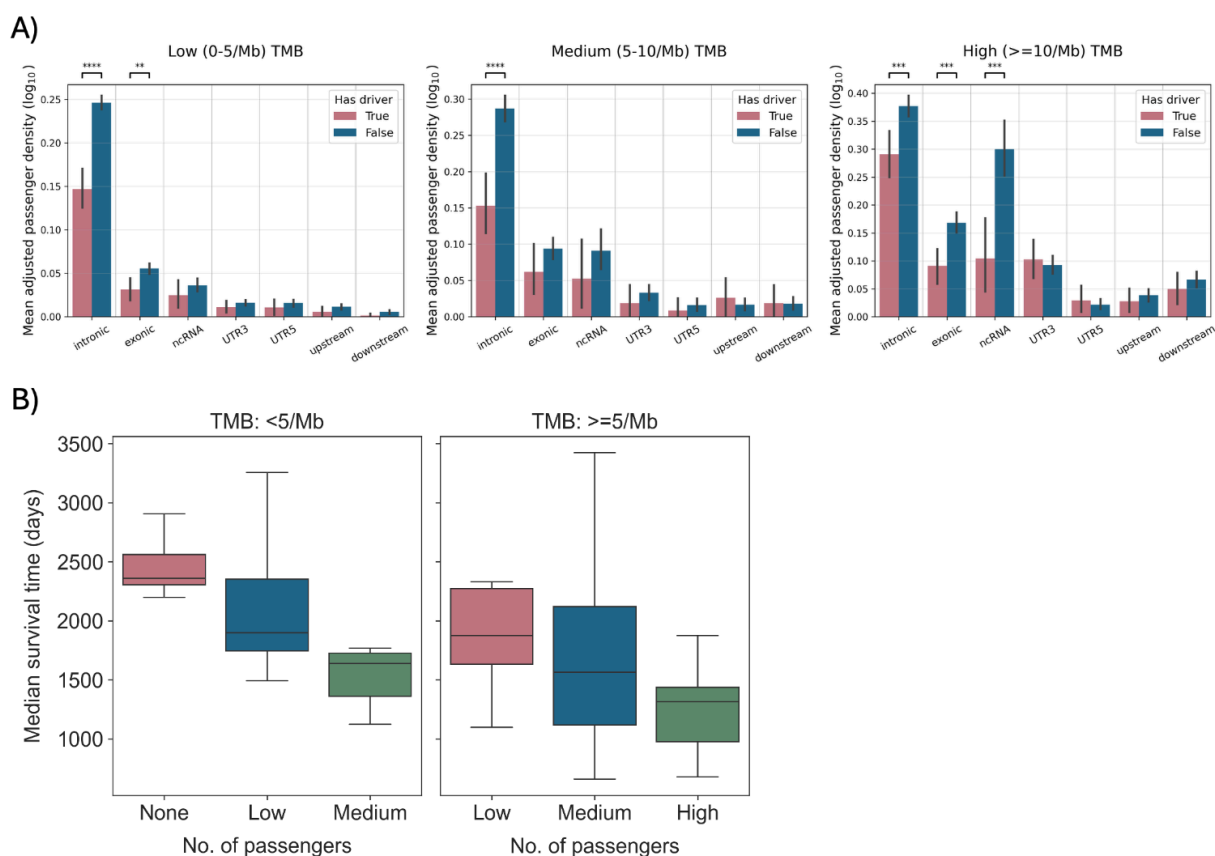


**Figure R3: Passenger density and overall survival stratified by genome-wide TMB. (A)** Passenger density in genic regions in groups of patients with and without driver mutations in the cancer gene, stratified by genome wide tumor mutation burden (TMB). **(B)** Median Survival of patients with no mutations, low, medium and high burden of passenger mutations in cancer genes, stratified by genome wide tumor mutation burden (TMB).

Other important issues.

3- The analysis where gene expression drops with multiplicity of mutations was unclear in two ways. First, are they indeed checking cases with many (sometimes 20+) mutations in the same gene \*in the same tumor\*? It is quite rare to have multiple somatic mutations per gene, and these cases are likely confined only to long genes and/or hypermutated tumors. If they indeed count 20+ per sample, we need to explore how this relates to mutation burden and spectra and gene length. Second, is the decreased expression level simply because of a higher chance of getting a nonsense mutation or a frameshift indel thereby triggering NMD?

We thank the reviewer for this comment. We confirm that there are several patients with more than 20 somatic mutations in a gene. The table below summarizes the number of samples containing at least one gene with a given range of passenger mutation counts. It also reports the number of unique genes observed with the given range of passenger mutation counts, along

with their corresponding average gene lengths. While we observe that average gene lengths tend to be higher in genes with a larger number of mutations, this pattern is not driven by a few unusually long genes but is instead observed across multiple genes of varying lengths.

| No. of passenger mutations in a gene | No. of samples | No. of genes | Gene Length Statistics (min, mean, max) |
|--------------------------------------|----------------|--------------|-----------------------------------------|
| 0                                    | 2,253          | 138          | 3,388 - 98,034.13 - 2,298,478           |
| 1                                    | 1,685          | 128          | 3,388 - 213,482.75 - 2,298,478          |
| 2                                    | 740            | 97           | 3,388 - 431,137.64 - 2,298,478          |
| 3                                    | 435            | 77           | 5,224 - 743,985.6 - 2,298,478           |
| 4                                    | 299            | 56           | 3,388 - 965,974.06 - 2,298,478          |
| 5                                    | 217            | 49           | 5,224 - 1,016,500.18 - 2,298,478        |
| 6–10                                 | 626            | 54           | 3,388 - 1,215,845.25 - 2,298,478        |
| 11–20                                | 525            | 47           | 3,388 - 1,178,945.56 - 2,298,478        |
| 20+                                  | 430            | 28           | 5,224 - 1,175,715.56 - 2,298,478        |

To further address this concern, we performed additional stratified analyses to disentangle the effects of gene length and tumor mutational burden (TMB). First, we stratified genes into 3 quartiles based on gene length (**Figure R4-A**). While the effects in quartile 1 are not consistent, we observe a consistent trend in quartiles 2 and 3, where increasing passenger mutation counts within a tumor suppressor gene is associated with reduced normalized gene expression. The monotonic decrease in expression within individual groups (Q2 and Q3) indicates that the observed effect is not solely driven by gene length.

Second, we stratified tumors by TMB into low (0-5 mutations/Mb), medium (5-10 mutations/Mb), and high ( $\geq 10$  mutations/Mb) categories (**Figure R4-B**). Within each TMB group, we still observe a decline in gene expression with increasing numbers of passenger mutations, demonstrating that the relationship is not merely a consequence of overall mutation burden or hypermutated samples.

These results have been included in supplementary material (**Supplementary Figure 9**) and the results section has been updated with the following text:

.....

To disentangle the effects of gene length and tumor mutational burden (TMB), we performed stratified analyses for tumor suppressors. When these genes were stratified by length quartiles, we observed a consistent decrease in normalized expression with increasing passenger mutation counts in tumor suppressor genes within the second and third quartiles, indicating that the effect is not solely attributable to gene length (**Supplementary Figure 9A**). Similarly, stratification by TMB revealed a progressive reduction in gene expression with increasing passenger mutations across low, medium, and high TMB groups (**Supplementary Figure 9B**), demonstrating that the observed relationship is not driven by overall mutation burden or hypermutated samples.

|||||

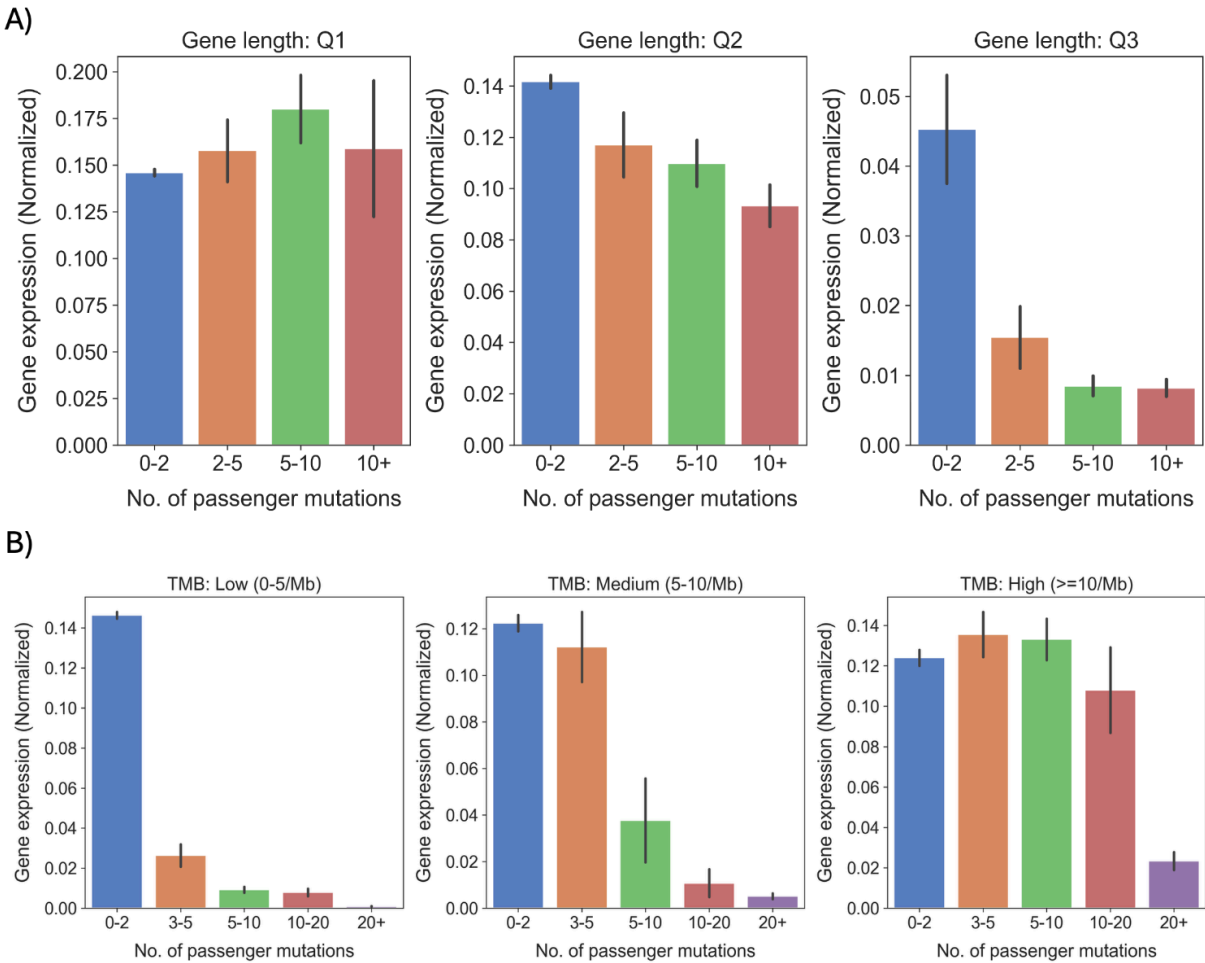


**Figure R4: Cumulative effects of passenger mutations on gene expression in Tumor Suppressor genes, stratified by gene length and TMB. (A)** Relationship between the number of passenger mutations in the absence of drivers and gene expression levels in tumor suppressor genes, stratified by gene length. **(B)** Relationship between the number of passenger mutations in the absence of drivers and gene expression levels in tumor suppressor genes, stratified by genome-wide tumor mutation burden (TMB).

To account for potential nonsense-mediated decay (NMD) effects, we removed data points (sample-gene pairs) containing nonsense mutations or frameshift indels and repeated the analysis. The results remained consistent, indicating that the observed decrease in expression with increasing mutation burden is not attributable to an increased likelihood of NMD. Figure 6A in the manuscript was updated accordingly as shown below (**Figure R5**). The Methods section was updated and the following text was added:

"To account for potential nonsense-mediated decay (NMD) effects, data points (sample-gene pairs) containing nonsense mutations or frameshift indels were removed."

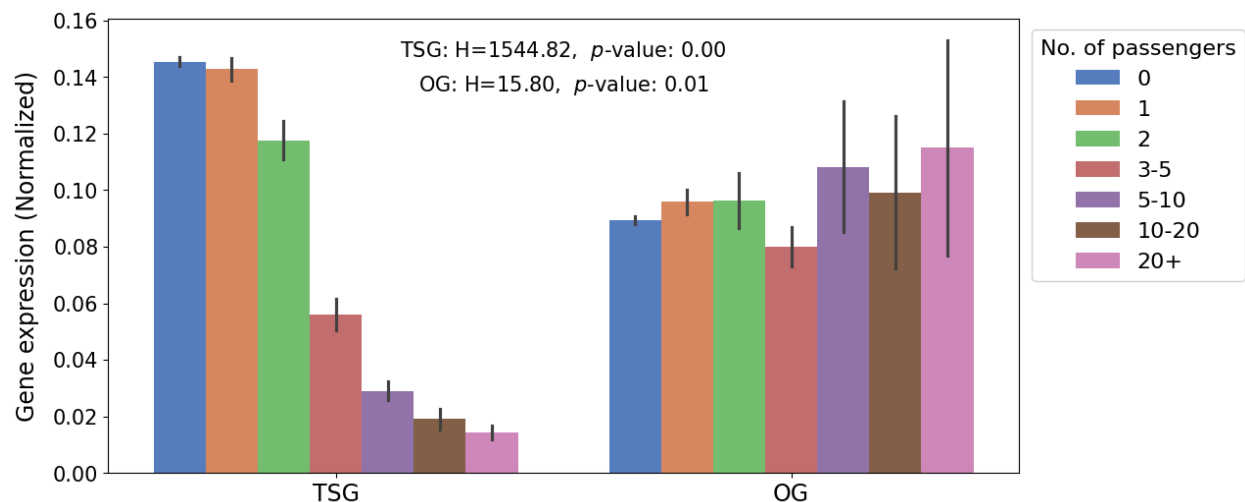


**Figure R5: Relationship between the number of passenger mutations and gene expression levels in cancer genes, in the absence of driver mutations.**

4- The definition of drivers (lines 445-448) seems to include flanking regions over-eagerly „examples without any identified driver mutations in the same gene, including promoter upstream regions (-3,000 base-pairs from the transcription start site) and downstream regions (3,000 base-pairs downstream from the transcription end site).“ 200 nt upstream of TSS would be the main promoter region, and up to 1.5k extended promoter (I have never seen 3k used). Downstream of TES, I am not aware of any mention of driver mutations there and I can't envision a mechanism how these mutations would act. I am not sure at all there is a rationale to excluding TES-downstream mutations.

We thank the reviewer for pointing this out. The mention of “3,000 base-pairs upstream and downstream” was a mistake in the text; we did not include these extended regions. In our analysis, we only included regions annotated as lying within the gene according to the PCAWG dataset. We have updated the Methods section to fix this.

For upstream regions, we defined the 1 kb immediately upstream of the transcription start site (TSS) as “upstream” according to ANNOVAR gene-based annotations. In the transcription factor analysis, we treat this 1 kb region as the promoter, and we have updated the manuscript wording to refer to it as the “extended promoter” for clarity. Similarly, the downstream region

refers to the 1 kb immediately downstream of the transcription end site (TES), which is labeled as “downstream” in ANNOVAR gene-based annotations.

5- CADD score is a little dated and not highly performant by today's standards. I understand the appeal in using the same score for coding and nc mutations, however at least for coding there are much better methods available (AlphaMissense, ESM, REVEL, Varsity etc).

In contrast, for splicing mutations Pangolin indeed is the state-of-the-art, so kudos for that. The Fig 4C plot contains a somewhat surprising result, which is that the predicted splicing effect of drivers (most of which, presumably, are selected via direct changes on aa sequence) seems to be similar to that of deleterious passengers. Can they comment on this, are there known driver hotspots with strong splicing effects, which could be a novel mechanism therefor?

We thank the reviewer for this valuable suggestion. We agree that for coding variants, more recent methods outperform CADD. We have now included an additional analysis of coding passenger mutations using AlphaMissense, which provides state-of-the-art pathogenicity predictions. We find that the results are consistent with our previous analysis using CADD scores. Specifically, we observe an enrichment of normalized AlphaMissense scores for passenger mutations in samples lacking driver mutations compared to those with identified drivers (**Figure R6**).

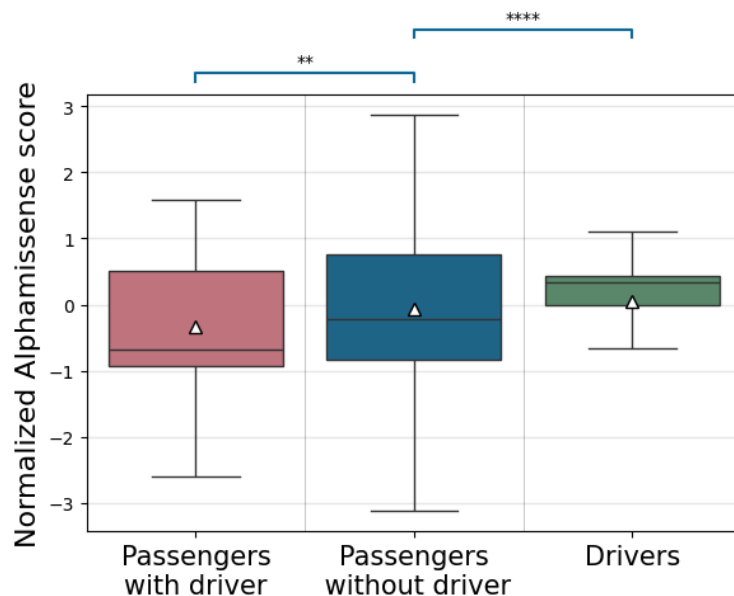


**Figure R6: Pan-cancer distribution of individual z-score normalized AlphaMissense scores for coding passenger mutations that emerge in the presence and absence of driver mutations.** Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

We have included this result in Figure 3B in the manuscript and added the following text in the Results and Methods section of the manuscript.

## Results section:

''''

These findings are consistent when using pathogenicity scores from Alphasense for protein-coding passenger mutations (Mann-Whitney U test,  $p$ -value<0.01; **Figure 3B**).

''''

## Methods section:

''''

For coding mutations, pathogenicity scores were also obtained from Alphasense using the predictions file AlphaMissense\_hg19.tsv.gz on Zenodo (DOI:10.5281/zenodo.8208688).

...

Pathogenicity scores from Alphasense were also normalized in a similar manner.

''''

We also thank the reviewer for highlighting the observation regarding splicing effects of known drivers in Fig 4C. In our dataset, we identified 203 driver mutations with predicted splicing effects, as illustrated below (**Figure R7**). Among these, 128 distinct driver mutations across 31 genes show strong predicted splicing disruption (Pangolin score > 0.5). Notably, *TP53* accounts for the largest fraction of these events, comprising approximately 35% of all driver mutations with strong predicted splicing effects. While several mutations are predicted to both weaken canonical splice sites and create alternative splice sites at different loci, loss-of-splicing effects predominate overall, accounting for 72.8% of predicted events compared with 27.2% gain-of-splicing events caused by known driver mutations.

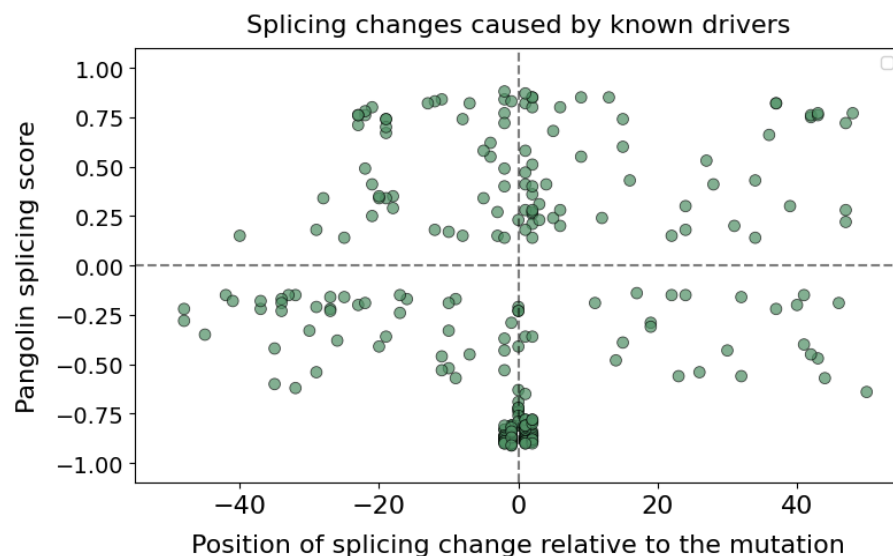


**Figure R7: Relative positions and predicted Pangolin scores for splicing changes caused by known driver mutations.**

6- For survival analysis, sample sizes are quite limited; these 710 samples are, I infer, the TCGA WGS subset within their PCAWG set. This makes one wonder why did they limit the analysis to the (fairly small) PCAWG, when there is more WGS available. For one, there is the Hartwig cohort; second, TCGA have released WGS mutation calls for several thousand samples via Genomic Data Commons. Having more samples would also stabilize their – currently quite messy and not showing clear trends – per-cancer-type plot in Fig 2E (note: this needs confidence intervals and/or a statistical test for every cell). I am not sure I would exactly call this „consistent enrichment“ (lines 149-150) as it is currently.

We thank the reviewer for this suggestion. Our original analysis was restricted to the PCAWG subset due to the comprehensive driver annotation and matched gene expression profile available for samples in that cohort. While we do not have access to the Hartwig cohort at this time, because we have encountered difficulties in procuring the data due to restrictions and contractual language requested by Hartwig that our institution is not permitted to sign, we have now included additional 8,042 samples from TCGA dataset to increase sample size. We have included these results in supplementary material as **supplementary figure 3**.

Consistent with results from the PCAWG dataset, passenger mutation density was found to be elevated in samples harboring hotspot mutations within the same gene compared to those without hotspots (Mann-Whitney U test,  $p$ -value < 0.0001; **Figure R8-A**). Passenger mutations were also more frequently observed than expected by chance (Chi-square test of homogeneity,  $p$ -value <  $10^{-6}$ ; **Figure R8-B**). When further stratifying by cancer subtype in the TCGA cohort, the differences in passenger densities remained consistent (Mann-Whitney U tests; **Figure R8-C**).

Kaplan-Meier analysis of overall survival further showed that patients harboring passenger mutations in the absence of driver mutations have significantly poorer survival outcomes (log-rank test,  $p_1 < 10^{-6}$ ; **Figure R8-D**), with survival progressively worsening as passenger mutation burden in cancer genes increases (log-rank tests,  $p_1 < 10^{-6}$ ,  $p_2 < 10^{-6}$ ; **Figure R8-E**). Overall, the results from the expanded TCGA cohort closely mirror our original findings.

Additionally, we agree with the reviewer, regarding Figure 2E, and have now removed it from the manuscript.

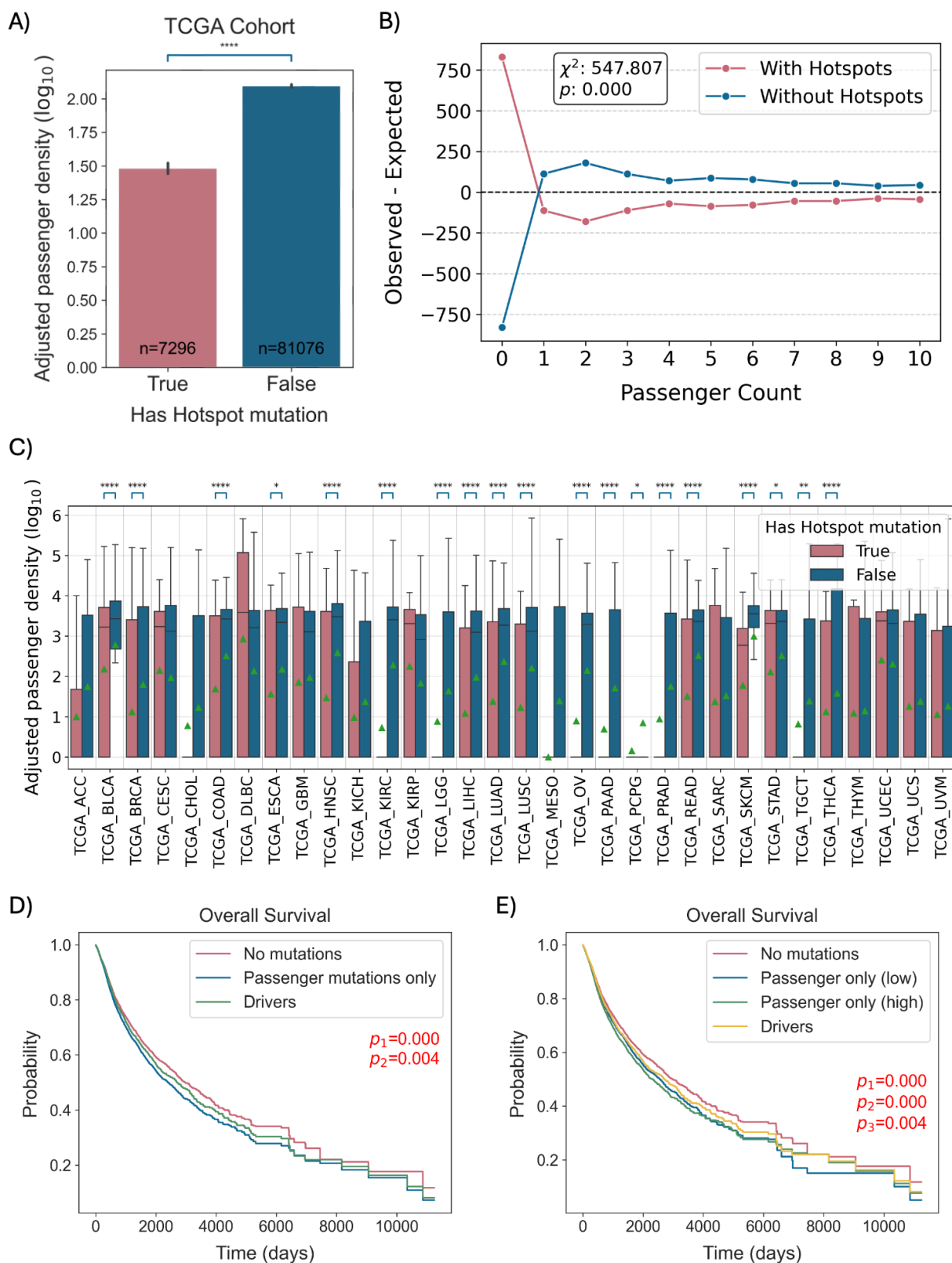


Figure R8: Results of core analyses reproduced in the TCGA dataset using cancer hotspots data

**for driver annotation. (A)** Mean adjusted passenger mutation density in samples with and without hotspot mutation in a cancer genes. **(B)** Difference in observed and expected number of samples with varying passenger mutation counts in a cancer gene. **(C)** Distribution of differences in mean passenger density in samples without and with driver mutations for each cancer gene, stratified by cancer subtype cohort. **(D)** Kaplan-Meier Survival plot for overall survival (OS) by mutational status of a cancer gene. **(E)** Kaplan-Meier Survival plot for overall survival (OS) stratified by passenger mutation burden in cancer genes. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

7- CNA seem to be fully ignored in the analyses, although they acknowledge at one point (for central nervous system cancer) they might explain some of the trends in their data. I think this is not a dealbreaker point because, if anything, ignoring CNA probably biases their results conservatively (as the CNA could substitute for some of the mini-drivers). That said I would have liked to see some quantitation of how big a role they think CNA would play. I understand the mutual-exclusivity analysis with CNA is complicated due to the co-occurrences CNA\*mutation in the same gene, as two-hit events, however e.g. the associations with survival or similar are more easily checkable.

We apologize for the lack of clarity, and thank the reviewer for this important observation. In our analyses, CNA identified by PCAWG as drivers were already included alongside other drivers when grouping samples as “with” or “without” drivers. If a CNV classified as a driver overlapped a cancer gene in a given patient, the corresponding patient–gene pair was considered to harbor a driver mutation in that gene. However, we acknowledge that we did not clearly state this in the manuscript and have now added this explanation in the methods. We provide the updated text below:

“””

If a CNV classified as a driver overlapped a cancer gene in a given patient, the corresponding patient–gene pair was considered to harbor a driver mutation in that gene.

“””

Additionally, we also incorporate the CNA burden to adjust passenger mutation density in various analyses to account for their effects. We have further elaborated the methods section to clarify this. The updated text is provided below:

“””

Mutational density was calculated separately for each group, adjusting for the gene length ( $l$ ), genome-wide mutation rates ( $N$ ), and copy number alteration burden ( $C$ ).

...

Let,

$n$  = No. of passenger mutations in the gene

$L$  = Length of the Genome

$N$  = Total no. of mutations across the genome

$l$  = Gene length

$C$  = CNA burden, i.e., fraction of genome length having copy number alterations

$M$  = Mutation density adjusted for genome-wide mutation rate and CNA burden

$$\text{Mutation rate} = \frac{N}{L}$$

$$\text{Expected rate in gene} = \text{Mutation rate} * \text{Gene length} = \frac{N * l}{L}$$

To quantify the excess of mutations in a gene beyond what is expected from the genome-wide mutation rate and gene length, we normalize the observed number of mutations by the expected mutation rate calculated above. This gives us,

$$\begin{aligned} \text{Mutation density after accounting for the expected mutation rate in the gene} &= \frac{n}{N * l / L} \\ &= \frac{n * L}{N * l} \end{aligned}$$

$$\text{Mutation density after adjusting for CNA burden, } M = \frac{n * L}{N * l * C}$$

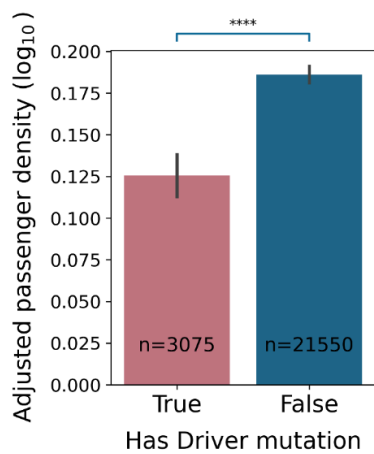
“””

Regarding the reviewer’s comment on the impact of CNVs on our conclusions, we performed our density analyses both unadjusted and adjusted for CNA burden. The results were consistent across both approaches (**Figure R9-A,C** and **Figure R9-B,D** respectively). For completeness, we present in the manuscript the analyses that incorporate CNV driver events as identified by the PCAWG consortium.

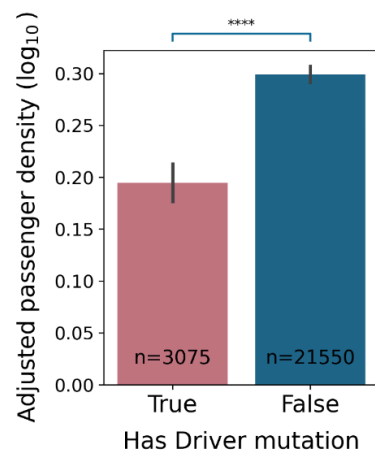
As suggested by the reviewer, we have now also performed an additional analysis to assess the association between copy number alteration (CNA) burden and overall survival in a pan-cancer cohort. CNA burden was modeled as a continuous variable in a Cox proportional hazards model, with hazard ratios reported per 10% increase in genome altered to evaluate the global, unadjusted impact of CNA burden on survival across tumors. We found that each 10% increase in CNA burden is associated with a 5% increase in the hazard of death (HR = 1.05, 95% CI = 1.01-1.08,  $p = 0.006$ ). This indicates a statistically significant but modest association between higher genomic instability and worse overall survival.

Further, when studying the Partial effects of the number of passengers as a covariate in the Cox proportional hazards model for progression-free survival and overall survival (**Figure 6D**), we already incorporate CNA burden as a covariate to account for these effects.

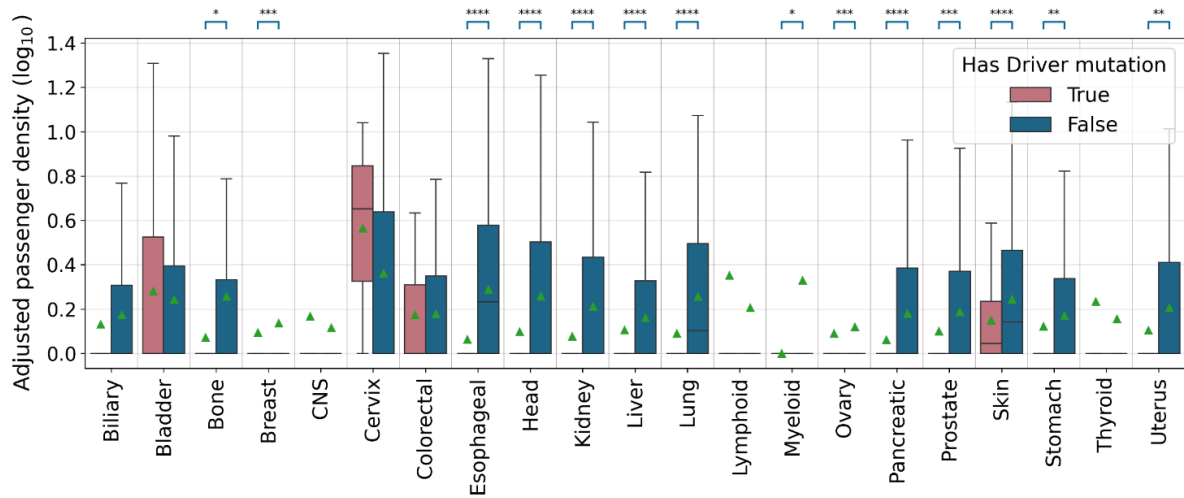
A) Without CNA Adjustment



B) With CNA Adjustment



C) Without CNA Adjustment



D) With CNA Adjustment

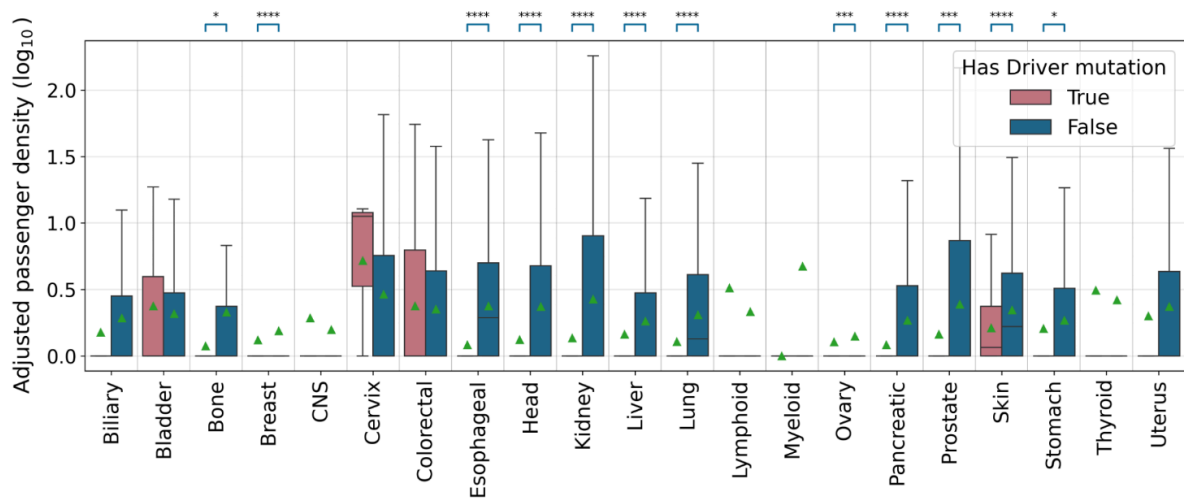


**Figure R9: Results of passenger density analysis with and without CNA burden adjustment. (A)** Mean adjusted passenger mutation density in samples with and without hotspot mutation in cancer genes, without CNA burden adjustment. **(B)** Mean adjusted passenger mutation density in samples with and without hotspot mutation in a cancer genes, with CNA burden adjustment. **(C)** Distribution of differences in mean passenger density in samples without and with driver mutations for each cancer gene, stratified by cancer subtype cohort, without CNA burden adjustment. **(D)** Distribution of differences in mean passenger density in samples without and with driver mutations for each cancer gene, stratified by cancer subtype cohort, with CNA burden adjustment. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

8- I don't quite see the point in the Fig. 2D analysis of mutation rates of different segments within a gene. I do see the interest in studying TSS proximal mutation rates (upstream: promoter; downstream: 5' UTR and sometimes intragenic enhancers) however the relevance of the rest of this analysis is less obvious. The up-and-down jumps are probably resulting from sparse data, although the „tube“ around the loess fit suggests otherwise; maybe larger bins would be more appropriate given the dataset size, and/or absolute-size bins anchored on the TSS or TES. Consider revising or dropping.

We thank the reviewer and agree that the results in Figure 2D were noisy. We have now removed this panel and replaced it with a pan-cancer passenger density analyses of the 8,042 additional WGS tumor samples from the TCGA dataset and the 570 WGS tumor samples from POG570 dataset (**Figure 2A**), demonstrating that the results are consistent across datasets.

9- The study introducing the SEISMIC method ( <https://doi.org/10.1038/s41467-022-34746-z> ) is conceptually related to their work, yet is not cited. Consider discussing.

We thank the reviewer; we have now added this study in our manuscript. We provide the updated text below:

“””

This view is also supported by a recent methodological tool, SEISMIC, which infers selection on somatic mutations by modeling tumor-type-specific background mutation rates and has revealed widespread signals of weak positive selection acting on mutations outside canonical driver events<sup>22</sup>. Kumar et al. previously conducted a comprehensive analysis of passenger mutations using the PCAWG datasets and identified weak driver-like effects among subsets of passenger mutations<sup>9</sup>. Building on this work, we demonstrate that weak effects are pervasive, and can accumulate within cancer genes to produce substantial, functionally relevant consequences, thereby revealing a collective mode of positive selection that was previously unappreciated.

“””

Reviewer #2 (Remarks to the Author):

The manuscript "The cumulative impact of passenger mutations on cancer development" reanalyses a large set of somatic mutations from public cancer genome datasets. It critically revises the notion of a passenger mutation as a neutral mutation without significant impact on cancer biology. The manuscript describes a number of statistical analysis, of which some present compelling evidence that there are weak driver mutations among the passengers, which, especially in absence of drivers, can contribute to explaining the genetic causes of cancer. In other cases it is not entirely clear if the observed effects are a consequence of statistical artifacts that need to be corrected for. While the manuscript addresses a very important question and discusses several interesting aspects, the reported results are difficult to interpret, as it is often unclear what and how many datapoints exactly are used (most figures do not report the group sizes, respectively, the  $n$ ). Also, the null hypothesis are often unclear or a baseline model is missing.

Before a publication a major revision is required.

We thank the reviewer for the helpful and constructive comments. We provide a point-by-point response below for each of the reviewers' comments.

Major comments:

1. It is not always clear what statistical hypotheses you are testing exactly. In line 107-110 you test differences on gene level for samples in which a passenger mutation co-occurs with a driver mutation against samples in which passenger mutations do not co-occur with driver mutations. How many observations do you have per group? Why do you compare frequencies instead of events?

We thank the reviewer for the comment. The specific hypothesis we tested in this analysis is whether the passenger mutation density in a cancer gene varies in the presence and absence of driver mutations in the gene. Formally,

**Null hypothesis ( $H_0$ ):** The mean passenger density across cancer genes is the same in observations (patient-gene pairs) with a driver mutation as in those without a driver mutation.

**Alternative hypothesis ( $H_1$ ):** The mean passenger density differs between the two groups.

We have updated the description (line 107-110) to clarify this. The updated text is as follows:

"To examine this, we tested whether the mean passenger density across cancer genes varies in observations (patient-gene pairs) with a driver mutation and in those without a driver mutation."

As shown in **Figure R10** below, in PCAWG cohort which is the primary cohort for our study, there are 3,075 observations in Group 1 (patient-gene pairs with a driver mutation in the gene) and 21,550 observations in Group 2 (patient-gene pairs without a driver mutation in the gene). Figure 1 in the manuscript has been revised to include this panel which shows the number of observations in each group (**Figure 1A**).

We compare passenger densities rather than just the event (passenger occurrence) to account for gene length being a potential cofounder, as longer genes would be more likely to accumulate mutations by chance.

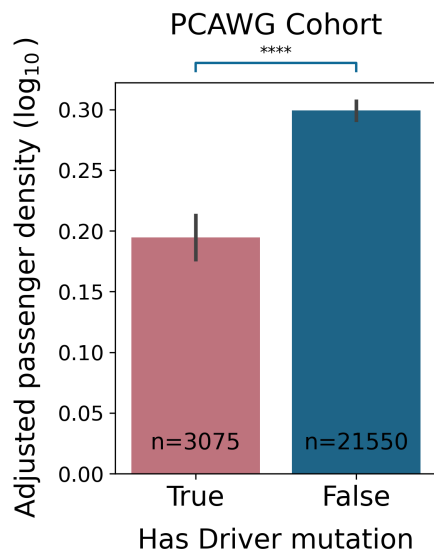


**Figure R10: Mean adjusted passenger mutation density in samples with and without driver mutation in cancer genes.** The number of data-points (patient-gene pairs) in each group are depicted as n. Mann-Whitney U tests with  $p$ -value is displayed as \*\*\*\* for  $p$ -value<0.0001.

2. As mutations in genes are rare events, it is likely that the majority of all mutated cancer genes have exactly one mutation. Therefore, it is not surprising that you rarely observe a co-occurrence between driver and passenger mutations. To understand the relevance of your observation, you have to compute a baseline model in which you test how often you observe a second passenger mutation in a gene and sample in which you have already observed one passenger, against passengers that remain solo. Compute and report on this baseline model.

We thank the reviewer for raising the concern that most mutated genes might carry only a single mutation, which could make passenger co-occurrences rare. As suggested, we computed a baseline model by counting the number of patient-gene pairs in which we observe exactly one passenger mutation versus those in which we observe two or more passenger mutations within the same cancer gene. We found that co-occurring passenger events are more common, as approximately 57% of passenger-mutated genes exhibit two or more passenger mutations, whereas only ~43% harbor a single passenger (**Figure R11**). This demonstrates that the co-occurrence of passenger mutations is widespread.

We have also included this figure in the manuscript as **supplementary figure 1A** and updated the results section with the following text:

.....

To rule out the possibility that the increased passenger density observed in the absence of driver mutations simply reflects the rarity of mutations in individual genes, we constructed a baseline model comparing patient-gene pairs with one versus multiple passenger mutations within the same cancer gene. We found that co-occurring passenger mutations are common, with ~57% of passenger-mutated genes harboring two or more passenger mutations and ~43% containing a single passenger mutation (**Supplementary Figure 1A**), demonstrating that passenger co-occurrence is widespread.

!!!!

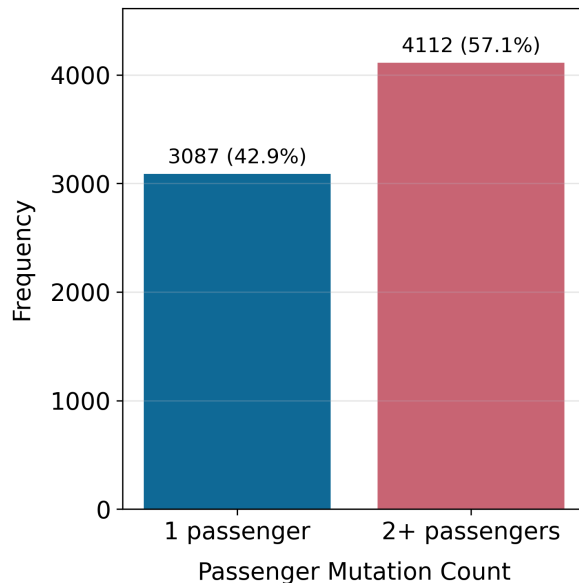


**Figure R11: Baseline model showing the prevalence of multiple passenger mutations within the same cancer gene, across patient-gene pairs in the PCAWG cohort.** Bar plot showing the proportion of patient-gene pairs harboring exactly one passenger mutation versus two or more passenger mutations.

3. The same holds true for the analysis in figure 2B). The method as described in lines 443-469 indicate that you do not account for the fact that the driver mutation should be counted as a mutation event as well. Instead, it is simply used to define two different groups. Under the null hypothesis that driver and passenger are the same you basically introduce an offset of one between the groups and with the consequence that the driver group is less often mutated. Compute a second model in which you replace “Passenger Count” by “Mutation Count” and reevaluate your analysis.

We thank the reviewer for this thoughtful suggestion. We believe, under the assumption that driver and passenger mutations arise independently, the expected number of passengers in a gene should be identical under our null hypothesis regardless of whether a driver is present, i.e.,  $E[\text{Passengers} \mid \text{Driver} = \text{True}] = E[\text{Passengers} \mid \text{Driver} = \text{False}]$ . This equality holds precisely because independence means that knowing a driver occurred provides no information about passenger accumulation. Our results indicate that this equality does not hold and, when considered alongside the other findings of the manuscript and after controlling for gene length

and genome-wide mutation rate, are consistent with positive selective pressure for passenger mutations in the absence of driver mutations.

4. Figure 2 C): Report the number of observations for each column. The reader cannot directly understand if some of the columns are supported by millions of observations, while others may only be backed by a single one. UTR5 mutation are, for instance, much rarer then intronic mutations.

We thank the reviewer for this comment. We have reported the number of observations for each column. For the density calculations, we include all samples, including those with zero mutations in a given region, to accurately reflect the mutational burden across the cohort. This ensures that rarer regions, such as 5' UTRs, are not overrepresented relative to more common regions like introns and that the absence of passenger mutations in a region is properly accounted for. Consequently, the total number of observations is the same across all regions, but here, we have now specifically reported the number of data points with at least one mutation in each region, as requested.

We have also updated the figure (now Figure 2E following panel rearrangement) and revised the corresponding figure caption in the manuscript accordingly. The updated plot and figure caption are shown below (Figure R12):

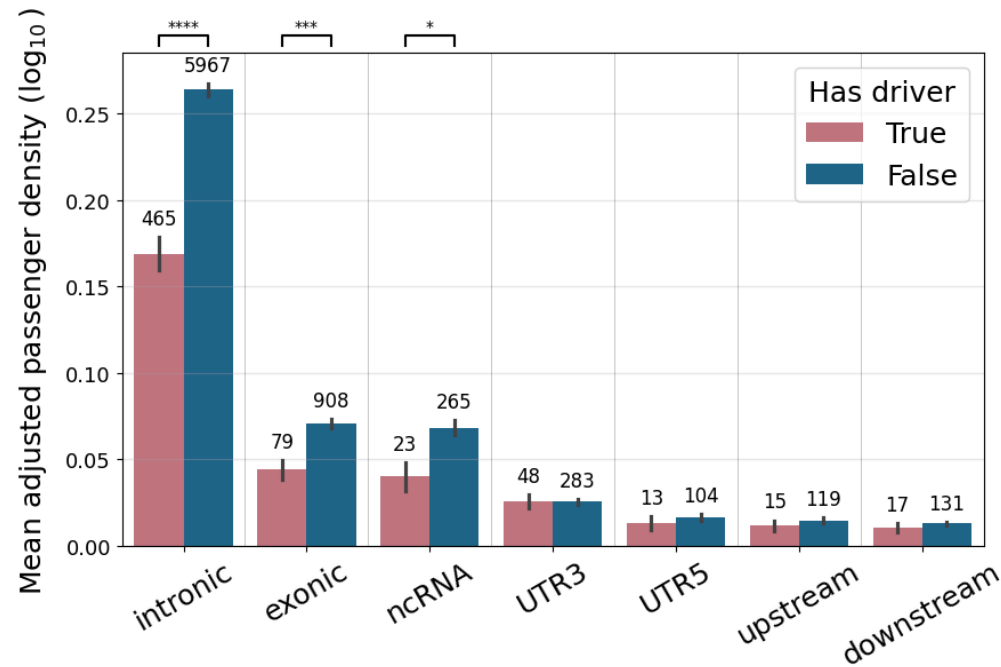


**Figure R12: Mean passenger mutation density in different genic regions.** The number of data points (patient-gene pairs) with non-zero mutations in each group is shown above the bars.

5. Figure 2 D) As discussed above, by not including the driver mutation itself you strongly reduce the mutation frequency. Crate a second version of the plot in which you display the

“Mean mutation density” on the y-axis instead of the “Mean passenger density”. Do you still observe a significant difference between the groups.

We thank the reviewer for the comment. Reviewer 1 suggested that we remove this due to the noisy data. We have now replaced this analysis with reproducing the density analysis in two additional datasets and presenting the mean passenger densities across all 3 cohorts as Figure 1A in the manuscript. The updated figure panel is shown below:

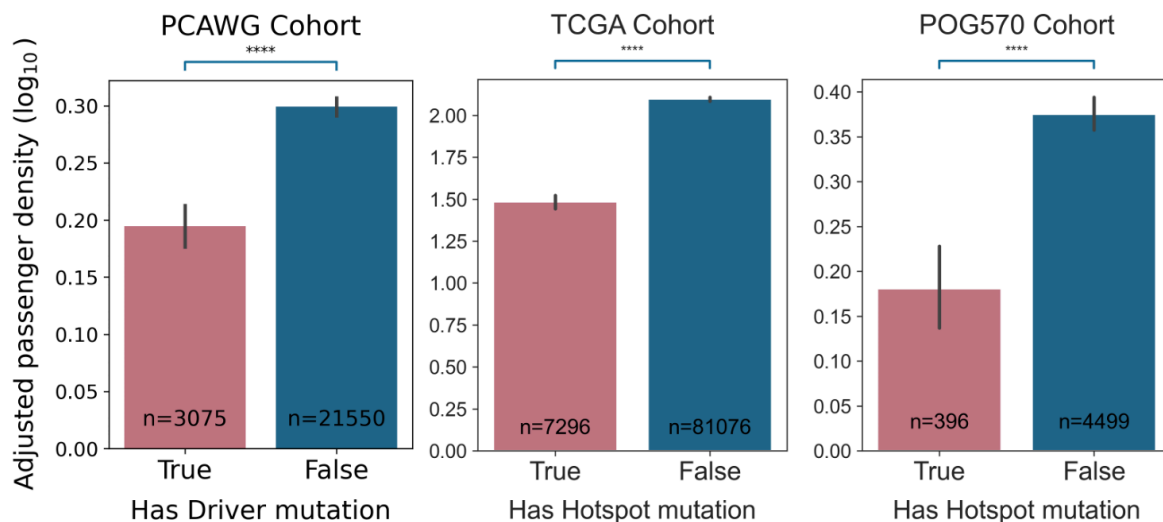


**Figure R13: Passenger mutation density in samples with and without driver mutations.** Mean passenger mutation density in groups of samples with and without driver mutation in a cancer gene presented for 3 different cohorts. The number of data points (patient–gene pairs) in each group is indicated as n. Statistical significance is displayed as \*\*\*\* for  $p$ -value < 0.0001. The Mann-Whitney U test was used to estimate the statistical difference between groups.

6. Figure 3 A)-D) indicate that there may be a subset of passenger mutations that qualifies as weak deleterious drivers mutations (blue curve above 0) in 3 A). This is not informative about the bulk of passenger mutations, but rather hints that the definition of driver mutation you are relying on is not perfect. Your argument that passenger mutations in general have a cumulative effect is therefore not strongly supported by these figures. Separating passenger mutations into real passengers and weak drivers would be a more logical conclusion.

We thank the reviewer for this thoughtful comment. In Figure 3A-D, we indeed observe that a subset of passenger mutations exhibits higher pathogenicity scores than the mean (blue curve above zero). While it is possible that the definition of driver mutation we use is not perfect (we address this below), in our manuscript, we explore this in the context of the alternative model we propose (Figure 1B) in which the effects of cancer mutations lie on a spectrum, with some passengers having weak driver-like effects. Importantly, these mutations appear to exert cumulative effects.

Regarding the reviewer's comment on cumulative effects, we clarify that Figure 3 is not intended to by itself demonstrate a cumulative impact of passenger mutations. Rather, it demonstrates that, contrary to the prevailing assumption that passengers are strictly neutral, the distribution of passenger pathogenicity scores shows non-neutral patterns suggestive of selection. This supports the idea that passenger mutations may have small but biologically meaningful impacts. The cumulative effects of passengers are demonstrated in Figure 6, which shows that accumulation of passenger mutations results in the inactivation of the tumor suppressor genes (**Figure 6A**) and that higher passenger burden in cancer genes is associated with progressively worse survival outcomes (**Figure 6B-D**).

Further, to address the concern that some of these observations arise from incomplete driver annotations, and to ensure these observations are not cohort- or annotation-specific, we repeated the core analyses using a more comprehensive and independent driver mutation set derived from cancerhotspots.org<sup>1</sup> (v2), which is based on recurrence patterns across approximately 25,000 tumor samples, as suggested by the reviewer. To further assess the robustness of our findings, this hotspot annotation was applied to an independent patient cohort, POG570<sup>2</sup> as well as the TCGA dataset. The hotspot mutations were used to define driver events and all remaining mutations classified as passengers. We reproduced the key findings from our original analysis, including the differential abundance of passenger mutations in the presence versus absence of driver events within cancer genes, as well as the association between passenger mutation burden and overall survival.

In both POG570 and TCGA cohorts, passenger mutation density was found to be elevated in samples harboring hotspot mutations within the same gene compared to those without hotspots (Mann-Whitney U test,  $p$ -value < 0.0001; **Figure R14-A** and **Figure R15-A**) and passenger mutations were more frequently observed than expected by chance (Chi-square test of homogeneity,  $p$ -value <  $10^{-6}$ ; **Figure R14-B** and **Figure R15-B**). When stratifying by cancer subtype in the TCGA cohort, the results remained consistent (Mann-Whitney U tests; **Figure R15-C**). We also find that patients with only passenger mutations in a cancer gene had significantly lower median survival times than patients without any mutations in the gene (Mann-Whitney U test,  $p$ -value < 0.05; **Figure R14-C**). Further analysis using the Kaplan-Meier estimator revealed that patients with passenger mutations in the absence of driver mutations also demonstrated significantly worse survival outcomes (log-rank test,  $p_1=0.022$ ; **Figure R14-D** and  $p_1<10^{-6}$ ; **Figure R15-D**), with the outcome worsening with increase in passenger burden in cancer genes (log-rank test,  $p_1<10^{-6}$ ,  $p_2<10^{-6}$ ; **Figure R15-E**).

These results demonstrate that our conclusions are robust to both the choice of driver mutation set and the patient cohort, and are not driven solely by potential limitations in the original PCAWG driver annotations. We have added these figures as **supplementary figures 2 and 3** and also added the following text in the manuscript:

#### **In Results Section:**

“““““

We also reproduced our key findings in the TCGA dataset and the POG570 cohort using an independent set of cancer hotspot mutations for driver annotation (**Supplementary Figure 2-3**), ensuring the robustness of our results.

...

To further assess robustness, we replicated the analysis in the TCGA dataset and the POG570 cohort using an independent cancer hotspot-based driver annotation. The results were highly concordant, consistently showing an enrichment in the mean adjusted passenger density in the absence of driver (or hotspot) mutations within a cancer gene (Mann-Whitney U test,  $p$ -value<0.0001; **Figure 2A**).

.....

## References:

1. Chang, M. T., Bhattarai, T. S., Schram, A. M., Bielski, C. M., Donoghue, M. T., Jonsson, P., ... & Taylor, B. S. (2018). Accelerating discovery of functional mutant alleles in cancer. *Cancer discovery*, 8(2), 174-183.
2. Pleasance, E., Titmuss, E., Williamson, L. et al. Pan-cancer analysis of advanced patient tumors reveals interactions between therapy and genomic landscapes. *Nat Cancer* 1, 452-468 (2020). <https://doi.org/10.1038/s43018-020-0050-6>

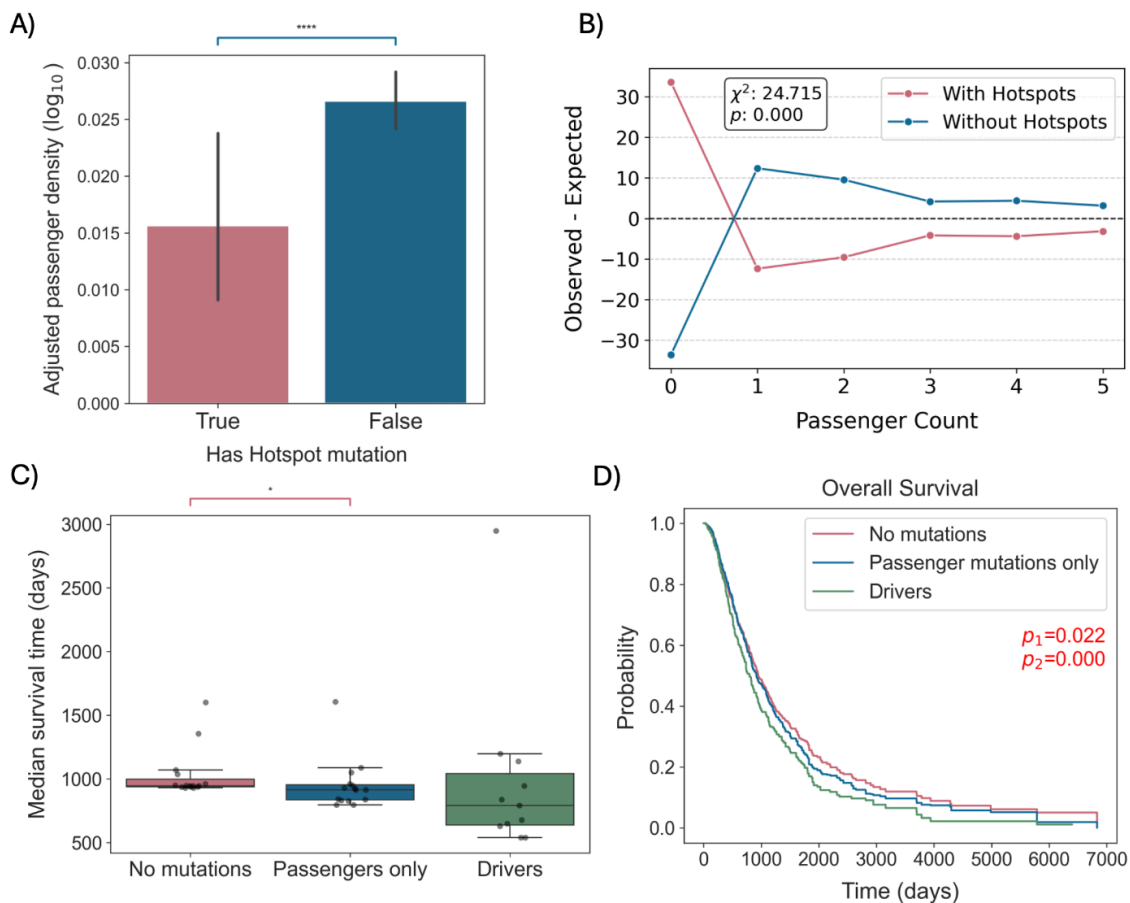


**Figure R14: Results of core analyses reproduced in the POG570 dataset using cancer hotspots data for driver annotation. (A)** Mean adjusted passenger mutation density in samples with and without hotspot mutation in a cancer genes. **(B)** Difference in observed and expected number of samples with varying passenger mutation counts in a cancer gene. **(C)** Distribution of median survival time (in days) of patients. Each data point corresponds to the median survival for a particular cancer gene. **(D)** Kaplan-Meier Survival plot for overall survival (OS) by mutational status of a cancer gene. Statistical significance in panels (A) and (B) is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

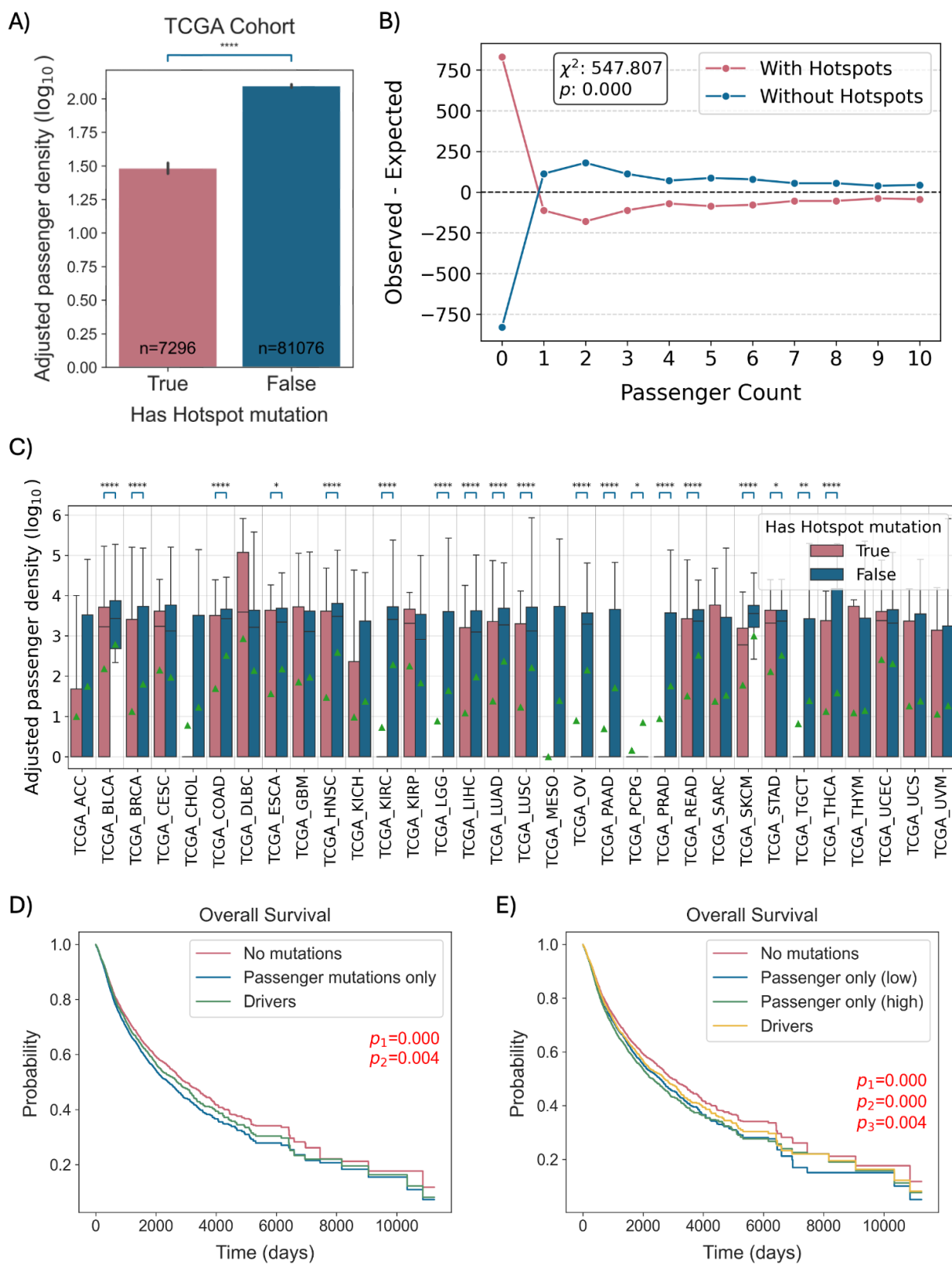


**Figure R15: Results of core analyses reproduced in the TCGA dataset using cancer hotspots data for driver annotation.** (A) Mean adjusted passenger mutation density in samples with and without hotspot mutation in a cancer genes. (B) Difference in observed and expected number of samples with varying passenger mutation counts in a cancer gene. (C) Distribution of differences in mean passenger density in samples without and with driver mutations for each cancer gene, stratified by cancer subtype cohort. (D) Kaplan-Meier Survival plot for overall survival (OS) by mutational status of a cancer gene. (E) Kaplan-Meier Survival plot for overall survival (OS) stratified by passenger mutation burden in cancer genes. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

7. You may want to explicitly state that a considerable portion of driver mutations are actually gain-of-function and thus cannot be characterized by the CADD score.

We thank the reviewer and agree with their point. We have now added the following to explicitly state that gain-of-function driver mutations may not be captured with CADD scores:

“However, we note that because many driver mutations are gain-of-function, deleteriousness scores like CADD may not capture their impact; we therefore use CADD mainly as a proxy for passenger-impact potential rather than driver strength.”

8. From the method section it is unclear if you used scaled or unscaled C-scores as input when you computed the z-score of the CADD score. Furthermore, you do not mention in the caption of Figure 3 that you do not directly show the CADD scores, but their z-scores. This is only indicated in the Method section lines 496–497. Clarify this to avoid confusion.

We thank the reviewer for pointing this out. We used raw (unscaled) CADD scores to compute the z-scores shown in Figure 3. We have updated both the Methods section and the Figure 3 legend to explicitly indicate that the figure displays z-scores of raw CADD scores, rather than the original CADD scores themselves.

The updated methods section text is as follows:

"To facilitate comparative analysis, we obtained raw CADD scores and z-score normalized them within each gene."

The revised figure caption with updates highlighted is as shown below:

“““““

Figure 3: Pathogenicity profiles of passenger mutations in the presence or absence of driver mutations across cancer types. (A) Pan-cancer distribution of individual **z-score normalized CADD scores** for passenger mutations that emerge in the presence and absence of driver mutations. (B) Pan-cancer distribution of individual z-score normalized Alphasense scores for coding passenger mutations that emerge in the presence and absence of driver mutations. (C) Distribution of **z-score normalized CADD scores** for passenger mutations in samples without driver mutation in a driver gene versus samples with driver mutation in the driver gene, by cancer tissue type. (D) Distribution of cumulative **z-score normalized CADD scores** of passenger mutations in samples without driver mutation in a driver gene versus samples with

driver mutation in the driver gene for each cancer subtype. (E) Distribution of **z-score normalized CADD scores** for passenger mutations in samples without driver mutation in a driver gene versus samples with driver mutation in the driver gene, in different genic regions. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001. The Mann-Whitney U test was used to estimate the statistical difference between groups. Effect sizes in (E) are quantified using Cohen's d.  $p$ -values were adjusted for multiple testing using the Benjamini-Hochberg procedure, and the corresponding FDR-adjusted  $q$ -values are reported.

“””

9. Figure 3 E) in lines 193-195 you mention effect sizes, but cite the Man-Whitney U test, which usually reports significance instead. Clarify were you compute p-values and were you report statistical effect size such as Cohens d. Optimally, report both. As your analysis is overpowered due to pooling mutations across different cohorts, the statistical effect size may be more informative than p-values. Also consider to corrected for multiple testing when the number of tests.

We thank the reviewer for this comment. For Figure 3E,  $p$ -values were computed using the Mann-Whitney U test, and multiple testing was performed using the Benjamini-Hochberg FDR correction. To address the reviewer's suggestion, we have now also included Cohen's d as a measure of effect size. We have updated the figure, figure legend and Methods section to indicate that both FDR-adjusted  $p$ -values and Cohen's d are reported. The revised panel is shown below:

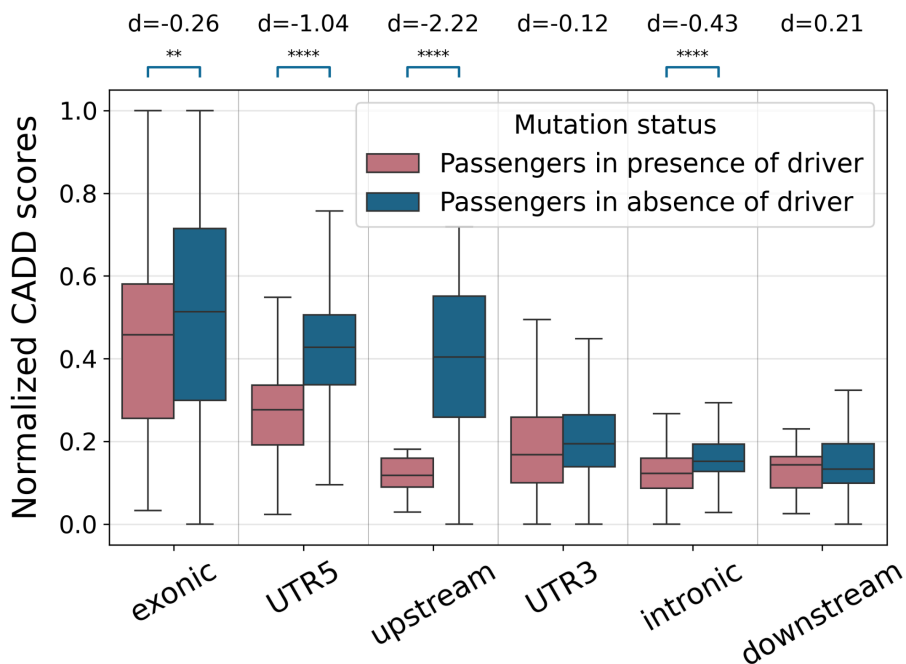


**Figure R16: Distribution of z-score normalized CADD scores for passenger mutations in samples without driver mutation in a driver gene versus samples with driver mutation in the driver gene, in different genic regions.**

The revised figure caption with updates highlighted is as shown below:

“”””

Figure 3: Pathogenicity profiles of passenger mutations in the presence or absence of driver mutations across cancer types. (A) Pan-cancer distribution of individual z-score normalized CADD scores for passenger mutations that emerge in the presence and absence of driver mutations. (B) Pan-cancer distribution of individual z-score normalized Alphasense scores for coding passenger mutations that emerge in the presence and absence of driver mutations. (C) Distribution of z-score normalized CADD scores for passenger mutations in samples without driver mutation in a driver gene versus samples with driver mutation in the driver gene, by cancer tissue type. (D) Distribution of cumulative z-score normalized CADD scores of passenger mutations in samples without driver mutation in a driver gene versus samples with driver mutation in the driver gene for each cancer subtype. (E) Distribution of z-score normalized CADD scores for passenger mutations in samples without driver mutation in a driver gene versus samples with driver mutation in the driver gene, in different genic regions. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001. The Mann-Whitney U test was used to estimate the statistical difference between groups. **Effect sizes in (E) are quantified using Cohen's d. In (C) and (E),  $p$ -values were adjusted for multiple testing using the Benjamini–Hochberg procedure, and the corresponding FDR-adjusted  $q$ -values are reported.**

“”””

Methods section was updated to include the following text:

"Multiple testing correction was performed using the Benjamini-Hochberg procedure, and effect sizes were quantified using Cohen's d."

We also note that throughout the manuscript, we perform multiple testing corrections, systematically.

10. Figure 4A) The text and the axis label indicate that TF binding data is shown in this panel, while the caption in line 251 refers to „Splicing effects of passenger mutations“. Also the headline of figure 4 only mentions splicing changes. Correct this.

We thank the reviewer for catching this error. The figure captions for Figures 4A and 4C have been corrected, and the Figure 4 title has been updated to include both splicing and transcription factor binding effects. The updated figure caption is as follows:

“”””

**Figure 4: Excess of mutations causing splicing changes or altering transcription factor binding affinity in cancer genes lacking driver mutations. (A)** Effect of passenger mutations in the presence or absence of driver mutations, and of driver mutations, on transcription factor binding affinity. **(B)** Differential effects on transcription factor binding affinity caused by passengers in the presence and absence of driver mutations. **(C)** Density of passenger and driver mutations associated with splicing changes, with passengers stratified by the presence or absence of drivers. **(D)** Effects of somatic mutations on splicing relative to the distance from the mutation, for individual cancer types. **(E)** Genes that contain passenger mutations causing large splicing changes (absolute Pangolin score  $\geq 0.8$ ) in different cancer types. **(F)** Distribution of large splicing changes (pangolin score  $\geq 0.8$ ) relative to the causative passenger and driver mutations for selected cancer types with high prevalence of splicing disruptive mutations. Statistical significance is displayed as \* for  $p\text{-value} < 0.05$ , \*\* for  $p\text{-value} < 0.01$ , \*\*\* for  $p\text{-value} < 0.001$ , and \*\*\*\* for  $p\text{-value} < 0.0001$ . In (A) and (B),  $p$ -values were adjusted for multiple testing using the Benjamini–Hochberg procedure, and the corresponding FDR-adjusted  $q$ -values are reported.

''''''

11. Figure 4B while the Method section states that you have computed adjusted  $p$ -values, the y-axis label states only  $p$ -value, and the caption does not provide additional information. Be precise to avoid confusion. Consider to use the term  $q$ -value as abbreviation for adjusted  $p$ -value.

We thank the reviewer for pointing this out. In Figure 4B, the values shown correspond to multiple-testing-adjusted  $p$ -values computed using the Benjamini–Hochberg procedure. As suggested, we have now updated the y-axis label to “ $q$ -value”, and clarified this terminology consistently throughout the Methods and Results sections. We have also included the following text in figure caption for clarification:

"In (A) and (B),  $p$ -values were adjusted for multiple testing using the Benjamini–Hochberg procedure, and the corresponding FDR-adjusted  $q$ -values are reported."

12. Regarding your observation that survival times correlate with passenger mutations in driver genes: Survival strongly correlates with overall mutation load in many cancer types. You need to convince the reader that your observation is not a consequence of that effect, as higher baseline mutation rate results in more mutations in all genes, including passenger mutations in driver genes. Explain how you correct for the background mutation rate in this analysis. Furthermore, report the numbers at risk.

We thank the reviewer for their comment. Increasing evidence has indicated that passenger mutations can play an active role in cancer progression. This is often measured by overall Tumor Mutational Burden (TMB). The current scientific consensus posits that the accumulation of passenger mutations present better targets to the immune system, resulting in a deleterious fitness effect to cancer cells (PMID: 28420421; PMID: 25838375). As a result, higher TMB is associated with improved clinical outcomes including improved response to immunotherapies

and other precision oncology methods (PMID: 23388632; PMID: 25277973). Contrary to this, our data demonstrate that when considering mutational burden in cancer genes, higher burden is associated with worse outcomes.

To control for TMB in the survival analysis, we include TMB as a covariate in the Cox proportional hazard models (**Figure 6D**). Specifically, the model incorporates the total number of driver mutations across the genome, total mutational burden, and CNA burden, in addition to age, sex, event occurrence, and time to event.

Further, we have now stratified the median survival analysis by low (<5 mutations/Mb) and high TMB ( $\geq 5$  mutations/Mb) to demonstrate that our results are consistent across patients with all TMB profiles (**Figure R3-B**). These results have been included in supplementary material as **Supplementary Figure 4B**, along with the following text in the Results section:

"This effect remained consistent when patients were stratified by tumor mutation burden (**Supplementary Figure 4B**)."

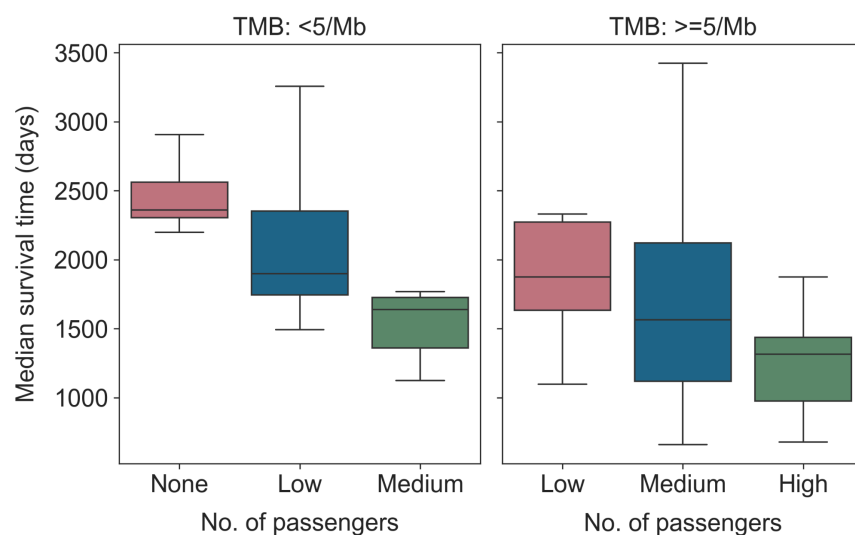


**Figure R17: Median Survival stratified by TMB:** Median Survival of patients with no mutations, low, medium and high burden of passenger mutations in cancer genes, stratified by genome wide tumor mutation burden (TMB).

In our survival analysis, each datapoint corresponds to patient-gene pairs rather than individual patients. As a result, the same patient can sometimes appear in multiple groups across different genes. So, for clarity, we report numbers at risk only in the per-gene analysis shown in Figure 5C. The updated plots with numbers at risk reported is shown below, and **Figure 5C** has been updated accordingly.

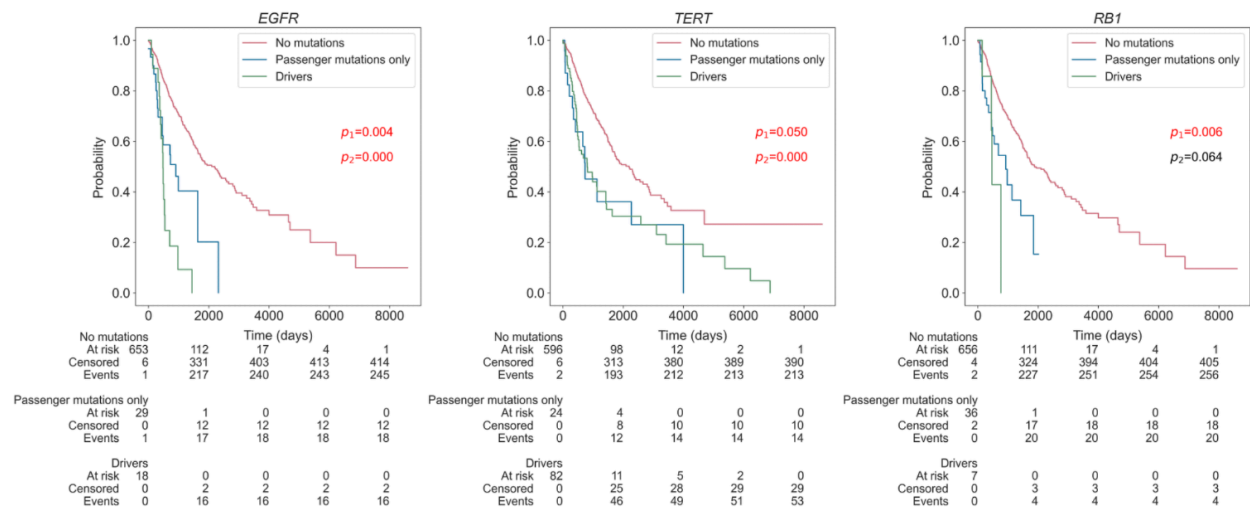


**Figure R18: Kaplan-Meier Survival analysis for genes where passenger mutations lead to a statistically significant decrease in the overall survival (OS) of the cancer patient.** Patients are grouped into 3 categories: patients without any mutations, patients with passenger mutations only, and patients with driver mutations. Numbers at risk are reported for each group.  $p_1$  refers to the  $p$ -value for groups "No mutation" vs "Passenger mutations only", and  $p_2$  refers to the  $p$ -value for groups "No mutation" vs "Drivers". Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

13. Figure 6: This analysis may, as well, simply reflect the higher mutation number in cancers with poor outcome. Test this, for instance, by comparing the genome-wide number of mutations in the three ARID2 groups. If passengers=0 has lower average counts then the other two groups, you need to correct for this factor.

We thank the reviewer for this comment. The analysis shown in Figure 6D was performed separately for each cancer type, thereby avoiding confounding due to differences in survival outcomes across cancer types. To ensure that the observed associations are not driven by differences in genome-wide mutation burden, we explicitly accounted for total mutation count in our analyses. Specifically, we used Cox proportional hazards models that included total mutation count as a covariate, together with age at diagnosis, sex, and gene-specific passenger burden (including ARID2 passenger counts), thereby testing the effect of passenger burden independent of genome-wide mutation burden and other covariates.

We have updated the methods for further clarification and provide the revised text here:

.....

Hazard modelling was performed using the lifelines library in Python, separately for each cancer type, using progression-free interval (PFI) and overall survival (OS) as clinical endpoints or events. We applied a Cox proportional hazards model to assess the association between survival and multiple covariates, including the presence of driver mutations in individual cancer

genes, the number of passenger mutations per gene, event occurrence, time to event, age at diagnosis, sex, the total number of driver mutations across the genome, the genome-wide mutation burden, and copy number alteration (CNA) burden. Genes with passenger mutations in at least ten samples were included in the model. Genes for which the number of passenger mutations was identified as a statistically significant covariate ( $p < 0.05$ ), after adjusting for multiple testing using the Bonferroni correction, were selected for further analysis of their partial effects independent of genome-wide mutation burden and other covariates. For genes exhibiting a wide range in the number of passenger mutations, the partial effects on survival were evaluated at the 0<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, and 100<sup>th</sup> percentiles of passenger mutation counts.

''''''

14. Your discussion raises a number of very valid points, but it lacks an analysis of how the driver mutation annotation you are using is actually defined. It should be very clear to everyone in the field that the biological impact of mutations is continuum ranging from neutral mutations, over mutations with weak impact to mutations with strong impact. Putting somewhere on this continuum a threshold and calling everything with stronger impact a driver mutation, is a semantic simplification. Some of your analysis reflect this fact, while others may require some additional corrections. Your discussion is hinting at these aspects (e.g. “challenge the canonical driver-passenger dichotomy” line 371-372), but by still using the term passenger mutation for everything that is not a driver, it is unclear if you address relevance to all passenger mutations or if you consider a set of “weak driver mutations” among them. Clarify your position, or at least discuss likelihood of these two interpretations.

We agree with the reviewer that driver annotation reflects a threshold placed on a continuum of mutational impact rather than a strict biological boundary. In this study, we adopt the PCAWG definition of drivers for consistency, while explicitly demonstrating that many mutations labeled as passengers exhibit measurable functional effects, consistent with weak or context-dependent drivers. We also demonstrate that these effects built up, suggesting a need to incorporate them in a combined model. Our analyses are therefore not intended to imply uniform neutrality of all passenger mutations, but rather to show that a subset of them can collectively exert driver-like effects. We have clarified this point in the Discussion to emphasize that our conclusions support a continuum-based model, where driver and passenger labels represent degrees of functional impact rather than distinct categories.

We have included the following text in Discussion:

''''''

Throughout this work, we use the term passenger mutations to denote variants not annotated as drivers in current catalogs, while recognizing that this group includes mutations spanning a continuum of functional impact, from near-neutral to weak or context-dependent drivers whose effects can accumulate and become biologically consequential.

''''''

We have also expanded our discussion relative to other related works and provide below:

“””

This view is also supported by a recent methodological tool, SEISMIC, which infers selection on somatic mutations by modeling tumor-type-specific background mutation rates and has revealed widespread signals of weak positive selection acting on mutations outside canonical driver events<sup>22</sup>. Kumar et al. previously conducted a comprehensive analysis of passenger mutations using the PCAWG datasets and identified weak driver-like effects among subsets of passenger mutations<sup>9</sup>. Building on this work, we demonstrate that weak effects are pervasive, and can accumulate within cancer genes to produce substantial, functionally relevant consequences, thereby revealing a collective mode of positive selection that was previously unappreciated.

“””

15. The definition of what mutation is a driver mutation and what mutation is a passenger mutation is most central to your manuscript. In the method section this definition is given as: „Driver mutations for each patients were obtained from the ICGC and TCGA datasets“ (line 438). This description is not sufficient to reproduce your analysis. Specify in the method section how you objectively define a driver mutation (or cite the exact publication you are relying on), and where exactly you obtained the mutation list from you are using (e.g. name of supplementary datafile, command you used in a data portal, etc.).

We thank the reviewer for highlighting this. We have now expanded the relevant Methods section to clarify how driver mutations were identified. The updated text in methods section is as follows:

“””

Driver mutations were defined based on the curated list published by the PCAWG<sup>11</sup> consortium. Specifically, we used the following data files from the ICGC/TCGA driver mutation pan-cancer compendium -

1. TableS3\_panorama\_driver\_mutations\_ICGC\_samples.public.tsv.gz (public)
2. TableS3\_panorama\_driver\_mutations\_TCGA\_samples.controlled.tsv.gz (controlled access)

In our analysis, mutations present in these files were classified as driver mutations, while all remaining somatic SNVs and indels were treated as passenger mutations. Additionally, when a CNV classified as a driver overlapped a cancer gene in a given patient, the corresponding patient-gene pair was considered to harbor a driver mutation in that gene.

“””

We have also now included clarification on the exactly how the data were obtained, and have included the following text in Methods section:

“””

The PCAWG data were originally obtained via the ICGC Data Portal website, which has since been decommissioned. However, the same PCAWG data remains accessible through the ICGC-ARGO, and can now be downloaded by following the instructions provided in the

ICGC-25K data access [documentation](#). The driver mutation data for the TCGA portion of the dataset is controlled-access and is available upon dbGaP authorization under accession phs000178.v11.p8.

""""

#### Reference:

1. The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. Nature 578, 82–93 (2020). <https://doi.org/10.1038/s41586-020-1969-6>

#### Minor comments:

- Figures panels are sometimes referenced in lower case and sometimes in upper case. A consistent labeling would be preferable.

We thank the reviewer for pointing this out. We have revised the manuscript to ensure that all figure panels are referenced consistently using uppercase lettering throughout the text.

#### Reviewer #2 (Remarks on code availability):

The manuscript needs to be edited to a sufficient quality level before I would consider to revise the code.

We thank the reviewer for this comment. The manuscript has now been substantially edited for clarity and overall quality, and the corresponding code and documentation has been updated to ensure it reflects the revised analyses and presentation.

#### Reviewer #3 (Remarks to the Author):

##### Summary

Somatic mutations play critical roles in cancer initiation, progression, and metastasis. Most research has concentrated on driver mutations that are believed to underlie the cancer phenotype, overlooking the more numerous passenger mutations which are generally considered biologically neutral. Nayak et al. provide a comprehensive pan-cancer analysis of passenger mutations from whole-genome sequencing data from 2,263 primary tumors across 31 cancer types, highlighting the cumulative importance of passenger mutations in the absence of bona fide driver mutations. They demonstrate that passenger mutations are significantly enriched in cancer genes that lack canonical driver mutations where they are associated with higher predicted pathogenicity. These mutations exert their impact through changes in transcription factor binding, splicing, and gene expression and their presence correlates with

poorer clinical outcomes. This work challenges the traditional binary classification of mutations into driver and passenger categories, proposing instead a continuum model that better captures the complexity of the impact of somatic mutations in cancer.

#### Strengths

1. The manuscript addresses an understudied aspect of cancer genomics by systematically characterizing the cumulative effect of passenger mutations in cancer genes.
2. The results provide valuable insights into mutation density distribution, functional pathogenicity, gene regulatory disruptions, and prognostic implications.
3. The findings emphasize the necessity of considering passenger mutations in cancer models, contributing to our understanding of cancer evolution and prognosis.

We would like to thank the reviewer for the overall positive assessment and helpful comments. We provide a point-by-point response below.

#### Major concerns

1. Can the authors please clarify the exact driver genes and mutations that are studied here? A representative example of a cancer gene with and without driver mutation in Figure 1 would greatly improve clarity of the approach.

We thank the reviewer for this comment. Driver genes for each PCAWG cancer cohort were obtained from the IntoGen database. Genes mutated in at least 2% of samples within a cohort were retained for analysis. From these, the top 10 most frequently mutated driver genes were selected for each cohort. In cases where multiple genes shared the same mutation frequency as the 10th-ranked gene, all were included. These details are described in the Methods section, and for additional clarity and reproducibility, the complete list of selected driver genes for each cohort is now provided in the Supplementary Files.

We have also expanded the relevant Methods section to clarify how driver mutations were identified. The updated text in methods section is as follows:

''''''

Driver mutations were defined based on the curated list published by the PCAWG <sup>11</sup> consortium. Specifically, we used the following data files from the ICGC/TCGA driver mutation pan-cancer compendium -

3. TableS3\_panorama\_driver\_mutations\_ICGC\_samples.public.tsv.gz (public)
4. TableS3\_panorama\_driver\_mutations\_TCGA\_samples.controlled.tsv.gz (controlled access)

In our analysis, mutations present in these files were classified as driver mutations, while all remaining somatic SNVs and indels were treated as passenger mutations. Additionally, when a CNV classified as a driver overlapped a cancer gene in a given patient, the corresponding patient-gene pair was considered to harbor a driver mutation in that gene.

''''''

Additionally, Figure 1 now depicts a representative cancer gene both with and without a driver mutation, highlighting how mutations were classified in our analysis. The updated Figure 1 (1C added) is shown below:

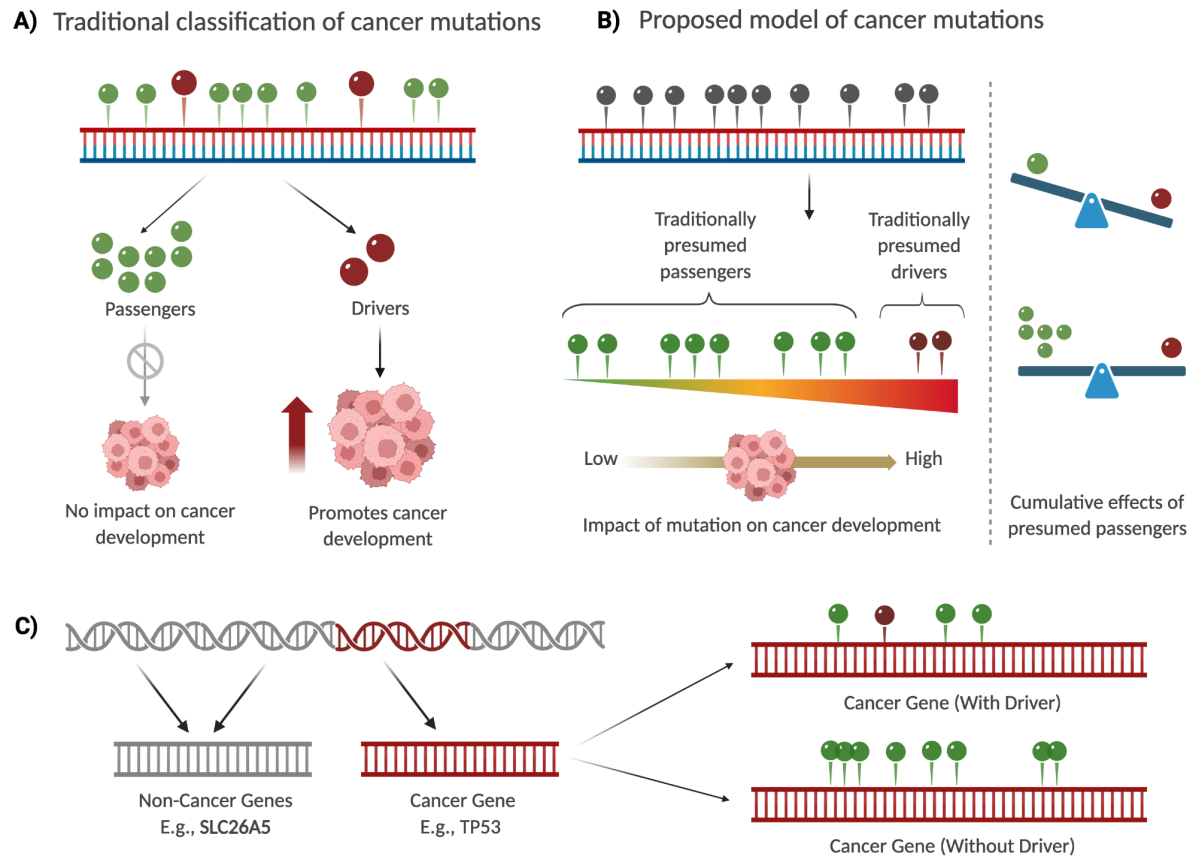


**Figure R19: Modelling the impact of passenger mutations in cancer. (A)** Traditional binary classification of cancer mutations as drivers and passengers. **(B)** Proposed model of cancer development with mutational effects on a spectrum. **(C)** Infographic illustrating the selection of cancer genes and the classification of mutations for our study.

2. Is it generally true that genes without known driver mutations have higher mutation rates, or is this effect driven by few distinct (and potentially very interesting) genes? One would assume that the cumulative passenger effect specifically affects tumor suppressor genes. Can the authors explore this by stratifying their analyses by known oncogenes and tumor suppressors? As they already show in Figure 6A, these gene classes exhibit different properties.

We thank the reviewer for this insightful question. To determine whether the observed increase in passenger mutation rates is driven by a small subset of genes or reflects a broader pattern, we stratified our passenger density analysis by cancer gene class, separating tumor suppressor genes (TSGs) and oncogenes (OGs). We computed, for each cancer gene, the difference in

mean passenger mutation density between samples without driver mutations and those with driver mutations. We find that this difference is predominantly positive for tumor suppressor genes, indicating that passenger mutation density is generally higher in the absence of drivers across many TSGs rather than being driven by a small number of outliers as shown in **Figure R20** here. Oncogenes however did not show any deviation, suggesting a potentially different relationship between passenger accumulation and driver status in this gene class.

This results supports the reviewer's comment that cumulative passenger effects might preferentially affect tumor suppressor genes, and complements the observed differences between oncogenes and tumor suppressor genes shown in Figure 6A. However, the consistency of this shift across many TSGs indicates a widespread phenomenon specific to tumor suppressors rather than the effect being driven by a few genes.

We have now included this plot in figure 2 as panel 2B and updated the results section to include the following text:

"We also examined whether this excess of passenger mutations differed between tumor suppressor genes and oncogenes and found that the increase was predominantly driven by tumor suppressor genes (Figure 2B). "

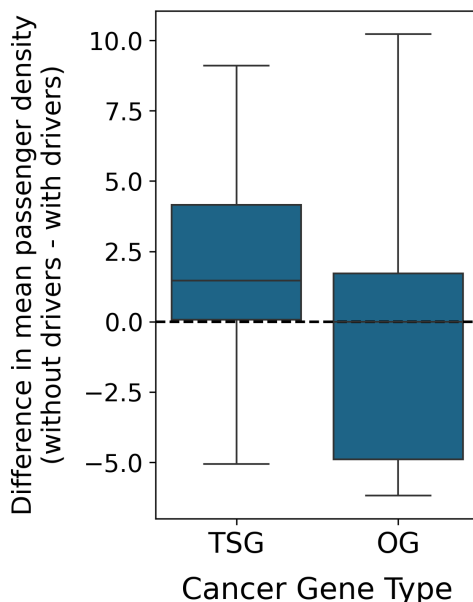


**Figure R20:** Distribution of differences in mean passenger density in samples without and with driver mutations for each cancer gene, stratified by cancer gene type.

3. Is the increased passenger effect potentially confounded by high-level copy number alterations of cancer genes, which may not carry driver mutations but can accumulate more passenger mutations because there are more copies in the cell?

We thank the reviewer for this comment. To account for the potential confounding effects of copy number alterations (CNAs), both due to multiple copies of cancer genes and their genome-wide influence on mutation rate, our passenger mutation density measures are adjusted for CNA burden. We define CNA burden as the fraction of the genome affected by copy number alterations, and passenger densities are divided by this value to correct for CNAs. This ensures that the observed increase in passenger density reflects the accumulation of passenger mutations beyond what would be expected due to the effects of copy number alterations. We have updated the Methods to clarify that this adjustment was performed. The revised text is provided below, along with relevant details from the Methods describing how the density measures were adjusted for CNAs.

''''

Mutational density was calculated separately for each group, adjusting for the gene length ( $l$ ), genome-wide mutation rates ( $N$ ), and copy number alteration burden ( $C$ ).

...

Let,

$n$  = No. of passenger mutations in the gene

$L$  = Length of the Genome

$N$  = Total no. of mutations across the genome

$l$  = Gene length

$C$  = CNA burden, i.e., fraction of genome length having copy number alterations

$M$  = Mutation density adjusted for genome-wide mutation rate and CNA burden

$$\text{Mutation rate} = \frac{N}{L}$$

$$\text{Expected rate in gene} = \text{Mutation rate} * \text{Gene length} = \frac{N * l}{L}$$

$$\begin{aligned} \text{Mutation density after accounting for the expected mutation rate in the gene} &= \frac{n}{N * l / L} \\ &= \frac{n * L}{N * l} \end{aligned}$$

$$\text{Mutation density after adjusting for CNA burden, } M = \frac{n * L}{N * l * C}$$

''''

4. Could the authors please comment in the discussion on negative selection of genes with a driver mutation, where additional passenger mutations might decrease fitness? (contrasting positive selection of passenger mutations).

Yes, we agree with the reviewer, this is plausible. We have updated the discussion to include this possibility and we provide the updated discussion text below:

'''''

In contrast to the positive selection observed for passenger mutations in cancer genes lacking canonical drivers, genes already harboring strong driver mutations may, in certain contexts, be

subject to negative selection against additional passenger mutations. Further mutational disruption in such genes could reduce cellular fitness, impair essential oncogenic functions, or impose an excessive deleterious load. This interpretation is consistent with our observation of a lower passenger mutation burden in driver-mutant cancer genes and suggests that the direction of selection acting on passenger mutations is contingent on the existing mutational landscape of each gene.

“””

5. Also please discuss the possibility that the catalog of “driver” mutations in PCAWG is incomplete. This may specifically be true for splice site mutations in PTEN and ARID genes, where the present study shows sometimes profound changes in the estimated effect of passenger mutations (which might, in fact, be missed drivers).

We acknowledge that the catalog of driver mutations from PCAWG might be incomplete. To address this concern, we repeated the core analyses using a more comprehensive and independent driver mutation set derived from cancerhotspots.org<sup>1</sup> (v2), which is based on recurrence patterns across approximately 25,000 tumor samples, as suggested by the reviewer. To further assess the robustness of our findings, this hotspot annotation was applied to an independent patient cohort, POG570<sup>2</sup> as well as the TCGA dataset. The hotspot mutations were used to define driver events and all remaining mutations classified as passengers. We reproduced the key findings from our original analysis, including the differential abundance of passenger mutations in the presence versus absence of driver events within cancer genes, as well as the association between passenger mutation burden and overall survival.

In both POG570 and TCGA cohorts, passenger mutation density was found to be elevated in samples harboring hotspot mutations within the same gene compared to those without hotspots (Mann-Whitney U test,  $p$ -value < 0.0001; **Figure R21-A** and **Figure R22-A**) and passenger mutations were more frequently observed than expected by chance (Chi-square test of homogeneity,  $p$ -value <  $10^{-6}$ ; **Figure R21-B** and **Figure R22-B**). When stratifying by cancer subtype in the TCGA cohort, the results remained consistent (Mann-Whitney U tests; **Figure R22-C**). We also find that patients with only passenger mutations in a cancer gene had significantly lower median survival times than patients without any mutations in the gene (Mann-Whitney U test,  $p$ -value < 0.05; **Figure R21-C**). Further analysis using the Kaplan-Meier estimator revealed that patients with passenger mutations in the absence of driver mutations also demonstrated significantly worse survival outcomes (log-rank test,  $p_1=0.022$ ; **Figure R21-D** and  $p_1<10^{-6}$ ; **Figure R22-D**), with the outcome worsening with increase in passenger burden in cancer genes (log-rank test,  $p_1<10^{-6}$ ,  $p_2<10^{-6}$ ; **Figure R22-E**).

These results demonstrate that our conclusions are robust to both the choice of driver mutation set and the patient cohort, and are not driven solely by potential limitations in the original PCAWG driver annotations. We have added these figures as **supplementary figures 2 and 3** and also added the following text in the manuscript:

**In Results Section:**

We also reproduced our key findings in the TCGA dataset and the POG570 cohort using an independent set of cancer hotspot mutations for driver annotation (**Supplementary Figure 2-3**), ensuring the robustness of our results.

...

To further assess robustness, we replicated the analysis in the TCGA dataset and the POG570 cohort using an independent cancer hotspot-based driver annotation. The results were highly concordant, consistently showing an enrichment in the mean adjusted passenger density in the absence of driver (or hotspot) mutations within a cancer gene (Mann-Whitney U test,  $p$ -value<0.0001; **Figure 2A**).

.....

## References:

1. Chang, M. T., Bhattarai, T. S., Schram, A. M., Bielski, C. M., Donoghue, M. T., Jonsson, P., ... & Taylor, B. S. (2018). Accelerating discovery of functional mutant alleles in cancer. *Cancer discovery*, 8(2), 174-183.
2. Pleasance, E., Titmuss, E., Williamson, L. et al. Pan-cancer analysis of advanced patient tumors reveals interactions between therapy and genomic landscapes. *Nat Cancer* 1, 452-468 (2020). <https://doi.org/10.1038/s43018-020-0050-6>

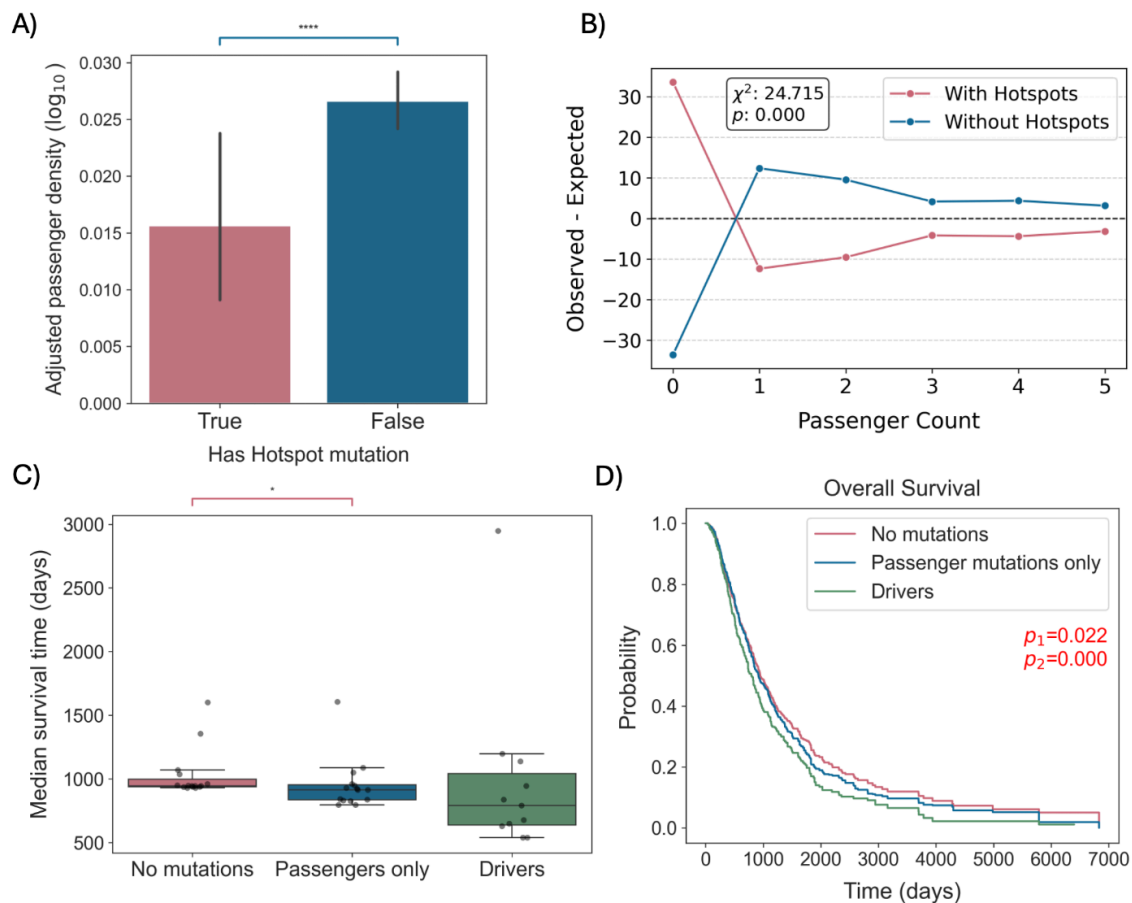


**Figure R21: Results of core analyses reproduced in the POG570 dataset using cancer hotspots data for driver annotation.** (A) Mean adjusted passenger mutation density in samples with and without hotspot mutation in a cancer genes. (B) Difference in observed and expected number of samples with varying passenger mutation counts in a cancer gene. (C) Distribution of median survival time (in days) of patients. Each data point corresponds to the median survival for a particular cancer gene. (D) Kaplan-Meier Survival plot for overall survival (OS) by mutational status of a cancer gene. Statistical significance in panels (A) and (B) is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

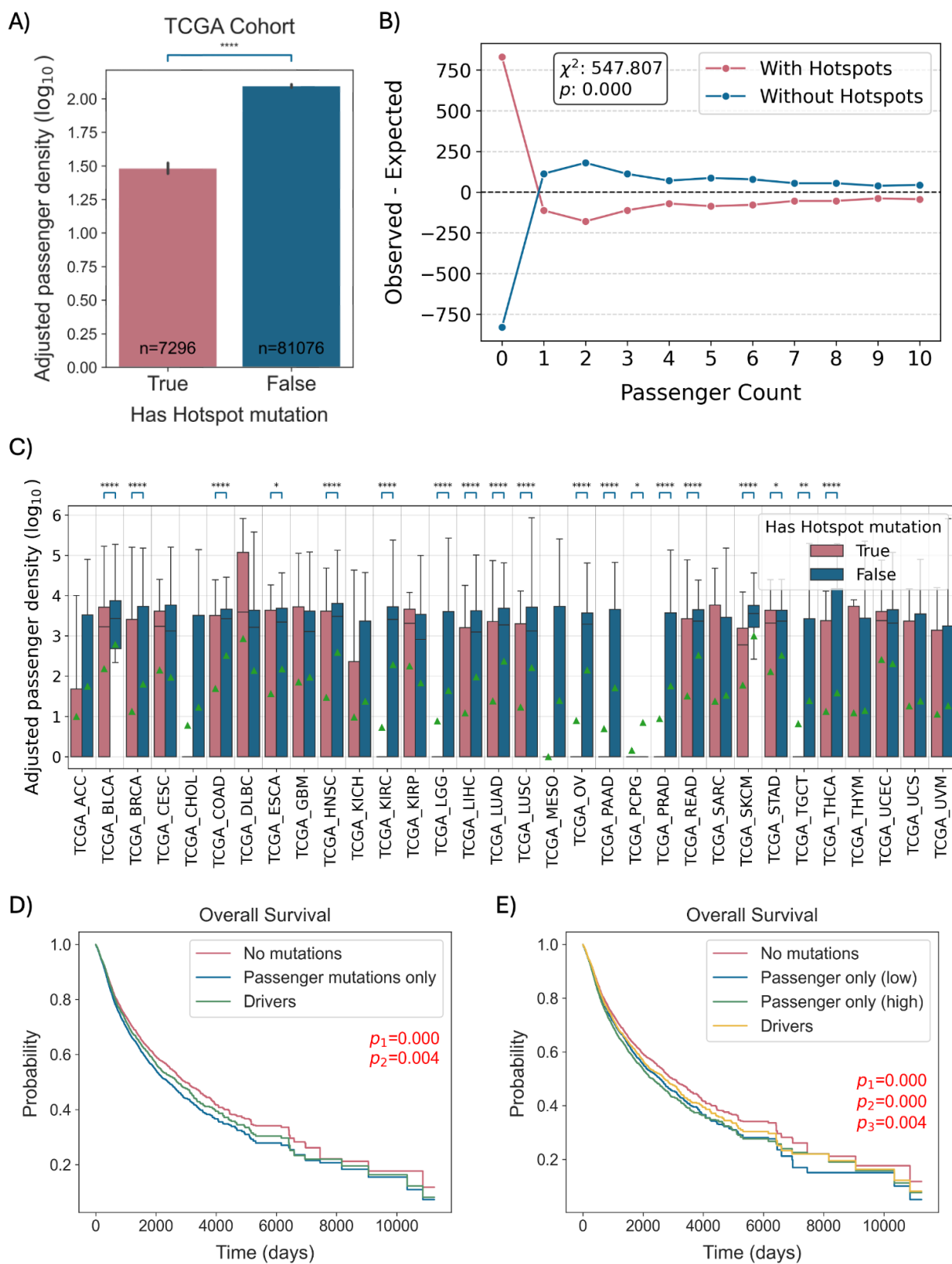


**Figure R22: Results of core analyses reproduced in the TCGA dataset using cancer hotspots data for driver annotation.** (A) Mean adjusted passenger mutation density in samples with and without hotspot mutation in a cancer genes. (B) Difference in observed and expected number of samples with varying passenger mutation counts in a cancer gene. (C) Distribution of differences in mean passenger density in samples without and with driver mutations for each cancer gene, stratified by cancer subtype cohort. (D) Kaplan-Meier Survival plot for overall survival (OS) by mutational status of a cancer gene. (E) Kaplan-Meier Survival plot for overall survival (OS) stratified by passenger mutation burden in cancer genes. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

We have included the following text in Discussion to acknowledge the limitations of our current definitions of drivers:

“”””

Throughout this work, we use the term passenger mutations to denote variants not annotated as drivers in current catalogs, while recognizing that this group includes mutations spanning a continuum of functional impact, from near-neutral to weak or context-dependent drivers whose effects can accumulate and become biologically consequential.

“”””

6. Is the transcription factor analysis mainly driven by oncogenes, tumor suppressor genes, or a few select cancer genes, which would be very interesting to see? Which of the 107 significantly disrupted TFs are expressed in the respective cancer types?

We thank the reviewer for this insightful question. To address it, we performed a stratified analysis of transcription factor (TF) binding disruption by gene type. We found that the observed differential TF binding disruption is primarily driven by mutations in upstream regulatory regions of oncogenes (**Figure R23-A**), whereas tumor suppressor genes showed a weaker effect (**Figure R23-B**).

We also extended this analysis to passenger mutations affecting splicing. Both oncogenes and tumor suppressor genes showed enrichment for splicing mutations in the absence of driver events. However, oncogenes more frequently harbored such mutations (**Figure R23-C**). These results suggest that mutations in oncogenes have a stronger influence on TF binding and splicing compared to tumor suppressor genes.

We have now included this result as **Supplementary Figure 5** and updated the Results section with the following text:

“”””

Stratification by gene class further reveals that this differential TF binding disruption is primarily driven by mutations in upstream regulatory regions of oncogenes (**Supplementary Figure 5A**), with a weaker effect observed for tumor suppressor genes (**Supplementary Figure 5B**), especially in oncogenes.

...

Stratification by gene class reveals that, although both oncogenes and tumor suppressor genes exhibit enrichment for splicing mutations in the absence of driver events, oncogenes more frequently harbor such mutations (**Supplementary Figure 5C**).

.....

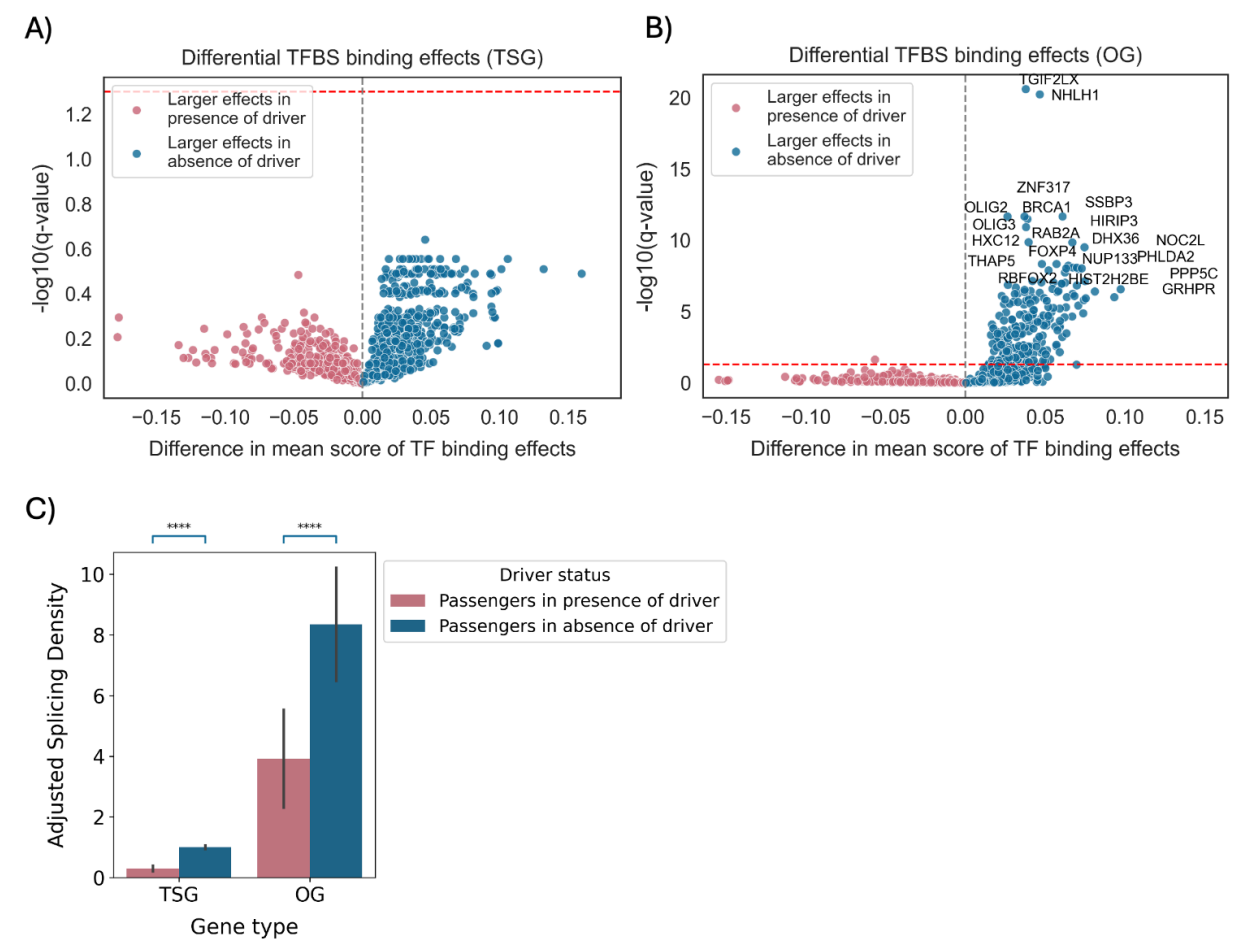


**Figure R23: Analysis of passenger mutations causing splicing changes or altering transcription factor binding affinity in cancer genes lacking driver mutations, stratified by tumor suppressor genes and oncogenes.** (A) Differential effects on transcription factor binding affinity caused by passengers in the presence and absence of driver mutations in Tumor Suppressor genes. (B) Differential effects on transcription factor binding affinity caused by passengers in the presence and absence of driver mutations in Oncogenes. (C) Density of passenger mutations associated with splicing changes in the presence and absence of drivers, stratified by tumor suppressor genes and oncogenes. In (A) and (B),  $p$ -values were adjusted for multiple testing using the Benjamini–Hochberg procedure, and the corresponding FDR-adjusted  $q$ -values are reported. Statistical significance in (C) is displayed as \* for  $p\text{-value}<0.05$ , \*\* for  $p\text{-value}<0.01$ , \*\*\* for  $p\text{-value}<0.001$ , and \*\*\*\* for  $p\text{-value}<0.0001$ .

Regarding TF expression in specific cancer types, our TF analysis was deliberately restricted to mutations in upstream regulatory regions. As a result, only a small number of samples harbor such mutations for any given gene-TF combination, limiting statistical power for cancer-specific analyses for all cancer-types. Here, we present a case study in Lymph-BNHL and

Skin-Melanoma, where sufficient number of upstream mutations allowed cancer-specific analysis.

We identified transcription factors that were differentially disrupted between groups of patients with and without driver mutations in cancer genes (adjusted  $q$ -value  $< 0.05$ ; mean difference in TF binding effects  $\geq 0.05$ ; **Figure R24-A,B**). Examination of their expression profiles revealed that most of these TFs are indeed expressed in the respective cancer samples (**Figure R24-C,D**), supporting the biological relevance of the observed disruptions. This suggests that passenger mutations in upstream regulatory regions can meaningfully impact TF activity in cancers.

We have added this result as **Supplementary Figure 6** and updated the Results section accordingly. The added text is as follows:

''''

To assess whether the differentially disrupted transcription factors (TFs) are expressed in patients, we performed a case study in Lymph-BNHL and Skin-Melanoma. We identified TFs that were significantly disrupted between patients with and without driver mutations in cancer genes (adjusted  $q$ -value  $< 0.05$ ; mean difference in TF binding effects  $\geq 0.05$ ; **Supplementary Figure 6A,B**). Analysis of their expression profiles showed that most of these TFs are actively expressed in the respective cancer samples (**Supplementary Figure 6C,D**), supporting the biological relevance of the observed disruptions. These findings indicate that passenger mutations in upstream regulatory regions can meaningfully affect TF activity in cancers.

''''

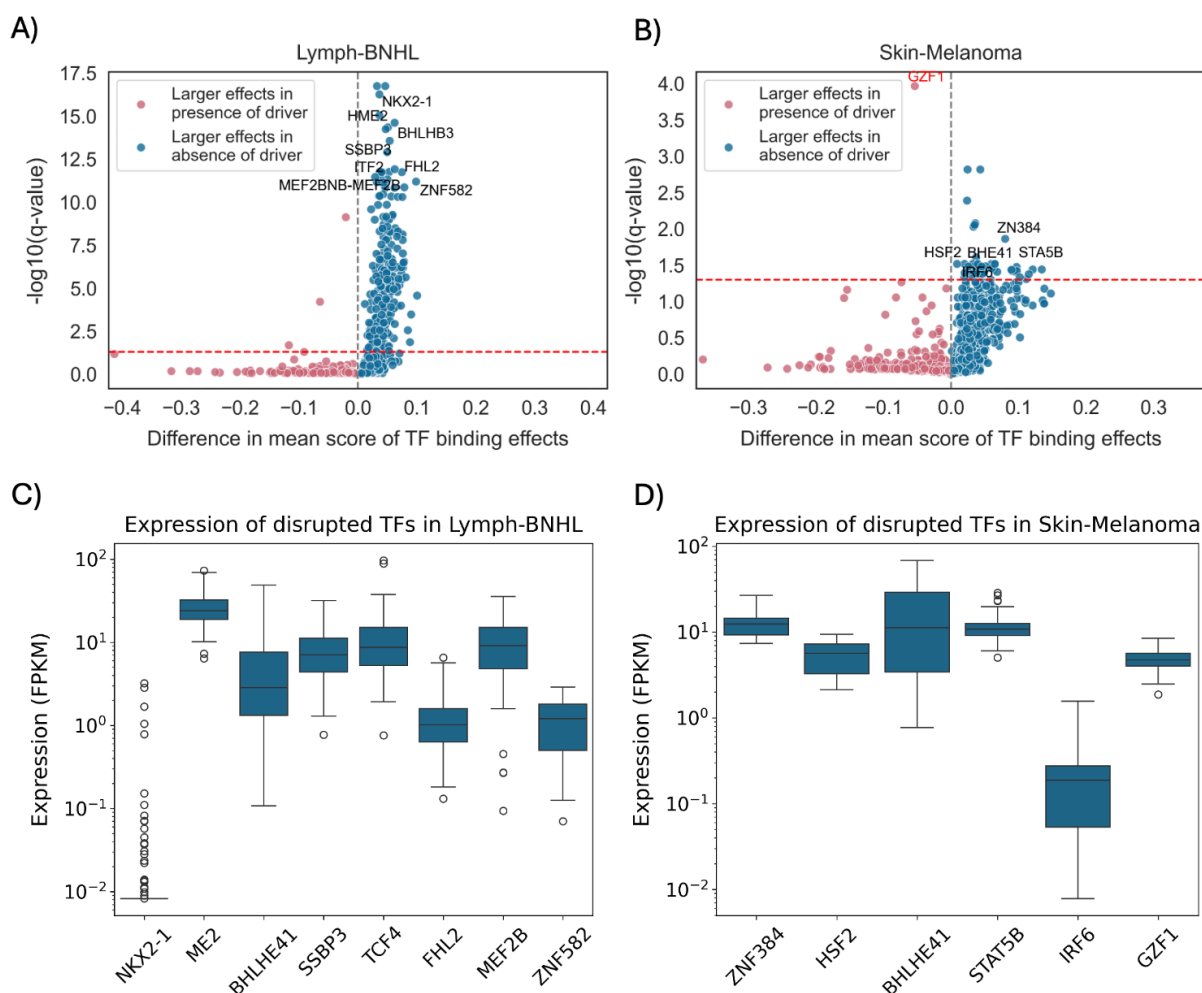


**Figure R24: Effects of passenger and driver mutations on splicing.** (A) Differential effects on transcription factor binding affinity caused by passengers in the presence and absence of driver mutations in Lymph-BNHL. (B) Differential effects on transcription factor binding affinity caused by passengers in the presence and absence of driver mutations in Skin-Melanoma. (C) Expression of differentially disrupted transcription factors in Lymph-BNHL. (D) Expression of differentially disrupted transcription factors in Skin-Melanoma. Only TFs with FDR adjusted  $q\text{-value} < 0.05$  and difference in mean score  $\geq 0.05$  between the two groups are shown here.

7. Cancer-Specific Survival Analyses: Survival outcomes vary strongly across cancer types. The current aggregated survival analysis potentially masks critical cancer-type-specific differences. Survival analyses stratified by individual cancer types would yield more precise, clinically relevant insights.

We thank the reviewer for this important comment. We agree that survival outcomes can vary substantially across cancer types and that cancer-specific analyses would, in principle, provide more clinically granular insights. We therefore evaluated survival analyses stratified by cancer-type. However, these analyses lacked sufficient statistical power due to limited sample

sizes within individual tumor types once patients were further stratified by driver status and passenger mutation burden. As a result, these analyses lacked sufficient power to yield robust or reproducible cancer-specific conclusions.

8. Related to survival analyses in Figure 5, do patients with only passenger mutations in *EGFR* have high-level amplifications that could contribute to the observed effect? Similarly, do the patients with *TERT* passenger mutations have CNVs, SVs, or activating promoter mutations?

We thank the reviewer for the comment. Copy number variants (CNVs), structural variants (SVs), activating promoter mutations, and other non-coding driver events were explicitly accounted for using the curated PCAWG driver mutation compendium. Events classified as drivers by PCAWG, including driver CNVs, were treated as driver mutations in our analysis. Accordingly, if passenger mutations occurred in samples harboring a driver CNV overlapping the gene under consideration, such cases were not classified as “passengers in the absence of drivers” and were excluded from the survival analysis.

Focusing specifically on *EGFR* and *TERT*, copy number alterations are present in samples across all three groups. To assess whether high-level amplifications could explain the observed survival differences in samples with passenger mutations only, we compared the CNA burden between samples with no mutations and those with passenger mutations but no driver events. A Mann-Whitney U test showed no statistically significant difference in CNA burden between these groups ( $p$ -value=0.14 for *EGFR*,  $p$ -value=0.42 for *TERT*), indicating that the observed survival effect is unlikely to be driven by CNVs alone.

9. Methodological Clarification and Comparison: Kumar et al. (2020) previously conducted a comprehensive analysis of passenger mutations using the same PCAWG datasets. Can the authors please discuss how their methodology and findings provide additional insights, particularly regarding the molecular characterization and modeling of passenger mutations?

We thank the reviewer for the comment. We agree with the importance of this previous work which we had already incorporated in our introduction. We have now expanded the Discussion to include a direct comparison with Kumar et al. (2020). Overall, our findings build upon this previous work, while additionally demonstrating that weak individual effects can accumulate, and thus should be looked at collectively rather than individually, and that these effects have direct functional consequences and implications for clinical biomarkers.

“””

Kumar et al. previously conducted a comprehensive analysis of passenger mutations using the PCAWG datasets and identified weak driver-like effects among subsets of passenger mutations<sup>9</sup>. Building on this work, here we demonstrate that weak effects are pervasive, are not isolated but can accumulate within cancer genes to produce substantial, functionally relevant consequences, thereby revealing a collective mode of positive selection that was previously unappreciated.

“””

10. Please provide additional information (such as the list of tumor-specific driver genes, Cox regression analysis etc) in supplementary tables. It looks like these may eventually be shared in the GitHub repository but are currently not available.

We thank the reviewer and agree. Additional information has now been included in the supplementary tables and deposited in the Zenodo at <https://doi.org/10.5281/zenodo.18226975>. All the code used for processing the data and generating the figures are available in our GitHub repository at [https://github.com/Georgakopoulos-Soares-lab/cancer\\_passengers](https://github.com/Georgakopoulos-Soares-lab/cancer_passengers).

11. Please ensure all files posted on GitHub comply with controlled and open access data rules.

We agree with the reviewer and we have been careful not to provide any sensitive patient information, such as individual mutations of patients. We have checked throughout our repository to comply.

#### Minor concerns

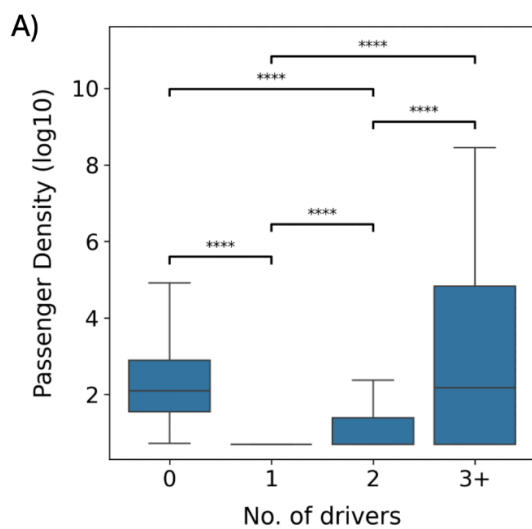
1. Multiple Driver Mutations Clarification: Could the presence of multiple driver mutations within a single patient influence the passenger mutation rate results? Clarifying this would improve the interpretation of mutation density analyses.

We have now examined how often multiple driver mutations in the same gene appear in cancer patients. In our cohort, among patients who harbor a driver mutation in a given cancer gene, the number of driver mutations per patient is generally low (mean = 1.25, median = 1). Consistent with this, most instances (patient-gene pairs) involve a single driver mutation, with only a small fraction of patient-gene pairs associated with multiple driver mutations (2 drivers: 2,557;  $\geq 3$  drivers:  $<1\%$ ). This indicates that when a driver mutation is observed in a given cancer gene, it typically corresponds to a single driver event rather than multiple drivers in the same gene.

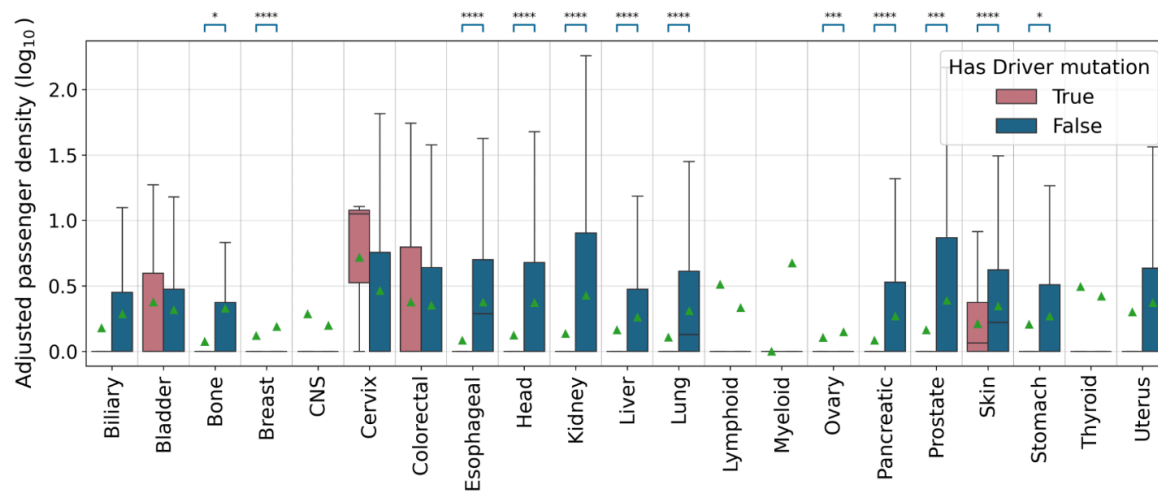
| No. of Drivers | Count |
|----------------|-------|
| 1              | 8,281 |
| 2              | 2,557 |
| 3+             | 94    |

We also compare the passenger mutation density in the case of 0, 1, 2 or 3+ driver mutation events in a gene. For genes with one or two driver events, we consistently observed higher passenger mutation density in the absence of drivers (Mann–Whitney U test, adjusted  $p$ -value  $< 0.0001$ , **Figure R25A**), but genes with three or more driver events did not reach statistical significance, likely due to limited sample size ( $<1\%$ ).

Accounting for the number of driver events strengthens the effect on passenger mutation density and increases statistical significance in several cancer types (**Figure R25B**). However, we did not apply this adjustment in the manuscript to avoid overstating the results.



B) Without Adjusting for Driver Count



After Adjusting for Driver Count

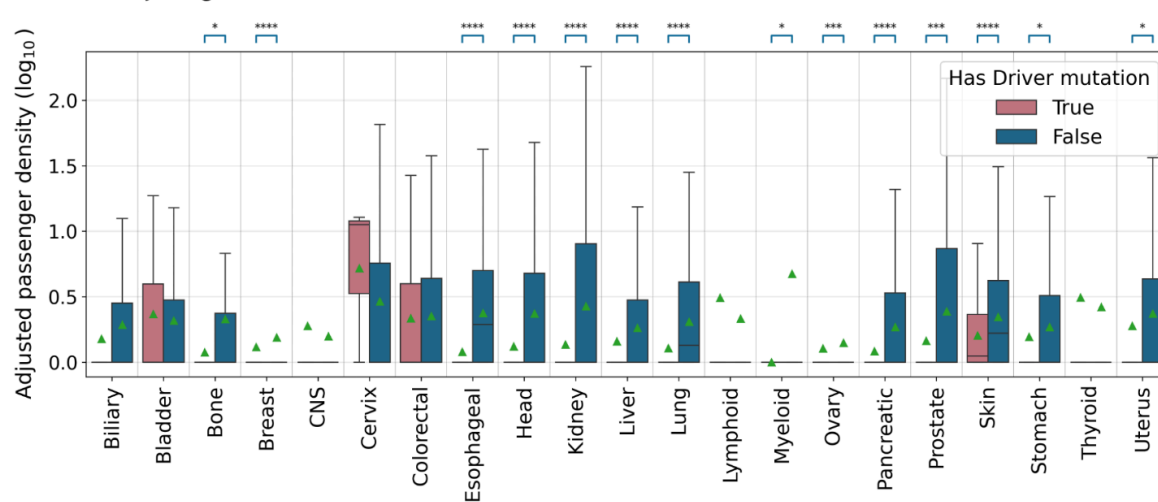


**Figure R25: Effect of number of driver mutations on the passenger density in the cancer gene.** (A) Association between passenger mutation density and the number of driver mutations per patient. (B) Passenger density results with and without adjustment for number of drivers in the gene. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001.

These results have been included as **Supplementary Figure 1B,C**, and the following text has been added to the Results section:

''''

We also examined whether multiple driver mutations within a single patient influence passenger mutation density. For genes with one or two driver events, passenger density was consistently higher in the absence of drivers (Mann–Whitney U test, adjusted  $p < 0.0001$ ; **Supplementary Figure 1B**), whereas genes with three or more drivers did not reach significance, likely due to limited sample size (<1%). Accounting for the number of driver events strengthens this effect and increases statistical significance in several cancer types (**Supplementary Figure 1C**); however, we did not apply this adjustment in the reported results to avoid overstating the effect.

''''

We also added the following text in the Methods section:

''''

To assess whether passenger mutation density was influenced by the number of driver events, we stratified patient-gene pairs by the number of driver mutations within the gene (0, 1, 2, or 3+). Passenger mutation densities were then compared between groups with and without driver mutations within each group using the Mann-Whitney U test, with  $p$ -values adjusted for multiple testing using Benjamini-Hochberg procedure. To assess the impact of adjusting for driver number on passenger density, density values in groups harboring driver mutations were normalized by the number of driver events per gene, while groups without driver mutations were left unadjusted. These adjusted analyses were performed for comparative purposes only and are presented in the **Supplementary Figure 1C**; they were not retained in the final Results to maintain a conservative interpretation.

''''

## 2. Figures:

a. Please clearly explain the meaning of symbols (e.g., green triangles in Figure 2a).

We apologize, and thank the reviewer for pointing this out. The green triangles in Figure 2A denote the mean values of the passenger mutation densities. We have now revised the figure caption to explicitly indicate this. The revised text is shown below:

"Figure 2: Passenger mutation density in samples with and without driver mutations. (A) Mean passenger mutation density in groups of samples with and without driver mutation in a cancer gene presented for 3 different cohorts. The number of data points (patient–gene pairs) in each group is indicated as n. (B) Distribution of differences in mean passenger density in samples without and with driver mutations for each cancer gene, stratified by cancer gene type. (C)

Distribution of passenger mutation density for each cancer type; green triangles indicate the mean values. (D) Difference in observed and expected number of samples with varying passenger mutation counts in a cancer gene. (E) Mean passenger mutation density in different genic regions. The number of data points (patient-gene pairs) with non-zero mutations in each group is shown above the bars. All mutation densities are reported as mutations per megabase (mutations/Mb). Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001. The Mann-Whitney U test was used to estimate the statistical difference between groups. In panels (C) and (E),  $p$ -values were adjusted for multiple testing using the Benjamini–Hochberg procedure, and the corresponding FDR-adjusted  $q$ -values are reported."

b. Address apparent discrepancies (such as absent boxes in breast cancer) despite statistical significance in Figure 2A.

We thank the reviewer for pointing this out. The apparent absence of boxes in certain categories, such as breast cancer in Figure 2A, is due to the majority of values in these groups being zero. As a result, the boxplot appears compressed, even though the underlying data are present and were used to compute statistical significance. We have now added the raw data used for generation of this plot on Zenodo at <https://doi.org/10.5281/zenodo.18226975>.

c. Figure 4A shows TF scores, but the legend describes a splicing analysis.

We thank the reviewer for catching this error. The figure captions for Figures 4A and 4C have been corrected, and the Figure 4 title has been updated to include both splicing and transcription factor binding effects. The updated figure caption is as follows:

''''

**Figure 4: Excess of mutations causing splicing changes or altering transcription factor binding affinity in cancer genes lacking driver mutations. (A)** Effect of passenger mutations in the presence or absence of driver mutations, and of driver mutations, on transcription factor binding affinity. **(B)** Differential effects on transcription factor binding affinity caused by passengers in the presence and absence of driver mutations. **(C)** Density of passenger and driver mutations associated with splicing changes, with passengers stratified by the presence or absence of drivers. **(D)** Effects of somatic mutations on splicing relative to the distance from the mutation, for individual cancer types. **(E)** Genes that contain passenger mutations causing large splicing changes (absolute Pangolin score  $\geq 0.8$ ) in different cancer types. **(F)** Distribution of large splicing changes (pangolin score  $\geq 0.8$ ) relative to the causative passenger and driver mutations for selected cancer types with high prevalence of splicing disruptive mutations. Statistical significance is displayed as \* for  $p$ -value<0.05, \*\* for  $p$ -value<0.01, \*\*\* for  $p$ -value<0.001, and \*\*\*\* for  $p$ -value<0.0001. In (A) and (B),  $p$ -values were adjusted for multiple testing using the Benjamini–Hochberg procedure, and the corresponding FDR-adjusted  $q$ -values are reported.

''''

d. Some panels appear to be swapped in Figure 3.

We apologize for this. We have now fixed the panels.

e. Please clarify the meaning of the “ncRNA” group for protein-coding cancer driver genes in Figure 3E.

We thank the reviewer for this comment. The “ncRNA” category in Figure 3E was defined using ANNOVAR gene-based annotations, which assign variants relative to overlapping transcripts. Variants labeled as “ncRNA” overlap transcripts lacking protein-coding annotation in the selected gene model. Because this label reflects annotation status rather than definitive coding potential, we have removed this category from the figure to avoid confusion and updated the manuscript accordingly.

3. Please explain why passenger mutations preferentially create new splice sites while driver mutations predominantly disrupt existing splice sites, potentially offering new mechanistic insight.

We thank the reviewer for this insightful question. Mutations that disrupt existing canonical splice sites often lead to severe loss-of-function effects, such as exon skipping, frameshifts, or premature truncation, and are therefore more likely to be strongly selected, recurrent, and classified as driver mutations in curated driver catalogs. In contrast, mutations that create alternative or cryptic splice sites typically result in partial or context-dependent splicing changes, rather than complete loss of gene function. These effects are often weaker and less recurrent across tumors, making them less likely to be classified as drivers and more likely to be annotated as passenger mutations. Nevertheless, when such mutations accumulate, their collective impact on transcription can become biologically meaningful. Thus, the observed preference of passenger mutations for creating new splice sites, versus drivers predominantly disrupting existing sites, reflects both differences in the magnitude of their functional effects and the criteria by which driver mutations are defined and selected.

We further illustrate the effects of passenger and driver mutations on splicing using the Borzoi<sup>1</sup> model. A passenger mutation in the *LRP1B* gene in esophageal cancer, occurring in the absence of a driver, generates a novel splice site (**Figure R26-A**), whereas a driver mutation in *FBXW7* in colorectal cancer disrupts an existing splice site (**Figure R26-B**).

These examples are included as **supplementary figure 7** and the Results section is updated to include the following text:

“”””

While passenger mutations were associated with both loss and gain in splicing (**Figure 4D-F**), a majority of the passenger mutations led to splicing gain, whereas drivers were found to predominantly disrupt existing splice sites (**Figure 4D**). Mutations that disrupt canonical splice sites often result in severe loss-of-function consequences, such as exon skipping, frameshifts,

or premature truncation, and are therefore more likely to be classified as driver mutations in curated catalogs. In contrast, passenger mutations more frequently generate alternative or cryptic splice sites, which tend to produce partial or context-dependent splicing alterations rather than complete loss of gene function. Although these effects are typically weaker and less recurrent, their accumulation can have biologically meaningful consequences. This distinction is illustrated using predictions from the Borzoi model, where a passenger mutation in *LRP1B* in esophageal cancer creates a novel splice site in the absence of a driver mutation (**Supplementary Figure 7A**), whereas a driver mutation in *FBXW7* in colorectal cancer disrupts an existing splice site (**Supplementary Figure 7B**).

''''

We also added the following text in the Methods section:

''''

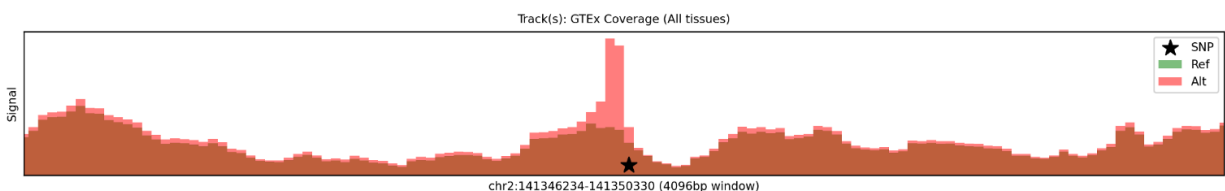
To illustrate the differential effects of passenger and driver mutations on splicing, we used Borzoi, a deep learning-based sequence model that predicts the cell-type-specific and tissue-specific RNA-seq coverage from DNA sequence. Two representative examples were selected: a passenger mutation in *LRP1B* from an esophageal cancer sample lacking driver mutations in the gene, and a driver mutation in *FBXW7* from a colorectal cancer sample. For each variant, reference and alternate genomic sequences centered on the mutation site were provided as input to Borzoi. Predicted GTEx RNA-seq coverage tracks for the wild-type and mutated sequences were visualized, and gradient-based contribution scores were compared between alleles to assess changes in splice site usage.

''''

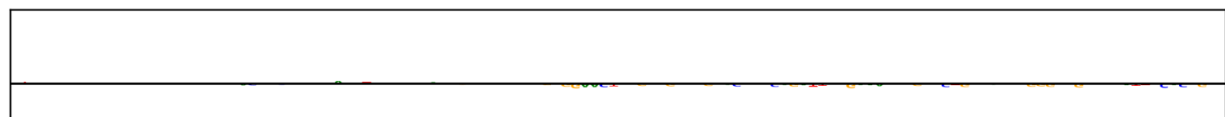
#### Reference:

1. Linder, J., Srivastava, D., Yuan, H. *et al.* Predicting RNA-seq coverage from DNA sequence as a unifying model of gene regulation. *Nat Genet* 57, 949–961 (2025). <https://doi.org/10.1038/s41588-024-02053-6>

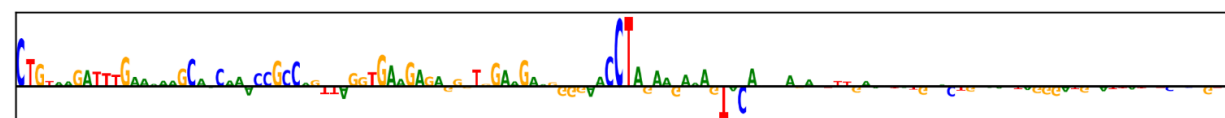
**A) Passenger mutation chr2:141348282 T>C (hg38) in LRP1B gene in Esophageal Cancer**



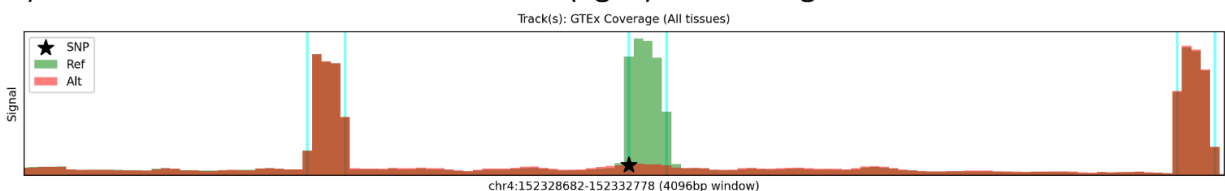
Wild Type



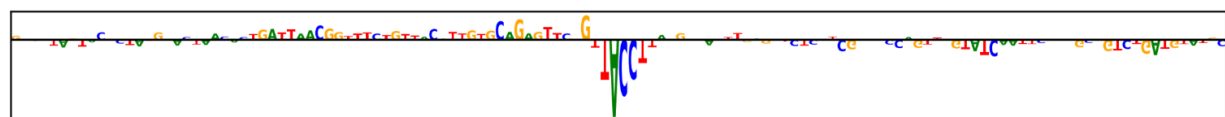
Mutant



**B) Driver mutation chr4:152330730 A>G (hg38) in FBXW7 gene in Colorectal Cancer**



Wild Type



Mutant



**Figure R26. Effects of passenger and driver mutations on splicing. (A)** A passenger mutation in the LRP1B gene in esophageal cancer, occurring without a driver mutation, generates a novel splice site. **(B)** A driver mutation in the FBXW7 gene in colorectal cancer disrupts an existing splice site. In both panels, GTEx RNA-seq coverage is shown along with contribution scores (gradient) from Borzoi, highlighting the impact of each mutation on splice site usage.

4. Please clarify whether the summarized CADD score is normalized for gene length.

We thank the reviewer for this comment. The CADD scores were normalized for each gene, but not specifically for gene length, as we do not expect gene length to affect the deleteriousness of the mutations. To account for differences in the range of CADD scores across genes, the scores

were rescaled for each gene using z-score normalization. We have now updated the manuscript to reflect this clarification:

“”

Because CADD is a per-variant deleteriousness score, we did not apply an explicit gene-length normalization to the summarized CADD burden.

“”

Reviewer #3 (Remarks on code availability):

Some directories mentioned in the GitHub instructions are not available (e.g. results). Due to time constraints, we did not attempt to run the code ourselves.

We thank the reviewer and apologize for the confusion. The results directory is generated when the code is executed. We have now updated the GitHub documentation to clarify this point. All the code used for processing the data and generating the figures are available in our GitHub repository at [https://github.com/Georgakopoulos-Soares-lab/cancer\\_passengers](https://github.com/Georgakopoulos-Soares-lab/cancer_passengers). Additional information has now been included in the supplementary tables and deposited in the Zenodo at <https://doi.org/10.5281/zenodo.18226975>.

Reviewer #4 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

**Reviewer #1 (Remarks to the Author)**

The authors have thoroughly addressed my remarks. The analysis has been improved by testing against a more comprehensive set of hotspots, and additionally by incorporating additional genome sequences/cohorts, thus results are more compelling with the extra data introduced in revision. The presentation has been improved by omitting some less informative parts, and literature context is now better covered. Overall this study is an important contribution towards understanding the prevalence of mini-drivers and mechanisms by which they act, the analyses are quite solid and presentation is good.

We thank the reviewer for their positive evaluation of our work and for acknowledging the improvements in analysis, presentation, and contextualization.

**Reviewer #2 (Remarks to the Author)**

My comments have been sufficiently addressed.

We thank the reviewer for confirming that their comments have been satisfactorily addressed.

**Reviewer #3 (Remarks to the Author)**

In the present manuscript, Nayak and colleagues have greatly improved their evidence for impact of passenger mutations in cancer, specifically through addition of the POG570 and TCGA datasets. This is now a well-thought out and presented story.

I just have a few minor comments:

We thank the reviewer for their thoughtful comments and for recognizing the improvements made through the inclusion of additional datasets and analyses. We address the remaining minor comments below:

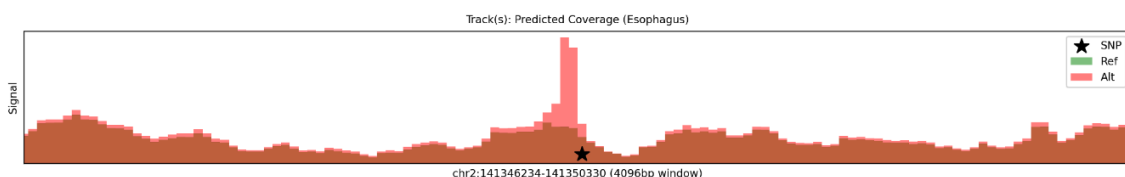
**1. In the Borzoi figure, the predicted track should not be labeled "GTEx" (as GTEx is observed data)**

We thank the reviewer for pointing this out. We have corrected the "GTEx" label to correctly indicate that these are predicted tracks.

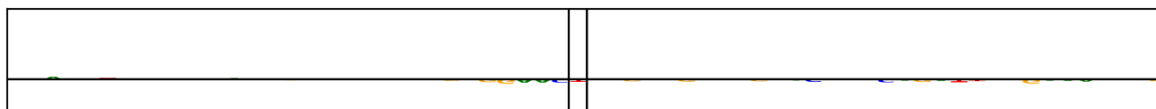
**2. In the same figure, it is hard to see whether the predicted sequences line up with the mutation position. Could you please highlight the bases in the sequence logo track that are mutated, maybe with a box that extends through all tracks?**

We thank the reviewer for this helpful comment. We have updated the figure to clearly highlight the mutated bases in the sequence logo tracks using a bounding box. In the coverage track, the SNP location is already indicated by a star marker. In addition, we have clarified in the figure caption that the coverage track spans a 4096 bp window centered on the mutation, whereas the sequence logo track is zoomed in to a 64 bp window centered on the mutation for visual clarity. The updated figure, incorporating these changes as well as those from Comment 1, is provided below for reference.

**A) Passenger mutation chr2:141348282 T>C (hg38) in LRP1B gene in Esophageal Cancer**



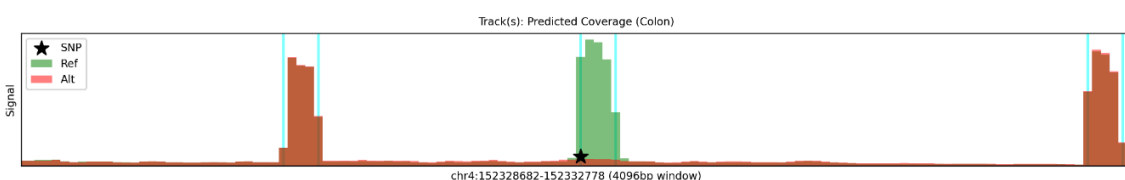
Wild Type



Mutant



**B) Driver mutation chr4:152330730 A>G (hg38) in FBXW7 gene in Colorectal Cancer**



Wild Type



Mutant

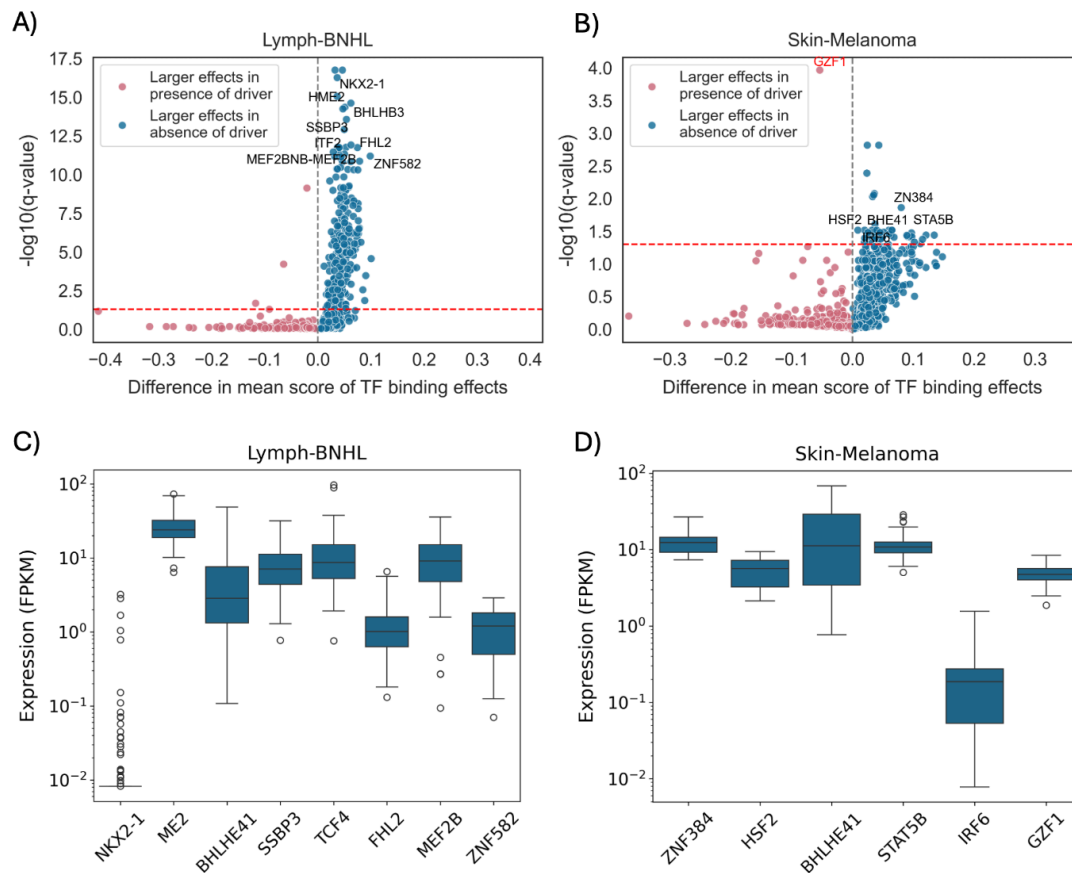


**Supplementary Figure 7: Effects of passenger and driver mutations on splicing. (A)** A passenger mutation in the LRP1B gene in esophageal cancer, occurring in the absence of a driver mutation, generates a novel splice site. **(B)** A driver mutation in the FBXW7 gene in colorectal cancer disrupts an existing splice site. In both panels, the top track shows the predicted RNA-seq coverage across a 4096 bp window centered on the SNP (indicated by a star), while the lower panels show Borzoi contribution scores (gradients) across a 64 bp window centered on the mutation (highlighted by a bounding box), for the wild-type and mutant sequences, respectively.

**3. In Supp figure 6C,D: "Expression of disrupted TFs":** The TFs that are not disrupted themselves, but rather their motifs upstream of cancer genes? (the title of the figure is confusing)

We thank the reviewer for this suggestion and agree that the original title could be confusing. We have revised the title to indicate only the cancer type due to space constraints and updated the figure caption to clarify that the transcription factors themselves are not

disrupted; rather, their upstream binding motifs in cancer genes are affected. The revised figure and caption are provided below.



**Supplementary Figure 6: Effects of passenger and driver mutations on splicing.** (A) Differential effects on transcription factor binding affinity caused by passengers in the presence and absence of driver mutations in Lymph-BNHL. (B) Differential effects on transcription factor binding affinity caused by passengers in the presence and absence of driver mutations in Skin-Melanoma. (C) Expression of transcription factors with differentially disrupted upstream binding motifs in Lymph-BNHL. (D) Expression of transcription factors with differentially disrupted upstream binding motifs in Skin-Melanoma. In A and B, each data point represents a transcription factor; p-values were computed using the two-sample t-test and adjusted for multiple testing using the Benjamini-Hochberg procedure; the corresponding FDR-adjusted q-values are reported. In C and D, only the transcription factors with FDR-adjusted q-value < 0.05 and a difference in mean score  $\geq 0.05$  between the two groups are shown here. Box plots show the median, interquartile range (IQR), and 1.5xIQR (whiskers).