Limited evidence of tumour mutational burden as a biomarker of response to immunotherapy

Carino Gurjao [1,2,3], Dina Tsukrov [3], Maxim Imakaev [3], Lovelace J. Luquette [4] and Leonid A. Mirny [3,2,1] *

1. Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA
2. Broad Institute of MIT and Harvard, Cambridge, MA, USA
3. Institute for Medical Engineering and Science, and Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA
4. Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA

Correspondence: Leonid Mirny leonid@mit.edu, 617-452-4862
Carino Gurjao cgurjao@mit.edu

Abstract

Cancer immunotherapy by immune checkpoint blockade (ICB) is effective for several cancer types [1], however, its clinical use is encumbered by a high variability in patient response. Several studies have suggested that Tumour Mutational Burden (TMB) correlates with patient response to ICB treatments [2–6], likely due to immunogenic neoantigens generated by novel mutations accumulated during cancer progression [7]. Association of TMB and response to checkpoint inhibitors has become widespread in the oncoimmunology field, within and across cancer types [7–11], and has led to the development of commercial TMB-based biomarker platforms. As a result, patient prioritization for ICB based on individual TMB level was recently approved by the FDA [12]. Here we revisit the association of mutational burden with response to checkpoint inhibitors by aggregating the largest pan-cancer dataset with more than 2500 ICB-treated patients with sequencing data and clinical annotation. Surprisingly, we find little evidence that TMB is predictive of patient response to immunotherapy. Our analysis suggests that previously reported associations arise from a combination of confounding disease subtypes and incorrect statistical testing. We show that using a TMB threshold for clinical decisions regarding immunotherapy could skew access to treatment for patients who may benefit from these therapies. Finally, we present a simple mathematical model that extends the neoantigen theory, is consistent with the lack of association between TMB and response to ICB and highlights the role of immunodominance. Our analysis calls for caution in the use of TMB as a biomarker and emphasizes the necessity of continuing the search for other genetic and non-genetic determinants of response to immunotherapy.

Immune checkpoint blockade (ICB) treatments such as anti-CTLA-4 and anti-PD1, which target regulatory pathways in T-lymphocytes to enhance anti-tumour immune responses, have already proven to elicit durable clinical responses for some patients [1,13–15]. However, the genetic determinants of response to immunotherapy have yet to be found. Several studies [2–6] suggested that Tumour Mutational Burden (TMB), computed as the total number of nonsynonymous somatic mutations, is correlated with response to immunotherapy in cancer. The underlying hypothesis posits that a fraction of nonsynonymous mutations become exposed as epitopes and constitute neoantigens, which can trigger an anticancer response by the immune system. The association between high mutational burden and response to immunotherapy, within and across cancer types [7–11], has been widely reported in the scientific literature and the media. As a result, TMB is currently discussed as the most clinically advanced biomarker of response to immune checkpoint blockade [16,17], and the FDA approved the use of TMB to identify patients most likely to derive clinical benefit [12]. These studies also triggered a search for inexpensive assays to evaluate TMB directly [18], as well as TMB-derived measures, such as neoantigens, neoepitopes, and mutation clonality [19], which are all currently under investigation to further stratify patients most likely to respond to immunotherapy. Our analysis focuses on TMB itself, as this is the most widely used and only FDA-approved measure.

TMB association with clinical benefit from immunotherapy

To evaluate the association of TMB with response to ICB, we aggregated and analyzed data for 882 immunotherapy patients with publicly available pre-treatment whole-exome sequencing data (referred below as CPI800+, **Table S1** and **Material and Methods**). We included patient-level data from an aggregate of early seminal studies [20] as well as recent clear cell renal cell cancer [21,22], non-small cell lung cancer [6], bladder cancer [23] and melanoma [2,24,25] ICB-treated cohorts. For every dataset examined, we retrieved TMB levels and survival data (Progression-Free Survival (PFS) or Overall Survival (OS)) for each patient. Cohorts provided response classification for most patients.

We also leveraged a very recent meta-analysis of 1283 patients (termed CPI1000+) who underwent immunotherapy [26] , have unified TMB (n=1083) and response definition, as well as survival measures for some patients (n=545 with OS data). Furthermore, we obtained gene panel data (MSK-IMPACT) for 1662 patients (**Table S1**) who underwent immunotherapy. To the best of our knowledge, together this dataset constitutes the largest pan-cancer aggregate of ICB-treated patients with sequencing and clinical data, which allow a robust unified statistical assessment of TMB as a predictor of ICB response

Consistent with published studies [6,20,21,24,27], we find (**Figure 1A** and **Figure S1**) that only the melanoma datasets (mel1 and mel2) and non-small cell lung cancer datasets (lung1 and lung2) yield a significant difference in TMB between responders and nonresponders ($p=8.3\times10^{-6}$ and $p=7.7\times10^{-3}$ for lung1 and lung2, $p=2.6\times10^{-2}$ and $p=4.1\times10^{-2}$ for mel1 and mel2, Mann-Whitney U test). Two of the three other cancer types analyzed – clear cell Renal Cell Carcinoma (ccRCC) and Head and Neck Squamous Cell Carcinoma (HNSCC) – showed no trend or an unexpected inverse one between TMB and response, although the association was non-significant (**Figure 1A**). The lack of associations between TMB and response to ICB in these two cancers was previously observed [22,28,29]. An identical analysis in CPI1000+ revealed similar results: there is no association between TMB and response in HNSCC and ccRCC, as well as breast cancer. Of note, CPI1000+ comprises colorectal cancer (CRC) specimens, and more bladder cancers (n=250) than our CPI800+ aggregate. Both showed an association between TMB and response ($p=0.044$ and $p=5.4\times10^{-7}$, respectively) (**Figure S1**).

All datasets showed a considerable overlap in TMB between responders and non-responders, as well as a large range of TMB values for the same cancer type. We thus tested (i) whether the association between TMB and response is confounded by cancer subtypes, i.e. due to the different response rates of cancer subtypes with

different TMB ranges, and (ii) whether TMB can be used as a biomarker to predict response to ICB.

We hypothesized that different cancer subtypes, with distinct TMB ranges and response rates, confound the observed increase in TMB in patients with clinical response to ICB. In particular, acral and mucosal melanomas are known to yield lower TMB and have a poorer prognosis [30]. Similarly, non-small cell lung tumours from smokers have higher TMB and published studies showed that ICB confers a survival advantage in smokers compared to never smokers [31]. Consistent with a previous study [24], we find that stratifying melanoma patients based on their disease subtype removes the association observed between TMB and clinical benefit in mel1 and mel2 (**Figure 1B**). However, stratifying non-small cell lung tumours based on the patient smoking status still showed a significantly higher TMB for responders versus non-responders, for smokers (current and former) ($p=1.3\times10^{-4}$ and $p=1.8\times10^{-2}$ for lung1 and lung2, Mann-Whitney U test)), but not among non-smokers. Other factors may contribute to a substantially better response of higher TMB in smoker patients. In particular, the presence of Chronic Obtrusive Pulmonary Disease (COPD) [32] could be a factor underlying the better response of high-TMB patients. While none of the ICB-treated cohort provided COPD status, several recent observations are consistent with the confounding role of COPD in response to high-TMB patients. First, COPD status is associated with increased survival after ICB [31,33]. Second, we find that TMB is significantly increased in COPD patients from TCGA (**Figure S2**). Third, the presence of EGFR mutation, that is infrequent in lung tumours of COPD patients [34,35], has been reported to correlate with poor response to immunotherapy [20]. Consistently, a large lung study that excluded patients with targetable EGFR mutation (KEYNOTE-189 [36], n=293) observed no association of high TMB with survival and clinical response to ICB. Together this suggests that COPD status can be a confounder that could explain higher TMB among individuals that respond to ICB; this hypothesis and the underlying biology can be tested by future studies.

We also revisit a meta-analysis that reported a positive correlation between response rates and TMB across different cancer types [10,11]. In that study, each cancer type is characterized by a median TMB and a median response rate; and splits melanoma and colorectal cancers – but not other cancers – into subtypes. We find that the correlation of the cancer-median TMB with the response rate reported in this study is driven solely by the TMB-response association of melanoma and colorectal cancers subtypes (**Figure S3**): when three points representing these subtypes were removed, the correlation becomes non-significant (p=0.10 for monotherapy, and p=0.21 for combination therapy). Thus, beyond subtypes with extreme differential response, no association between TMB and response rate across different cancer types is present in available data.

Overall, the evidence of an association between TMB and response to ICB relies largely on data for two cancer types: melanoma and non-small cell lung cancer. However, melanoma is confounded by subtypes and lung cancer requires more data and COPD stratification to validate the use of TMB. Crucially, an elevated TMB among responders does not imply the suitability of TMB for patient classification and treatment prioritization. We illustrate this particular point in the two following sections, by (i) investigating whether a significant cutoff can be established that would identify a high TMB group with significantly increased survival and (ii) assessing the overall accuracy of TMB as a classifier of response to ICB [10].

TMB association with survival post-immunotherapy

To evaluate the use of TMB for prioritizing patients, we tested whether it is possible to find a significant TMB cutoff that can separate patients into groups with significantly distinct survival rates. Strikingly, plots of survival versus TMB (**Figure 2A** and **Figure S4**) do not show a visible correlation or TMB cutoff that could differentiate longer and shorter surviving patients. Nevertheless, several studies established such TMB

thresholds [3,5] and reported a seemingly statistically significant difference in survival between patients below and above the threshold. One caveat of this approach is that it suffers from inherent multiple hypothesis testing made when the TMB thresholds have been selected among numerous possible values. This inherent multiple hypothesis testing would require further correction of the p-values; a step that is missing in all of the studies. However, standard approaches (e.g., Bonferroni correction, FDR correction) for multiple hypotheses testing would be too stringent because the hypotheses (i.e., the choices for TMB cutoff) are not independent.

Hence, we used a randomization analysis to address this limitation. This approach is similar to known multiple hypothesis testing methods [37,38] and earlier statistical studies that examined associations between dose and response in epidemiological studies [39]. We define the optimal TMB threshold as that which maximizes the difference in survival (i.e. minimizes the logrank p-value, a standard survival analysis test) between groups above and below the threshold. First, the optimal TMB threshold and its p-value (p_{real}) was found for the original data. Next, we randomly shuffled TMB among patients, while keeping survival and censored labels unchanged, and found the optimal TMB threshold and its p-value (p_{shuf}) for each randomized data. Finally, the p-value corrected for multiple hypotheses is derived by repeating the shuffling 1000 times and computing the fraction of shufflings where $p_{shuf} < p_{real}$ (**Figure 2B)**.

Applied to the melanoma and lung cancer datasets (**Figure 2C** and **Figure 2D**), we find that the majority (~60-70%) of randomly shuffled datasets produced p_{shuf} below the standard 0.05 threshold, emphasizing the need for multiple hypothesis correction. Overall, this correction for multiple hypothesis testing reveals non-significant p-value for all datasets. In particular for lung cancer, for which we previously observed a significant association between TMB and clinical benefit, we obtain a corrected p-value of 0.06 among smokers for lung1 and 0.23 for lung2. Of note, lung1 cohort that has p-value is close to significance contains 50% more EGFR patients than lung2, further suggesting

that the observed weak association might be due confounding effects to EGFR mutation status and/ or COPD status.

We also ran our analysis using OS (for datasets where both are available: mel1, mel2 and lung1) instead of PFS as an endpoint and showed comparable results, suggesting that survival definitions do not drive the results of our analysis (**Figure S5**). A similar analysis for individual cancer types (bladder, melanoma and non-small cell lung cancer; **Figure S6**) from CPI1000+ shows the lack of significant TMB threshold that can differentiate patients with significantly distinct survival rates.

We further obtained consistent results for 1662 patients of MSK-IMPACT cohort treated with ICB but genotyped with gene panels rather than whole-exome sequencing (**Figure S6**). Most of the 10 cancer types tested had a non-significant p-value including colorectal cancer (p=0.088) and melanoma (p=0.093) that have marginally significant p-values, except for non-small cell lung cancer (p=0.034). This study did not provide additional information such as tumour location for melanoma, Microsatellite Instability (MSI) status for colorectal cancer, or COPD for non-small cell lung tumours, which can confound the association of TMB with response [24,40].

Taken together, our analysis shows that no single TMB cutoff can significantly distinguish a high TMB group with presumed increased post-treatment survival. Even though survival is the intuitive way of assessing benefit to immunotherapy, it does not take into account other objective measures used for clinical categorization of response (e.g. changes in lesion size, assessment of pathological lymph nodes, etc). Thus, we next investigated the predictiveness of TMB in regards to standard clinical metrics of response.

TMB as a biomarker of response to immunotherapy

The key component for validating a biomarker is acceptable classification accuracy, i.e. the biomarker's capacity to correctly classify a patient's response [41]. ROC curves

(**Figure 3A**) provide a comprehensive view of specificity and sensitivity over all possible cutoffs and show the lack of a clear TMB cutoff that could be used in the clinic. The Area Under the Curve (AUC) is an aggregate measure of performance and is low in most datasets: mel1 and mel2 yield AUC of 0.62 and 0.59, and lung2 has an AUC of 0.68. Lung1, however, has the highest AUC of 0.85, which, as we show below, is still insufficient to select patients for ICB. ROC curve analysis on CPI1000+, with unified TMB, also shows a similarly poor AUC of 0.6 (**Figure S7**).

Next, we computed the proportion of misclassified patients based on the recent FDA approval of 10 mutations/Mb threshold to select patients for ICB (**Figure 3B** and **Figure 3C).** We find that, on average across the non-small cell lung cancer and melanoma datasets, 62% of responders were below the treatment prioritization threshold and 19% of non-responders were above. While these misclassification rates were vastly different between datasets, current efforts that focus on harmonizing TMB estimates across testing laboratories and pipelines are limited by the poor predictive power of TMB. Indeed, our ROC analysis shows that even the optimal cutoff (Youden index associated cutoff) for each dataset would result in an average 25% of responders below the treatment prioritization threshold and thus discouraged from receiving a potentially efficacious and life-extending treatment (**Figure 3C**). As such, the main challenge in using TMB in the clinic does not reside in harmonizing the values but in poor association between TMB and response to treatment.

TMB and cancer immunogenicity

Neoantigen theory is widely used to argue that cancers with high TMB are more likely to elicit an immune response after ICB. Although our results show the lack of such dependence, we demonstrate that the effect we observe can nevertheless be explained by a simple mathematical model of neoantigens and immunogenicity.

Our model (**Materials and Methods**) aims to explain (i) the lack of association between TMB and response; and (ii) the response by cancers with even very low TMB. In our model, each mutation has a probability $P_{immunogenic}$ to become immunogenic, i.e. to be expressed and presented as an epitope, to interact with the major histocompatibility complex, and to trigger an immune response (**Figure 4**). To include possible limited sensitivity of the immune system, we further require that at least k_{crit} such mutations are present to mount an immune response (for k_{crit}=1, a single mutation that becomes immunogenic triggers a full response). The components of our model are illustrated in **Figure 4A** and further explained in **Materials and Methods.**

Figure 4B shows the probability of eliciting a response ($P_{immune\ response}$) as a function of TMB for a range of $P_{immunogenic}$ and k_{crit} values. Our model has two regimes: If individual mutations are unlikely to be immunogenic ($P_{immunogenic}$<0.1, **Fig S8**), the response rate increases gradually with TMB, as widely expected but inconsistent with observed clinical data. On the contrary, if single mutations are likely to be immunogenic $P_{immunogenic}$>0.1, the probability of response saturates for TMB $\gtrsim$ 10, making most tumours equally likely to respond to ICB, as we observed above. For $P_{immunogenic}$ in the range estimated in silico [42] (0.22 for weak binders to T cells, and 0.64 for strong binders) and $k_{crit} \approx 1\text{-}2$, the probability of eliciting a response quickly approaches 1 for TMB$\gtrsim$10 and stays constant and independent of TMB. (**Materials and Methods).** The model further suggests that for the regime consistent with the data ($P_{immunogenic}$=0.2-0.6; $k_{crit} \approx 1\text{-}2$) (i) >90% of tumours with as little as 10 non-synonymous mutations are immunogenic; (ii) when 90% of tumours are immunogenic they have on average as few as 2 immunogenic mutation. These results are consistent with recently observed immunodominance hierarchies of the T cell responses [43]: low TMB tumours can mount responses as robust as high TMB tumours since only a small subset of neoantigens are targeted by T cells.

Taken together, our model and analysis of the available data indicate that cancer with even very few mutations can be immunogenic, suggesting that patients with low TMB might also mount robust immune responses, as has been recently shown for pediatric patients with acute lymphoblastic leukaemia [43].

Discussion

Tumour Mutational Burden, a measure of the total somatic nonsynonymous mutations in a tumour, recently became a popular biomarker of response to ICB, notably because of its relative simplicity to assess.

However, this paradigm is largely based on a series of early papers that examined response in melanoma and lung cancer that we show here to be potentially confounded by tumour subtype. Several recent studies have also reported poor association of TMB with response for specific cancer types, and highlighted TMB and its expression/presentation-based derivatives as problematic for clinical cohort classification [27]. In particular for melanoma, recent analyses [24] and our results indicate that the site location can explain the observed association between TMB and response to ICB. For lung cancer, our analysis points to the possibility that co-occurrence with COPD may explain the association between TMB and response to ICB among smokers. Overall we demonstrate that while most cohorts and cancer types show the lack of association of TMB and response or survival, the remaining statistical signal in some cohorts can arise due to confounders such as clinical subtypes. Future studies can examine the underlying biology -including TMB and neoantigens- explaining the better responses to ICB in certain clinical and cancer subtypes. .

Critically, even if responders show significantly elevated TMB, such associations do not imply the suitability of TMB as a biomarker of response. In particular, we show that no TMB cutoff can distinguish groups of patients with significantly different survival rates.

Besides, we show that TMB has poor accuracy as a classifier of response, even in the best-case scenario (Youden optimal cutpoint). This result challenges a recent FDA approval of TMB for prioritizing patients for ICB. If implemented, such TMB-based clinical decision making would deprive many patients who can benefit from ICB from receiving a life-extending treatment.

A recent ICB clinical trial that used FDA-approved TMB threshold (KEYNOTE-158) [44] has focused on rare cancers, excluding melanoma and lung cancer. While claiming a higher response rate among high-TMB patients, the trial observed little, if any, difference in overall survival of high-TMB and other patients, putting in question the clinical use of TMB-based prioritization.

We also put forward a simple model that reconciles our findings with the neoantigen theory. Our model shows that if each mutation has a high chance of triggering an immune response, then only a few new mutations make a cancer immunogenic, consistent with the observed immunodominance when the immune response is mounted against only a few of the neoantigens. This result is also consistent with the observed lack of association between antigen density and T-cell presence previously reported [45].
Moreover, our model suggests that most cancers are immunogenic, arguing that failures of ICB likely arise due to factors independent of cancer immunogenicity. Quantitative measurements [46] and modelling of neo-antigenic effects can deepen our understanding of cancer development and response to immunotherapy.

Although attractive and scalable, TMB does not consider the effect of specific mutations (missense, frameshift etc), their presentation and clonality [19], nor the state of the tumour, its microenvironment, and interactions with the immune system that can be integrated into potentially better predictors of response to ICB [47,48]. For the biology of oncoimmunity, our analysis suggests that, contrary to the neoantigen theory, cancer

immunogenicity does not increase with the growing load of neoantigens, and that clinical subtypes can underlie better response to ICB.

Altogether, our analysis indicates that low TMB should not be used to deprive otherwise eligible patients of immunotherapy treatment, and stimulates further research into other determinants of response to immunotherapy.

Material and Methods

Immunotherapy study population

CPI800+ was formed of eight independent WES cohorts (n=882, detailed in **Table S1**). The TMB and clinical annotations were not modified from the original studies. Post ICB sequenced samples were excluded from our analysis. In addition, gene panel datasets (n=1662, detailed in **Table S1**) were identified from cbioportal [49].

TCGA data

Lung cancer TCGA data were also retrieved from cbioportal [49], and additional clinical annotations were downloaded from The Cancer 3′ UTR Atlas [50]. COPD status was assessed based on the standard spirometric classification, i.e. post-bronchodilator ratio of forced expiratory volume in one second (FEV1) and forced vital capacity (FVC) below 70%.

Statistical analysis

We used R version 3.6.2 to perform statistical analyses. Two-group comparisons were evaluated by a two-sided Mann–Whitney U test unless otherwise indicated. $P < 0.05$ was considered statistically significant.

Code availability

The R code and data used to reproduce the analysis and figures from the paper are available on GitHub https://github.com/mirnylab/TMB_analysis

Model of cancer immunogenicity

Response to ICB treatment requires that the cancer is immunogenic and that immunotherapy can mount the immune response to this immunogenic cancer: $P_{response} = P_{immune\ response} * P_{therapy}$. The probability that immunotherapy works, given that the cancer is immunogenic, $P_{therapy}$, depends on the specifics of treatment and other physiological

variables, so we'll assume it to be constant. The probability of being immunogenic, $P_{immune\ response}$, however, depends on the ability of mutations to trigger the immune response. Assume that every nonsynonymous mutation has the probability $P_{immunogenic}$ (noted below as p) to be expressed and presented as an epitope, to interact with the major histocompatibility complex, and to trigger an immune response.

In the scenario of immunodominance, in which the immune response is mounted against only a few of the neopeptides, only $k \leq k_{crit}$ such mutations are sufficient to mount an immune response. Hence, the probability of being immunogenic is the probability of having at least k_{crit} presented and immunogenic mutations out of TMB:

$$P_{immune\ response} = \sum_{k=k_{crit}}^{TMB} p^k (1-p)^{TMB-k} C^k_{TMB} \approx 1 - \sum_{k=0}^{k_{crit}-1} \text{Poisson}(k, TMB*p),$$

where $p = P_{immunogenic}$.

In the case of $k_{crit}=1$, even a single mutation, if immunogenic, can trigger a response yielding $P_{immune\ response} = 1-(1-p)^{TMB} \approx 1-\exp(-TMB*p)$. It is easy to see that for this case $P_{immune\ response}$ saturates at p*TMB~1. Thus to achieve approximately constant $P_{immune\ response}$ for TMB>10-20, one needs p>0.1 for $k_{crit}=1$. Achieving a similar effect for $k_{crit}>1$ (i.e. $P_{immune\ response}$ that doesn't depend on TMB for TMB>10-20) requires even higher p>0.2. Moreover for $k_{crit}=1$, one can estimate the expected number of immunogenic mutations (p*TMB) present when 90% of cancers are immunogenic: $0.9 = P_{immune\ response} \approx 1-\exp(-TMB*p)$, gives p*TMB=2.3. I.e. irrespective of specific value of p, when 90% of cancers are immunogenic they carry only ~2 immunogenic mutations.

Furthermore, our model also explains a puzzling observation that immunoediting, i.e. negative selection against immunogenic mutations, is inefficient, allowing tumours to accumulate a high TMB [51]. Indeed, once a cancer accumulates mutations making it immunogenic, additional mutations incur no additional selective disadvantage i.e. show "the epistasis of diminishing return", and hence accumulate as neutral or weakly damaging passenger mutations [52–54] Moreover, according to this argument, cancer would have to develop means to suppress the immune response early in its development, a prediction that can be tested in future studies of cancer clonal evolution

Acknowledgement

We are grateful for Krupa Thakkar, Kevin Litchfield and Charles Swanton for running our analysis on their recent “CPI1000+” meta-aggregate of immunotherapy patients [26].

We are thankful for many productive and exciting discussions of this work to Christopher McFarland, Johannes Berg, Donate Weghorn, Eli van Allen, Kenneth Kehl, Shamil Sunyaev, Martha Luksza, Michael, Lassig, Boris Reizis, Virginia Savova, Gregory Kryukov, Sebastian Amigorena, Sean McGrath, Baptiste Boisson, Toni Choueiri, Paul B Robbins and all members of the Mirny lab. We are grateful to the organizers and participants of “Physicists working on Cancer” workshop at the Weizmann Institute of Science, Schwartz/Reisman Institute for Theoretical Physics, particularly, Eytan Domany, Herbert Levine, Caterina La Porta and Stefano Zapperi. We are also grateful to the organizers and participants of the workshop “Tumors and Immune Systems: From Theory to Therapy” seminar at Institut d'Etudes Scientifique de Cargèse, particularly, Alexandra Walczak, Thierry Mora, Vassili Soumelis, Paul Thomas and Jason George.

LJL was supported by the Training Program in Bioinformatics and Integrative Genomics (NIH T32HG002295, PI: P.Park). This project grew from the qualifying examination problem in Bioinformatics and Integrative Genomics given by LAM to LJL. We acknowledge support of the MIT-France Seed Fund, and The Chicago Region Physical Science Oncology Center (PS-OC, National Cancer Institute U54CA193419).

References

1. Ribas, A. & Wolchok, J. D. Cancer immunotherapy using checkpoint blockade. *Science* **359**, 1350–1355 (2018).
2. Hugo, W. *et al.* Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **168**, 542 (2017).
3. Snyder, A. *et al.* Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014).
4. Van Allen, E. M. *et al.* Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
5. Rizvi, N. A. *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015).
6. Hellmann, M. D. *et al.* Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. *Cancer Cell* **33**, 843–852.e4 (2018).
7. Chan, T. A. *et al.* Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann. Oncol.* **30**, 44–56 (2019).
8. Cristescu, R. *et al.* Pan-tumor genomic biomarkers for PD-1 checkpoint blockade–based immunotherapy. *Science* vol. 362 eaar3593 (2018).
9. Samstein, R. M. *et al.* Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* **51**, 202–206 (2019).
10. Yarchoan, M., Hopkins, A. & Jaffee, E. M. Tumor Mutational Burden and Response

Rate to PD-1 Inhibition. *N. Engl. J. Med.* **377**, 2500–2501 (2017).

11. Osipov, A. *et al.* Tumor Mutational Burden, Toxicity, and Response of Immune Checkpoint Inhibitors Targeting PD(L)1, CTLA-4, and Combination: A Meta-regression Analysis. *Clin. Cancer Res.* (2020) doi:10.1158/1078-0432.CCR-20-0458.
12. Food and Drug Administration. Highlights of prescribing information: KEYTRUDA. *https://www.accessdata.fda.gov/drugsatfda_docs/label/2020/125514s068lbl.pdf*.
13. Postow, M. A., Callahan, M. K. & Wolchok, J. D. Immune Checkpoint Blockade in Cancer Therapy. *J. Clin. Oncol.* **33**, 1974–1982 (2015).
14. Hodi, F. S. *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. *N. Engl. J. Med.* **363**, 711–723 (2010).
15. Dang, T. O., Ogunniyi, A., Barbee, M. S. & Drilon, A. Pembrolizumab for the treatment of PD-L1 positive advanced or metastatic non-small cell lung cancer. *Expert Rev. Anticancer Ther.* **16**, 13–20 (2016).
16. Sharma, P., Hu-Lieskovan, S., Wargo, J. A. & Ribas, A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell* **168**, 707–723 (2017).
17. Heeke, S. & Hofman, P. Tumor mutational burden assessment as a predictive biomarker for immunotherapy in lung cancer patients: getting ready for prime-time or not? *Translational Lung Cancer Research* **7**, 631 (2018).
18. Sicklick, J. K. *et al.* Molecular profiling of cancer patients enables personalized combination therapy: the I-PREDICT study. *Nat. Med.* **25**, 744–750 (2019).
19. McGranahan, N. & Swanton, C. Neoantigen quality, not quantity. *Sci. Transl. Med.*

11, (2019).

20. Miao, D. *et al.* Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat. Genet.* **50**, 1271–1281 (2018).
21. Miao, D. *et al.* Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* **359**, 801–806 (2018).
22. Braun, D. A. *et al.* Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat. Med.* **26**, 909–918 (2020).
23. Snyder, A. *et al.* Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: An exploratory multi-omic analysis. *PLoS Med.* **14**, e1002309 (2017).
24. Liu, D. *et al.* Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat. Med.* **25**, 1916–1927 (2019).
25. Riaz, N. *et al.* Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **171**, 934–949.e16 (2017).
26. Litchfield, K. *et al.* Meta-analysis of tumor and T cell intrinsic mechanisms of sensitization to checkpoint inhibition. (2020).
27. Wood, M. A., Weeder, B. R., David, J. K., Nellore, A. & Thompson, R. F. Burden of tumor mutations, neoepitopes, and other variants are weak predictors of cancer immunotherapy response and overall survival. *Genome Med.* **12**, 33 (2020).
28. Patel, N. A., Vokes, N. I., Elmarakeby, H., Hanna, G. J. & Van Allen, E. M. Abstract 5859: Genomic correlates of response to immune checkpoint inhibitors in advanced

head and neck squamous cell carcinoma. *Molecular and Cellular Biology / Genetics* (2020) doi:10.1158/1538-7445.am2020-5859.

29. Motzer, R. J. *et al.* Avelumab plus axitinib versus sunitinib in advanced renal cell carcinoma: biomarker analysis of the phase 3 JAVELIN Renal 101 trial. *Nat. Med.* (2020) doi:10.1038/s41591-020-1044-8.
30. Hayward, N. K. *et al.* Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175–180 (2017).
31. Mark, N. M. *et al.* Chronic Obstructive Pulmonary Disease Alters Immune Cell Composition and Immune Checkpoint Inhibitor Efficacy in Non-Small Cell Lung Cancer. *Am. J. Respir. Crit. Care Med.* **197**, 325–336 (2018).
32. Wang, W. *et al.* Impact of COPD on prognosis of lung cancer: from a perspective on disease heterogeneity. *Int. J. Chron. Obstruct. Pulmon. Dis.* **13**, 3767–3776 (2018).
33. Salehi-Rad, R. & Dubinett, S. M. Understanding the Hurdles in Lung Cancer Immunotherapy in the Context of Chronic Obstructive Pulmonary Disease. *American journal of respiratory and critical care medicine* vol. 198 835–837 (2018).
34. Lim, J. U. *et al.* Chronic Obstructive Pulmonary Disease-Related Non-Small-Cell Lung Cancer Exhibits a Low Prevalence of EGFR and ALK Driver Mutations. *PLoS One* **10**, e0142306 (2015).
35. Suzuki, M. *et al.* Molecular Characterization of Chronic Obstructive Pulmonary Disease-Related Non-Small Cell Lung Cancer Through Aberrant Methylation and Alterations of EGFR Signaling. *Annals of Surgical Oncology* vol. 17 878–888

(2010).

36. Garassino, M. C. *et al.* Evaluation of blood TMB (bTMB) in KEYNOTE-189: Pembrolizumab (pembro) plus chemotherapy (chemo) with pemetrexed and platinum versus placebo plus chemo as first-line therapy for metastatic nonsquamous NSCLC. *J. Clin. Orthod.* **38**, 9521–9521 (2020).
37. Westfall, P. H. & Stanley Young, S. *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. (John Wiley & Sons, 1993).
38. Dudoit, S., Shaffer, J. P. & Boldrick, J. C. Multiple Hypothesis Testing in Microarray Experiments. *Stat. Sci.* **18**, 71–103 (2003).
39. Shlyakhter, A., Mirny, L., Vlasov, A. & Wilson, R. Monte Carlo modeling of epidemiological studies. *Human and Ecological Risk Assessment: An International Journal* **2**, 920–938 (1996).
40. Xiao, Y. & Freeman, G. J. The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. *Cancer discovery* vol. 5 16–18 (2015).
41. Pepe, M. S., Feng, Z., Janes, H., Bossuyt, P. M. & Potter, J. D. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J. Natl. Cancer Inst.* **100**, 1432–1438 (2008).
42. Turajlic, S. *et al.* Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *Lancet Oncol.* **18**, 1009–1021 (2017).
43. Zamora, A. E. *et al.* Pediatric patients with acute lymphoblastic leukemia generate

abundant and functional neoantigen-specific CD8+ T cell responses. *Sci. Transl. Med.* **11**, (2019).

44. Marabelle, A. *et al.* Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol.* **21**, 1353–1365 (2020).
45. Spranger, S. *et al.* Density of immunogenic antigens does not explain the presence or absence of the T-cell–inflamed tumor microenvironment in melanoma. *Proceedings of the National Academy of Sciences* vol. 113 E7759–E7768 (2016).
46. Sarkizova, S. & Hacohen, N. How T cells spot tumour cells. *Nature* vol. 551 444–446 (2017).
47. Łuksza, M. *et al.* A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* **551**, 517–520 (2017).
48. Gurjao, C. *et al.* Intrinsic Resistance to Immune Checkpoint Blockade in a Mismatch Repair–Deficient Colorectal Cancer. *Cancer Immunol Res* **7**, 1230–1236 (2019).
49. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, l1 (2013).
50. Feng, X., Li, L., Wagner, E. J. & Li, W. TC3A: The Cancer 3' UTR Atlas. *Nucleic Acids Res.* **46**, D1027–D1030 (2018).
51. Efremova, M. *et al.* Targeting immune checkpoints potentiates immunoediting and changes the dynamics of tumor evolution. *Nat. Commun.* **9**, 32 (2018).

52. McFarland, C. D., Mirny, L. A. & Korolev, K. S. Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 15138–15143 (2014).

53. Persi, E., Wolf, Y. I., Leiserson, M. D. M., Koonin, E. V. & Ruppin, E. Criticality in tumor evolution and clinical outcome. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E11101–E11110 (2018).

54. Weghorn, D. & Sunyaev, S. Bayesian inference of negative and positive selection in human cancers. *Nat. Genet.* **49**, 1785–1788 (2017).

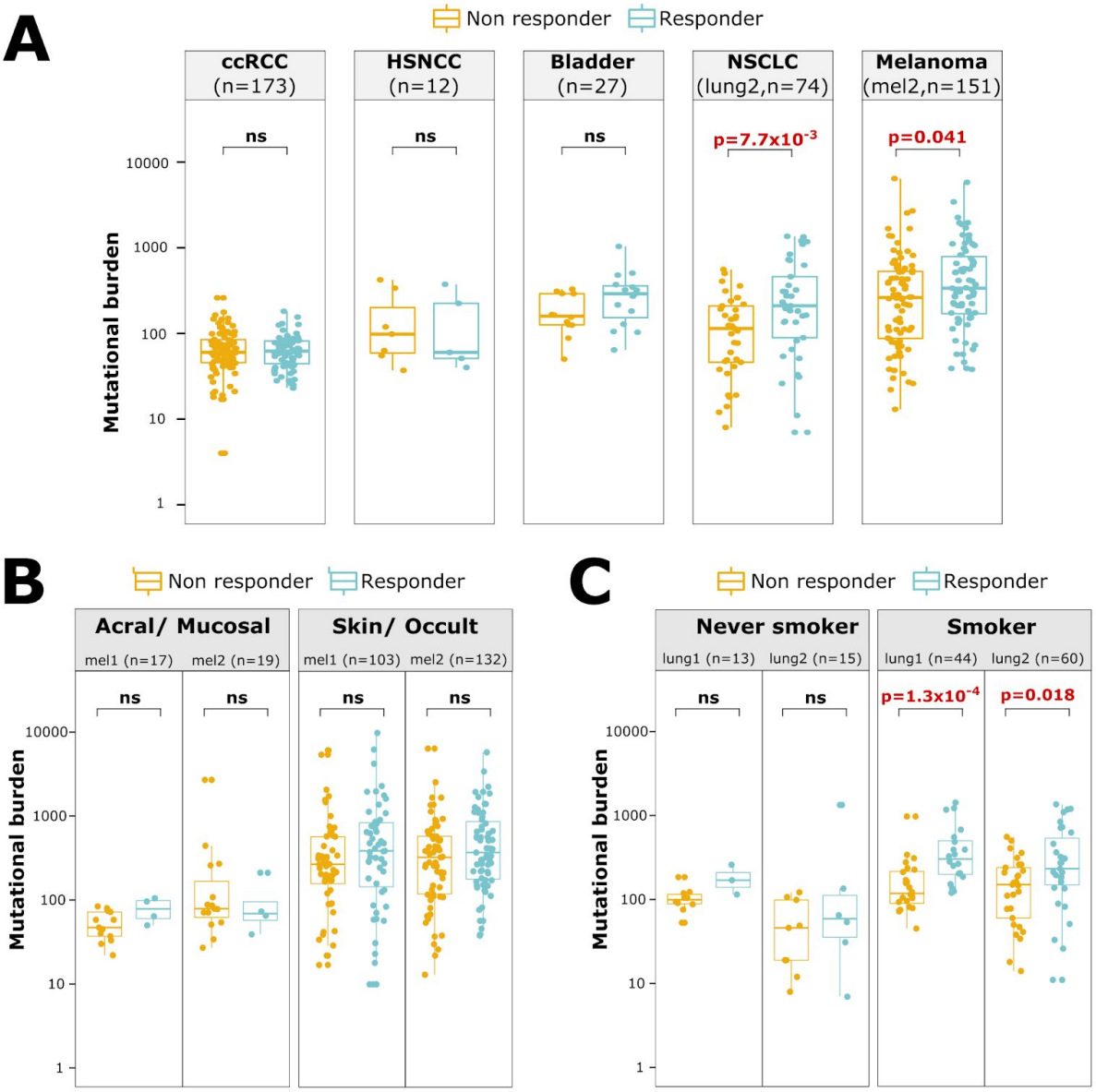

Figure 1: TMB association with clinical benefit from ICB across cancers
(A) Association of TMB with response to ICB across five cancer types from CPI800+ (the largest cohorts of each cancer type are plotted here, the others are shown in Figure S1A). Only melanoma and non-small cell lung cancer have a significantly different TMB between responders and non-responders. ccRCC: clear cell Renal Cell Carcinoma, HNSCC: Head and Neck Squamous Cell Carcinoma; NSCLC: Non Small Cell Lung Carcinoma **(B)** Association of TMB with response to ICB for specific melanoma subtypes. When split into subtypes, TMB does not associate with response to ICB. **(C)** Association of TMB with response to ICB for subtypes of non-small cell lung cancer patients. When split into subtypes, TMB associates with response to ICB only among smokers.

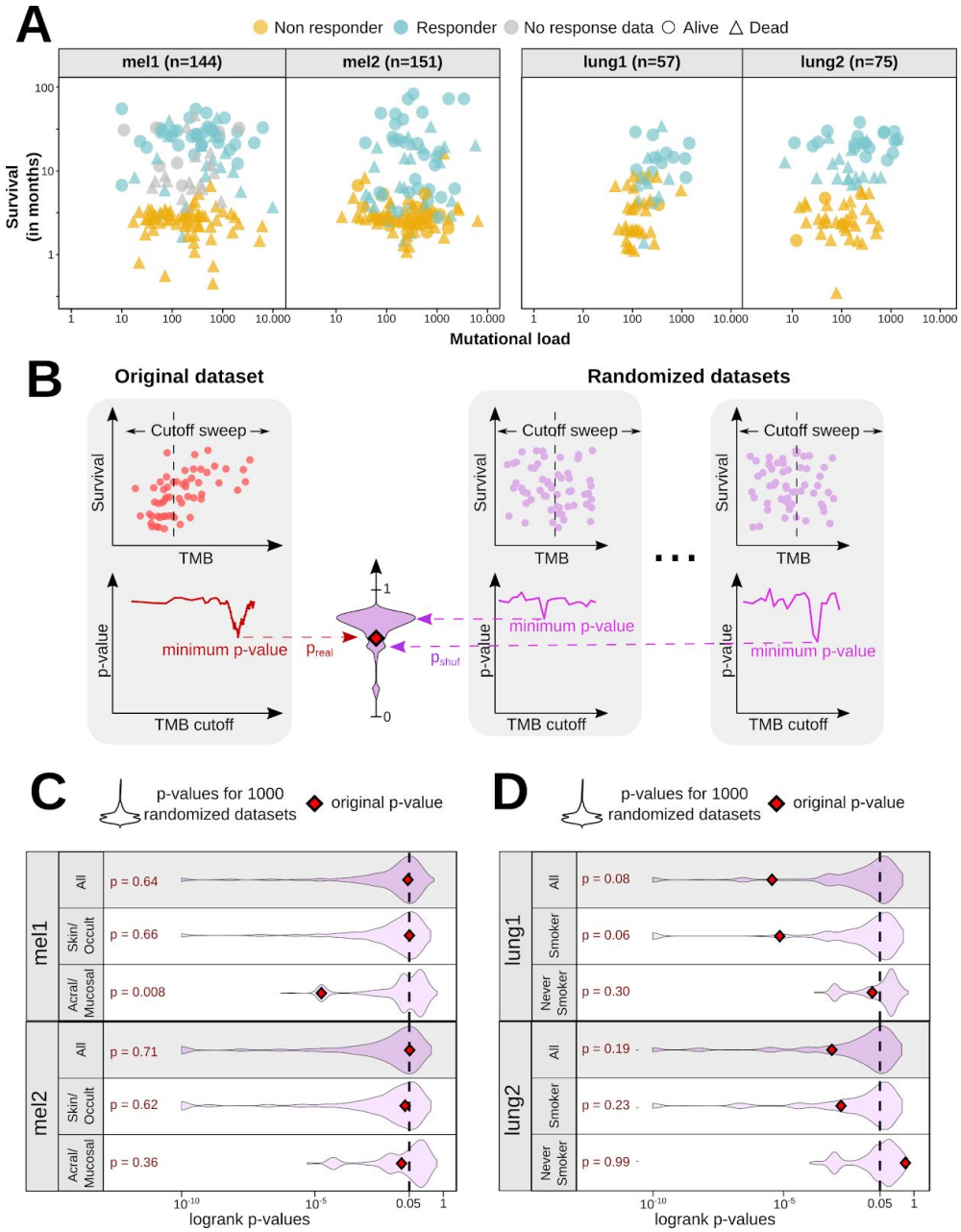

Figure 2: TMB association with progression-free survival post-immunotherapy **(A)** Plots of progression-free survival and TMB for melanoma and lung cancer ICB cohorts show the lack of correlation or of an obvious TMB cutoff. (Similar plots by cancer subtypes are shown in Figure S3) **(B)** Overview of the randomization analysis. Left: the optimal cutoff is found to maximize the difference between survival between groups above and below the cutoff (i.e to minimize the logrank p-value, yielding p_{real}). Right: the same procedure for shuffled data yields p_{shuf}. The fraction of $p_{shuf}<p_{real}$ produces a p-value corrected for multiple hypothesis testing for non-independent tests. **(C)** Results of the randomization analysis in the melanoma cohorts and stratification by subtypes (p-values < 10^{-10} not shown) **(D)** Randomization analysis results in the lung cancer cohorts and stratification by subtypes (p-values < 10^{-10} not shown).

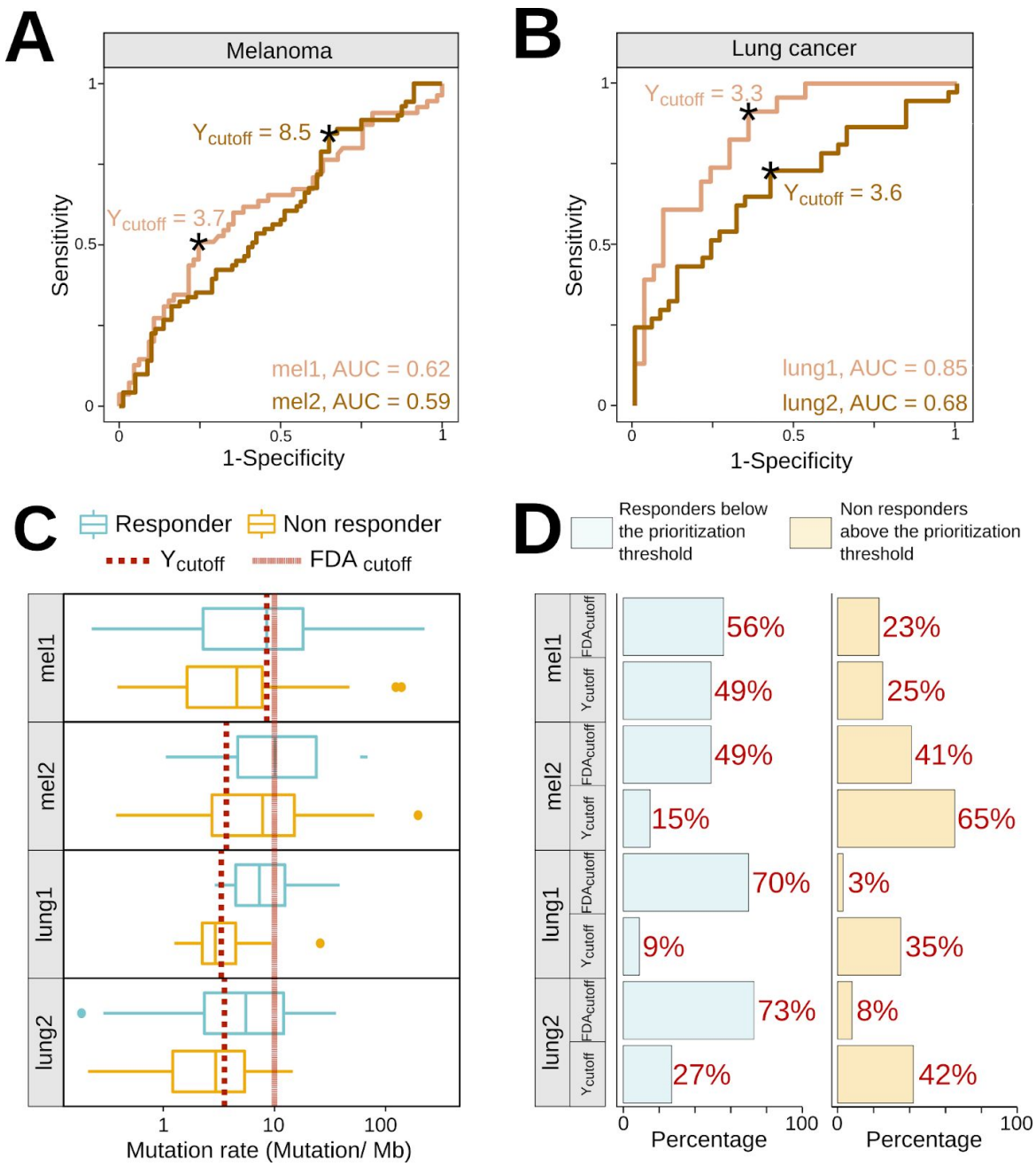

Figure 3: TMB as a biomarker of response to immunotherapy

(A)(B) ROC curves for the melanoma and lung cancer cohorts. Youden index associated cutoffs are also plotted. **(C)** Boxplots of nonsynonymous mutation rates across responders and non responders in the melanoma and lung cancer cohorts. The FDA-approved cutoff (10 mutations/Mb) and the best cutoff (Youden index associated cutoff) are shown by vertical lines. **(D)** Proportion of misclassified patients based on the FDA-approved cutoff, as well as the Youden index cutoff for each dataset. The use of either cutoff leads to substantial fraction of misclassified patients (potential responders below the treatment cutoff, or non-responders above the cutoff).

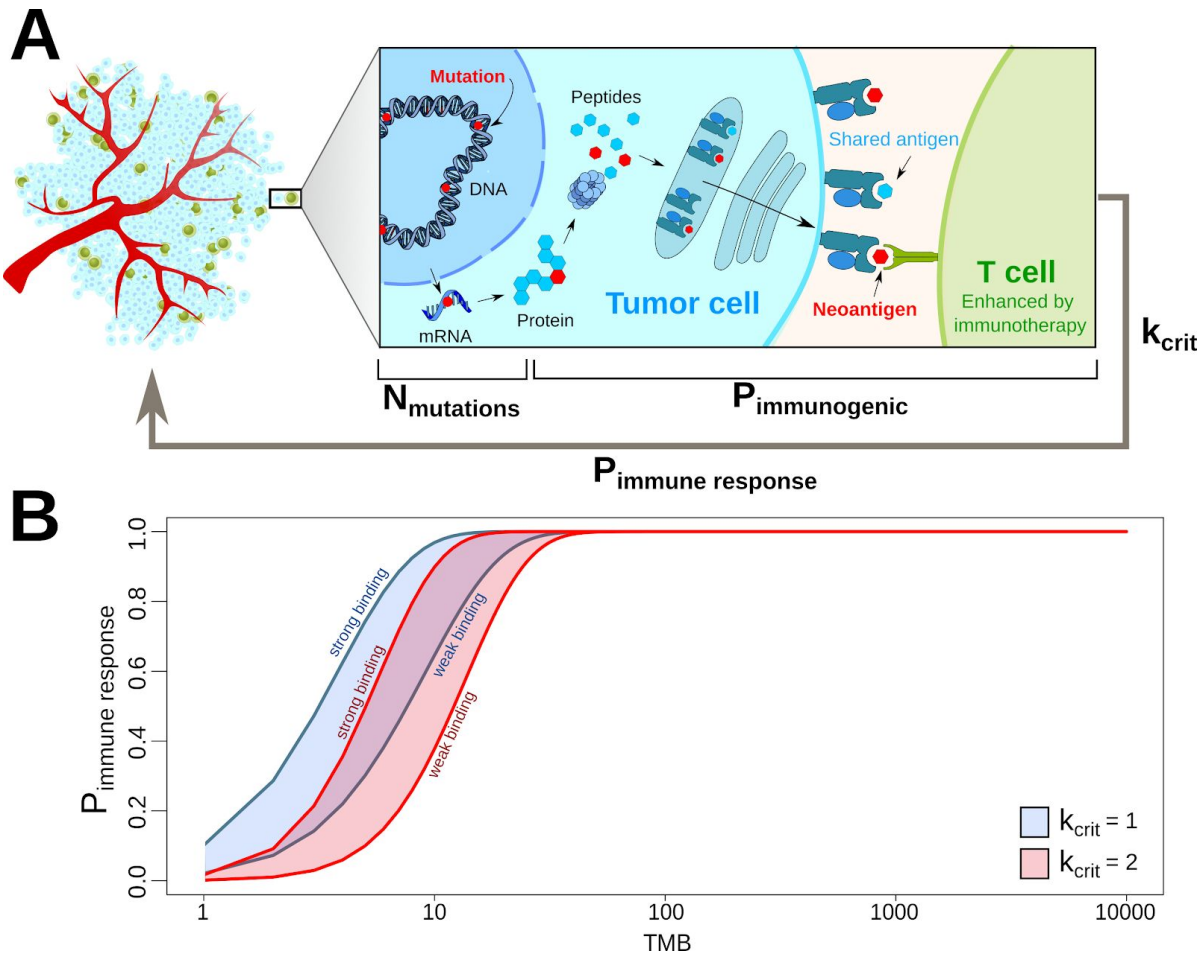

Figure 4: TMB and cancer immunogenicity

(A) Our model of cancer immunogenicity coarse-grains several cellular processes into the probability that a mutation becomes immunogenic ($P_{immunogenic}$). If the number of immunogenic mutations reaches k_{crit}, the cancer triggers an immune response **(B)** The probability of immune response $P_{immune\ responce}$ as a function of TMB for a range of k_{crit} and $P_{immunogenic}$. Rapid saturation of $P_{immune\ responce}$. TMB requires low k_{crit} and sufficiently high $P_{immunogenic}$>0.1 (see **Materials and Methods**).

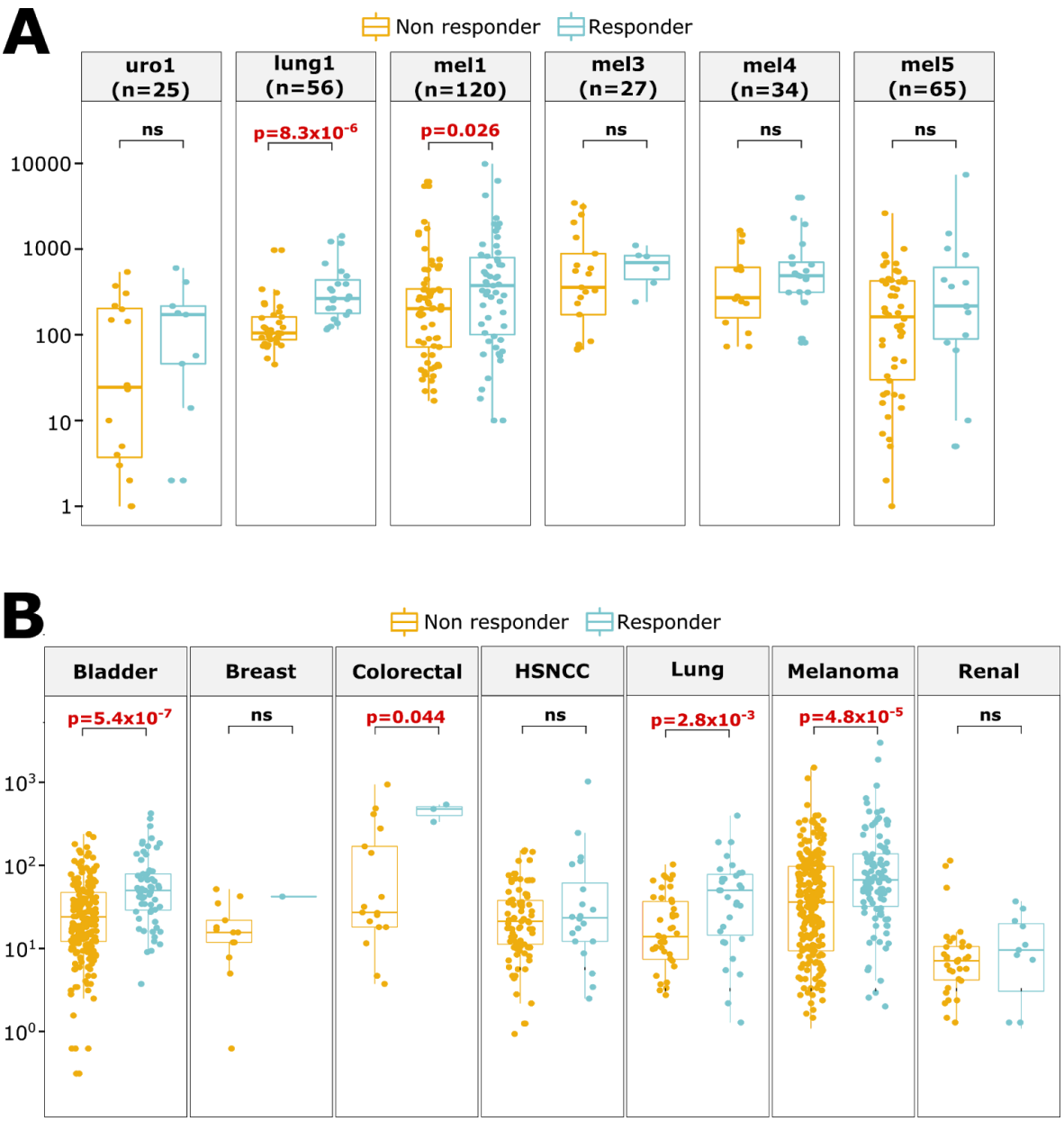

Figure S1: TMB association with clinical benefit from ICB across cancers
(A) Association of TMB with response to ICB across three cancer types from CPI800+. Only melanoma and non-small cell lung cancer have a significantly different TMB between responders and non-responders. **(B)** Association of TMB with response to ICB across seven cancer types from CPI1000+. Melanoma, non-small cell lung, bladder and colorectal cancer have a significantly different TMB between responders and non-responders.

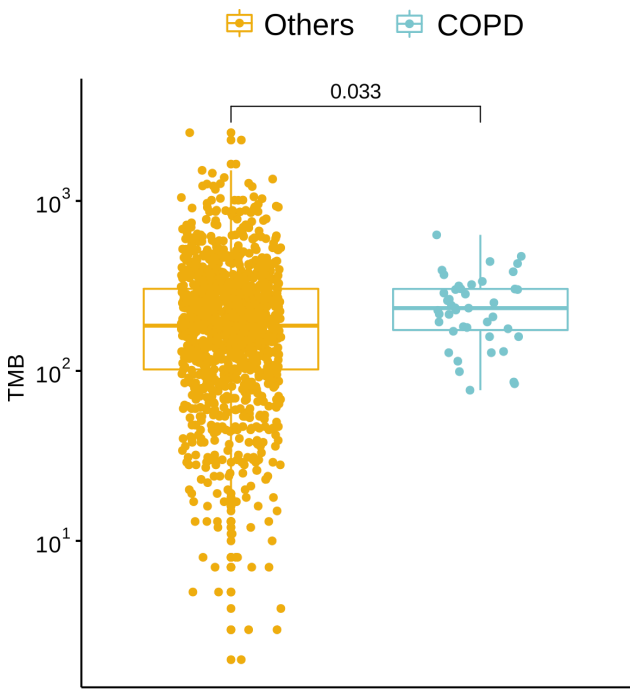

Figure S2: Chronic obtrusive pulmonary disease status and TMB
Association of TMB with COPD in TCGA. Of the 83 patients with COPD data, 43 were diagnosed with COPD. Here we compare COPD patients (n=43) to the rest of the cohort (n=1101).

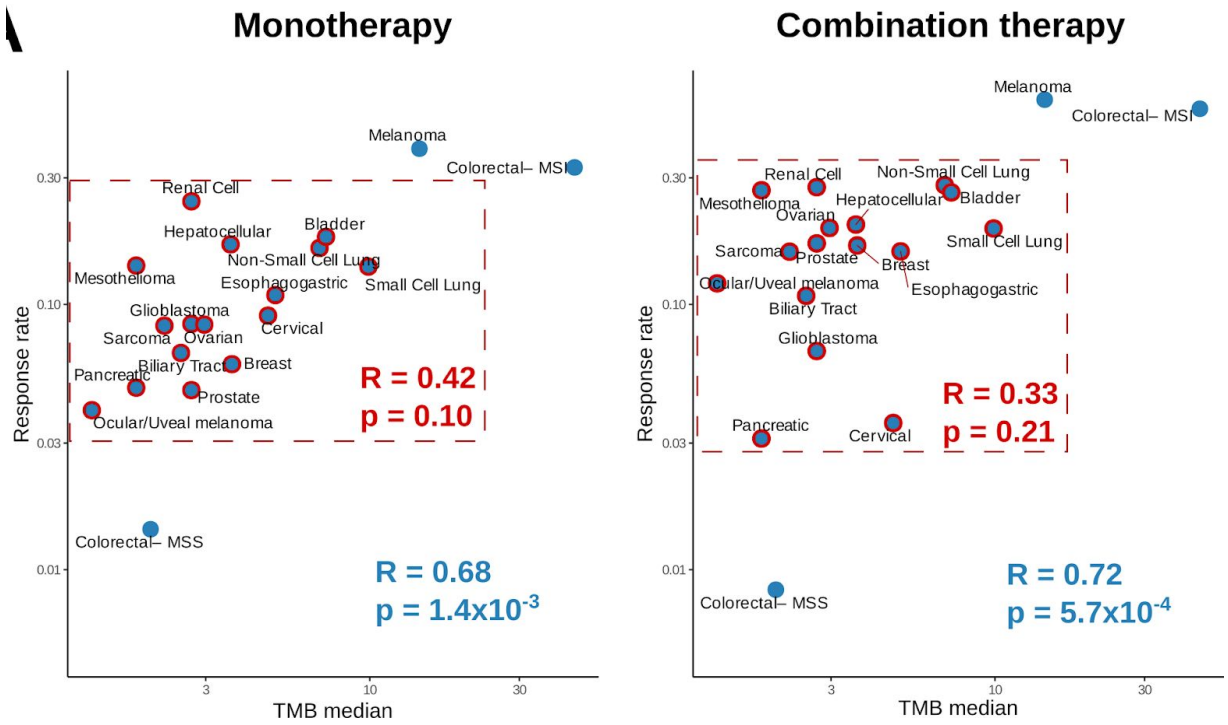

Figure S3: Correlation between response rates and TMB across cancer types
Median TMB in 19 cancer types of patients who underwent immunotherapy treatment (monotherapy or combination therapy). Pearson R correlation coefficient and p-value was calculated for all patients (in blue) and a subset of patients (red box and values) after removing melanoma and colorectal cancers.

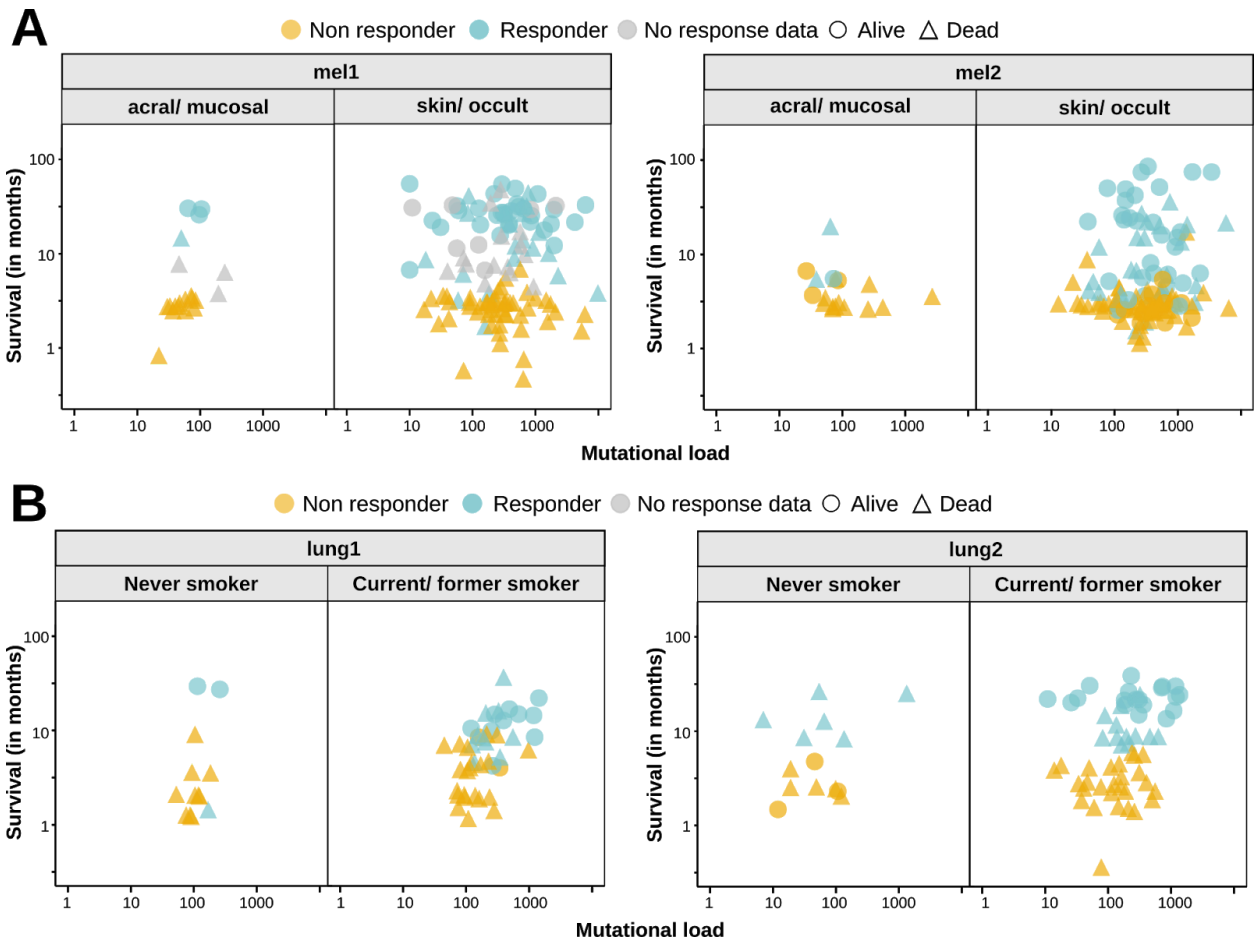

Figure S4: TMB association with progression-free survival post-immunotherapy
(A) (B) Plots of progression-free survival and TMB for melanoma and lung cancer ICB cohorts labeled by cancer subtype, showing the lack of correlation or of an obvious TMB cutoff.

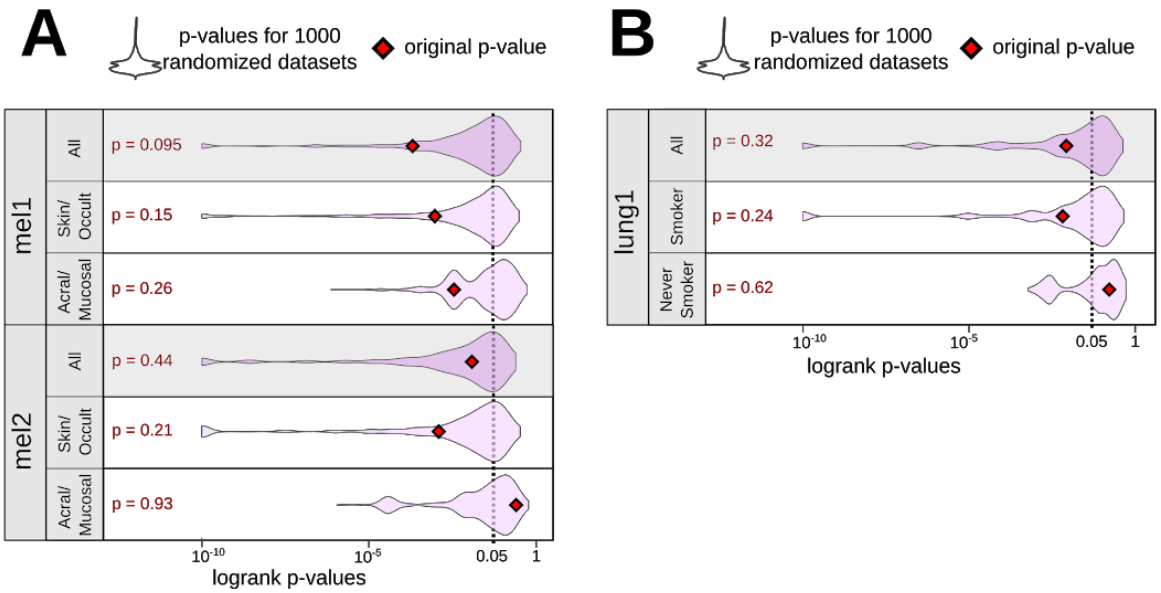

Figure S5: TMB association with overall survival post-immunotherapy
(A) Randomization analysis results in mel1 and mel2 and stratification by subtypes (p-values < 10^{-10} not shown) **(B)** Randomization analysis results in lung1 and stratification by subtypes (p-values < 10^{-10} not shown). When corrected for multiple hypotheses all cohorts fail to provide a statistically significant cutoff.

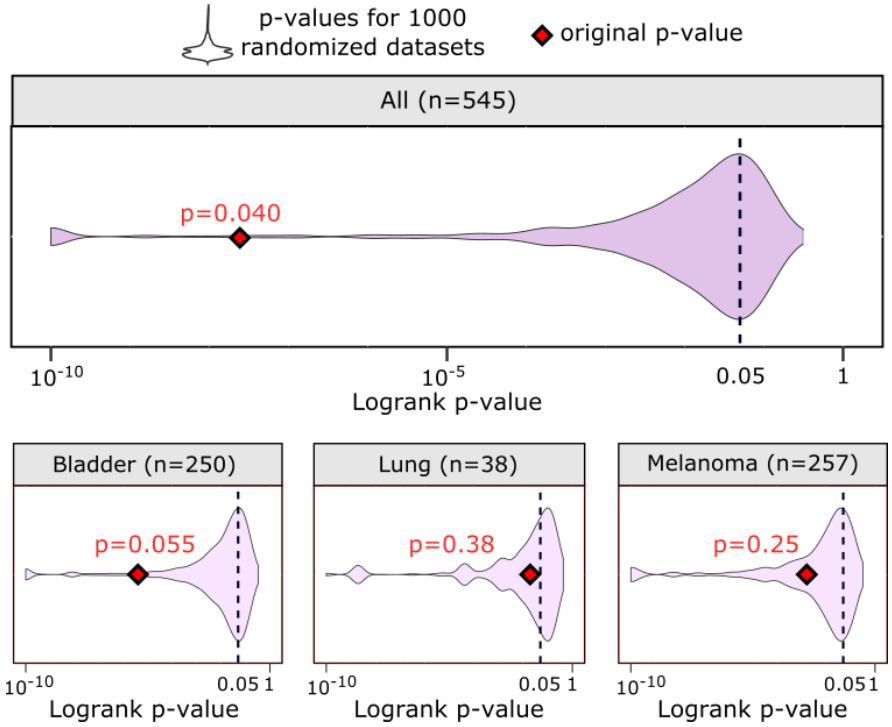

Figure S6: TMB association with overall survival post-immunotherapy
Results of the randomization analysis in CPI1000+ (p-values < 10-10 not shown).When cancer types of CPI1000+ were combined, a nominally significant p-value (p=0.04) arises, likely due to cancer types with different TMB ranges showing significantly different survival rates to ICB.

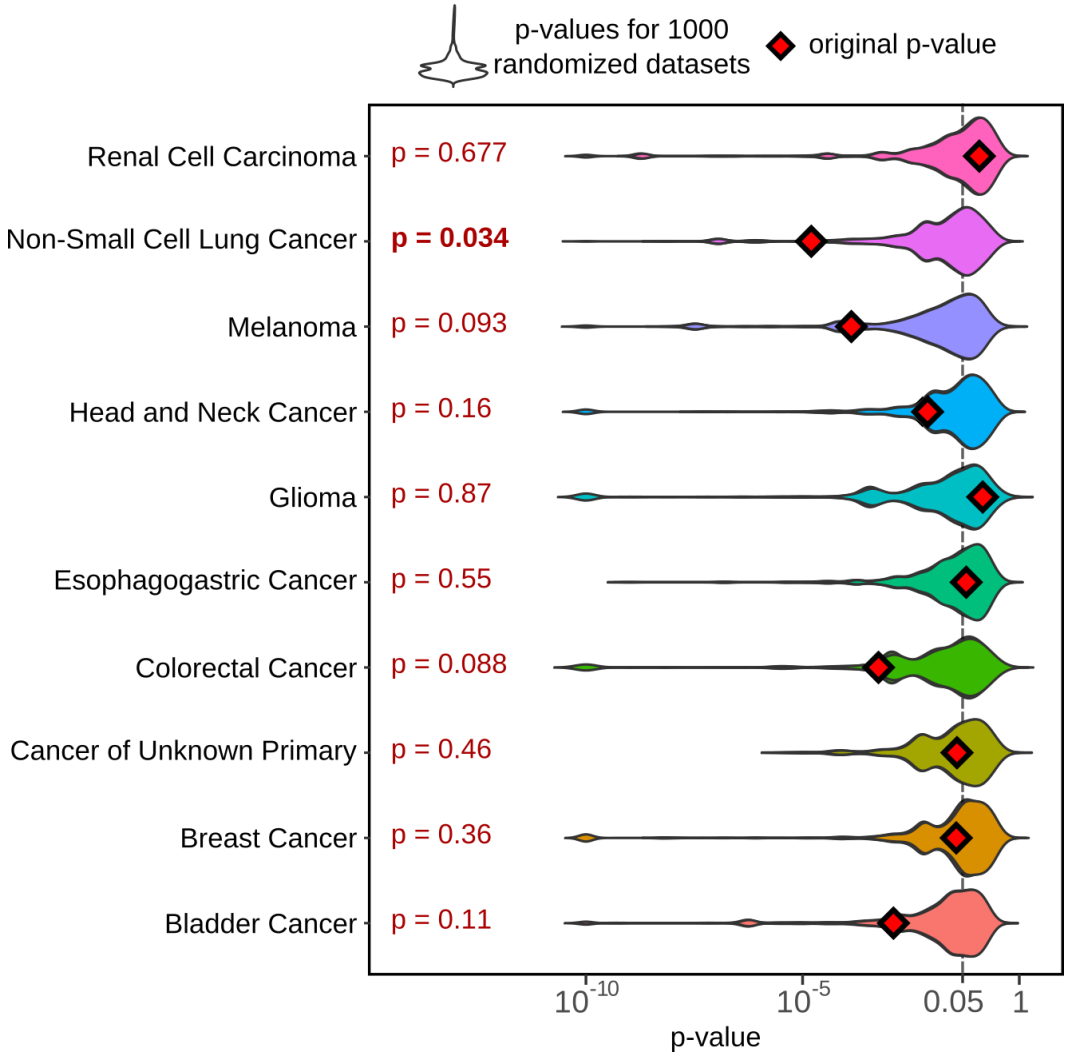

Figure S7: TMB association with overall survival post-immunotherapy
Randomization analysis results in multiple cancer types with MSK-IMPACT targeted next-generation sequencing data (p-values < 10^{-10} not shown)

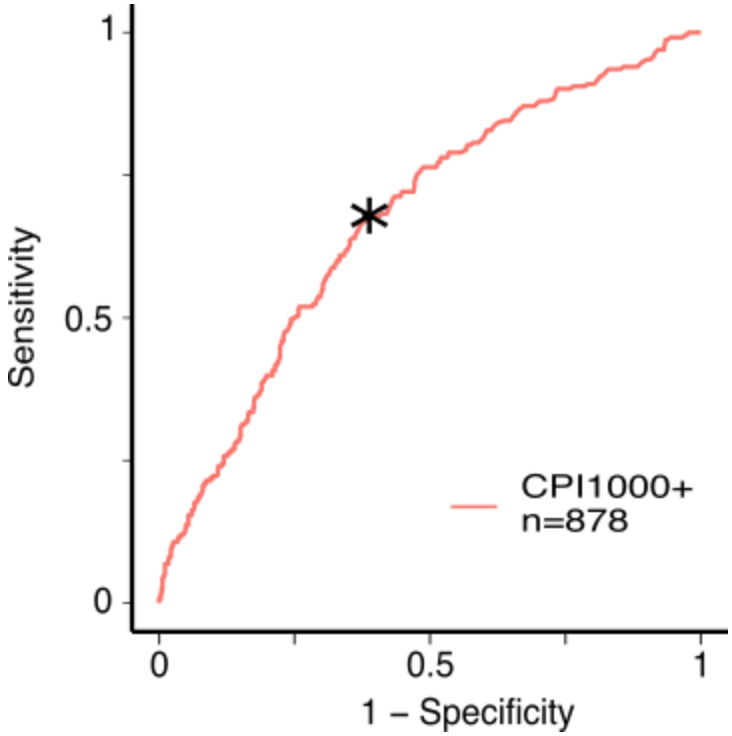

Figure S8: TMB as a biomarker of response to immunotherapy
ROC curve for CPI1000+. The Youden index associated cutoffs is also plotted.

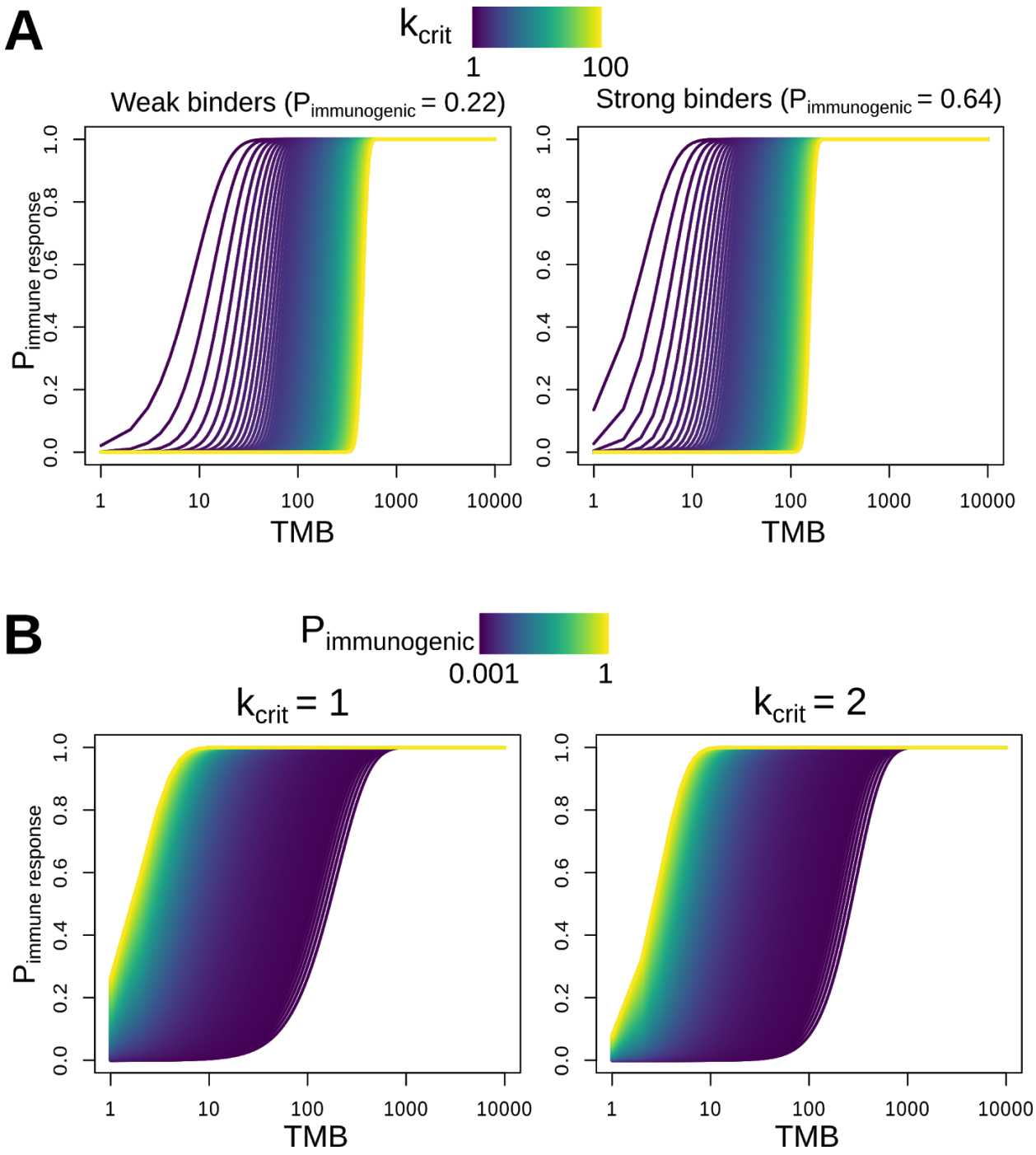

Figure S9: Components of cancer immunogenicity
(A) Probability of eliciting an immune response for a range of k_{crit} values
(B) Probability of eliciting an immune response for a range of $P_{immunogenic}$ values