



Published in final edited form as:

Nat Cancer. 2020 January ; 1(1): 99–111. doi:10.1038/s43018-019-0008-8.

Multimodal genomic features predict outcome of immune checkpoint blockade in non-small-cell lung cancer

Valsamo Anagnostou^{1,2,9,*}, Noushin Niknafs^{1,9}, Kristen Marrone^{1,2}, Daniel C. Bruhm¹, James R. White¹, Jarushka Naidoo^{1,2}, Karlijn Hummelink³, Kim Monkhorst³, Ferry Lalezari³, Mara Lanis¹, Samuel Rosner¹, Joshua E. Reuss¹, Kellie N. Smith^{1,2}, Vilmos Adleff¹, Kristen Rodgers⁴, Zineb Belcaid¹, Lamia Rhymee¹, Benjamin Levy^{1,2}, Josephine Feliciano^{1,2}, Christine L. Hann^{1,2}, David S. Ettinger^{1,2}, Christos Georgiades⁵, Franco Verde⁶, Peter Illei^{1,7}, Qing Kay Li⁷, Alexander S. Baras⁷, Edward Gabrielson⁷, Malcolm V. Brock⁴, Rachel Karchin^{1,8}, Drew M. Pardoll^{1,2}, Stephen B. Baylin¹, Julie R. Brahmer^{1,2}, Robert B. Scharpf¹, Patrick M. Forde^{1,2}, Victor E. Velculescu^{1,2,7,8,*}

¹The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA ²The Bloomberg-Kimmel Institute for Cancer Immunotherapy, Johns Hopkins University School of Medicine, Baltimore, MD, USA ³Antoni van Leeuwenhoek Nederlands Kanker Instituut, Amsterdam, the Netherlands ⁴Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD, USA ⁵Department of Radiology and Surgery, Johns Hopkins University School of Medicine, Baltimore, MD, USA ⁶Department of Radiology, Johns Hopkins University School of Medicine, Baltimore, MD, USA ⁷Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, USA ⁸Institute for Computational Medicine, Johns Hopkins University, Baltimore, MD, USA ⁹These authors contributed equally: Valsamo Anagnostou, Noushin Niknafs

Abstract

Despite progress in immunotherapy, identifying patients that respond has remained a challenge. Through analysis of whole-exome and targeted sequence data from 5,449 tumors, we found a significant correlation between tumor mutation burden (TMB) and tumor purity, suggesting that low tumor purity tumors are likely to have inaccurate TMB estimates. We developed a new method to estimate a corrected TMB (cTMB) that was adjusted for tumor purity and more accurately predicted outcome to immune checkpoint blockade (ICB). To identify improved predictive markers together with cTMB, we performed whole-exome sequencing for 104 lung tumors treated with ICB. Through comprehensive analyses of sequence and structural alterations, we discovered a significant enrichment in activating mutations in receptor tyrosine kinase (RTK)

Reprints and permissions information is available at www.nature.com/reprints.

*Correspondence and requests for materials should be addressed to V.A. or V.E.V. vanagno1@jhmi.edu; velculescu@jhmi.edu. Author contributions

V.Anagnostou and V.E.V. conceived and designed the study. V.Anagnostou, K.Ma., J.N., K.H., K.Mo., F.L., M.L., S.R., J.E.R., K.N.S., V.Adleff, K.R., Z.B., L.R., B.L., J.F., C.L.H., D.S.E., C.G., F.V., P.I., A.S.B., E.G., M.V.B., J.R.B. and P.M.F. acquired data. V.Anagnostou, N.N., D.C.B., J.R.W., Q.K.L., R.K., D.M.P., S.B.B., J.R.B., R.B.S., P.M.F. and V.E.V. analyzed and interpreted data. All authors have approved the submitted version and are accountable for their contributions.

Extended data is available for this paper at <https://doi.org/10.1038/s43018-019-0008-8>

Supplementary information is available for this paper at <https://doi.org/10.1038/s43018-019-0008-8>.

genes in nonresponding tumors in three immunotherapy treated cohorts. An integrated multivariable model incorporating cTMB, RTK mutations, smoking-related mutational signature and human leukocyte antigen status provided an improved predictor of response to immunotherapy that was independently validated.

The intersection of cancer genomics with new immunologic approaches is revolutionizing cancer therapy. Antitumor responses are integrally connected to tumor genomics as neoantigens stemming from somatic mutations seem to shape immune responses and drive clinical benefit from immunotherapy in several tumor types including non-small-cell lung cancer (NSCLC)¹⁻⁴. A high TMB has been associated with benefit from ICB across tumor types^{5,6}. Despite the value of TMB in predicting response and survival in the context of ICB, there are tumors with a high TMB that do not respond and conversely there are tumors with low TMB that benefit from immunotherapy. It has become apparent that understanding the nuances of TMB in terms of clonal composition, contribution of mutation types and mutational signatures, and enrichment in oncogenic drivers will be critical for developing a robust predictor of response to ICB. Moreover, tissue-based TMB estimates may be challenging in low tumor purity samples and in tumors with a higher intratumoral heterogeneity^{7,8}. Adding to these complexities, the estimation of tumor purity itself may be challenging as pathologic assessments are frequently imprecise and have limited reproducibility⁹. These limitations are reflected in the current National Comprehensive Cancer Network guidelines, where the use of TMB as a predictive biomarker is limited by lack of calibration and harmonization across multiple next-generation sequencing (NGS) platforms.

Response to immunotherapy is orchestrated by immune-related pathways, with the antigen presentation machinery playing a chief role as mutation-associated neoantigens (MANAs) are presented on major histocompatibility complex-I (MHC-I) molecules to CD8⁺ T cells and trigger an antitumor immune response that translates to clinical benefit. Genetic variation in the antigen presentation machinery, both at a germline as well as a somatic level may therefore modulate an effective antitumor immune response^{10,11}. Previous efforts have largely focused on single biomarkers of response to immunotherapy, highlighting the unmet need for the development of integrated nuanced molecular models. Here, we report an integrated approach where an improved measure for TMB, corrected for tumor purity, is combined with human leukocyte antigen (HLA) class I genetic variation, genomic alterations in RTK genes and genome-wide mutational signatures to capture the multifaceted nature of the tumor-immune system crosstalk and more accurately predict outcome for immunotherapy.

Results

Cohorts and clinical response assessment.

We evaluated 3,788 tumor samples from The Cancer Genome Atlas (TCGA) and 1,661 tumor samples from a published cohort of immunotherapy treated patients⁶ to investigate the complexities of TMB estimates derived from whole exome and targeted sequencing, respectively. In parallel, we performed genome-wide sequence and structural analyses for a

cohort of 104 NSCLC patients treated with ICB (cohort 1; Supplementary Table 1 and Methods). NGS data from two published cohorts of NSCLC patients treated with ICB (cohorts 2 and 3, see Methods for detailed description) were obtained and analyzed to validate key findings from cohort 1 (refs. ^{1,2,12}). Durable clinical benefit was defined as complete, partial response or stable disease with a duration >6 months from the time of treatment initiation (Supplementary Table 1). Progression-free survival (PFS) and overall survival (OS) were used to determine clinical outcome (see Methods).

Limitations of TMB as a predictive biomarker of response to immunotherapy.

To determine how tumor purity may affect TMB we performed simulation analyses and evaluated the tumor purity needed to accurately determine TMB in the setting of different clonal composition backgrounds. Through these analyses we determined the tumor purity required to establish reliable TMB assessments and that observed TMB (obsTMB) also depends on intratumoral clonal heterogeneity (Fig. 1a–d). At the lower spectrum of tumor purity, TMBs of clonally heterogeneous TMB-high and clonally homogeneous TMB-low tumors become indiscernible, underlining the need to correct TMB for tumor purity (Fig. 1a–d).

To substantiate these findings, we analyzed tumor whole-exome sequencing data from 3,788 TCGA samples from seven tumor types (bladder carcinoma, breast carcinoma, colon adenocarcinoma, head and neck squamous cell carcinoma (HNSCC), kidney clear cell carcinoma (KIRC), NSCLC and melanoma) and found a correlation between obsTMB and tumor purity, with a lower number of alterations observed in samples with low tumor purity (Extended Data Fig. 1). We found that the correlation between obsTMB and tumor purity was particularly pronounced for samples with tumor purity below 50% (Pearson's $R = 0.33$, $P = 0.05$, $R = 0.48$, $P = 0.0003$, $R = 0.18$, $P = 0.09$ and $R = 0.39$, $P = 0.002$ for HNSCC, KIRC, lung adenocarcinoma and lung squamous cell carcinoma, respectively; Fig. 1e–h). Given that targeted NGS approaches enable deeper sequencing coverage and may therefore mitigate the effect of tumor purity on analysis of low tumor purity highly clonally heterogeneous tumors, we evaluated the correlation between tumor purity and TMB in a large cohort of tumors sequenced with targeted NGS⁶. Consistent with our findings, we identified a significant correlation between tumor purity and obsTMB estimates, particularly in NSCLC (Fig. 1i and Extended Data Fig. 1), suggesting that tumor purity remains a limiting factor for accurately estimating TMB even in the setting of higher sequencing depth.

To more carefully examine TMB and identify other biomarkers of response to ICB, we performed whole-exome sequencing on 104 matched tumor/normal pairs from NSCLC patients treated with ICB. Eighty-nine cases that passed strict quality control measures (Methods) were further analyzed (cohort 1, Supplementary Tables 1–3) and a published cohort of 34 NSCLC patients treated with anti-PD1 blockade¹ (cohort 2) was analyzed to validate findings from cohort 1. We employed both a mutant allele frequency (MAF)-based as well as a copy number-based approach to determine tumor purity as we have previously described¹³ (Methods). Similar to the findings from the NSCLC TCGA cohort (Fig. 1g,h and Extended Data Fig. 2), in the two immunotherapy treated NSCLC cohorts, the

correlation between obsTMB and tumor purity was particularly pronounced for tumor purity less than 30% ($P=0.008$ and $P=0.08$ for overall comparisons of TMB across tumor purity tiers for cohort 1 and cohort 2, respectively; see Extended Data Fig. 2). These findings suggested that obsTMB values may largely deviate from the true TMB in low purity tumors.

Corrected TMB more accurately predicts outcome of ICB.

To overcome this limitation of obsTMB measurements, we developed an approach to estimate corrected TMB (cTMB) values for each tumor based on tumor purity. We first simulated 20,000 tumors with various levels of intratumoral heterogeneity, TMB and depth of coverage using a reference set from TCGA. We then used in silico dilutions of these simulated tumors to model the obsTMB resulting from characterization of each simulated tumor sample at various levels of tumor purity (Methods). For each simulated tumor we generated a correction factor for different purity tiers (Fig. 2a and Supplementary Table 4). An analysis of the obsTMB in cohort 1 revealed that, consistent with previous findings^{1,4,14}, patients with durable clinical benefit to ICB had significantly higher obsTMB compared to patients with nondurable clinical benefit (Mann–Whitney U test $P=0.002$, false discovery rate (FDR)-adjusted $P=0.012$, Supplementary Table 5). However, there was a substantial overlap in the range of obsTMB between the two groups, and obsTMB marginally predicted OS (log-rank $P=0.048$, Fig. 2b). Using our developed correction factors for different purity tiers, we determined cTMB values for the TCGA tumors (Extended Data Fig. 3a) and the tumors in our cohort (Extended Data Fig. 3b). cTMB resulted in better prognostication of outcome (log-rank $P=0.014$, Fig. 2c), suggesting that the obsTMB may be largely underestimated in low tumor purity samples and result in misclassification of patients with these tumors.

We then sought to determine the value of cTMB in tumors sequenced by targeted NGS and followed a similar approach to generate correction factors for each tumor purity tier. This analysis showed that accuracy of TMB estimates from targeted NGS drops significantly for tumor purity less than 20% (Fig. 2d). Reanalysis of 1,661 tumor samples treated with ICB⁶, revealed that using cTMB estimates improved OS prognostication (log-rank $P=2.1 \times 10^{-7}$ for TMB estimates from targeted NGS and log-rank $P=2.8 \times 10^{-8}$ for cTMB, Fig. 2e,f).

Genomic features of response to ICB.

We further refined our approach by interrogating mutational signatures, as smoking-related C>A transversions have been identified in NSCLC patients with clinical benefit from ICB^{14,15}. Given the aforementioned challenges of tumor purity, we evaluated the number of mutations needed to accurately estimate the contribution of the C>A rich molecular smoking signature. We performed in silico dilution experiments of whole-exome mutational profiles of 985 TCGA NSCLC tumors and found that a minimum of 20 nonsynonymous mutations would be required to predict the presence of a dominant smoking signature (Extended Data Fig. 4a,b). An analysis of tumor samples with at least 20 mutations revealed an enrichment of the molecular smoking signature in patients with durable clinical benefit (Mann–Whitney U test $P=0.003$, FDR-adjusted $P=0.027$, Fig. 3 and Supplementary Table 5). The molecular smoking signature more accurately predicted OS than obsTMB (log-rank $P=0.031$, Extended Fig. 4c), suggesting that the smoking-associated mutational processes are

the likely cause of high TMB and, therefore, for samples with low tumor purity, mutational signatures could serve as a proxy for TMB. Given the causal relationship of tobacco exposure with clonal mutations, we assessed the correlation between the molecular smoking signature and ICB response after adjusting for clonal TMB. The association of the molecular smoking signature with clinical outcome remained significant after clonal TMB adjustment (logistic regression, $P = 0.035$), suggesting that the molecular smoking signature may be informative in cases where TMB is not (Extended Data Fig. 4d).

We subsequently sought to identify genomic alterations in driver genes that were selectively associated with responding or nonresponding tumors after accounting for the mutation load of a given tumor. Such an adjustment is crucial given the higher probability of passenger mutations in driver genes in tumors with a high TMB. We found a significant enrichment in activating mutations in RTK genes in patients who did not derive durable clinical benefit from ICB (Fisher's exact $P < 0.001$, FDR-adjusted $P = 0.002$, Fig. 3 and Supplementary Table 5). The RTK superfamily of cell-surface receptors serve as mediators of cell signaling by extra-cellular growth factors and these oncogenes are activated by point mutations, amplifications (*FGFR1*, *IGF1R*) or both (*EGFR*, *ERBB2*, *MET*)¹⁶. EGFR exon 19 in-frame deletions (745KELREA>T, E746_A750del, L747_T751del), exon 20 in-frame insertions (N771_H773dup) and exon 21 point mutations (L858R) as well as ERBB2 exon 19 (E770_A771insAYVM) and exon 20 (776G>VC) in-frame insertions were exclusively found in nonresponding tumors in cohort 1 (Fig. 3 and Supplementary Table 3). Similarly, *EGFR*, *ERBB2*, *MET* and *IGF1R* amplifications were only observed in nonresponding tumors, and *FGFR1* amplifications were detected in two nonresponding tumors and one responding tumor (Fig. 3 and Supplementary Table 6). The distribution of activating RTK mutations was independent of TMB (Mann–Whitney U test $P = 0.33$ for TMB differences between RTK⁺ and RTK⁻ tumors) and remained significantly associated with clinical response to ICB after correction for TMB (logistic regression $P = 0.006$, Supplementary Table 7). RTK mutations strongly correlated with outcome even after correcting for both molecular smoking signature (logistic regression, $P = 0.007$) and smoking history (logistic regression, $P = 0.017$). RTK downstream signaling may promote intrinsic resistance to ICB through inhibitory effects on T cell recruitment and function¹⁷. Consistent with this notion, we found a significantly lower CD8⁺ T cell infiltration in tumors with activating RTK mutations (CD8⁺ T cell density of 7.36 ± 2.5 versus 15.16 ± 2.5 for tumors with and without activating RTK mutations, $P = 0.036$), indicating that RTK signaling may be linked to intratumoral T cell depletion. RTK activating mutations conferred reduced survival (log-rank $P = 0.005$, Extended Data Fig. 5a) and we validated these observations in cohort 2, where an enrichment of activating mutations in RTK genes was found in nonresponding tumors and resulted in worse PFS (log-rank $P = 0.009$, Extended Data Fig. 5b,c). We further examined this finding in a third independent cohort of 240 NSCLC patients treated with ICB², where tumors had been analyzed with targeted NGS. The data from this cohort confirmed our findings and revealed that RTK activating mutations in *EGFR*, *ERBB2*, *MET*, *FGFR1* and *IGF1R* were enriched in nonresponding tumors (Fisher's exact $P = 0.027$) and that RTK alterations were associated with shorter PFS (log-rank $P = 0.035$, Extended Data Fig. 5d).

Recurrent alterations in *ARID1A* were found in patients from cohort 1 with durable clinical benefit (Mann–Whitney U test $P = 0.005$, FDR-adjusted $P = 0.024$), with a trend toward

statistical significance after correction for TMB ($P=0.062$, Fig. 3 and Supplementary Table 7). *ARID1A* deficiency has been shown to impair mismatch repair¹⁸, which would in turn cause an increased mutation burden and drive responses to ICB. *KEAP1* mutations, in particular inactivating mutations with concurrent loss of the wild-type allele, were more commonly found in patients with nondurable clinical benefit, however, this observation did not reach statistical significance ($P=0.074$, Fig. 3 and Supplementary Table 7). A homozygous deletion in *PTEN* was found in one patient with a short-lived response to ICB and *MDM2/MDM4* amplifications were identified in three patients with nondurable clinical benefit (Fig. 3), consistent with previously described mechanisms of resistance^{19,20}. We did not detect any loss-of-function mutations in *JAK1* or *JAK2* (ref. ²¹) or an enrichment of cooccurring *KRAS* and inactivating *STK11* mutations²² in nonresponding tumors (Fig. 3). Additionally, we observed homozygous deletions in IFN- γ pathway genes²³ but their frequency was similar in responding and nonresponding tumors, and these deletions cooccurred with loss of the *CDKN2A* tumor suppressor gene in all but two of the nine cases in which they were present (Extended Data Fig. 6a–c). *CDKN2A* and the group of IFN- γ pathway genes are on chromosome 9p 917Kb apart, and therefore IFN- γ deletions may be cooccurring passengers in the setting of a driver *CDKN2A* deletion. Genomewide copy number analyses were employed to investigate differences in tumor aneuploidy (Methods), however, we did not find any significant differences in the fraction of the genome with allelic imbalance or loss of heterozygosity (LOH) between patients with durable and nondurable clinical benefit (Extended Data Fig. 6d and Supplementary Tables 8 and 9). As PD-L1 expression is routinely used as a predictive biomarker for ICB therapy, we assessed PD-L1 expression by immunohistochemistry in cohort 1, however, did not identify any differences in PD-L1 expression between responding and nonresponding tumors (Mann–Whitney U test $P=0.357$, Supplementary Table 1).

In parallel, we followed a pathway-focused approach to identify enrichment or mutual exclusivity of genomic alterations in oncogenic processes or signaling pathways. We specifically considered DNA damage repair genes and the WNT- β -catenin pathway, given the linkage of the former to mutation accumulation and response to immunotherapy^{24,25} and the association of the latter with T cell exclusion²⁶. We identified one responding TMB-high tumor with biallelic inactivation of *MLH1*, but did not identify an overall enrichment in deleterious somatic DNA repair gene mutations in responding tumors (Extended Data Fig. 7a). Similarly, we detected a gain-of-function *CTNNB1* hotspot mutation in a nonresponding tumor but no additional differences in activating mutations in the WNT pathway between responders and nonresponders (Extended Data Fig. 7a).

The premise of immune targeted therapy relies on MANA-specific immune responses²⁷, but computationally predicted MANA load has been tightly correlated with TMB and has similar predictive value^{1,4,28}. In line with previous studies we found a strong correlation between TMB and predicted MANA load ($R=0.93$, $P<0.001$, Supplementary Table 10). As only a small fraction of predicted MANAs are indeed immunogenic²⁹, we focused on neoantigens that have predicted MHC affinities ≤ 50 nM and for which the corresponding wild-type peptide does not bind MHC class I (affinity $\leq 1,000$ nM) hypothesizing that these ‘fit’ neoantigens are most likely to be identified as nonself by the immune system and potentiate an antitumor immune response³⁰. A higher number of fit MANAs was found in

responding versus nonresponding tumors, however, fit neoantigen load was less predictive of outcome compared to TMB (Mann–Whitney U test $P=0.01$, FDR-adjusted $P=0.05$; Fig. 4 and Supplementary Table 5). We further focused on neoantigens stemming from frameshift alterations, as conceptually these could generate multiple immunogenic neoantigens. Responding tumors showed a trend for a higher number of predicted MANAs derived from frameshift mutations (Mann–Whitney U test $P=0.08$, Fig. 4). We then studied the potential of hotspot mutations in driver and other genes to generate fit MANAs as such alterations may be less likely to be eliminated as a means of immune escape. A subset of clonal hotspot frameshifts and in-frame indels generated fit MANAs (Extended Data Fig. 7b), however, patients harboring fit hotspot MANAs did not have a favorable outcome (log-rank $P=0.1$). While these findings do not provide definitive evidence that clonal hotspot MANAs are indeed immunogenic, they provide the foundation for further exploration of hotspot peptide-specific T cell responses, especially when derived from frameshift mutations^{25,31}.

Genetic variation in antigen presentation machinery affects response to ICB.

Antigen presentation deficiency may lead to immune escape through both HLA class I germline homozygosity and somatic LOH^{11,32}. In cohort 1, 22 cases were homozygous for at least one HLA class I locus in their germline, and somatic HLA LOH occurred in 27 tumors (Fig. 5a and Supplementary Table 11). Mutations in HLA class I genes were rare (only seen in three cases), consistent with previous studies³³. Through analysis of 3,601 TCGA samples, we did not find an enrichment for LOH in chromosome 6p—that contains the HLA class I loci—compared to background arm-level allelic imbalance in NSCLC. The degree of 6p LOH was higher in NSCLC compared to other tumor types ($P<0.001$, Extended Data Fig. 8). The β_2 -microglobulin locus was frequently lost by LOH, however, we did not detect any concurrent inactivating mutations, rendering this an infrequent mechanism of immune evasion in cohort 1 (Fig. 5a). Conceptually, tumors with increased TMB would be more likely to be recognized by the immune system but may overcome this evolutionary disadvantage through HLA loss and diminished neoantigen presentation. While germline HLA zygosity was not correlated with TMB for most of the tumors examined, combined germline and tumor HLA status was correlated with TMB such that tumors with a lower TMB harbored a more intact antigen presentation capacity (Extended Data Fig. 8). We did not find any association between TMB and HLA class I supertypes, and germline HLA class I variation was not associated with outcome (Extended Data Fig. 9).

HLA class I germline zygosity and somatic HLA class I LOH events were combined to determine the effect of unique number of HLA class I alleles on response to ICB. Tumors with reduced antigen presentation potential were linked to worse outcome to ICB (Fig. 5b) and when antigen presentation capacity and TMB were combined, NSCLCs with low TMB and reduced antigen presentation potential had a significantly shorter OS (log-rank $P=0.01$, Fig. 5c). Tumors with lower antigen presentation capacity showed a significantly lower level of CD8⁺ T cell density (Mann–Whitney U test $P=0.005$, Fig. 5d), suggesting that these tumors may present a less diverse neoantigen repertoire to cytotoxic T cells and therefore have the potential to become partially invisible to the immune system. Furthermore, cases with maximal HLA class I heterozygosity were found to have a less clonal T cell receptor

(TCR) repertoire ($P=0.01$, Extended Data Fig. 9), suggesting that HLA variation determines the selection and clonal expansion of neoantigen-specific T cells.

Integrated model of response to ICB.

We then assessed correlations among genomic and immune features and identified four clusters of inter-correlated features centered on RTK mutations, HLA genetic variation, tumor aneuploidy and cTMB (Extended Data Fig. 10a). Given the importance of cTMB, RTK mutations and HLA genetic variation from the single feature analysis as well as the additive predictive value of the molecular smoking signature, we combined cTMB, molecular smoking signature, RTK activating mutations and HLA genetic variation in a multi-parameter predictor of outcome (Fig. 6a). We applied multivariate Cox proportional hazards regression analysis to evaluate the combined contribution of these molecular features in predicting OS in cohort 1 (Supplementary Table 12), followed by independent validation of the model in cohort 2. A risk score was calculated as the exponential of the sum of product of mean-centered covariate values and their corresponding coefficient estimates and used to classify patients in high and low risk groups (Methods). Patients classified in the high risk category had a significantly shorter OS compared to patients at low risk for disease progression (median OS 13 months versus not reached, log-rank $P=0.0001$, HR = 3.29, 95% CI: 1.77–6.14; Fig. 6b) and these findings were supported by a separate analysis of PFS in cohort 2 (median PFS 3 versus 8 months, log-rank $P=0.017$, HR = 2.73, 95% CI: 1.15–6.45; Fig. 6c). These findings were replicated when only baseline tumors treated with single or dual ICB were considered (Extended Data Fig. 10b,c).

Discussion

Individual biomarkers of response to immunotherapy such as PD-L1 expression and TMB have modest predictive use across a plethora of studies. Through our analyses we showed that the complexities of the predictive value of TMB may be in part attributed to tumor purity and developed a new approach to generate cTMB values that more accurately predicted outcome for ICB. Low tumor purity, mainly due to sampling, may greatly affect TMB assessments, resulting in falsely low TMB in low tumor cellularity samples, especially for tumors with a higher fraction of subclonal mutations⁷. These findings are of particular importance for metastatic NSCLC where most tumor samples are obtained by bronchoscopy or core needle biopsies and are therefore subject to tumor purity limitations. While targeted NGS may alleviate the tumor purity effect given the higher coverage compared to whole-exome sequencing, our findings suggest that TMB values should only be interpreted after taking into consideration the tumor purity of the sample analyzed in both settings.

While tumor molecular features such as TMB, PD-L1 expression, and host immune infiltrates predict in part which patients will benefit from immune targeted therapies, identifying patients with primary resistance is equally important. Potential mechanisms of primary resistance span a broad range from tumor genomic features^{2,14,19,22,34}, expression of immune checkpoint molecules³⁵, insufficient cytotoxic T cell infiltration³⁶, transcriptomic signatures^{37,38} and antigen presentation defects^{11,19,22,32}. However, each mechanism explains a fraction of primary resistance and has not been validated in

independent cohorts. We report here a significant enrichment in activating RTK genomic alterations in nonresponding tumors that identified patients with an inferior outcome from ICB in three independent NSCLC cohorts. While activating *EGFR* and *MET* mutations have been described in patients that do not benefit from ICB^{2,14,39}, our study demonstrates that activating genomic alterations in RTK genes including *EGFR*, *HER2*, *MET*, *FGFR1* and *IGF1R* are linked with primary resistance to ICB independent of mutation burden. RTKs transduce signals through the mitogen-activated protein kinase and PI3K-AKT-mTOR pathways, which have been implicated in regulation of immune responses in the tumor microenvironment such that oncogenic signaling through mitogen-activated protein kinase and PI3K renders the microenvironment less permeable to cytotoxic T cells^{19,40}.

Through the observations in this study, we have combined key molecular features into a predictive classifier for NSCLC patients treated with ICB. Previous attempts to combine biomarkers have focused on a limited number of features such as TMB and chromosomal imbalance⁴¹, TMB and immune cell gene expression profiles³⁷ or HLA variation and TMB^{11,32}. Our multivariable model incorporates an improved measure of TMB through correction of tumor purity, RTK mutations, molecular smoking signature and HLA genetic variation, highlighting the need for development of integrative platforms that capture the complexities of the cancer-immune system crosstalk. Our work may be limited by the sample size, the heterogeneity of our cohort and use of PFS as an endpoint for cohorts 2 and 3. We believe that PFS is a reasonable surrogate for OS, especially for heavily pretreated NSCLC and large-scale validation studies of prospectively collected cohorts will further confirm these approaches.

Our analyses provide insights into mechanisms of response to ICB and lay the groundwork for improved integrated molecular markers of outcome for these therapies. Moving forward, we envision integrative predictive models of response that are not limited only to static features of pretreatment tumors but also encompass dynamic biomarkers⁴² that capture tumor evolution under the selective pressure of immunotherapy.

Methods

Cohort characteristics.

We obtained matched tumor-normal exome sequencing data from 3,788 patients in TCGA (<http://cancergenome.nih.gov>), as outlined in the TCGA publication guidelines <http://cancergenome.nih.gov/publications/publicationguidelines>, focusing on tumors that would be relevant for immunotherapy. Cohort 1 consisted of 104 NSCLC patients treated with ICB at Johns Hopkins Sidney Kimmel Cancer Center and the Nederlands Kanker Instituut. Of these, 15 cases were not included in the final analyses because of tumor purity <10% or absence of matched normal samples. Of the 89 eligible patients, 46 were male and average age was 65 years. The studies were approved by the Institutional Review Board and patients provided written informed consent for sample acquisition for research purposes. Clinical characteristics for all patients are summarized in Supplementary Table 1. Exome data from a published cohort of NSCLC patients treated with PD1 blockade (cohort 2) were obtained and analyzed to validate key findings from cohort 1 (refs. ^{1,12}). A publicly available cohort of 240 NSCLC patients treated with ICB² was obtained through CBioPortal for Cancer

Genomics² and used to validate the association of RTK mutations with outcome (cohort 3). A publicly available cohort of 1,661 tumors analyzed by targeted NGS was obtained through the CBioPortal for Cancer Genomics⁶ to validate the correlation between TMB and tumor purity in the setting of higher sequencing depth. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Treatment and assessment of clinical response.

Eighty patients were treated with anti-PD1 therapy, seven patients received combination anti-PD1 and anti-CTLA4 therapy and two patients were treated with chemotherapy and anti-PD1 therapy. Given the challenges with conventional radiologic response assessments that may underestimate the unique patterns and timing of response to immune targeted therapies⁴³, we defined response as durable clinical benefit if complete, partial response or stable disease was achieved with a duration >6 months. Responding and nonresponding tumors, therefore refer to durable clinical benefit and nondurable clinical benefit, respectively. Response assessment was not evaluable for two patients. PFS and OS were defined as the time elapsed between the date of treatment initiation and the date of disease progression or death from disease, or the date of death, respectively. Ultimately, OS was used to determine long-term outcome for cohort 1. OS was evaluable for all 89 patients in cohort 1; OS was not available for cohorts 2 and 3; therefore, PFS was used. While OS remains the gold standard for evaluation of outcome after immunotherapy, PFS is a reasonable surrogate that is recognized as such in NSCLC by regulatory authorities, especially for heavily pretreated NSCLC cohorts with a short expected OS and low likelihood of clinical benefit post immunotherapy. Response assessments and outcome for cohort 1 are shown in detail in Supplementary Table 1.

Sample preparation and whole-exome sequencing.

Whole-exome sequencing was performed on preimmunotherapy tumor and matched normal samples, with the exception of three cases for which tumor from the time of resistance to therapy was analyzed (Supplementary Table 1). Tumor formalin-fixed paraffin-embedded (FFPE) blocks were retrieved and underwent pathological review for confirmation of lung cancer diagnosis and assessment of tumor cellularity. Slides from each FFPE block were macrodissected to remove contaminating normal tissue. Matched normal samples were provided as peripheral blood. DNA was extracted from patients' tumors and matched peripheral blood using the Qiagen DNA FFPE and Qiagen DNA blood mini kit, respectively (Qiagen). Fragmented genomic DNA from tumor and normal samples used for Illumina TruSeq library construction (Illumina) and exonic regions were captured in solution using the Agilent SureSelect v.4 kit (Agilent) as previously described¹³. Paired-end sequencing, resulting in 100 bases from each end of the fragments for the exome libraries was performed using Illumina HiSeq 2500 instrumentation (Illumina). The mean depth of total and distinct coverage for the pretreatment tumors were $\times 231$ and $\times 144$, respectively (Supplementary Tables 2, 3 and 6).

Primary processing of exome data and identification of putative somatic mutations.

Somatic mutations, consisting of point mutations, insertions, and deletions across the whole exome were identified using the VariantDx custom software for identifying mutations in

matched tumor and normal samples as previously described^{13,44}. Somatic sequence alteration calls for cohort 1 are listed in Supplementary Table 3. Somatic mutations were annotated against the set of mutations in COSMIC (v.84) database, and the number of samples with identical amino acid change were reported. Mutations were characterized as hotspots when the same amino acid change was reported in at least ten tumor samples in COSMIC v.84 database. Missense mutations were evaluated for their potential as cancer drivers by CHASMplus⁴⁵. We performed an unbiased enrichment analysis for sequence and focal copy number alterations across the exome in responding versus nonresponding tumors. All genes or gene groups with a mutation frequency above 5% were included in the differential enrichment analysis.

For the TCGA cohort, WES-derived somatic mutation calls from the TCGA PanCancer Atlas MC3 project were retrieved from the NCI Genomic Data Commons (<https://gdc.cancer.gov/about-data/publications/mc3-2017>)⁴⁶. Mutation calls in cohort 2 were obtained from reanalysis of the original calls¹² and consequence prediction was performed using CRAVAT⁴⁷. TMB scores for the cohort of 1,661 tumors were retrieved from the original publication and refer to the total number of somatic mutations identified normalized to the exonic coverage of the targeted panel used in megabases⁶.

Neoantigen prediction and feature characterization.

To assess the immunogenicity of somatic mutations, exome data combined with each individual patient's MHC class I haplotype were applied in a neoantigen prediction platform as previously described (ImmunoSelect-R, Personal Genome Diagnostics)¹³. For single base substitutions, ImmunoSelect-R performs a comprehensive assessment of paired somatic and wild-type peptides 8–11 amino acids in length at every position surrounding a somatic mutation. In the case of frameshifts, all peptides 8–11 amino acids encompassing the new protein sequence resulting from the frameshift alteration were considered (Supplementary Table 10). To infer germline HLA genotypes, whole-exome sequencing data from paired tumor/normal samples were first aligned to a reference allele set, which was then formulated as an integer linear programming optimization procedure to generate a final genotype by OptiType v.1.0 (ref. 48). The HLA genotype served as input to netMHCpan to predict the MHC class I binding potential of each somatic and wild-type peptide (half maximal inhibitory concentration (IC₅₀) in nM)⁴⁹. Peptides were further evaluated for antigen processing (netCTLpan⁵⁰) and were classified as cytotoxic T lymphocyte epitopes (E) or nonpeptides (NA). Paired somatic and wild-type peptides were assessed for self-similarity based on MHC class I binding affinity⁵¹. Neoantigen candidates meeting an IC₅₀ affinity <5,000 nM were subsequently ranked based on MHC binding and T cell epitope classifications. A single MANA per mutation was selected based on their MHC affinity and neoantigen candidates with an MHC affinity <500 nM were further selected to estimate the neoantigen tumor burden and used for downstream analyses. We further characterized MANAs based on their immunogenic potential by selecting neopeptides with high MHC affinity for which their wild-type counterpart predicted not to bind MHC class I molecules (fit MANA: MHC affinity for mutant peptide <50 nM and for wild-type peptide >1,000 nM). For MANAs stemming from frameshift mutations, we considered the length of the resulting protein until a stop codon was reached, as a longer amino acid sequence would have the

potential to generate more immunogenic neoantigens. We subsequently filtered out sequences more prone to undergo nonsense mediated decay⁵², during this process aberrant transcripts are typically removed at the messenger RNA level and therefore would not stand a chance of occurring despite the presence of bioinformatic predictions.

The percentage of frameshift mutations undergoing nonsense mediated decay is shown in Supplementary Table 10. Frameshift MANAs were interrogated for similarity to microbial and viral antigens by matching the peptide sequence to peptides in the Immune Epitope Database (IEDB, www.iedb.org), requiring a match of >80% for identity and >75% for length.

Mutational signatures.

We extracted mutational signatures based on the fraction of coding point mutations in each of 96 trinucleotide contexts and estimated the contribution of each signature to each tumor sample using the `deconstructSigs` R package⁵³. To evaluate the impact of the total number of observed single base substitutions on detection of a smoking signature within a tumor sample, we performed in silico dilution experiments using somatic mutation data from 985 NSCLC samples from the TCGA PanCancer Atlas MC3 project. A total of 76 tumors (64 lung adenocarcinomas (LUAD) and 12 lung squamous cell carcinomas (LUSC), with average patient pack years of 43.8 and 32.8, respectively) with mutational loads >250 (requiring a minimum 10% MAF and at least four variant supporting reads per mutation) and a detected smoking signature with >75% contribution were diluted in silico by subsampling to lower mutation counts from 5 up to 100. For each round of subsampling, tumor mutations were reevaluated for a smoking signature using the `deconstructSigs` package. We then assessed reductions in the smoking signature and overall percentage deviation from the original smoking signature percentage contribution in the sample.

Copy number analyses, tumor purity and ploidy assessment.

The somatic copy number profile and the extent of aneuploidy in each tumor were estimated using whole-exome sequencing data as follows. First, the relative copy number profile of each tumor sample was determined by evaluating the number of reads mapping to exonic and intronic regions (bins) of the genome while correcting them for confounding factors such as region size, GC content and sequence complexity. The corrected density profile in each tumor sample was then compared to a reference generated by processing a panel of normal samples in a similar manner to define log copy ratio values that reflect the relative copy number profile of each genomic region⁵⁴. Next, circular binary segmentation⁵⁵ (CBS) was applied to bin-level copy ratio values to reduce the inherent noise associated with stochastic read count variation and to enable accurate assessment of copy number breakpoints; that is, boundaries between genomic segments with distinct somatic copy number. Finally, a genome-wide analysis of segmental copy ratio values combined with minor allele frequency of heterozygous single nucleotide polymorphisms (SNPs) overlapping the segments, implemented as an in-house pipeline, yielded an estimate of tumor purity and ploidy. In brief, the model exhaustively evaluated all plausible combinations of tumor purity and ploidy and returned the optimal combination of the two parameters using a maximum likelihood approach. The performance of this platform was

compared against FACETS⁵⁶ on a collection of 97 NSCLC tumors and the two methods provided similar estimates of tumor purity ($r = 0.94$, $P < 2.2 \times 10^{-16}$) and ploidy ($r = 0.66$, $P = 1.489 \times 10^{-13}$). The estimated purity and ploidy of the tumor sample were subsequently used to determine the allele specific copy number of genome segment by selecting the combination of total and minor copy number that best approximate the segment's log copy ratio and average minor allele frequency. For a segment with total copy number (n) and minor copy number (n_B), the expected levels for log copy ratio ($\log R$) and minor allele frequency (M) can be calculated as

$$\log R = \log\left(\frac{n\alpha + 2(1-\alpha)}{\phi\alpha + 2(1-\alpha)}\right)$$

$$M = \frac{n_B\alpha + (1-\alpha)}{n\alpha + 2(1-\alpha)}$$

where α represents tumor purity and ϕ denotes tumor ploidy.

Focal amplifications and homozygous deletions were determined as segments of the genome with length ≥ 3 Mbp and total copy number greater than or equal to three times ploidy of the genome (amplification), or total copy number of zero (deletion, Supplementary Table 6). To increase the specificity of our approach, a set of blacklisted regions was created from a panel of 96 healthy control samples. For each healthy sample, a weighted mean and weighted standard deviation was calculated from segment means obtained from the circular binary segmentation algorithm on copy ratio values, weighted by the number of bins supporting each segment. Genomic intervals in each healthy sample with a segment mean greater than three standard deviations away from the mean were added to the blacklist. Focal alterations where $>50\%$ of the segment overlapped a blacklisted region in at least two healthy control samples were dropped. In addition, segments supported by fewer than five bins and also segments from GC-rich and GC-poor regions of the genome where more than 50% of bins supporting a segment had a GC content of less than 35% or greater than 70% were excluded.

We calculated several measures of tumor aneuploidy including the fraction of the genome with LOH (complete loss of the minor allele) and allelic imbalance (inequality of major and minor allele copy number) (Supplementary Tables 8 and 9). In each tumor sample, the modal copy number was determined as the most prevalent total copy number value across the genome. The fraction of the genome with total copy number different from this modal value was calculated and referred to as a nonmodal copy number fraction. This measure of aneuploidy is equal to zero for a euploid genome, and increases as the tumor genome accumulates copy number aberrations. Finally, we determined the fraction of the genome at each observed total copy number value, and applied the concept of entropy from information theory to quantify the amount of uncertainty in the assignment of total copy number for each genomic segment⁵⁷. Genome copy number entropy is at its minimum when the entire genome is at a single total copy number, and reaches its maximum when all the observed total copy number levels represent equal fractions of the genome; for example, 25% of the genome at $n = 1, 2, 3$ and 4.

For a subset of cases ($n = 14$ in cohort 1 and $n = 11$ in cohort 2) where the pipeline could not determine the purity and ploidy due to low tumor purity, technical noise or copy number heterogeneity, a mutation-based measure of tumor purity based on the median of mutant allele fractions was used to derive an approximate measure of tumor purity. Tumor purity estimates from copy number analysis above were combined with these mutation-based estimates to define the ‘adjusted tumor purity’ measure.

Evaluation of tumor purity in TCGA samples.

Consensus tumor purity estimates from four independent methods were obtained for TCGA samples⁵⁸. We restricted our analysis to 3,788 TCGA samples from seven tumor types (BLCA, BRCA, COAD, HNSCC, KIRC, LUAD, LUSC and SKCM) that had both MC3 mutation calls and a consensus tumor purity estimate. For each cancer type, we computed the Pearson correlation between the total number of mutations called in each sample and tumor purity (Extended Data Fig. 1). Tumor purity for the cohort of 1,661 tumors were retrieved from the original publication⁶.

Mutation clonality assessment.

MAF, ploidy and purity were incorporated to estimate mutation cellular fraction that is the fraction of cancer cells that harbor a specific mutation (Supplementary Table 3). SCHISM⁵⁹ was applied to determine the mutation cellular fraction based on the observed variant allele frequency, estimated copy number and sample purity by following an approach similar to that previously described¹³. Briefly, the expected MAF (V_{exp}) of a mutation with mutation cellular fraction (CF) present in m copies (mutation multiplicity), at a locus with total copy number in the tumor sample (n_T) and total copy number in the matched normal sample (n_N), with purity (α) can be calculated as

$$V_{\text{exp}} = \frac{mCF\alpha}{\alpha n_T + (1 - \alpha)n_N}$$

where m indicates multiplicity; that is, the number of mutant copies present in the cancer cells. A confidence interval for variable V_{exp} can be derived based on the observed distinct mutant counts and distinct coverage assuming a binomial process. Substitution of this value in the above equation resulted in a confidence interval for the product of the two unknown variables (m and CF). Finally, the following set of rules were applied to determine the mutation cellular fraction: (1) For clonal mutations (CF = 1), the product $m \times$ CF only assumes integer values; therefore, if the confidence interval includes an integer value, that value is equal to the multiplicity of the mutation and the mutation is clonal (CF = 1). (2) For mutations where the upper bound of the confidence interval for $m \times$ CF is below 1, multiplicity is assumed to be 1. If the point estimate for CF is within a tolerance threshold (0.25) of 1.0, the mutation is assumed to be clonal and CF is substituted by 1.0. Otherwise, the mutation is deemed subclonal. (3) For mutations where the confidence interval for $m \times$ CF does not encompass an integer number and the entire interval exceeds 1.0, it is plausible to assume a multiplicity greater than 1.0. In this case, the multiplicity is set to smallest integer value such that the confidence value for CF falls within the expected interval of [0, 1]. This procedure results in a point estimate for CF. Similar to (2), if the point estimate is

within a tolerance threshold (0.25) of 1.0, the mutation is assumed to be clonal and CF is substituted for 1.0; otherwise, the mutation is considered subclonal.

As mentioned above, to determine somatic variants and eliminate false positive calls, we employed a MAF threshold of 10%. While, this approach ensures the inclusion of high confidence mutation calls, its stringency may result in an underestimate of subclonal mutations especially in the setting of low tumor purity.

Limitations of TMB assessment.

The impact of tumor purity and intratumoral heterogeneity on the accuracy of TMB estimates was evaluated in a simulation experiment (Fig. 1). The experiment modeled two tumor samples with distinct subclonal composition, and assessed their estimated TMB at tumor purity levels ranging from 20 to 100% in 10% increments. The first simulated tumor with TMB of 265 contained four mutation clusters at cellular fractions 1.00 ($n = 100$), 0.70 ($n = 50$), 0.40 ($n = 40$) and 0.2 ($n = 75$). The second simulated tumor with TMB of 150 contained two mutation clusters at cellular fractions 1.00 ($n = 100$) and 0.50 ($n = 50$). At each level of tumor purity, the following process was repeated in ten replicates to estimate the obsTMB. Distinct coverage (c) of each mutation was determined as:

$$c \sim \Gamma(\beta \mu_C, \beta)$$

where μ_C is the mean distinct coverage of the sample, and was set to set to 200. The rate parameter β determined the variance of base-level coverage in the sample, and was set to 0.013 based on evaluation of coverage distribution in 100 tumor samples. Distinct mutant read count (m) were generated by assuming a draw from a binomial distribution with probability of success set to the expected mutation allele frequency (v_{exp}) given the purity of the tumor sample (α) and cellular fraction of the mutation (CF), assuming absence of somatic copy number alterations at the mutation loci as follows:

$$v_{\text{exp}} = \frac{\alpha \times \text{CF}}{2}$$

$$m \sim \text{binom}(c, v_{\text{exp}})$$

$$\hat{v} = \frac{m}{c}$$

Mutations with simulated distinct coverage $c \geq 10$, distinct mutant read count $m \geq 3$ and observed allele frequency $\hat{v} \geq 10\%$ were determined to be present, and were tallied up to derive the obsTMB. The obsTMB was calculated in each replicate, and the mean was reported (Fig. 1).

Correction of TMB for tumor purity.

We generated cTMB values based on obsTMB and tumor purity as follows. Given our findings that low tumor purity can limit the detection of subclonal mutations and skew the estimates of clonal composition, we first established the level of intratumor heterogeneity in a set of TCGA NSCLC cancers with high tumor purity. Purity, ploidy, and allele specific copy number profiles of the tumor samples based on analysis of SNP6 copy number array

data were obtained from Synapse (<https://www.synapse.org/#!Synapse:syn1710464.2>). A set of 31 NSCLC samples with tumor purity of at least 80% and tumor ploidy in the range of [1.5, 5.0] was selected, where highly confident mutation calls (MC3 set) were available, and somatic copy number profile was determined. To closely match this analysis pipeline to that used for whole-exome samples in the study cohorts, we filtered the mutation calls to those with a minimum MAF of 10% and minimum distinct mutant read count of 3. We estimated the cellular fraction of mutations in each tumor as described above, and determined the fraction of clonal mutations. This analysis revealed a low level of intratumor heterogeneity in untreated lung tumors, as we observed clonal mutation fraction of 70% or above in all but two of the 31 tumors analyzed. Given the small number of lung tumors where the clonal composition could be accurately determined, we identified an additional group of samples to supplement the original set. We identified 704 highly pure (purity $\geq 80\%$) tumors with available mutation and copy number data from the TCGA project in tumor types other than NSCLC, and characterized them in terms of clonal composition. We derived an estimate for the clonal composition of each tumor defined as the frequency of observed mutations in cellular fraction bins of width 0.05 spanning the [0,1] interval, and used these estimates as a basis to model mutation cellular fraction values in the simulation experiment. We further filtered this set to ensure that their level of intratumor heterogeneity matches that of NSCLC tumors by requiring clonal mutation fraction of 70% or above. The clonal composition from this reference combined set of NSCLC ($n = 29$) and other ($n = 577$) tumors with high clonal fraction ($\geq 70\%$) was used to model mutation cellular fraction in the following simulation experiment.

We subsequently simulated 20,000 in silico tumor samples, where the true TMB of each tumor was determined by sampling from the distribution of TMB in TCGA NSCLC samples. The mean sample sequence depth of coverage (C) was set to follow a normal distribution with $\mu = 150$ and $\sigma = 10$. The clonal composition of each tumor was specified by randomly sampling from the reference set. The cancer cell fraction of mutations in each tumor were determined by sampling from a multinomial distribution with p parameters set to match the tumor's clonal composition.

Next, following the approach outlined above, we determined the obsTMB at tumor purity values ranging from 10–100% for each tumor sample. At each level of tumor purity and for each tumor sample, the ratio of true to obsTMB was determined. The median of this ratio across the simulated tumors was considered as a multiplicative correction factor used to transform the obsTMB to a value referred to as cTMB that more closely approximates the true TMB. The median and 95% confidence interval of the correction factor (r) calculated at different levels of tumor purity (α) from the simulation experiment are reported in Supplementary Table 4.

$$cTMB = r(\alpha) \times \text{obsTMB}$$

We applied this approach to the tumor samples in cohort 1 and estimated the cTMB and its 95% confidence interval (Extended Data Fig. 5).

Given the increased application of targeted sequencing to estimate TMB of clinical samples, we performed a similar simulation experiment to delineate the extent to which TMB estimates for targeted sequencing samples would be affected by tumor purity. The higher depth of coverage in targeted sequencing allows detection of mutations at lower allelic fractions while avoiding false positives, which enables more sensitive detection of clonal mutations in low tumor purity samples and subclonal mutations in samples of moderate to high purity. To capture these differences in sensitivity, we applied the following modifications to the outline described above. First, in determining the extent of intratumor heterogeneity, MC3 mutation calls from TCGA were not filtered based on MAF or distinct mutant read count. Second, the set of highly pure TCGA tumors across cancer types were not filtered by requiring the minimum fraction of clonal mutations, but were sampled such that their distribution of clonal mutation fractions matches the distribution observed in all 31 TCGA high purity lung tumors. Third, the mean depth of coverage for in silico tumor samples is found by sampling from the distribution of sample depth of coverage in a recent study of ICB treated patients profiled by targeted NGS⁶. Finally, in analysis of simulated mutation read counts, the minimum MAF required was lowered 2% and the minimum mutant read count to two to determine the mutation as present.

HLA genetic variation.

OptiType v.1.2. was used to determine HLA class I haplotypes⁴⁸. The highly polymorphic nature of the HLA loci limits the accuracy of sequencing read alignment and somatic mutation detection by conventional methods. Therefore, a separate bioinformatic analysis using POLYSOLVER³³ was applied to detect and annotate the somatic mutations in class I HLA genes. HLA class I haplotypes derived from application of OptiType v.1.2 to TCGA RNA-seq samples were retrieved from Genomic Data Commons⁶⁰ (<https://gdc.cancer.gov/about-data/publications/panimmune>). To assess the possibility of loss of germline alleles in tumor, allele specific copy number profiles of the tumor samples from analysis of SNP6 copy number array data were obtained from Synapse (www.synapse.org/#!/Synapse:syn1710464). LOH of each HLA gene was determined by considering the minor allele copy number of the overlapping genomic region (minor copy number = 0 indicated complete loss of minor allele). Individual HLA-I alleles were classified into discrete supertypes, based on similar peptide anchorbinding specificities⁶¹.

Evaluation of somatic HLA loss.

Given the essential role of MHC class I molecules in presentation of neoantigens and initiation of a cascade of events that leads to antitumor immune response, we determined their maintenance or loss in tumor by applying the bioinformatics tool LOHHLA using default program settings³². LOHHLA determines allele specific copy number of HLA locus by realignment of NGS reads to patient-specific HLA reference sequences, and correction of the resulting coverage profile for tumor purity and ploidy. At each HLA locus heterozygous in germline, LOH was declared if the copy number for one of the two alleles was below 0.5, and there was a statistically significant difference between the log copy ratio of the two alleles ($P_{\text{Val_unique}} < 0.01$). The unique number of class I HLA alleles in tumor was calculated by subtracting the number of germline heterozygous alleles with somatic LOH from the total number of unique alleles in germline.

TCR sequencing.

TCR clones were evaluated in tumor tissue by NGS. DNA from tumor samples was isolated by using the Qiagen DNA FFPE kit. TCR- β CDR3 regions were amplified using the survey ImmunoSeq assay in a multiplex PCR method using 45 forward primers specific to TCR V β gene segments and 13 reverse primers specific to TCR J β gene segments (Adaptive Biotechnologies)⁶². Productive TCR sequences were further analyzed. For each sample, a clonality metric was estimated to quantitate the extent of mono- or oligo-clonal expansion by measuring the shape of the clone frequency distribution. Clonality values range from 0 to 1, where values approaching 1 indicate a nearly monoclonal population (Supplementary Table 13).

Immunohistochemistry and Interpretation of PD-L1 and CD8 staining.

Immunolabeling for CD8-PD-L1 dual detection was performed on formalin-fixed, paraffin embedded sections on a Ventana Discovery Ultra autostainer (Roche Diagnostics). Briefly, following deparaffinization and rehydration, epitope retrieval was performed using Ventana Ultra CC1 buffer (Roche Diagnostics) at 96 °C for 64 min. Sections were subsequently incubated with the primary mouse antihuman CD8 antibody, (1:100 dilution, clone m7103, Dako) at 36 °C for 60 min, followed by incubation with an antimouse HQ detection system (Roche Diagnostics) and application of the Chromomax DAB IHC detection kit (Roche Diagnostics). Following CD8 detection, primary and secondary antibodies from the first round of staining were stripped on board using Ventana Ultra CC2 buffer at 93 °C for 8 min. An anti-PD-L1 primary antibody (1:100 dilution; E1L3N clone, Cell Signaling Technologies) was applied at 36 °C for 20 min. PD-L1 primary antibodies were detected using an antirabbit HQ detection system (Roche Diagnostics) followed by Discovery Purple IHC detection kit (Roche Diagnostics), counterstaining with Mayers hematoxylin, dehydration and mounting. A minimum of 100 tumor cells were evaluated per specimen; only membranous staining was considered specific and further interpreted. PD-L1 protein expression was evaluated based on the intensity of staining on a 0 to 3+ scale, and the percentage of immune-reactive tumor cells. CD8-positive lymphocyte density was evaluated per $\times 20$ high power field.

Statistics and reproducibility.

The experiments described in this study were not randomized and the investigators were blinded to clinical outcomes during experiments. No statistical method was used to predetermine sample size. Fifteen cases, where excluded due to tumor purity <10% or absence of a matched normal sample. Differences between responding and nonresponding tumors were evaluated using chi-squared or Fishers exact test for categorical variables and the Mann–Whitney test for continuous variables. The Pearson correlation coefficient (R) was used to assess correlations between continuous variables and the Spearman rho coefficient was calculated for nonparametric correlations. Analysis of variance and Kruskal–Wallis one-way analysis of variance were performed to assess differences in continuous features among groups. P values were corrected using the Benjamini–Hochberg procedure and the associated FDR values were calculated. Tumors were classified based on their nonsynonymous sequence alteration load in high and low mutators, using the second tertile

as a cut-off point. The median point estimate and 95% CI for PFS and OS were estimated by the Kaplan-Meier method and survival curves were compared by using the nonparametric log-rank test. Univariate Cox proportional hazards regression analysis was used to determine the impact of individual parameters on OS. A multivariable Cox proportional hazards model was employed using cTMB, RTK mutations, smoking mutational signature and number of HLA germline alleles for cohort 1. A risk score reflecting the relative hazard was calculated as the exponential of the sum of the product of mean-centered covariate values and their corresponding coefficient estimates for each case. The second tertile of the risk score was used to classify patients in high risk (top 33.3%) and low risk (bottom 66.6%) groups. For validation of the model performance, a Cox proportional hazards model was trained on cohort 1 and the coefficients were estimated. This trained model was then applied to samples in cohort 2 and assessed risk score was calculated as described above. All *P* values were based on two-sided testing and differences were considered significant at *P* < 0.05. Statistical analyses were done using the SPSS software program (v.25.0.0 for Windows) and R v.3.2 and higher, (<http://www.R-project.org/>).

Reporting Summary.

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

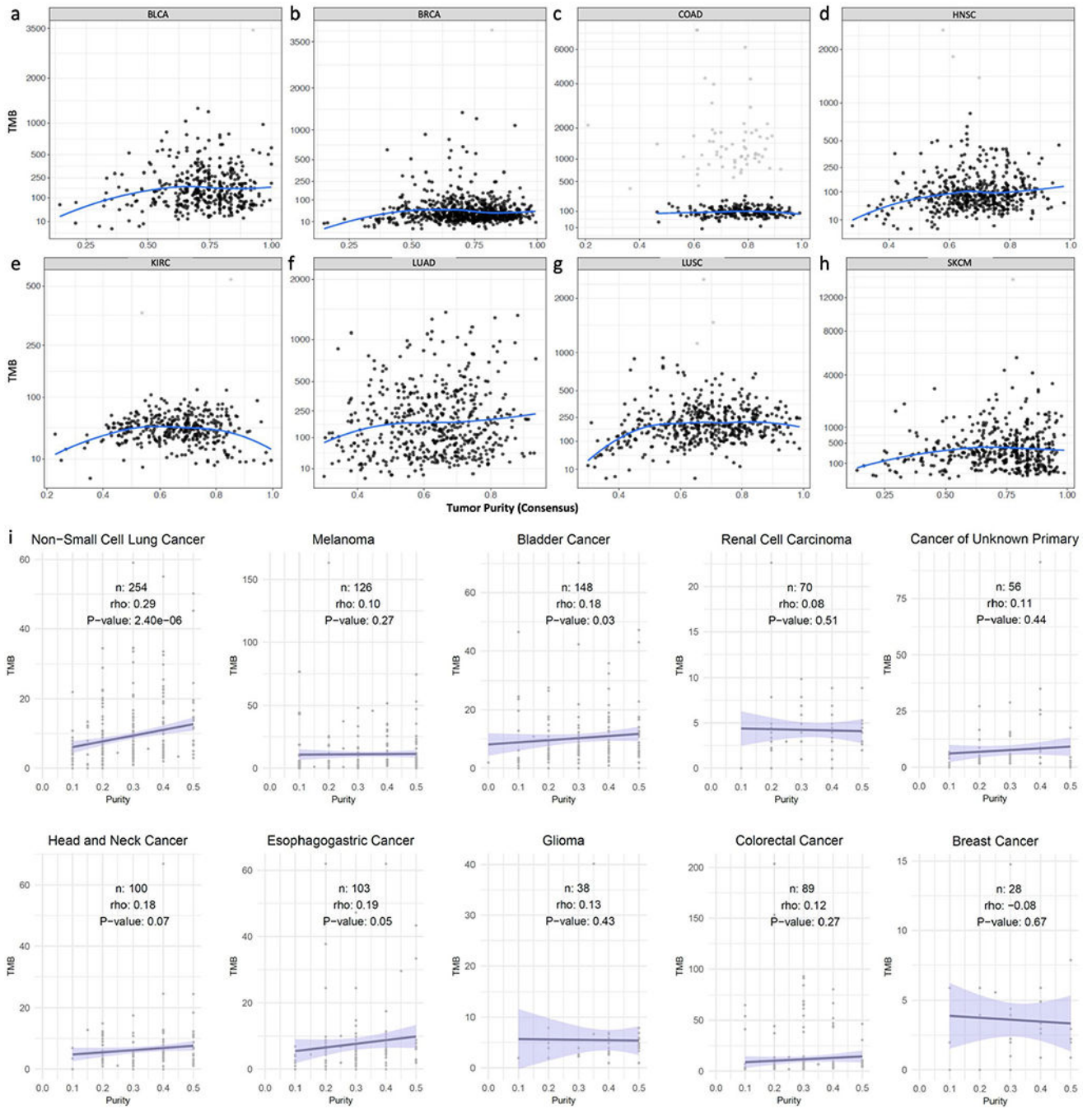
Whole-exome sequencing data that support the findings of this study have been deposited in the database of Genotypes and Phenotypes (dbGaP) and the European Genome-phenome Archive under accession codes [phs001940.v1.p1](#) and [EGAS00001003892](#), respectively (cohort 1). Source data for the TCGA tumor samples were retrieved from <http://cancergenome.nih.gov>. WES-derived somatic mutation calls from the TCGA PanCancer Atlas MC3 project were retrieved from the NCI Genomic Data Commons⁴⁶ (<https://gdc.cancer.gov/about-data/publications/mc3-2017>). Previously published genomic data, reanalyzed here, were obtained from dbGaP under accession code [phs000980.v1.p1](#) (cohort 2) and from the CBioPortal for Cancer Genomics^{2,6} (https://www.cbioportal.org/study/summary?id=nsclc_pd1_msk_2018 and https://www.cbioportal.org/study/summary?id=tmb_mskcc_2018). All other data supporting the findings of this study are available from the corresponding author upon reasonable request.

Code availability

Somatic mutations were identified using the VariantDx software⁴⁴. Missense mutations were evaluated for their potential as cancer drivers by CHASMplus (v.1.0.0)⁴⁵ with consequence prediction using CRAVAT (v.4.3)⁴⁷. HLA genotyping was performed using OptiType (v.1.2)⁴⁸. MHC class I prediction of neoantigens used netMHCpan⁴⁹ and netCTLpan⁵⁰. Mutational signatures were quantified using the deconstructSigs R package (v.1.8.0). Somatic copy number profiles were estimated using CNV-kit⁵⁴, FACETS (v.0.5.0)⁵⁶ and as previously described¹³. POLYSOLVER was applied to detect somatic mutations in class I HLA genes³³. LOHHLA was applied to determine loss of HLA genes³². Further statistical

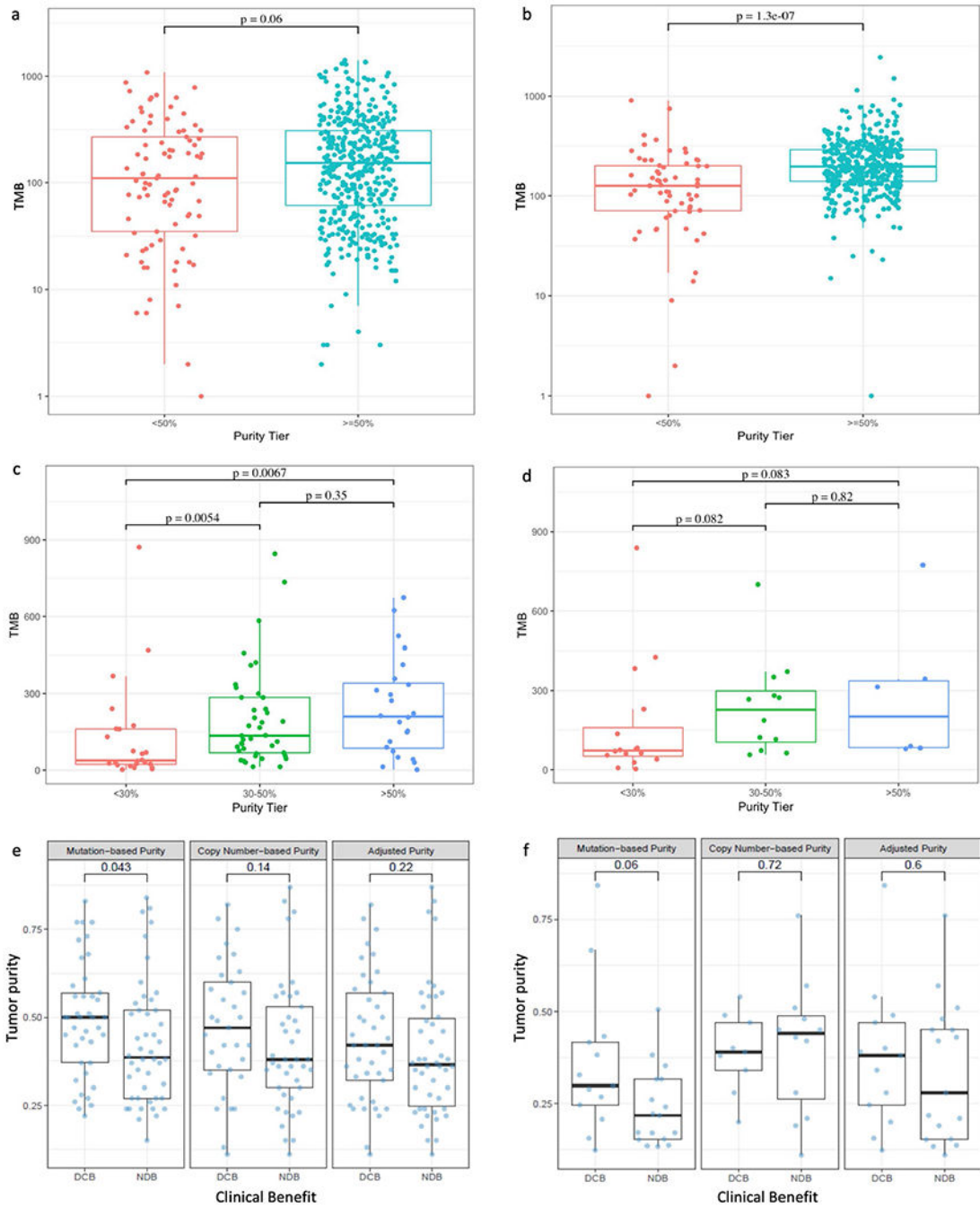
analyses were performed using the SPSS software program (v.25.0.0) and R v.3.2 and higher.

Extended Data



Extended Data Fig. 1 | Correlations between tumor purity and tumor mutation burden estimates from whole exome and targeted next-generation sequencing. (a-h) Paired tumor-normal whole exome sequencing data from 3,788 TCGA samples from 8 tumor types (BLCA, N=405 tumors, BRCA, N=778 tumors, COAD, N=287 tumors,

HNSCC, N=504 tumors, KIRC, N=367 tumors, LUAD, N=508 tumors, LUSC, N=477 tumors, SKCM, N=462 tumors) for which immune targeted approaches have shown promising clinical benefit were analyzed and revealed a positive correlation between obsTMB and tumor purity in bladder cancer (Pearson's $R=0.083$, $p=0.095$), lung adenocarcinoma (Pearson's $R=0.092$, $p=0.038$), and lung squamous cell carcinoma (Pearson's $R=0.089$, $p=0.051$). A loess model was fitted to the mutation sequence data for each tumor type. For colon adenocarcinoma, tumor samples with a hypermutator genotype are shown in gray and were excluded from the model fit. Pearson's correlation coefficient (R) was used to assess correlations between continuous variables and p values are based on two-sided testing. (i) TMB scores derived from targeted sequencing and tumor purity assessments were retrieved from a published cohort of 1,661 tumors treated with immune checkpoint blockade (Samstein et al., *Nat Genet*, 2019) and non-parametric correlations were evaluated by the Spearman's rho statistic. A significant correlation between obsTMB and tumor purity was identified for NSCLC (N=254 tumors, Spearman $\rho=0.29$, $p=2.4 \times 10^{-6}$), bladder cancer (N=148 tumors, Spearman $\rho=0.18$, $p=0.03$), esophagogastric cancer (N=103 tumors, Spearman $\rho=0.19$, $p=0.05$) and head and neck cancer (N=100 tumors, Spearman $\rho=0.18$, $p=0.07$). The purple solid line indicates a linear regression fit and the purple shaded area represents the estimated 95% confidence interval of the probability density of TMB at each level of purity.



Extended Data Fig. 2 | Tumor purity correlates with TMB estimates in the immunotherapy treated cohorts.

The correlations of tumor purity with tumor mutational burden and clinical response in 957 TCGA NSCLC samples and in immunotherapy cohorts 1 and 2 are shown in a-f. (a-b) TCGA lung adenocarcinomas-LUAD (N=508 tumors; a) and lung squamous cell carcinomas-LUSC (N=477 tumors; b) with a higher degree of normal contamination had a significantly lower TMB compared to tumors with a tumor purity > 50% (Mann-Whitney U test $p=0.06$ and $p=1.3e^{-7}$ for LUAD and LUSC respectively). (c-d) In cohorts 1 and 2, the

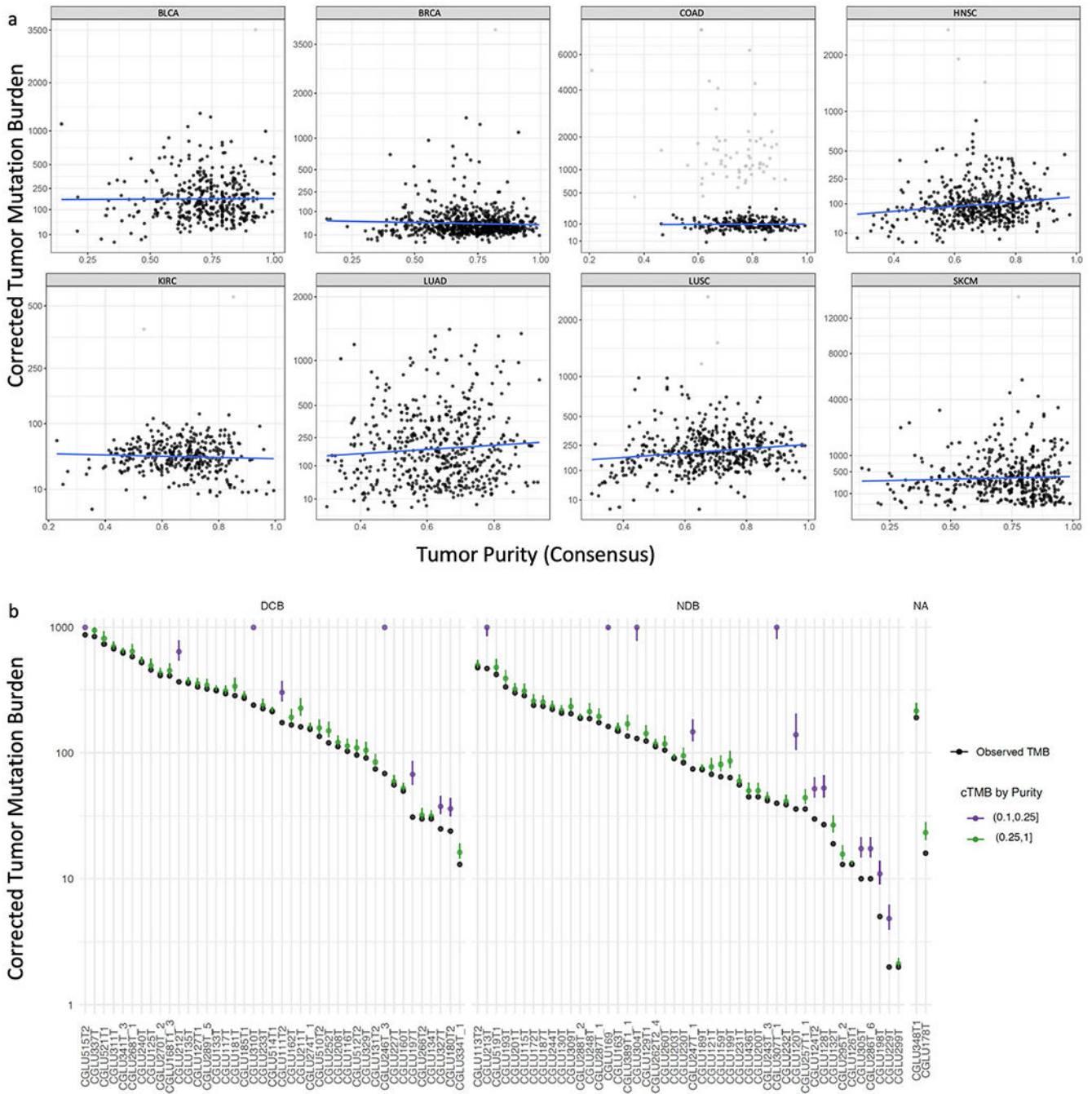
correlation between obsTMB and tumor purity was particularly pronounced for tumor purity less than 30% (Mann Whitney U test $p=0.008$ and $p=0.08$ for overall comparisons of TMB across tumor purity tiers for cohort 1 (N=87 tumors; c) and cohort 2 respectively (N=30 tumors; d). (e-f) Tumor purity was associated with clinical benefit from ICB when mutation-based purity was used, which is most likely attributed to the contribution of TMB in the mutation-based purity calculation; however no difference in tumor purity was found between responding and non-responding tumors when copy-number based tumor purity (N=74 tumors) and adjusted tumor purity (N=87 tumors) was used in cohort 1 (Mann Whitney U test $p=0.14$ and Mann Whitney U test $p=0.22$ respectively; e) and cohort 2 (Mann Whitney U test $p=0.72$ and Mann Whitney U test $p=0.6$ respectively; f). Seventy four and twenty one tumor samples from cohort 1 and 2 respectively had evaluable copy-number based tumor purity (Methods). The vertical line in each whisker plot signifies the median and the ends the upper and lower quartile, while the whiskers extend 1.5 times the interquartile range from the hinges. P values are based on two-sided testing.

Author Manuscript

Author Manuscript

Author Manuscript

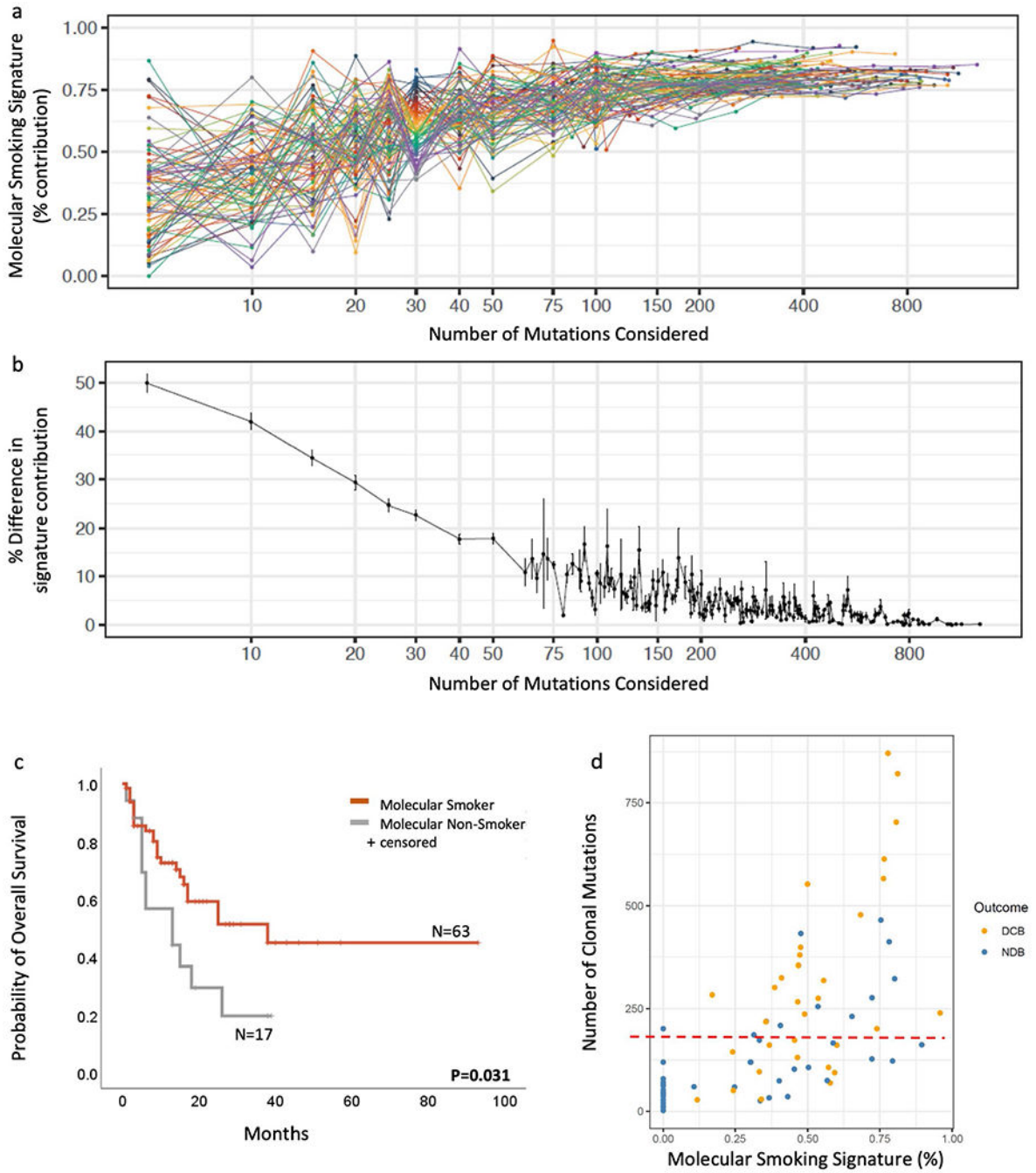
Author Manuscript



Extended Data Fig. 3 | Effect of TMB correction for the TCGA and immunotherapy-treated cohorts.

(a) Corrected TMB was replotted with tumor purity for TCGA samples to visualize whether our approach generated overcorrected TMB values on the lower end of tumor purity (N=3,788 tumors, BLCA N=405 tumors, BRCA N=778 tumors, COAD N=287 tumors, HNSCC N=504 tumors, KIRC N=367 tumors, LUAD N=508 tumors, LUSC N=477 tumors, SKCM N=462 tumors). We did not observe such phenomenon for all tumor types evaluated, suggesting that our correction factors did not generate high TMB-low tumor purity outliers. Additionally, we found that after using our TMB correction approach, there was now no

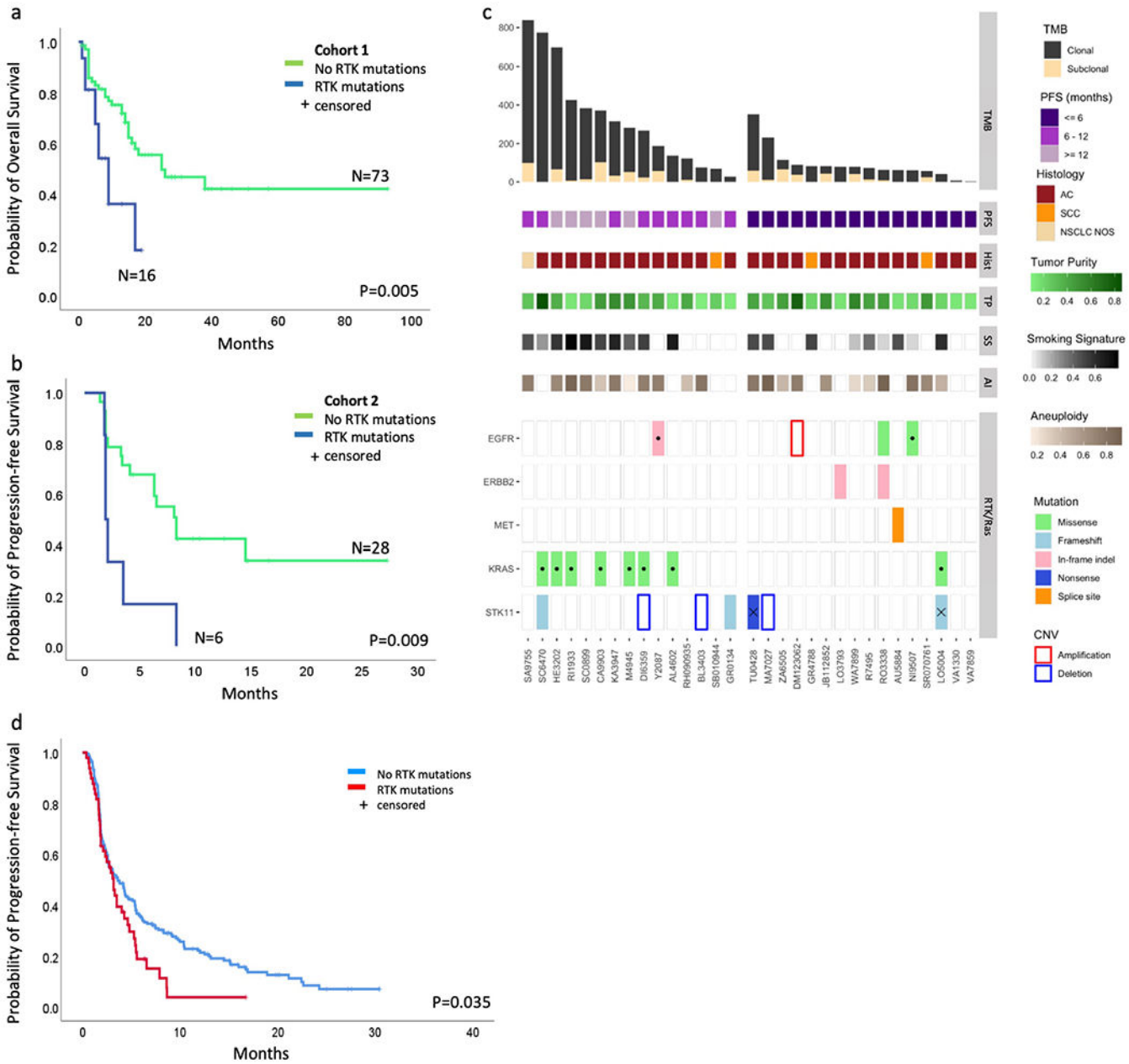
longer a positive correlation between cTMB and tumor purity (BLCA Pearson R 0.023 p=0.64, BRCA R=-0.013 p=0.71, COAD R=0.019 p=0.0.77, HNSCC R=0.031 p=0.48, KIRC R=0.015 p=0.78, LUAD R=0.071 p=0.11, LUSC R=0.069 p=0.13, SKCM R=0.048 p=0.31). Blue line indicates a linear fit. The Pearson correlation coefficient (R) was used to assess correlations between continuous variables and p values are based on two-sided testing. **(b)** Effect of TMB correction for tumor purity in cohort 1. Distribution of observed (black circles) and corrected TMB for patients in cohort 1 (N=89 patients) are shown for each tumor purity tier. Corrected TMB values are denoted by purple circles for tumor purity 0.1-0.25 and green circles for tumor purity >0.25, error bars represent 95% confidence intervals. cTMB values are capped at 1000. The second tertile of cTMB was used as a cut-off point to classify tumors as high vs low TMB categories in order to avoid bias related to multiple cutoff point selection. After correcting for tumor purity in cohort 1, ten patients were reclassified in TMB categories. As would be expected with lower tumor purity, there were four patients with tumor purity <20% that switched class from the low observed TMB to the high corrected TMB category but there was also one patient with tumor purity >20% that switched from low observed TMB to high corrected TMB category and also five patients with higher tumor purity (range 46-82%) that switched class from the high observed TMB to the low corrected TMB category. Switching between both TMB categories would be expected as our approach generates corrected TMB estimates that better capture the true TMB distribution and that in turn would affect the percentile threshold that defines high TMB more accurately. NA; radiographic response non evaluable.



Extended Data Fig. 4 | Impact of molecular smoking signature on outcome for immune checkpoint blockade and association with TMB.

(a-b) *In silico* dilution experiments of single base substitutions to evaluate the power to accurately determine contribution of a dominant mutation signature. Mutation signature analyses were performed on whole exome data from 985 NSCLC tumors (N=508 lung adenocarcinomas and N=477 squamous cell carcinomas) obtained through TCGA. Seventy-six NSCLC tumors (N=64 lung adenocarcinomas and N=12 squamous cell carcinomas) had a tumor mutation load >250 and a molecular smoking signature >75% and were further selected for an *in silico* dilution series. (a) Mutation counts were diluted from maximum

count to a minimum of 5 using random resampling (3 re-sampling replicates per dilution level), to evaluate consistency and divergence in the predicted presence of a smoking signature. Connected points in (a) indicate mean values per case per dilution series. On average, 20 mutations were sufficient to predict the presence of a smoking signature at a 50% level. (b) Mutational load below 20 mutations lead to a 30% difference from the original contribution of the C>A transversion rich signature value and therefore represents a threshold beyond which, there is a significant deviation from accurately determining a dominant mutation signature. Center values indicate the mean values per dilution level and error bars indicate standard error of the mean. (c) Patients with a molecular smoking mutational signature derived durable benefit from immune checkpoint blockade (N=80, log rank $p=0.031$). TMB did not fully explain the association between molecular smoking signature and ICB response. (d) The number of clonal mutations was plotted against the percent contribution of the molecular smoking signature for each tumor (N=74 tumors in total, N=33 responding tumors and N=41 non-responding tumors). Each dot represents a tumor and each tumor is color coded by response, with responding tumors showed in orange and non-responding tumors showed in blue. While there is a correlation between the number of clonal mutations and a dominant molecular smoking signature, there are tumors with a low clonal TMB (dotted red line indicates a TMB threshold used to classify TMB in high/low categories) but a dominant molecular smoking signature which derive benefit from immune checkpoint blockade. The median point estimate and 95% confidence intervals for survival were estimated by the Kaplan-Meier method and survival curves were compared by using the nonparametric log rank test. Log rank p values are based on two-sided testing.



Extended Data Fig. 5 | Impact of RTK mutations on outcome and genomic drivers associated with response to immune checkpoint blockade in cohort 2.

(a-b) Activating RTK mutations identified a group of patients with dismal prognosis in both cohort 1 (N=89 patients, log rank p=0.005) and cohort 2 (N=34 patients, log rank p=0.009).

(c) In cohort 2 (N=31), responding tumors had a higher total and clonal TMB compared to non-responding tumors (Mann Whitney U test p=0.006 and p=0.004 respectively).

Progression-free survival, histology and tumor purity are shown as separate panels. Patients with a clinical response had a higher contribution of the molecular smoking signature (N=15 patients, Mann Whitney U test p=0.054). There were no differences in tumor aneuploidy between responders and non-responders (N=74 patients, Mann Whitney U test p=0.72). A significant enrichment in RTK activating mutations, including point mutations and

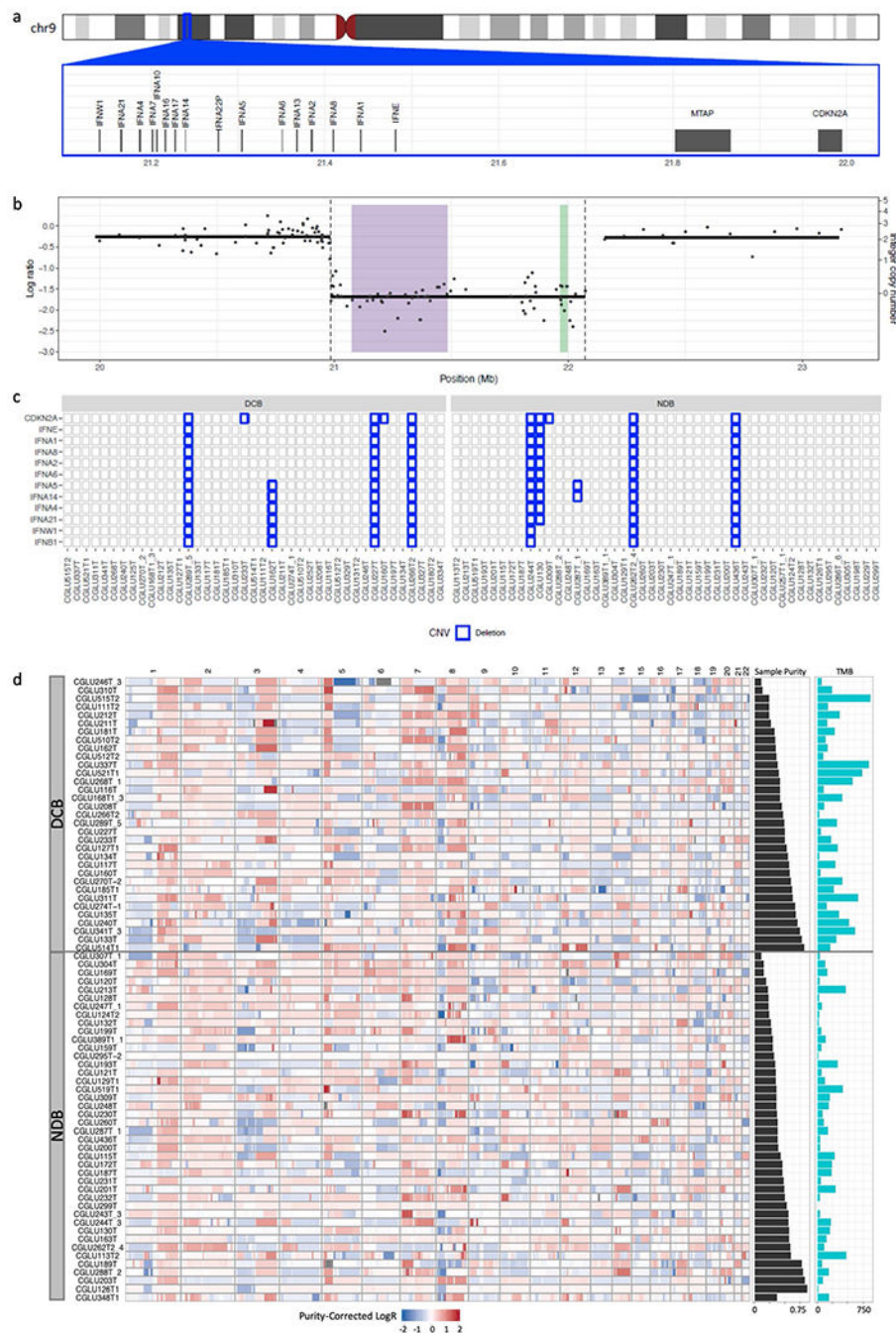
amplifications in *EGFR*, amplifications in *ERBB2* and *MET* exon 14 skipping, was found in non-responding tumors (N=89 total patients, chi-squared p=0.056). (d) A third cohort of NSCLC patients treated with ICB was obtained from CBioportal (N=240 patients) as described in detail in the Methods section; for this cohort, sequence and copy number alterations, as well as outcome information were publicly available. Patients with activating RTK mutations in *EGFR*, *ERBB2*, *MET*, *FGFR1* and *IGF1R* had a significantly shorter progression-free survival (log rank p=0.035). We defined response as durable clinical benefit if complete, partial response or stable disease was achieved with a duration >6months. The median point estimate and 95% confidence intervals for survival were estimated by the Kaplan-Meier method and survival curves were compared by using the nonparametric log rank test. CNV; copy number variation. cTMB; corrected TMB, RTK; receptor tyrosine kinase, Hist; histotype, TP; tumor purity, SS; smoking signature, AI; allelic imbalance. Dots indicate hotspot mutations, and × denotes loss of heterozygosity of the wild type allele. P values are based on two-sided testing.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



Extended Data Fig. 6 | Copy number analyses for NSCLC tumors in cohort 1.

Given that deletions in IFN- γ genes have been described as a potential mechanism of intrinsic resistance to immunotherapy, we investigated whether there is an enrichment in IFN- γ related gene copy number variation in non-responding tumors. **(a)** A cluster of IFN- γ related genes (*IFNE*, *IFNA1*, *IFNA2*, *IFNA4*, *IFNA5*, *IFNA6*, *IFNA8*, *IFNA14*, *IFNA21*, *IFNB1*) and *IFNB1* is located on chromosome 9 (p21.3), in close proximity to the *CDKN2A* locus. **(b)** The locus that contains both the IFN- γ related genes and *CDKN2A* was frequently found to be deleted; an example of such homozygous deletion is shown for

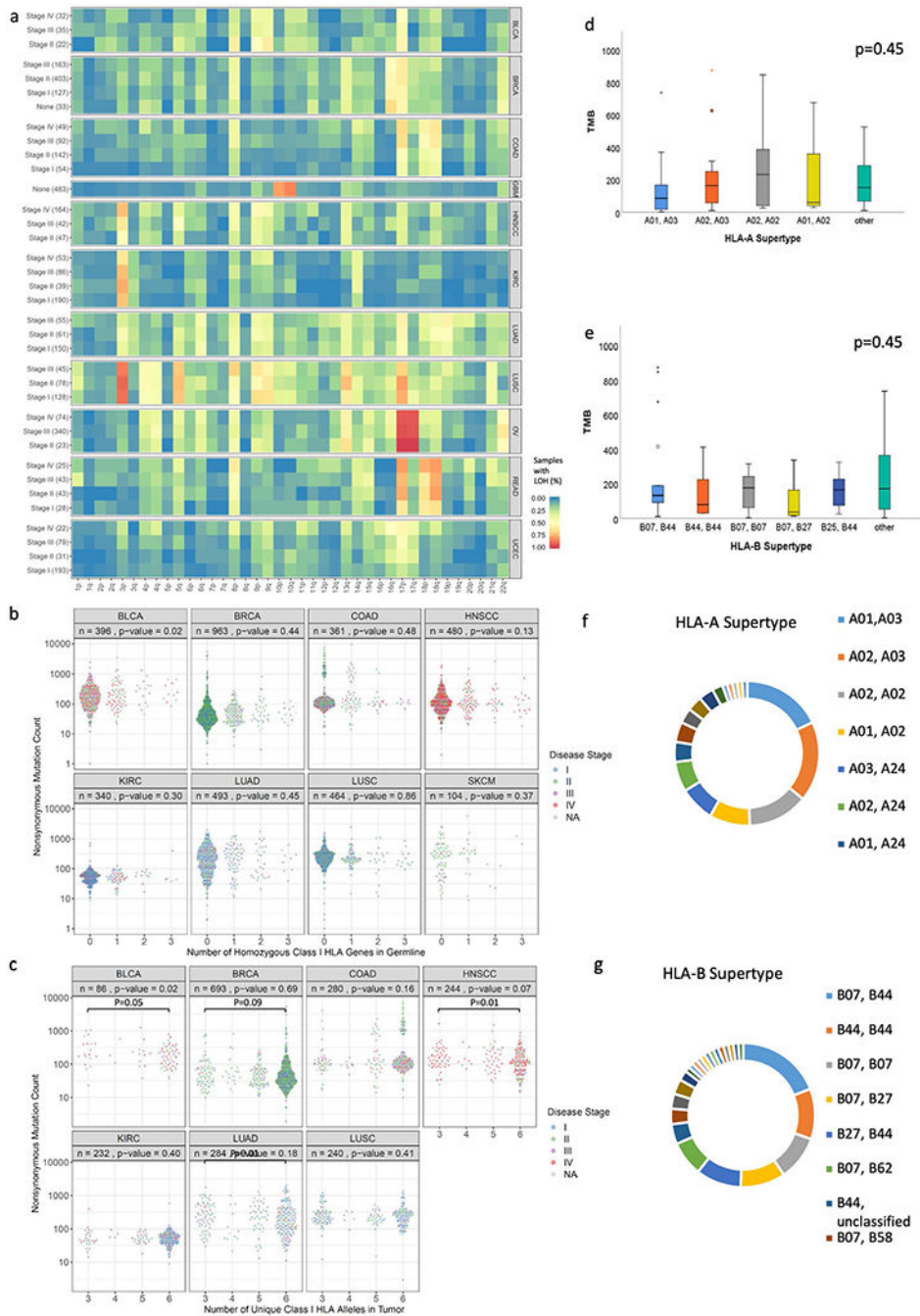
case CGLU262. The vertical axes denote the relative copy ratio (\log_2 scale), and the integer copy number levels assigned to genomic bins (circles) and segments. Purple and green boxes mark the coordinates of IFN gene cluster and *CDKN2A*, respectively. (c) The frequency of homozygous deletions in *IFNE*, *IFNA1*, *IFNA2*, *IFNA4*, *IFNA5*, *IFNA6*, *IFNA8*, *IFNA14*, *IFNA21*, *IFNW1* and *IFNB1* was similar in responding and non-responding tumors and more importantly, these deletions co-occurred with *CDKN2A* loss in 86% of the cases, whereas *CDKN2A* deletions also occurred independently. Given that, *CDKN2A* and the group of IFN- γ pathway genes lie on chromosome 9p within a span of 917 Kb, IFN- γ deletions may be passengers in the setting of a driver *CDKN2A* deletion. Seventy five tumors had evaluable copy number estimates (Methods). (d) A genome-wide analysis of copy number profiles revealed genomic regions with copy number gains and losses and was used to determine the extent of tumor aneuploidy. The relative copy ratio (LogR) values quantifying the abundance of each genomic region compared to the genome average (ploidy) are shown after correction for tumor purity in responding (N=33 tumors) and non-responding tumors (N=41 tumors). Red and blue shades indicate copy gains and losses, respectively, whereas white marks copy neutral regions. There was no significant difference in the degree of aneuploidy assessed by the fraction of genome with allelic imbalance between the two groups (Mann Whitney U test $p=0.367$, FDR-corrected $p=0.65$). We defined response as durable clinical benefit (DCB) if complete, partial response or stable disease was achieved with a duration >6months. P values are based on two-sided testing. CAN; copy number aberration, NDB; non durable clinical benefit.



Extended Data Fig. 7 | Pathway enrichment analysis for DNA damage repair genes and the WNT pathway and distribution of hotspot mutations and associated potentially immunogenic MANAs in cohort 1.

(a) We investigated co-occurrence of mutations in DNA damage repair genes involved in base excision repair (DDR-BER), DNA damage sensing (DS), the Fanconi anemia pathway (FA), homologous recombination (DDR-HR), mismatch repair (DDR-MMR), nucleotide excision repair (DDR-NER), non-homologous end joining (DDR-NHEJ) and translesion DNA synthesis (DDR-TLS). Mutations were characterized by consequence (missense, frameshift, nonsense, splice site, in-frame) and recurrence (hotspots) and loss of

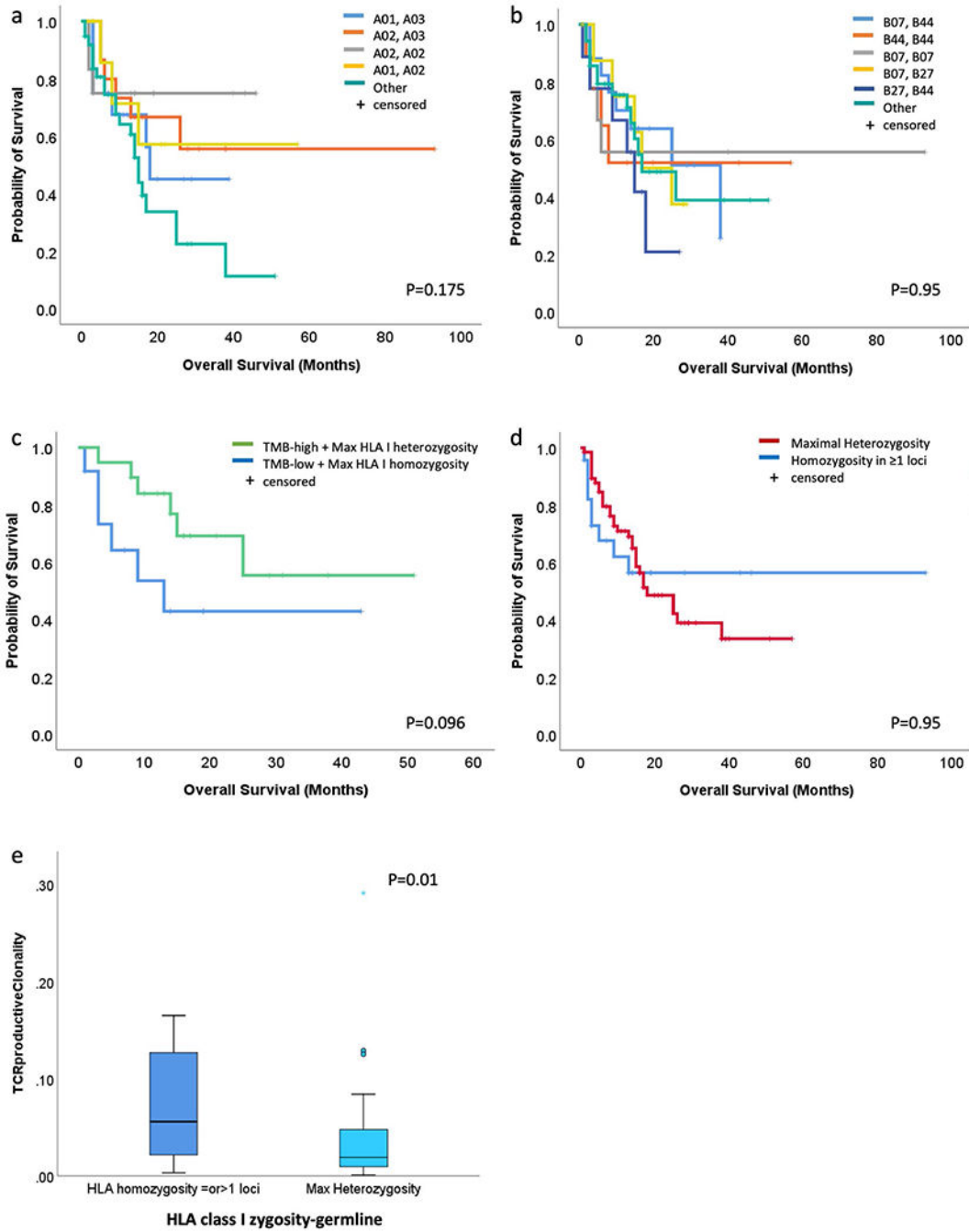
the wild type allele was considered in case of truncating mutations (biallelic inactivation). A similar analysis was performed for genes involved in the WNT pathway. A high TMB tumor with biallelic inactivation of *MLH1* and a tumor with a gain-of-function beta-catenin hotspot mutation were identified among responders (N=41 patients) and non-responders (N=46 patients) respectively, with no additional significant differences in genomic alterations in the DDR and WNT pathways between responders and non-responders. We defined response as durable clinical benefit (DCB) if complete, partial response or stable disease was achieved with a duration >6months. **(b)** The number of mutations with at least one fit MANA (determined as neopeptides with a predicted MHC affinity < 50nM for which the wild type peptides has a predicted MHC affinity of > 1000nM) in each tumor, divided by clonality and hotspot status is shown in the top distribution graph. Clinical response and overall survival are shown in the middle panel. Clonal hotspot frameshifts and in-frame insertions and deletions in *ANTRX2*, *TP53*, *EGFR*, *ASXL1*, *NOTCH2*, *ZFP36L2*, *FAM171B*, *SLC35F5*, *CD93* and *SLAMF1* generated fit MANAs shown in the lower panel. There was no difference in the fraction of clonal fit MANAs between responding (N= 33 tumors) and non-responding (N=41 tumors, Mann Whitney U test p=0.65) tumors. We defined response as durable clinical benefit (DCB) if complete, partial response or stable disease was achieved with a duration >6months. NDB: No durable benefit.



Extended Data Fig. 8 | HLA genetic variation and TMB correlations.

(a) We investigated whether there is an enrichment for chromosome 6p which contains the HLA class I loci for 3,767 tumors from TCGA (BLCA N=89 tumors, BRCA N=737 tumors, COAD N=348 tumors, GBM N=483 tumors, HNSCC N=269 tumors, KIRC N=370 tumors, LUAD N=287 tumors, LUSC N=257 tumors, OV N=456 tumors, READ N=146 tumors, UCEC N=325 tumors). LOH events in NSCLC were compared to the background arm-level allelic imbalance of the same tumor type and across tumor types. Chromosome 6p losses were not more frequent compared to other chromosomal arm level deletions (on the contrary

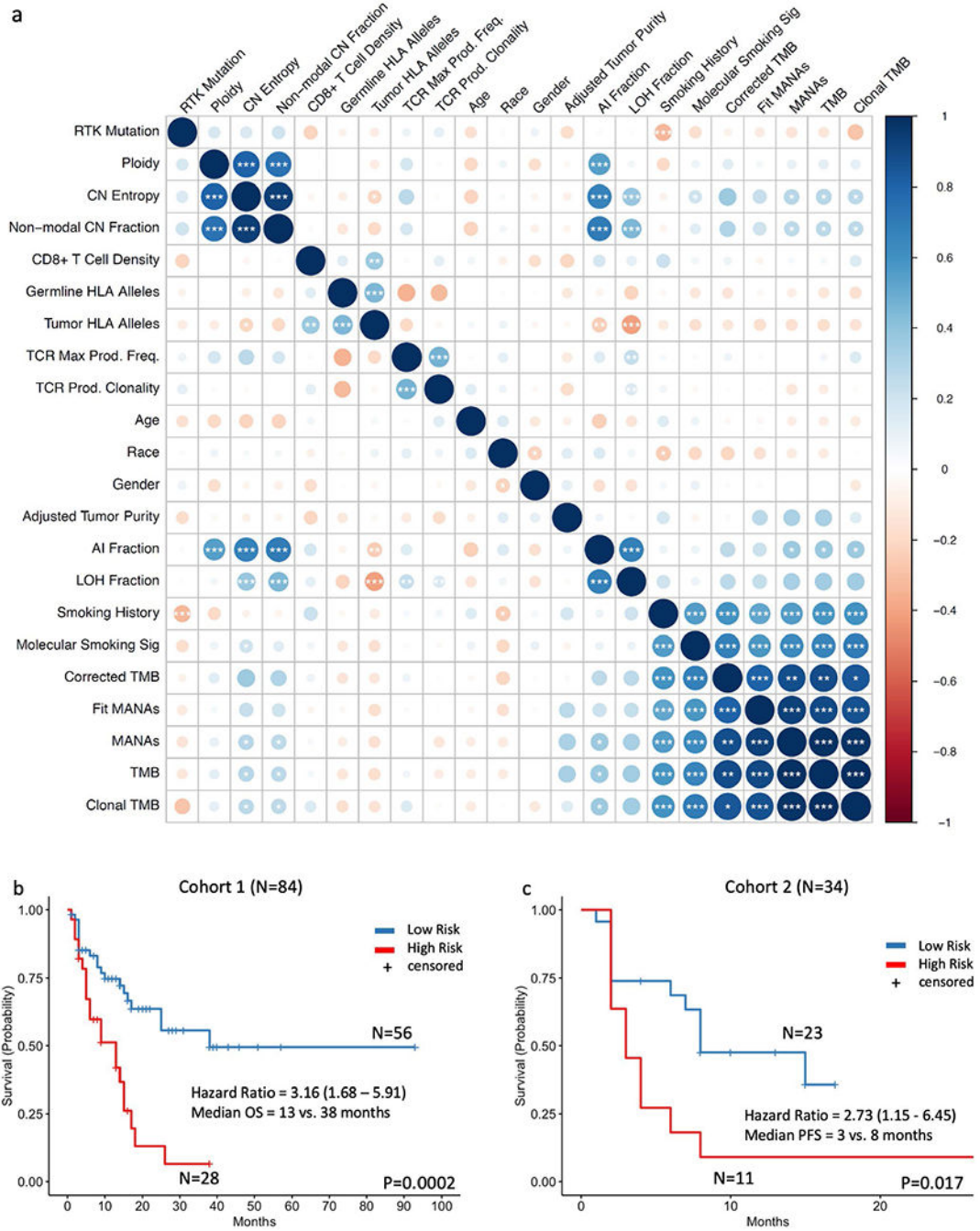
the degree of chromosome 6p LOH was lower compared to other chromosomal arms deletions in lung tumors, N=544 tumors, Fisher's exact $p=0.037$). In contrast, when chromosome 6p LOH events were compared between lung tumors and 9 tumor types (BLCA, BRCA, COAD, GBM, HNSCC, KIRC, OV, READ and UCEC), we found that LOH events involving chromosome 6p that contains the HLA class I loci are more frequent in lung cancer (NSCLC N=533, other tumors N= 3223, 17.3% vs. 8.2%, Fisher's exact $p=6.8e^{-10}$), without any evidence for positive selection of these events in advanced stage disease. Analysis of variance (ANOVA) was applied to assess the correlation of germline homozygosity in HLA class I genes with tumor mutation burden in 6 tumor types (BLCA, BRCA, COAD, HNSCC, KIRC, LUAD and LUSC, total N=3,601). **(b)** Germline HLA zygosity was not correlated with TMB for the vast majority of tumors examined with the exception of bladder cancer (N=396 tumors, ANOVA $p=0.02$). **(c)** Germline and tumor HLA class I status was combined to determine the number of unique HLA class I alleles in each tumor. The number of unique HLA class I alleles appeared to correlate with TMB such that tumors with a higher number of unique HLA class I alleles in the tumor harbored a lower non-synonymous mutation load for BLCA (N=86 tumors, ANOVA $p=0.02$) and HNSCC (N=244, ANOVA $p=0.07$). When tumors heterozygous for all three HLA class I loci (6 HLA class I alleles) were compared to tumors that were homozygous in all three HLA class I loci (3 HLA class I alleles), a lower TMB was noted in the tumors with the more intact antigen presentation capacity (N=86 tumors, Wilcoxon $p=0.05$ for BLCA, N=693 tumors $p=0.09$ for BRCA, N=244 tumors $p=0.01$ for HNSCC, N=284 tumors $p=0.01$ for LUAD). **(d-g)** HLA class I distribution by supertype and association with TMB. Individual HLA-I alleles were classified into discrete superotypes, based upon similar peptide-anchor-binding specificities. **(d)** TMB did not differ among different HLA-A superotypes (N=89 patients, Kruskal-Wallis $p=0.45$). **(e)** Similarly, there were no differences in TMB among different HLA-B superotypes (N=89 patients, Kruskal-Wallis $p=0.45$). **(f)** HLA-A supertype distribution for cases in cohort 1. **(g)** HLA-B supertype distribution for cases in cohort 1. Center values of the box plots in (d) and (e) indicate median values and error bars denote 95% confidence intervals. P values are based on two-sided testing.



Extended Data Fig. 9 | Impact of HLA genetic variation on survival.

(a-d) Association of HLA class I haplotypes and outcome. There was no association between HLA class I haplotypes and overall survival (N=89 patients, log rank p=0.175; a). The same observations held true for HLA-B haplotypes (N=89 patients, log rank p=0.95; b). (c) There was a trend towards longer overall survival for TMB high tumors with maximal germline HLA class I heterozygosity compared to tumors with low TMB and maximal germline HLA class I homozygosity (N=31 patients, log rank p=0.096). (d) Maximal germline heterozygosity vs. homozygosity with more than one HLA loci was not associated

with outcome (N=89 patients, log rank p=0.95). (e) Cases with maximal germline HLA class I heterozygosity were found to have a less clonal TCR repertoire (N=60 patients, Kruskal-Wallis p=0.01). Center values of the box plots in (e) indicate median values and error bars denote 95% confidence intervals. P values are based on two-sided testing.



Extended Data Fig. 10 | Correlations among genomic and immune features and multivariable model for prediction of outcome to immune checkpoint blockade, excluding patients on chemotherapy and tumors collected at the time of resistance in cohort 1.

(a) Relationships among different features were assessed by the Spearman's rho statistic for cohort 1 (N=89 patients) and p values were corrected for multiple comparisons. Four clusters of parameters were identified, related to RTK mutations, HLA genetic variation, tumor aneuploidy and TMB. The color of each dot refers to the Spearman rho coefficient value (darkest blue being 1 and darkest red being -1) and the size of each dot is proportional to the strength of the correlation. Three stars indicate an FDR adjusted p value of <0.05, two stars indicate an FDR adjusted p value of <0.1 and one star denotes an FDR adjusted p value of <0.2. (b-c) Corrected TMB, RTK mutations, molecular smoking signature and HLA germline variation were combined in a multivariable Cox proportional hazards regression model and a risk score was calculated for each case based on the weighted contribution of each parameter. Among 84 patients in total, patients with a higher risk score (N=28 patients) had a significantly shorter overall survival in cohort 1 (13 months vs. 38 months, HR=3.16, 95% CI: 1.68-5.91, log rank p=0.0002; b). The model was trained in cohort 1, fixed and applied in cohort 2 revealing a significantly shorter progression-free survival for high risk patients in cohort 2 (3 months vs. 8 months, HR=2.73, 95% CI 1.15-6.45, log rank p=0.017; c). The second tertile of the risk score was used to classify patients in high risk (top 33.3%, N=28 patients for cohort 1 and N=11 patients for cohort 2) and low risk (bottom 66.6%, N=56 patients for cohort 1 and N=23 patients for cohort 2) groups. The median point estimate and 95% confidence intervals for survival were estimated by the Kaplan-Meier method and survival curves were compared by using the nonparametric log rank test. Log rank p values are based on two-sided testing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank members of our laboratories for critical review of the manuscript. This work was supported in part by US National Institutes of Health grants (nos. CA121113 to V.E.V. and V.A., CA006973 to D.M.P. and V.E.V. and CA180950 to V.E.V.), the Commonwealth Foundation (to V.E.V. and S.B.B.), the Bloomberg-Kimmel Institute for Cancer Immunotherapy (to V.A., P.M.F., J.R.B., D.M.P. and V.E.V.), the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation (to V.E.V. and S.B.B.), the Eastern Cooperative Oncology Group—American College of Radiology Imaging Network (to V.A.), MacMillan Foundation (to V.A.), the V Foundation (to V.A., V.E.V. and S.B.B.), Swim Across America (to V.A.), the William R. Brody Faculty Scholarship (to R.K.), the SU2C-ACS Lung Cancer Dream Team (to P.M.F. and E.G.), PRIME Oncology (to J.N.), the SU2C DCS International Translational Cancer Research Dream Team Grant (no. SU2C-AACR-DT1415 to V.E.V.), the SU2C-LUNGevity-American Lung Association Lung Cancer Interception Dream Team (to J.B. and V.E.V.), the VARI-SU2C Epigenetics Dream Team Grant (to S.B.B.), the Allegheny Health Network—Johns Hopkins Research Fund (to V.A. and V.E.V.), the LUNGevity Foundation (to V.A. and P.M.F.), the Jim Toth Foundation (to J.B. and S.B.B.), the Rising Tide Foundation (to J.B. and S.B.B.) and the Mark Foundation for Cancer Research (to V.E.V.). Stand Up To Cancer is a program of the Entertainment Industry Foundation administered by the American Association for Cancer Research.

Competing interests

V.A. receives research funding from Bristol-Meyers Squibb. P.M.F. receives research funding for clinical trials from Bristol-Myers Squibb, AstraZeneca/MedImmune, Kyowa, Corvus and Novartis and is a consultant/advisory board member for AbbVie, AstraZeneca, Bristol-Myers Squibb, Boehringer Ingelheim, Celgene, EMD Serono, Inivata, Janssen, Lilly, Merck, and Novartis. J.R.W. is founder and owner of Resphera Biosciences, and is a consultant for Personal Genome Diagnostics. J.N. receives research funding from Merck, AstraZeneca/MedImmune, Kyowa and Calithera and is a consultant/advisory board member of Bristol-Myers Squibb, AstraZeneca/MedImmune and Takeda. C.L.H. is a consultant/advisory board member for AbbVie, Bristol-Myers Squibb and Genentech, receives research funding from Merrimack, GlaxoSmithKline, AbbVie, Bristol-Myers Squibb, and GlaxoSmithKline. B.L. is a consultant/advisory board member of AstraZeneca, Celgene, Genentech Rouche, Eli Lilly and Takeda and

receives research funding from Celgene and Boehringer Ingelheim. J.R.B. receives research grants from Bristol-Myers Squibb, AstraZeneca/MedImmune and Merck and is a consultant/advisory board member for Bristol-Myers Squibb, AstraZeneca/MedImmune and Merck. M.V.B. is a paid consultant for Cepheid Inc. J.I. owns LumaCyte, LLC. Stock. D.M.P. receives research support from Bristol-Myers Squibb, Compugen and Potenza Therapeutics and is a consultant for Aduro Biotech, Amgen, AstraZeneca/MedImmune, Bayer, Compugen, DNatrix, Five Prime, GlaxoSmithKline, ImmuneXcite, Jounce Therapeutics, Neximmune, Pfizer, Rock Springs Capital, Sanofi and Vesuvius/Tizona. V.E.V. is a founder of Personal Genome Diagnostics, a member of its Scientific Advisory Board and Board of Directors, and owns Personal Genome Diagnostics stock, which are subject to certain restrictions under university policy. V.E.V. is an advisor to Takeda Pharmaceuticals. Within the last 5 years, V.E.V. has been an advisor to Daiichi Sankyo, Janssen Diagnostics and Ignyta. The terms of these arrangements are managed by Johns Hopkins University in accordance with its conflict of interest policies.

References

1. Rizvi NA et al. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124–128 (2015). [PubMed: 25765070]
2. Rizvi H et al. Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-ligand 1 (PD-L1) blockade in patients with non-small-cell lung cancer profiled with targeted next-generation sequencing. *J. Clinical Oncol* 36, 633–641 (2018). [PubMed: 29337640]
3. Hellmann MD et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *New Eng. J. Med* 378, 2093–2104 (2018). [PubMed: 29658845]
4. Hellmann MD et al. Genomic features of response to combination immunotherapy in patients with advanced non-small-cell lung cancer. *Cancer Cell* 33, 843–852 (2018). [PubMed: 29657128]
5. Yarchoan M, Hopkins A & Jaffee EM Tumor mutational burden and response rate to PD-1 inhibition. *New Eng. J. Med* 377, 2500–2501 (2017). [PubMed: 29262275]
6. Samstein RM et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet* 51, 202–206 (2019). [PubMed: 30643254]
7. Raphael BJ, Dobson JR, Oesper L & Vandin F Identifying driver mutations in sequenced cancer genomes: computational approaches to enable precision medicine. *Genome Med.* 6, 5 (2014). [PubMed: 24479672]
8. Shi W et al. Reliability of whole-exome sequencing for assessing intratumor genetic heterogeneity. *Cell Rep.* 25, 1446–1457 (2018). [PubMed: 30404001]
9. Viray H et al. A prospective, multi-institutional diagnostic trial to determine pathologist accuracy in estimation of percentage of malignant cells. *Arch. Pathol. Lab. Med* 137, 1545–1549 (2013). [PubMed: 24168492]
10. Gettinger S et al. Impaired HLA class I antigen processing and presentation as a mechanism of acquired resistance to immune checkpoint inhibitors in lung cancer. *Cancer Disc.* 7, 1420–1435 (2017).
11. Chowell D et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* 359, 582–587 (2018). [PubMed: 29217585]
12. Wood DE et al. A machine learning approach for somatic mutation discovery. *Science Transl. Med* 10, eaar7939 (2018).
13. Anagnostou V et al. Evolution of neoantigen landscape during immune checkpoint blockade in non-small cell lung cancer. *Cancer Disc.* 7, 264–276 (2017).
14. Miao D et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat. Genet* 50, 1271–1281 (2018). [PubMed: 30150660]
15. Forde PM et al. Neoadjuvant PD-1 blockade in resectable lung cancer. *New Engl. J. Med* 378, 1976–1986 (2018). [PubMed: 29658848]
16. Krause DS & Van Etten RA Tyrosine kinases as targets for cancer therapy. *New Engl. J. Med* 353, 172–187 (2005). [PubMed: 16014887]
17. Liu C et al. BRAF inhibition increases tumor infiltration by T cells and enhances the antitumor activity of adoptive immunotherapy in mice. *Clinical Cancer Res.* 19, 393–403 (2013). [PubMed: 23204132]

18. Shen J et al. ARID1A deficiency promotes mutability and potentiates therapeutic antitumor immunity unleashed by immune checkpoint blockade. *Nat. Medicine* 24, 556–562 (2018).
19. Peng W et al. Loss of PTEN promotes resistance to T cell-mediated immunotherapy. *Cancer Disc.* 6, 202–216 (2016).
20. Kato S et al. Hyperprogressors after immunotherapy: analysis of genomic alterations associated with accelerated growth rate. *Clinical Cancer Res.* 23, 4242–4250 (2017). [PubMed: 28351930]
21. Zaretsky JM et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *New Engl. J. Med* 13, 2016 (2016).
22. Rizvi NA et al. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124–128 (2015). [PubMed: 25765070]
23. Gao J et al. Loss of IFN-gamma pathway genes in tumor cells as a mechanism of resistance to anti-CTLA-4 therapy. *Cell* 167, 397–404 (2016). [PubMed: 27667683]
24. Knijnenburg TA et al. Genomic and molecular landscape of DNA damage repair deficiency across the Cancer Genome Atlas. *Cell Rep.* 23, 239–254 (2018). [PubMed: 29617664]
25. Le DT et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 357, 409–413 (2017). [PubMed: 28596308]
26. Spranger S, Bao R & Gajewski TF Melanoma-intrinsic beta-catenin signalling prevents anti-tumour immunity. *Nature* 523, 231–235 (2015). [PubMed: 25970248]
27. Schumacher TN & Schreiber RD Neoantigens in cancer immunotherapy. *Science* 348, 69–74 (2015). [PubMed: 25838375]
28. Van Allen EM et al. Genomic correlates of response to CTLA4 blockade in metastatic melanoma. *Science* 350, 207–211 (2015). [PubMed: 26359337]
29. Bassani-Sternberg M et al. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nature Commun.* 7, 13404 (2016). [PubMed: 27869121]
30. Luksza M et al. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* 551, 517–520 (2017). [PubMed: 29132144]
31. Smith KN et al. Persistent mutant oncogene-specific T cells in two patients benefitting from anti-PD-1. *J. Immunother. Cancer* 7, 40 (2019). [PubMed: 30744692]
32. McGranahan N et al. Allele-specific HLA loss and immune escape in lung cancer evolution. *Cell* 171, 1259–1271 (2017). [PubMed: 29107330]
33. Shukla SA et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat. Biotechnol* 33, 1152–1158 (2015). [PubMed: 26372948]
34. Shin DS et al. Primary resistance to PD-1 blockade mediated by *JAK1/2* mutations. *Cancer Disc.* 7, 188–201 (2016).
35. Koyama S et al. Adaptive resistance to therapeutic PD-1 blockade is associated with upregulation of alternative immune checkpoints. *Nature Commun.* 7, 10501 (2016). [PubMed: 26883990]
36. Tumeh PC et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515, 568–571 (2014). [PubMed: 25428505]
37. Cristescu R et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 362, eaar3593 (2018). [PubMed: 30309915]
38. Hugo W et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* 165, 35–44 (2016). [PubMed: 26997480]
39. Sabari JK et al. PD-L1 expression, tumor mutational burden, and response to immunotherapy in patients with *MET* exon 14 altered lung cancers. *Ann. Oncol* 29, 2085–2091 (2018). [PubMed: 30165371]
40. Hu-Lieskovan S et al. Improved antitumor activity of immunotherapy with BRAF and MEK inhibitors in BRAF(V600E) melanoma. *Science Trans. Med* 7, 279ra241 (2015).
41. Roh W et al. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Science Trans. Med* 9, eaah3560 (2017).
42. Anagnostou V et al. Dynamics of tumor and immune responses during immune checkpoint blockade in non-small cell lung cancer. *Cancer Res.* 79, 1214–1225 (2019). [PubMed: 30541742]

43. Anagnostou V et al. Immuno-oncology trial endpoints: capturing clinically meaningful activity. *Clinical Cancer Res.* 23, 4959–4969 (2017). [PubMed: 28864724]
44. Jones S et al. Personalized genomic analyses for cancer mutation discovery and interpretation. *Science Trans. Med* 7, 283ra253 (2015).
45. Rizvi NA et al. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124–128 (2015). [PubMed: 25765070]
46. Ellrott K et al. Scalable open science approach for mutation calling of tumor exomes using multiple genomic pipelines. *Cell Syst.* 6, 271–281 (2018). [PubMed: 29596782]
47. Masica DL et al. CRAVAT 4: cancer-related analysis of variants toolkit. *Cancer Res.* 77, e35–e38 (2017). [PubMed: 29092935]
48. Szolek A et al. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics* 30, 3310–3316 (2014). [PubMed: 25143287]
49. Nielsen M & Andreatta M NetMHCpan-3.0; improved prediction of binding to MHC class I molecules integrating information from multiple receptor and peptide length datasets. *Genome Med.* 8, 33 (2016). [PubMed: 27029192]
50. Stranzl T, Larsen MV, Lundegaard C & Nielsen M NetCTLpan: pan-specific MHC class I pathway epitope predictions. *Immunogenetics* 62, 357–368 (2010). [PubMed: 20379710]
51. Kim Y, Sidney J, Pinilla C, Sette A & Peters B Derivation of an amino acid similarity matrix for peptide: MHC binding and its application as a Bayesian prior. *BMC Bioinformatics* 10, 394 (2009). [PubMed: 19948066]
52. Balasubramanian S et al. Using ALoFT to determine the impact of putative loss-of-function variants in protein-coding genes. *Nature Commun.* 8, 382 (2017). [PubMed: 28851873]
53. Rosenthal R deconstructSigs: identifies signatures present in a tumor sample. R version 1.8.0 <https://CRAN.R-project.org/package=deconstructSigs> (2016).
54. Talevich E, Shain AH, Botton T & Bastian BC CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol* 12, e1004873 (2016). [PubMed: 27100738]
55. Olshen AB, Venkatraman ES, Lucito R & Wigler M Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 5, 557–572 (2004). [PubMed: 15475419]
56. Shen R & Seshan VE FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* 44, e131 (2016). [PubMed: 27270079]
57. Shannon CE A mathematical theory of communication. *Bell Syst. Tech. J* 27, 379–423 (1948).
58. Aran D, Sirota M & Butte AJ Systematic pan-cancer analysis of tumour purity. *Nature Commun.* 6, 8971 (2015). [PubMed: 26634437]
59. Niknafs N, Beleva-Guthrie V, Naiman DQ & Karchin R Subclonal hierarchy inference from somatic mutations: automatic reconstruction of cancer evolutionary trees from multi-region next generation sequencing. *PLoS Comput. Biol* 11, e1004416 (2015). [PubMed: 26436540]
60. Thorsson V et al. The immune landscape of cancer. *Immunity* 48, 812–830 e814 (2018). [PubMed: 29628290]
61. Sidney J, Peters B, Frahm N, Brander C & Sette A HLA class I supertypes: a revised and updated classification. *BMC Immunol.* 9, 1 (2008). [PubMed: 18211710]
62. Carlson CS et al. Using synthetic templates to design an unbiased multiplex PCR assay. *Nature Commun* 4, 2680 (2013). [PubMed: 24157944]

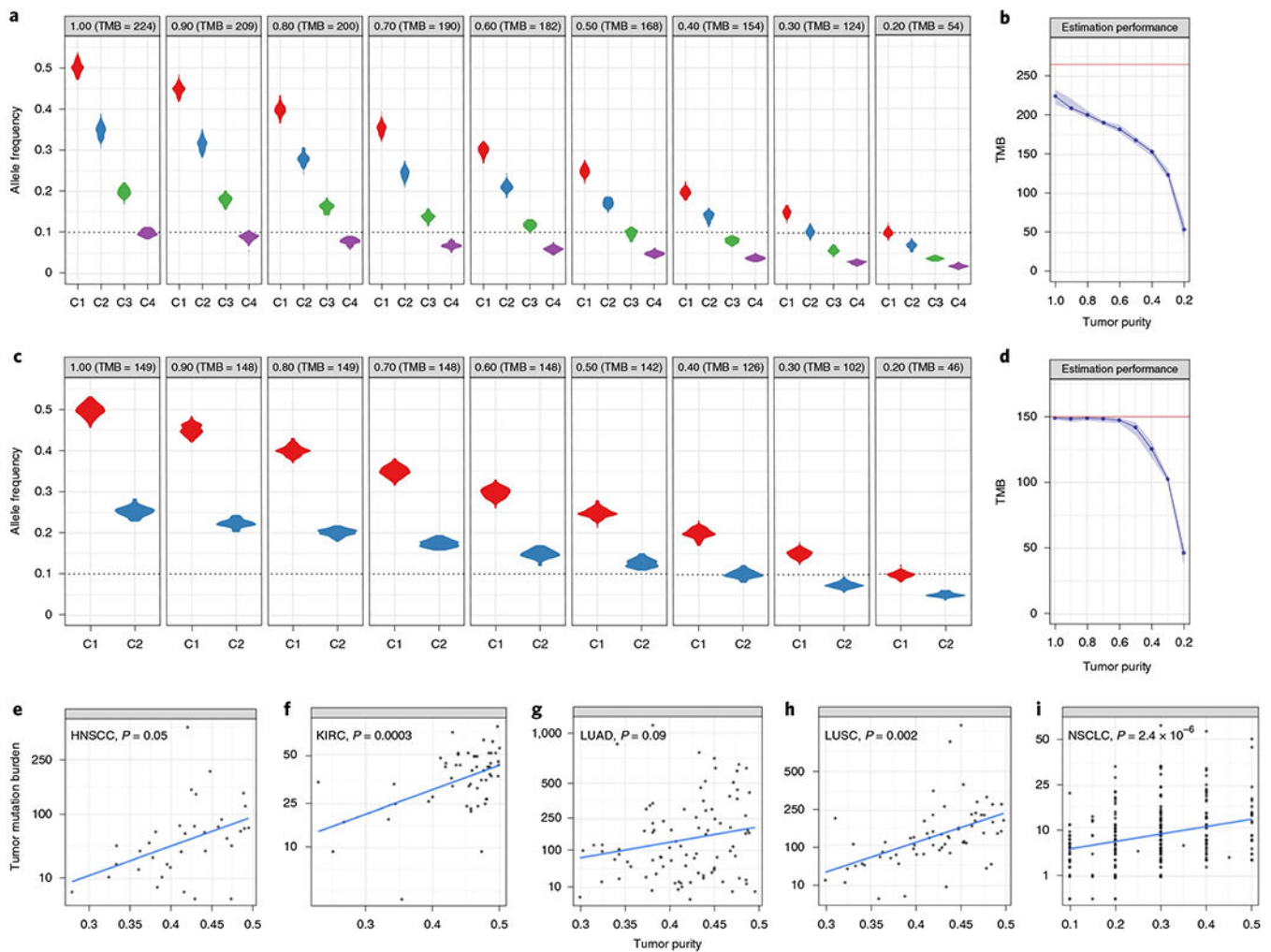


Fig. 1 | Evaluation of the impact of tumor purity and clonal heterogeneity on TMB estimates. **a–d**, Mutation burden was estimated for two in silico tumor samples, a high mutator with high intratumoral clonal heterogeneity (**a,b**) and a low mutator with low intratumoral heterogeneity (**c,d**), across a wide range of tumor purity values (0.2–1.0, shown in the header of each graph). **a**, MAF distributions are shown as violin plots for a simulated tumor with true TMB of 265 and four mutation clusters (C1 with $N = 100$ clonal mutations at cellular fraction (CF) = 1.00, C2 with $N = 50$ mutations at CF = 0.70, C3 with $N = 40$ mutations at CF = 0.40 and C4 with $N = 75$ mutations at CF = 0.20) at different tumor purity levels. The number of data points in each violin plot element represents the number of mutations in each mutation cluster. As tumor purity decreases from 1 (100% cancer) to 0.2 (20% cancer), the number of detected mutations decreases. The dotted line indicates a MAF of 10%, which is the threshold used for somatic mutation calling. **b**, Power of detection of different subclones decreased with decreasing tumor purity, resulting in a decline in TMB estimation accuracy. The blue line and ribbon mark the mean and range of estimated TMB across ten simulated tumor sample replicates, while the red dotted line indicates the true TMB of the tumor. **c**, MAF distributions are shown as violin plots for a simulated homogeneous tumor with true TMB of 150 and two mutation clusters (C1 with $N = 100$ clonal mutations at CF = 1.00 and

C2 with $N=50$ mutations at $CF=0.50$) at different tumor purity levels. Each violin plot element represents the number of mutations in each mutation cluster ($N=100$ mutations for C1 and $N=50$ mutations for C2). **d**, Estimated TMB for the tumor in **c** at each purity level shows that TMB estimates remain accurate for lower tumor purity tiers compared to the more heterogeneous tumor in **a**. As tumor purity decreases below 40%, TMB estimates converge. Panel headers indicate tumor purity and estimated TMB in **a** and **c**, and cellular fraction refers to the fraction of cancer cells harboring a mutation. **e-i**, Analysis of paired tumor-normal whole-exome sequencing data from TCGA samples with tumor purity less than 50% revealed a positive correlation between TMB and tumor purity in head and neck cancer ($N=35$ tumors, Pearson's $R=0.33$, $P=0.05$) (**e**), kidney clear cell carcinoma ($N=54$ tumors, Pearson's $R=0.48$, $P=0.0003$) (**f**), lung adenocarcinoma (LUAD; $N=89$ tumors, Pearson's $R=0.18$, $P=0.09$) (**g**) and lung squamous cell carcinoma (LUSC; $N=63$ tumors, Pearson's $R=0.39$, $P=0.002$) (**h**). TMB scores derived from targeted sequencing were highly correlated with tumor purity assessments ($N=254$ tumors, Spearman's $\rho=0.29$, $P=2.4 \times 10^{-6}$) (**i**). A linear model was fitted to the mutation sequence data for each tumor type, the Pearson correlation coefficient (R) was used to assess correlations between continuous variables, the Spearman rho coefficient was calculated for nonparametric correlations, and P values are based on two-sided testing.

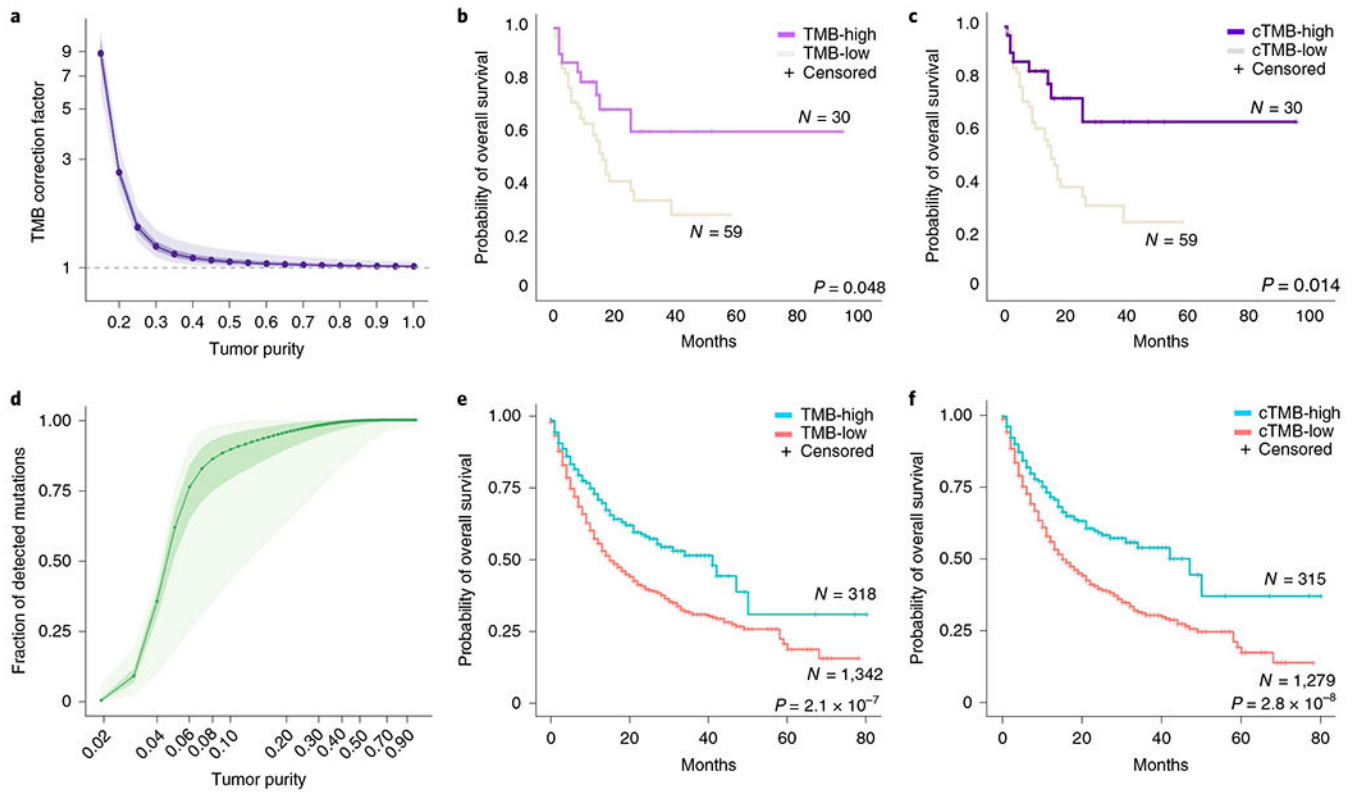


Fig. 2 |. Impact of cTMB on outcome for ICB.

a, Through simulation analyses of exome data ($N = 20,000$ simulations), we developed correction factors for different tumor purity values, and we determined cTMB values for the tumors in cohort 1. Dark purple shaded region represents the interquartile range, and light purple shaded region marks the 95% confidence interval. **b,c**, In cohort 1, $N = 89$ patients, patients with higher obsTMB (using the second tertile as a threshold) had marginally longer OS (log-rank $P = 0.048$, **b**) but the association of TMB with OS became more significant after TMB was corrected for tumor purity (log-rank $P = 0.014$, **c**). **d**, Correction factors were subsequently generated for obsTMB from targeted NGS simulated data ($N = 20,000$ simulations). The fraction of observed to true TMB (fraction of detected mutations) drops for tumor purity less than 20%. The 95% confidence intervals are shown in light green, whereas the dark green area denotes the interquartile range and dots represent the median. **e,f**, Reanalysis of $N = 1,661$ tumor samples analyzed with targeted NGS and treated with ICB, revealed an improved survival curve separation when cTMB was used (log-rank $P = 2.1 \times 10^{-7}$ for TMB (**e**) and log-rank $P = 2.8 \times 10^{-8}$ for cTMB (**f**)). The median point estimate and 95% confidence intervals for survival were estimated by the Kaplan-Meier method, and survival curves were compared using the nonparametric log-rank test. The log-rank P values are based on two-sided testing.

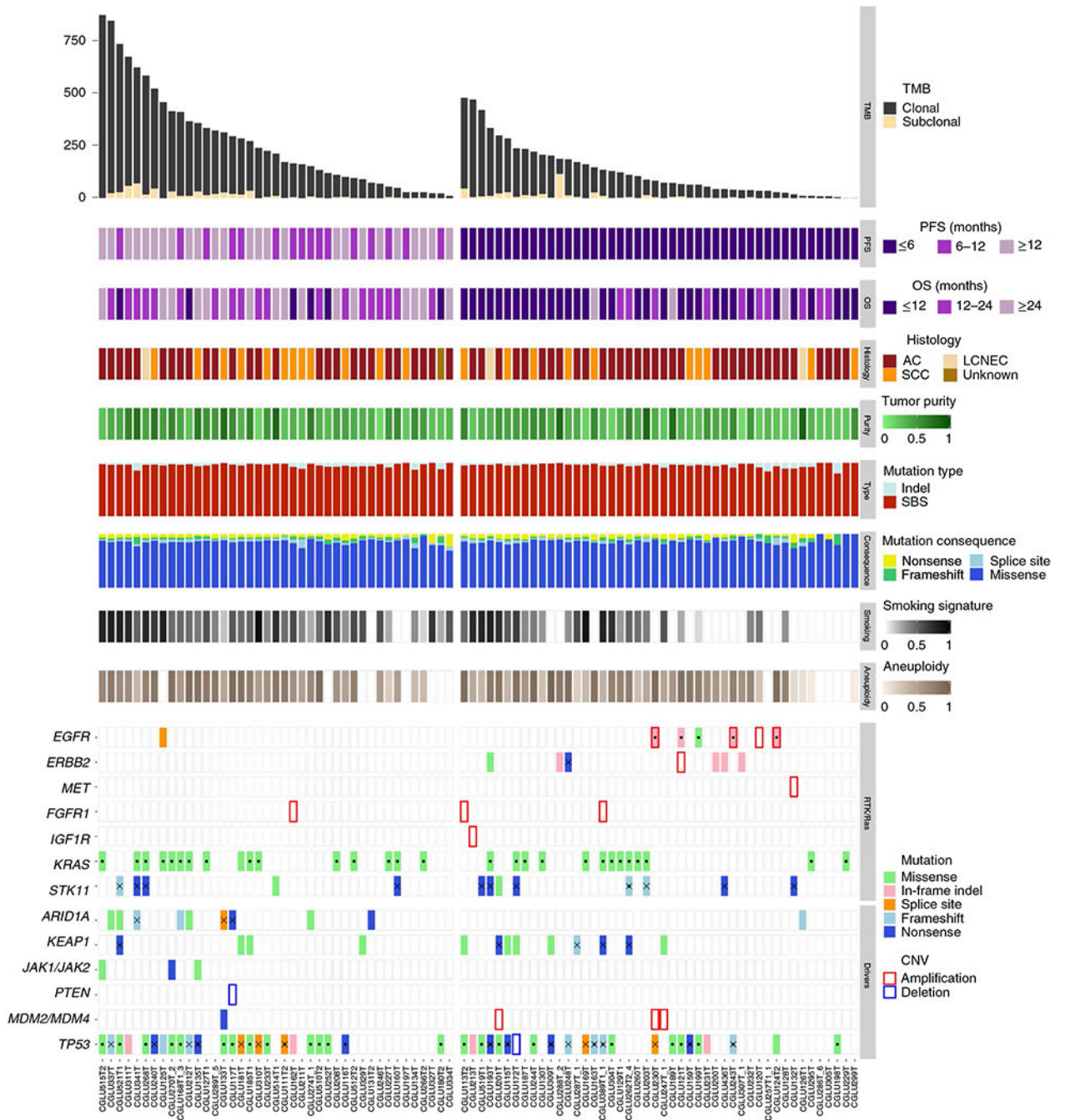


Fig. 3 | Genomic drivers associated with response to ICB.

Responding tumors had a higher total and clonal obsTMB compared to nonresponding tumors in cohort 1 ($N = 87$ patients, Mann–Whitney U test $P = 0.002$, FDR-adjusted for multiple comparisons $P = 0.012$; and Mann–Whitney U test $P = 0.0003$, FDR-adjusted $P = 0.005$, respectively), however, there was considerable overlap in the TMB range between responding and nonresponding tumors. There were no differences in tumor purity and tumor aneuploidy between responding and nonresponding tumors. Overall, a higher number of single-base substitutions and indels were found in responding tumors, which was largely

driven by their higher TMB. An enrichment in the C>A transversion-rich molecular smoking signature was found in patients with durable clinical benefit (Mann–Whitney U test, $P=0.003$, FDR-adjusted $P=0.027$). Activating mutations in RTK genes (*EGFR* and *ERBB2* point mutations and amplifications, *MET* amplification, *FGFR1* amplification and *IGF1R* amplification) were found to cluster in patients that did not derive durable clinical benefit from ICB (Fisher’s exact $P=0.0002$, FDR-adjusted $P=0.003$) independent of TMB (logistic regression TMB-adjusted $P=0.006$). Recurrent genomic alterations in *ARID1A*, including truncating mutations in the setting of LOH of the wild-type allele, were predominantly found in patients with durable clinical benefit (Fisher’s exact $P=0.005$, FDR-adjusted $P=0.024$, TMB-adjusted $P=0.062$). A trend toward enrichment in *KEAP1* mutations, especially in the context of biallelic inactivation was found in patients without durable clinical benefit (Fisher’s exact $P=0.24$, TMB-adjusted $P=0.074$). We did not detect any loss-of-function mutations in *JAK1* or *JAK2* or an enrichment of cooccurring *KRAS* and inactivating *STK11* mutations in nonresponding tumors. A homozygous deletion in *PTEN* was found in a patient with a short-lived response to ICB, and *MDM2/MDM4* amplifications were identified in three nonresponders. AC, adenocarcinoma; SCC, squamous cell carcinoma; LCNEC, large cell neuroendocrine carcinoma; SBS, single base substitution; CNV, copy number variation. Dots represent hotspot mutations, and \times denotes loss of heterozygosity of the wild-type allele.

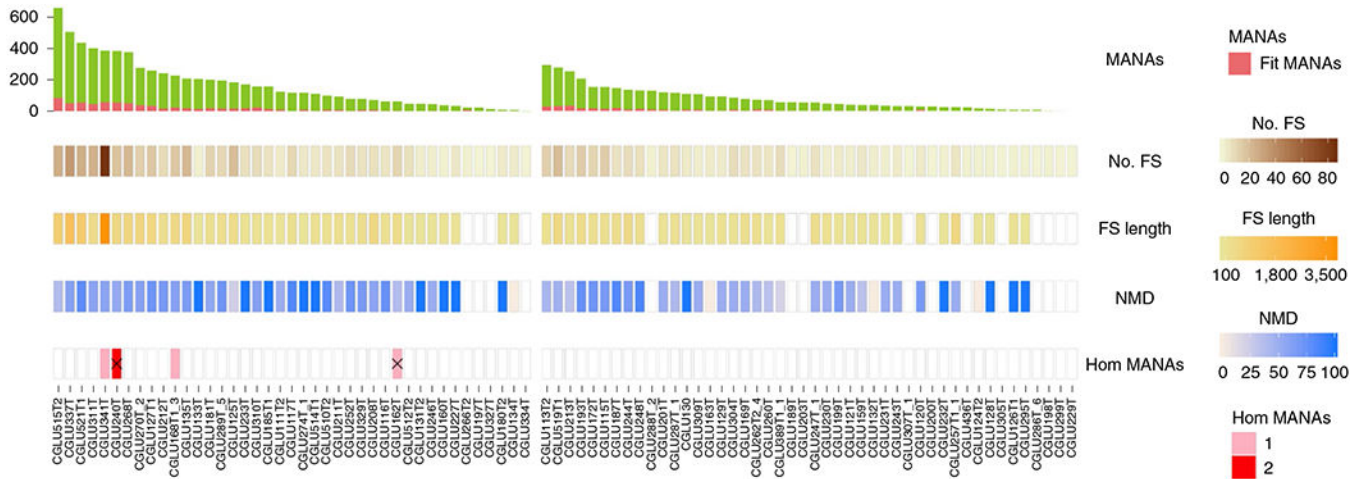


Fig. 4 |. MANA characteristics for NSCLC tumors in cohort 1.

The distributions of total MANA load and fit MANA load are shown in the top panel ($N=87$ patients). Responding tumors harbored a higher load of fit MANAs (determined as neopeptides with a predicted MHC affinity <50 nM for which the wild-type peptides has a predicted MHC affinity of $>1,000$ nM) compared to nonresponding tumors (Mann–Whitney U test $P=0.01$). MANAs derived from frameshift mutations were compared between responders and nonresponders after filtering out those most likely to undergo nonsense mediated decay; a higher MANA load stemming from frameshifts was found in responders (Mann–Whitney U test $P=0.08$). The cumulative length of frameshifts until reaching a stop codon was assessed after correcting for nonsense mediated decay and TMB; no differences were found between responding and nonresponding tumors. Neopeptides RLDGHTSL, FYSRAPEL and HRHPPVAL stemming from frameshift mutations in *SH2D7*, *ADAMTS12* and *KLHL42*, found in three responding tumors, were strongly similar to *Mycobacterium leprae*, *Mycobacterium tuberculosis* and HHV5 antigens, respectively. FS, frameshift; NMD, nonsense mediated decay; Hom, homologous. P values are based on two-sided testing.

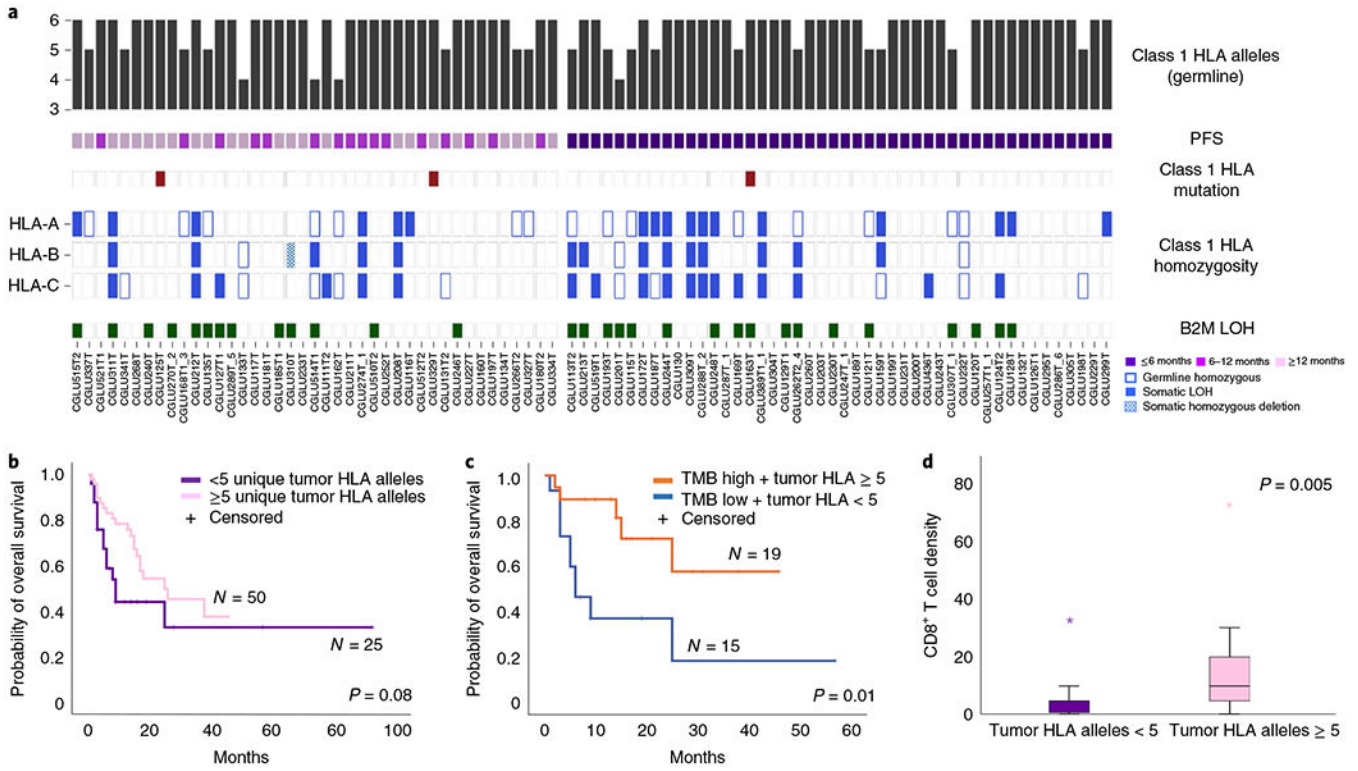


Fig. 5 | HLA class I genetic variation and association with response to ICB.

a, There were no differences in the number of germline HLA class I alleles or the degree in homozygosity found between responders and nonresponders. HLA class I somatic mutations were infrequent. HLA class I germline zygosity and somatic HLA class I LOH events were combined to calculate the unique number of HLA class I alleles on cancer cells. We identified one tumor with homozygous loss of HLA-B in patient CGLU310, who achieved durable clinical benefit from anti-PD1 therapy without evidence of disease progression 14 months after treatment initiation, suggesting that response may be attributed to NK-cell-mediated cell lysis in the setting of HLA class I homozygous deletion. There was no evidence of biallelic inactivation of β 2-microglobulin in cohort 1. **b**, Tumors with reduced antigen presentation potential (<5 unique tumor HLA class I alleles, $N = 75$ tumors) were linked to worse outcome (log-rank $P = 0.08$). **c**, This observation was more prominent when the number of HLA class I alleles in the tumor was combined with TMB. Patients with low TMB and reduced antigen presentation potential ($N = 34$ patients) had a significantly shorter OS (log-rank $P = 0.01$). **d**, Tumors with lower antigen presentation capacity showed a significantly lower level of CD8⁺ T cell density ($N = 57$ tumors, Mann–Whitney $P = 0.005$). Center values indicate median values, and error bars denote 95% confidence intervals. The median point estimate and 95% confidence intervals for survival were estimated by the Kaplan-Meier method, and survival curves were compared by using the nonparametric log-rank test. Mann–Whitney U tests and log-rank P values are based on two-sided testing.

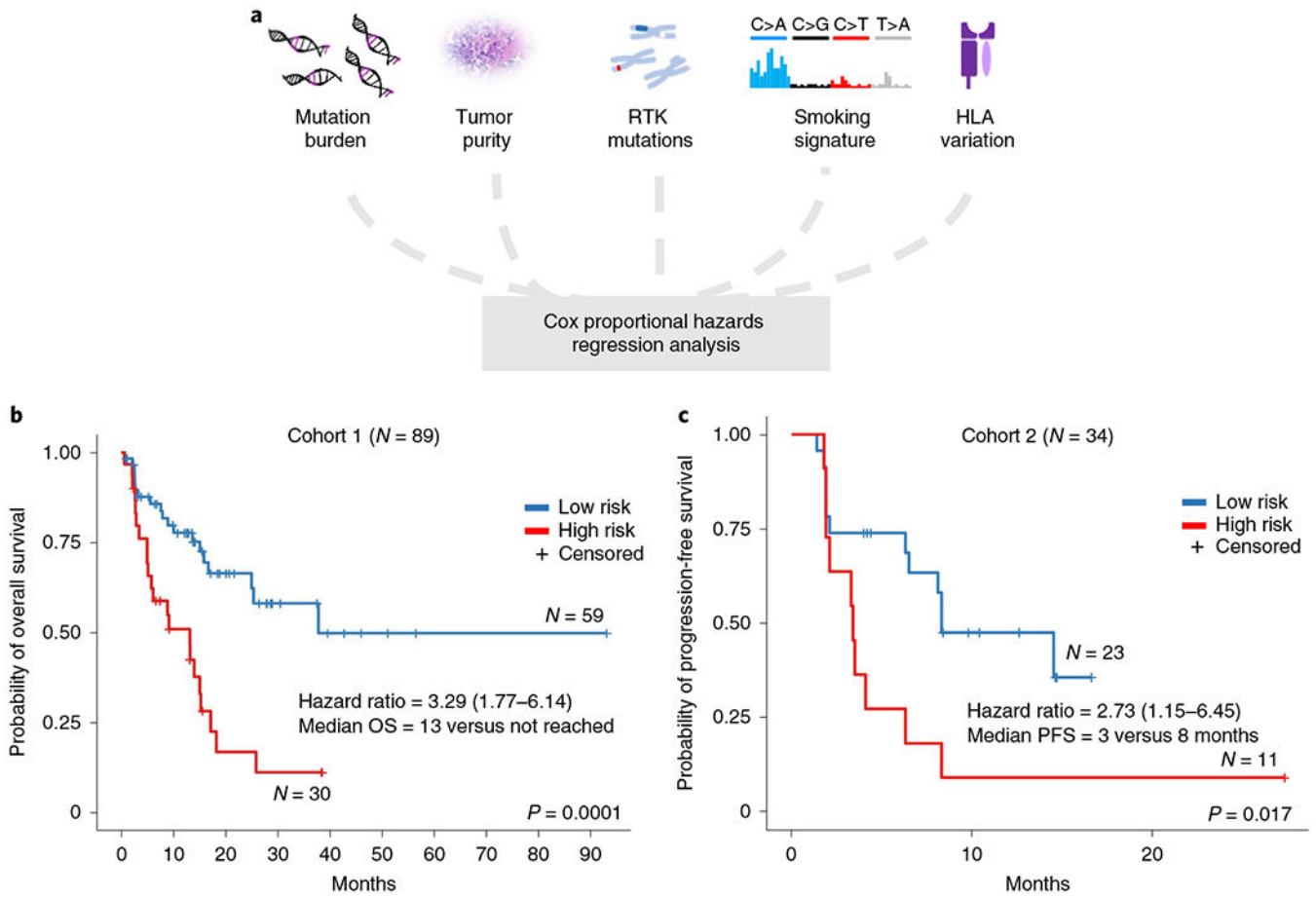


Fig. 6 |. Multivariable model for prediction of outcome to ICB.

a, cTMB, RTK mutations, molecular smoking signature and HLA germline variation were combined in a multivariable Cox proportional hazards regression model, and a risk score was calculated for each case based on the weighted contribution of each parameter. The second tertile of the risk score was used to classify patients into high risk (top 33.3%, $N = 30$ patients for cohort 1 and $N = 11$ patients for cohort 2) and low risk (bottom 66.6%, $N = 59$ patients for cohort 1 and $N = 23$ patients for cohort 2) groups. **b,c**, Patients with a higher risk score had a significantly shorter OS in cohort 1 (OS 13 months versus not reached, HR = 3.29, 95% CI: 1.77-6.14, log-rank $P = 0.0001$) (**b**) and PFS in cohort 2 (PFS 3 months versus 8 months, HR = 2.73, 95% CI = 1.15-6.45, log-rank $P = 0.017$) (**c**). The median point estimate and 95% confidence intervals for survival were estimated by the Kaplan-Meier method and survival curves were compared by using the nonparametric log-rank test. The log-rank P values are based on two-sided testing.