

Article

Not peer-reviewed version

AI Model for Predicting Anti-PD1 Response in Melanoma Using Multi-Omics Biomarkers

[Axel Gschwind](#) and [Stephan Ossowski](#) *

Posted Date: 13 January 2025

doi: 10.20944/preprints202501.0904.v1

Keywords: artificial intelligence; immune checkpoint inhibitors; anti-PD1; melanoma; genomics; transcriptomics; multi-omics; least absolute shrinkage and selection operator (LASSO); SHAP values; explainable machine learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

AI Model for Predicting Anti-PD1 Response in Melanoma Using Multi-Omics Biomarkers

Axel Gschwind ^{1,2} and Stephan Ossowski ^{1,2,*}

¹ Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany

² Institute for Bioinformatics and Medical Informatics, Universität Tübingen, Tübingen, Germany

* Correspondence: stephan.ossowski@med.uni-tuebingen.de

Simple Summary: Many patients with skin cancer are treated with immune checkpoint inhibitors. The treatment can trigger an immune response that attacks cancer cells. This is achieved by inhibiting immune checkpoints, which act as an “off-switch” of the immune system. To date, medical doctors have selected patients with a high number of DNA mutations for this kind of therapy, a biomarker called tumor mutation burden (TMB). Still, only a minority of patients benefits from the treatment, even if they are selected based on high TMB. In our study, we first collected a large cohort of patients with melanoma treated with immune checkpoint inhibitors. We then derived additional tumor biomarkers from the sequencing data of the patients. We combined these new biomarkers into an artificial intelligence model to better predict whether a patient will respond to the therapy. Our results showed that combining multiple biomarkers can improve the treatment selection for patients with melanoma by providing more accurate outcome predictions.

Abstract: Background: Immune checkpoint inhibitors (ICIs) have demonstrated significantly improved clinical efficacy in a minority of patients with advanced melanoma, whereas non-responders potentially suffer from severe side effects and delays in other treatment options. Predicting the response to anti-PD1 treatment in melanoma remains a challenge because the current FDA-approved gold standard, the nonsynonymous tumor mutation burden (nsTMB), offers limited accuracy. **Methods:** In this study, we developed a multi-omics-based machine learning model that integrates genomic and transcriptomic biomarkers to predict the response to anti-PD1 treatment in patients with advanced melanoma. We employed least absolute shrinkage and selection operator (LASSO) regression with 49 biomarkers extracted from tumor-normal whole-exome and RNA sequencing as input features. The performance of the multi-omics AI model was thoroughly compared to that of nsTMB alone, and to models that use only genomic or transcriptomic biomarkers. **Results:** We used publicly available DNA and RNA-seq datasets of melanoma patients for the training and validation of our model, forming a meta-cohort of 449 patients for which the outcome was recorded as RECIST score. The model substantially improved the prediction of anti-PD1 outcomes compared to nsTMB alone. Using SHAP values, we demonstrated the explainability of the model's predictions on a per-sample basis. **Conclusion:** We demonstrated that models using only RNA-seq or multi-omics biomarkers outperformed nsTMB in predicting the response of melanoma patients to ICI. Furthermore, our AI improves clinical usability by providing explanations of its predictions on a per-patient basis. Our findings underscore the utility of multi-omics data for selecting patients for treatment with anti-PD1 drugs. However, to train clinical-grade AI models for routine applications, prospective studies collecting larger melanoma cohorts with consistent application of exome and RNA sequencing are required.

Keywords: artificial intelligence; immune checkpoint inhibitors; anti-PD1; melanoma; genomics; transcriptomics; multi-omics; least absolute shrinkage and selection operator (LASSO); SHAP values; explainable machine learning

Introduction

Immune checkpoint inhibitors (ICIs) have revolutionized the therapy of solid cancers[1]. ICIs targeting the immune checkpoints cytotoxic T-lymphocyte associated protein (*CTLA4*) or programmed cell death protein 1/ligand 1 (*PD1/PDL1*) have become an integral part of modern cancer treatment. Anti-PD1 drugs like pembrolizumab or nivolumab have demonstrated clinical efficacy, leading to a substantial clinical benefit for a subset of patients with advanced melanoma[2,3].

Despite these efforts, the majority of patients selected for ICI treatment do not show any benefit from the therapy[4]. There is only one FDA-approved biomarker for the selection of patients for ICI in melanoma: the nonsynonymous tumor mutation burden (nsTMB). A higher nsTMB is associated with better anti-PD1 response rates, with a commonly applied threshold of 10 mutations per megabase[5]. However, nsTMB alone is insufficient to predict therapeutic outcomes accurately. For instance, objective responses to pembrolizumab were observed in only 29% of patients with high nsTMB compared to 6% of patients with low nsTMB in a recent study[6]. Moreover, patients undergoing anti-PD1 treatment could suffer from potentially life-threatening adverse events, highlighting the urgent need for more reliable prediction methods[7]. Many other non-approved genomic and transcriptomic biomarkers have been shown to correlate with ICI outcomes in various cancer types, suggesting a potential avenue for improved outcome prediction[8]. Most of these biomarkers can be derived from next-generation sequencing (NGS) data, particularly whole-exome sequencing (WES) and bulk RNA sequencing (RNA-seq)[9,10].

Given the complex tumor-immune system interactions, we hypothesized that a multi-omics approach might be a more accurate predictor of anti-PD1 response than single biomarkers. Thus, we developed a comprehensive approach that integrates both genomic and transcriptomic biomarkers as features of a multi-omics-based machine-learning model. Our model employs least absolute shrinkage and selection operator (LASSO) regression to enhance the accuracy of anti-PD1 outcome predictions in melanoma[11]. We utilized WES to generate genomic biomarkers from paired tumor and normal tissues, and RNA-seq to generate transcriptomic biomarkers of the tumor tissue. In total, our LASSO approach initially incorporated 49 multi-omics biomarkers, covering various aspects of tumor biology, including well-established nsTMB, neoantigens, resistance mechanisms, tumor immune microenvironment (immune cell infiltration), germline factors, and additional relevant metrics.

We obtained publicly available datasets of anti-PD1 treated melanoma patients to assemble a meta-cohort combining WES and RNA-seq, which we used to train and test the machine-learning model. We classified patients as therapy responders or non-responders according to the well-established Response Evaluation Criteria In Solid Tumors (RECIST)[12]. This binary classification facilitated the comparison of therapeutic outcomes between different studies.

Finally, we thoroughly evaluated our model using several performance metrics to compare our LASSO approach with the current gold standard nsTMB. Our findings showed that the integrated multi-omics approach substantially improved ICI prediction by combining several biomarkers. Notably, a model based on transcriptomic (RNA-seq) biomarkers alone performed at least as good or better than the genome-based nsTMB. A complementary analysis proved the general explainability of the AI predictions in a clinical context.

Materials and Methods

Patient Collective

We compiled a cohort from several previously published independent studies that investigated the response to anti-PD1 treatment in patients with metastatic melanoma. Raw sequencing data (paired tumor-normal WES and RNA-seq) were obtained from the Sequence Read Archive (SRA) and from the database of Genotypes and Phenotypes (dbGaP) after approval by the corresponding data access committees[13,14]. Patients were classified as responders or non-responders according to the

RECIST criteria. Patients with progressive disease, stable disease, or mixed responses were classified as non-responders, whereas those with partial or complete responses were classified as responders.

After quality control, the cohort reached a total size of 449 WES patients, of which 192 were responders and 257 were non-responders. For the RNA-seq samples, we obtained a cohort size of 308, of which 134 were responders and 174 were non-responders. Intersecting WES and RNA-seq data were available for 246 patients, of whom 110 were responders and 136 were non-responders. Figure 1(a) provides a detailed overview of the cohort. Each sub-cohort is denoted by the name of the first author of its seminal publication: Amato, Cristescu, Gide, Hugo, Liu, Pyke, Riaz, Wolchok[15-22].

AI models require a test dataset that contains data unseen by the model during training. Thus, we randomly split the data into a training- and a test dataset, resulting in 46 samples as the test dataset. We randomly selected test samples for which both WES and RNA-seq data were available. The split was stratified according to overall response and nsTMB. The final training-test split is illustrated in Figure 1(b).

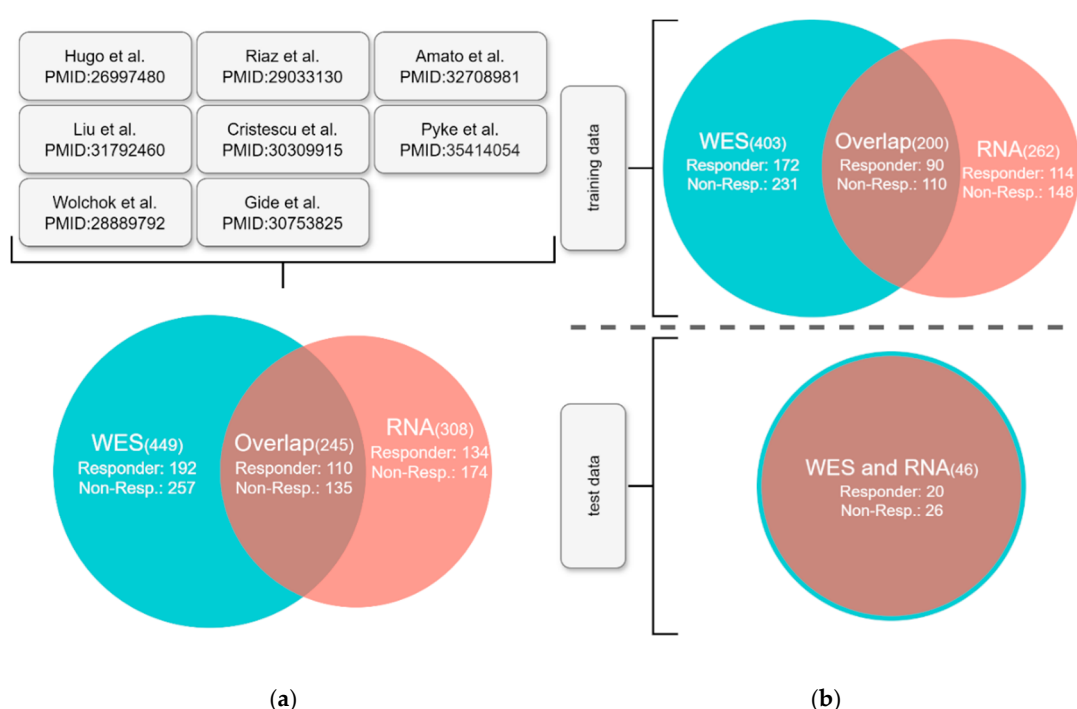


Figure 1. Overview of the overall cohort after quality control. The numbers in brackets denote the corresponding sample counts. (a) We integrated sequencing data from nine studies. We denote each study by the name of the first author of its publication. RNA-seq data were not available for every WES, and vice versa; however, we achieved an overlap of 245 samples. (b) The cohort was randomly split into a training- and test data set.

Machine Learning Model

Here, we used LASSO regression to predict response to anti-PD1 therapy. LASSO performs both regularization and feature selection by setting unimportant features to exactly zero. An important criterion for choosing LASSO as our machine learning (ML) model is its ability to handle multicollinearity in the input features, which is crucial for biological data[11]. Figure 2 shows an outline of our model design. To exploit the entire training dataset, we created one LASSO model exclusively using DNA biomarkers, and another using only RNA-seq features. The contribution of each biomarker was estimated using permutation feature importance[23]. Every feature from the DNA and the RNA model with non-zero coefficients and a feature importance of > 0.0055 was then used as an input feature for a multi-omics LASSO model, leading to a total of 10 features for the multi-omics model. This model combined DNA and RNA features and was trained on the overlap of samples for which both DNA and RNA-seq data were available.

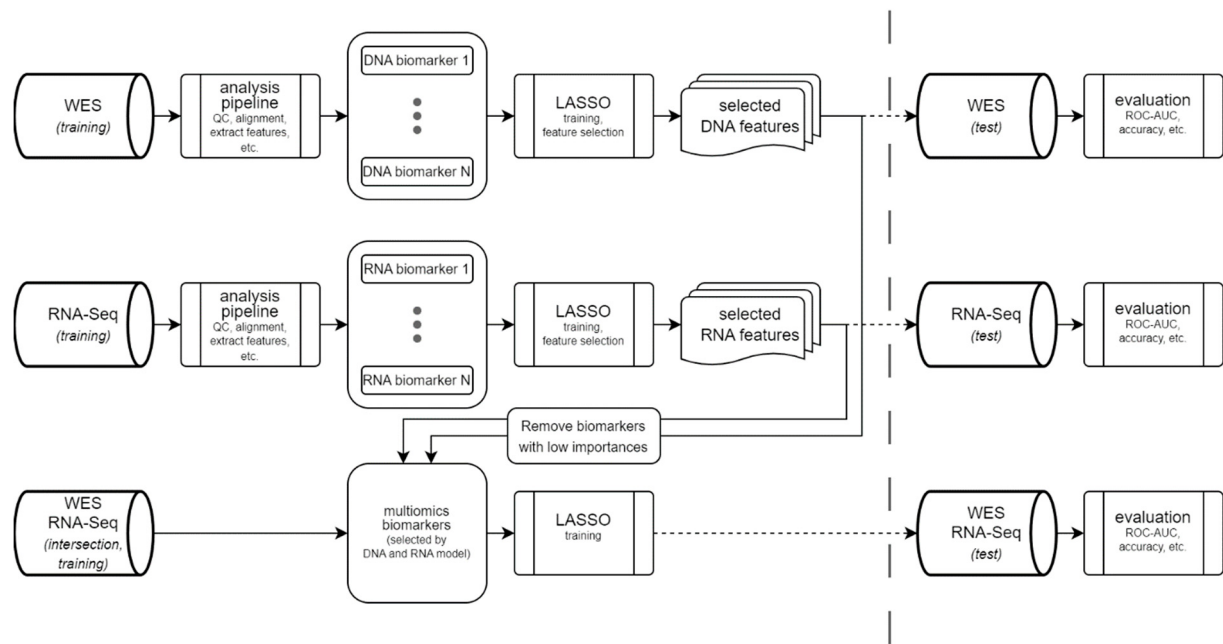


Figure 2. Flowchart of the AI models. First, we trained both the LASSO model using DNA data and the LASSO model using RNA data. The biomarkers selected that way were then used to train a multi-omics model using the overlap between all training samples for which both WES and RNA-seq data were available. We only used biomarkers with a permutation feature importance of > 0.0055 for the multi-omics model, resulting in 10 features. All three models were then evaluated in the test set using different performance measures, such as ROC AUC.

We used 40 biomarkers as features of the DNA model. The following WES-derived metrics, describing both somatic and germline properties as well as resistance mechanisms, were used:

- Somatic features: *BRAF*:V600E status, tumor purity, nsTMB, insertion and deletion (indel) burden, frameshift indel (fsIndel) burden, in-frame indel burden, splice burden, missense burden, synonymous burden, multiple amino acid burden, frameshift mutation proportion, nonsynonymous to synonymous substitution ratio (dN/dS ratio), copy number variant (CNV) burden, deletion burden, neoantigen burden, maximum neoantigen-binding affinity, mean differential agretopicity index (DAI), median DAI, maximum DAI, upper decile DAI, maximum recognition potential, maximum HEX alignment score, and maximum dissimilarity score.
- Germline features: mean HLA evolutionary divergence (HED), HLA-B27 supertype, HLA-B44 supertype, HLA-B62 supertype, homozygous HLA-B, and homozygous HLA-C.
- Mechanisms of ICI resistance (enriched in non-responders): alterations in *B2M*, alterations in *TP53*, alterations in *STK11*, alterations in *PTEN*, alterations in *KRAS*, alterations in *MDM2*, alterations in *MDM4*, and alterations in *EGFR*.
- Mechanisms of ICI response (enriched in responders) were sparse in the training dataset and merged into one biomarker: alterations in *JAK1/JAK2*, deletions at chr6p21.3 (this locus contains HLA class I-related genes), alterations in *CTNNB1*

The RNA-based ML model was trained using nine input features. We included lymphocyte infiltration, T cell receptor (TCR) α chain entropy, TCR β chain entropy, immunoglobulin heavy chain (IGHG) entropy, interferon- γ to immunosuppression (IFNG-IMS) ratio, in-frame fusion neopeptide count, gene expression profile (GEP) of HLA class I genes, GEP of *PDL1*, and GEP of *B2M*. Implementation details and a thorough discussion of every biomarker used as an input feature in both the DNA and RNA model are provided in Appendix A.

We used LASSO as implemented in LassoCV in the Python library scikit-learn 1.3.2[24]. We performed z-transformation with StandardScaler and trained each model using five-fold cross-validation. Receiver operating characteristic (ROC) area under the curve (AUC) values were used as

the main measures of the predictive power of the DNA-, RNA-, and multi-omics models. To further analyze model performance, we stratified the predicted probabilities into predicted responders and non-responders. We determined the threshold for these binary predictions by maximizing Youden's index for the training dataset, thereby ensuring an optimal trade-off between specificity and sensitivity[25]. Using this cutoff, we classified the model predictions as true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Machine learning models often suffer from a lack of explainability, which makes it difficult to determine the features that influence a specific prediction. Although we calculated permutation feature importance scores for our models, they are not meaningful on a per-sample basis, which is, however, crucial in clinical decision-making. Shapley Additive exPlanations (SHAP) is an approach that addresses this issue and provides feature contributions for predictions on a per-sample basis[26]. Therefore, we applied SHAP to evaluate the explainability of the three models and thereby their clinical applicability. This helps clinicians comprehend the model's prediction based on the biomarkers on which the model relied the most for a given case.

Bioinformatics Pipeline

Basic Analysis Pipeline

We used the megSAP pipeline for the basic analysis of both WES and RNA-seq data. megSAP is publicly available at GitHub (<https://github.com/imgag/megSAP>). Briefly, read alignment against GRCh38 was performed using BWA mem2 for WES samples. Subsequently, Manta 1.6.0 and Strelka 2.9.10 were used to call somatic variants (structural variants (SVs), single nucleotide variants (SNVs), and small indels)[27]. CNVs were called using ClinCNV 1.18.3[28]. All variants were annotated using VEP v109 and the COSMIC Cancer Mutation Census (CMC) v99[29,30]. For RNA-seq samples, reads were aligned against GRCh38 using STAR 2.7.10b. Read counts were then quantified using subread featureCounts 2.0.4 and normalized to transcripts per million (TPM) by megSAP[31]. In-frame and frameshift gene fusions were identified using Arriba 2.4.0. HLA genotyping was performed on germline WES and RNA-seq samples using our in-house Python3 implementation of hla-genotyper[32]. Results were visually inspected using the clinical decision support system GSvar (<https://github.com/imgag/ngs-bits>).

Target Regions

We constructed intersecting target regions for all studies included in our cohort. In short, we determined the region covered by at least 20x for every normal sample. Based on this information, we included a base in the target region if it was covered in at least half of the samples. The final target region was then constructed by intersecting it with the exonic coordinates ± 2 bases obtained from ensemble v109. These target regions were used to calculate the coverage files for the CNV calling by ClinCNV and for quality control. For any other biomarker that required a target region (e.g., TMB calculation), we used the intersection between the target regions of all sub-cohorts to avoid batch effects.

Quality Control

For paired tumor-normal WES, we discarded all samples that did not meet the minimum coverage requirements of 60x for the tumor sample and 20x for the normal sample. Samples were also excluded if known SNPs found in the tumor and the normal samples did not show a correlation of at least 0.8, indicating that both samples might have originated from different individuals, that is, sample confusion. Multiple sequenced samples from the same tumor were available for a few patients in some studies. In these cases, we excluded all samples except those with the highest SRA number.

On the level of somatic variants (SNVs and small indels), we excluded all variants that did not fulfill the following quality criteria: i) depth in tumor and normal sample $> 20x$, ii) tumor variant allele frequency (VAF) $\geq 5\%$, iii) normal VAF $< 0.17 \times$ tumor VAF, and iv) the variant was detected in

at least three or more reads. Copy number variants were discarded if ClinCNV reported a log likelihood of < 100.

Results

AI Model for prediction of Anti-PD1 Response

We trained a machine learning model predicting therapy response using a meta-cohort of 449 cases from eight studies, for which response to anti-PD1 was available as RECIST classification. Our cohort consists of a total of 449 WES samples and 308 RNA-seq samples. However, not every WES sample has a paired RNA-seq sample and vice versa. There are 204 cases with only Whole Exome Sequencing (WES), 63 cases with only RNA-seq data, and 245 cases with both data types available. Considering this heterogeneity in the OMICs data availability, we trained three models: 1) DNA-biomarker model, 2) RNA-biomarker model, and 3) multi-omics model. To use the same sub-cohort for evaluating all three models, we only chose cases for the test set that had both WES and RNA-seq data available. This test set consisted of otherwise randomly selected 20% of the cases and was not used for model training. An overview of the cohort and the split of the cohort into training and test sets is shown in Figure 1.

An overview of the training and testing procedures is shown in Figure 2. Briefly, 49 DNA- or RNA-based biomarkers (see Methods for a full list) were extracted from short read NGS data using the NGS analysis platform megSAP (see Methods for details). Next, we trained the DNA model using all cases of the training set with available WES data and the RNA model using all cases of the training set with available RNA data using LASSO. For both models, we performed feature selection (Figure 3(a) and Figure 3(b)) and trained a LASSO model for cases with both WES and RNA-seq data using only features with a feature importance of more than 0.0055 in one of the two single-data models (Figure 3(c); see Methods for details on model training). Finally, we compared the performance of the three models on the training and the test set and compared them with the well-established nsTMB.

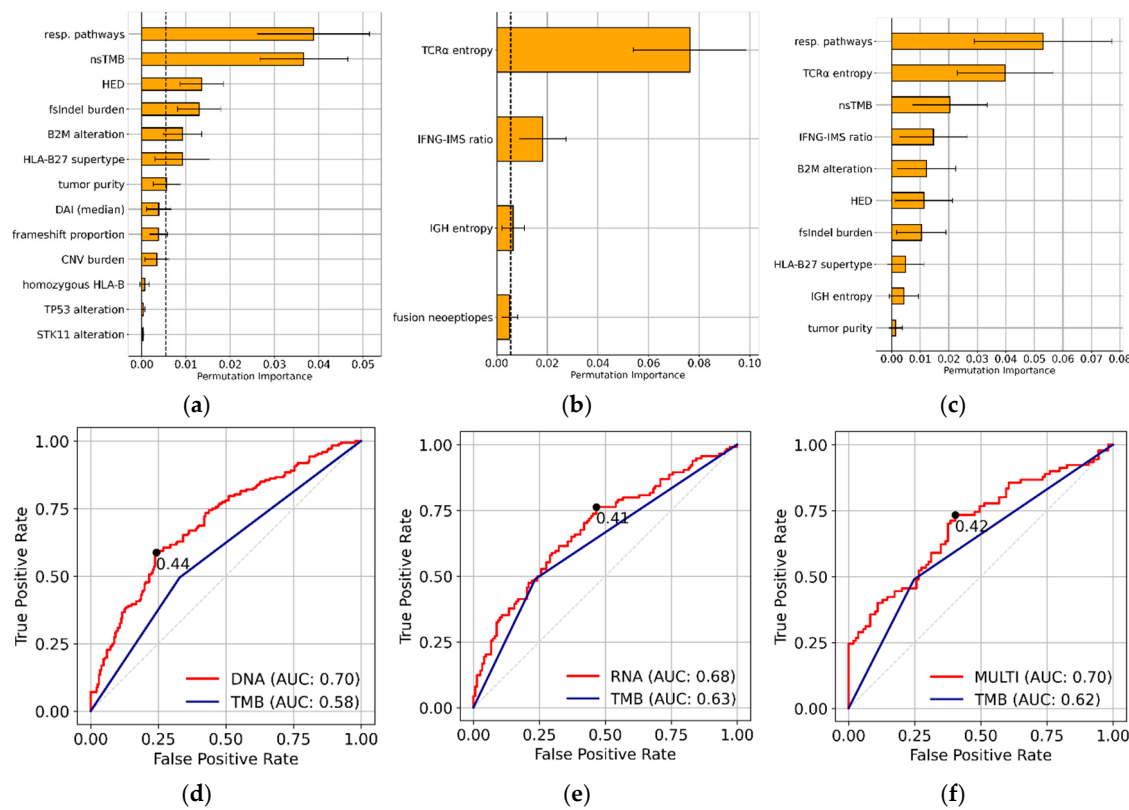


Figure 3. Feature permutation importance scores of the DNA model (a), the RNA model (b), and the multi-omics model (c). The dotted lines at 0.0055 indicate the cutoff used to include a feature in the multi-omics model. The

error bars represent the standard deviation for each biomarker. (d)-(f) show ROC curves for all three models. The red line represents the ROC curve for the LASSO model. The blue line represents the ROC curve for nsTMB using the FDA-approved threshold of 10 mutations per Mb. The black dots denote the optimal probability cutoffs determined by maximizing Youden's index. All diagrams in this figure were calculated exclusively using the training cohort. Note that the ROC AUC of nsTMB for the RNA model (e) was calculated using only the subset of samples for which paired WES samples were available.

Feature Selection and Feature Importance

First, we trained the DNA and RNA LASSO models on their corresponding biomarkers and training set using five-fold cross-validation. We obtained moderate LASSO regularization parameters of 0.021 for the DNA, and 0.024 for the RNA model, respectively. These parameters indicate that both models discarded some less important features but retained important features. As expected, the coefficients (impact on the model) of several biomarkers were shrunk to zero, such as the *BRAF* mutation status for the DNA model or mean HLA expression for the RNA model. Hence, we retained 13 features with non-zero coefficients for the DNA model and four non-zero coefficients for the RNA model. We then evaluated the feature permutation importance scores of the non-zero LASSO coefficients (Figure 3(a)-(b)). Every biomarker above the permutation feature importance threshold of 0.0055 was used as an input feature for the multi-omics model. This further reduced the feature count of the multi-omics model to seven DNA and three RNA features. The final multi-omics model incorporated the following 10 initial genomic and transcriptomic biomarkers: response pathway, nsTMB, HED, fsIndel burden, defects in *B2M*, HLA-B27 supertype, tumor purity, TCR α chain entropy, IFNG-IMS ratio, and IGH entropy. The multi-omic model had a small LASSO regularization parameter of 0.002 after five-fold cross-validation, indicating that most of the features were retained. Indeed, the multi-omics model selected all the initial input features and resulted in ten non-zero (DNA and RNA) coefficients (Figure 3(c)).

Model Performance on Training and Test Set

Next, we calculated the ROC curves using the response probabilities of the models for the training set, with the actual response data (responders and non-responders) serving as the ground truth. All three models outperformed the gold-standard nsTMB in terms of ROC AUC (Figure 3(d)-(f)). The DNA model achieved a ROC AUC of 0.70 versus 0.58 (nsTMB), the RNA model a ROC AUC of 0.68 versus 0.63 (nsTMB), and the multi-omics model a ROC AUC of 0.70 versus 0.62 (nsTMB) on the training dataset. The ROC AUC for the nsTMB of the RNA model was calculated using only the subset of samples for which paired WES data were available. The distribution of the prediction probabilities varied across the three models (Figure S1), emphasizing the necessity of model-specific probability cutoff values. By maximizing Youden's index, we identified optimal probability thresholds of 0.44 for the DNA model, 0.41 for the RNA model, and 0.42 for the multi-omics model, respectively. Subsequently, these thresholds were used to classify the model predictions as TP, TN, FP, and FN. Next, we evaluated the DNA, RNA, and multi-omics models using the independent test set. Figure 4(a)-(c) show the performance of the models in the form of ROC curves. All three models outperformed nsTMB alone in terms of ROC AUCs: 0.63 versus 0.60 for the DNA model, 0.62 versus 0.60 for the RNA model, and 0.64 versus 0.60 for the multi-omics model. These results imply that the model training did not overfit.

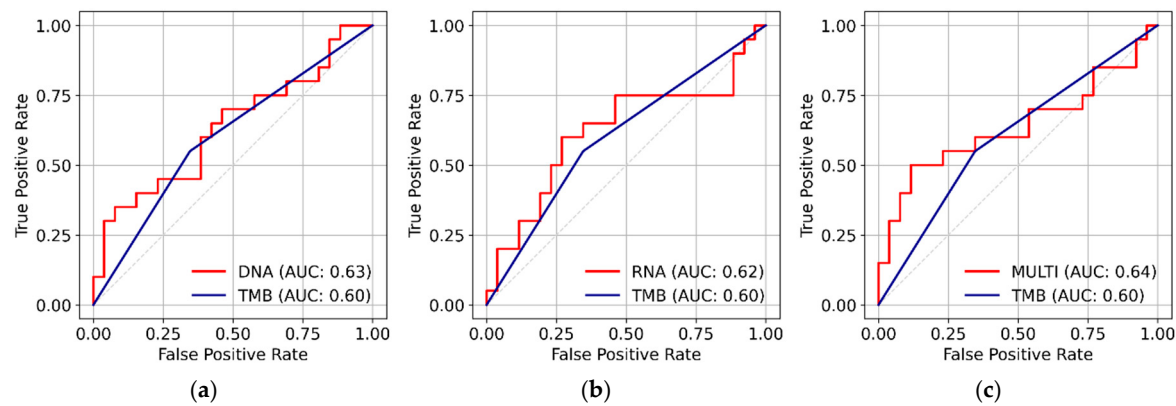
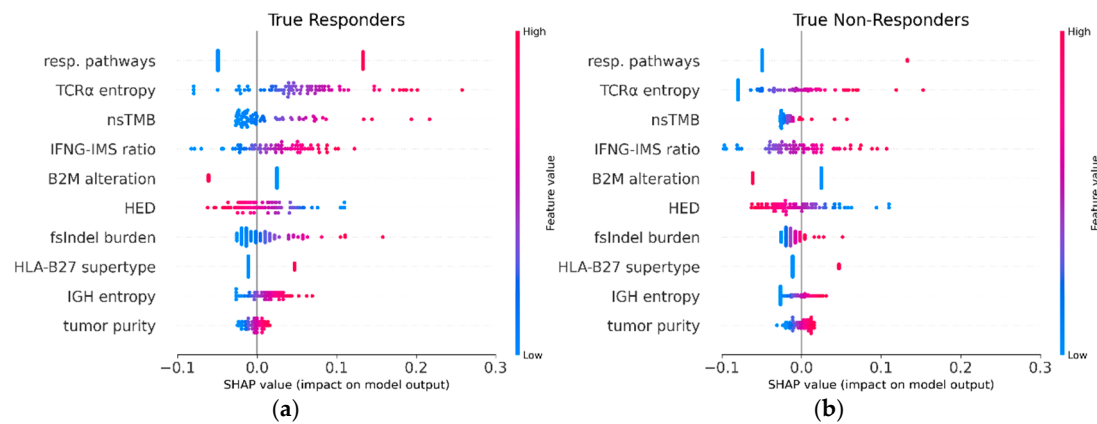


Figure 4. Model performance in the test set. (a)-(c) show ROC curves of the DNA model (a), the RNA model (b), and the multi-omics model (c). None of the models exhibited overfitting. The ROC AUC of the LASSO models was higher than the ROC AUC of the nsTMB alone.

SHAP values

Subsequently, we applied the SHAP methodology to the complete cohort analyzed using the multi-omics model to identify typical biomarkers contributing to the prediction of TP (correctly identified responders) and TN (correctly identified non-responders). SHAP values estimate the impact of each feature on the prediction on a per-sample basis. Figure 5(a)-(b) show summary plots (“bee swarm”) of the SHAP values for the multi-omics model for true responders and true non-responders. Each dot represents the contribution of a biomarker to a specific model prediction. Positive SHAP values contribute to the prediction of responders, whereas negative SHAP values indicate a negative contribution to the prediction of non-responders. The dot color denotes the original input value of each feature; high input values are plotted in red and low input values are plotted in blue. This allows us to determine the direction in which the features influence a prediction. For some of the biomarkers, we observed clear differences between the two groups. For instance, true non-responders had very low or zero TCR α chain entropy values that contributed to the prediction of TN. On the other hand, high TCR α chain entropy values contributed to the prediction of TP. To quantify such discriminative biomarkers between TP and TN, we calculated the differences in mean SHAP value for each feature (Figure 5(c)). We tested the significance of the differences using a two-sided Mann-Whitney U test with Bonferroni correction, resulting in seven statistically significant discriminative biomarkers between true responders and non-responders: TCR α chain entropy, response pathways, nsTMB, IFNG-IMG ratio, fsIndel burden, IGH entropy, and HED.



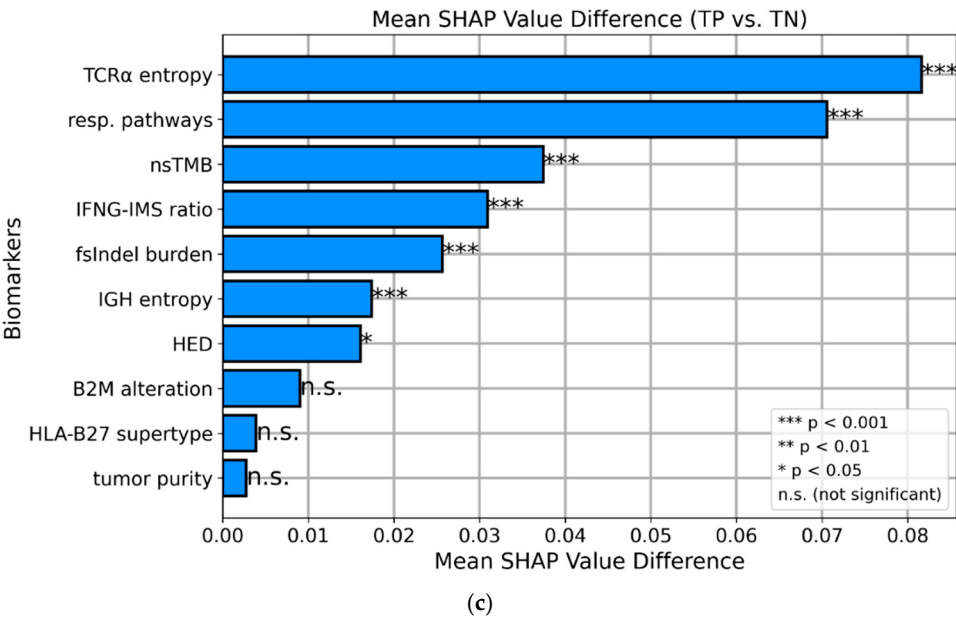


Figure 5. Summary plots (“bee swarm”) of the SHAP values for the multi-omics model using the overall cohort. True responders (a) were compared with true non-responders (b). Each dot represents the contribution of a feature to the prediction for a single sample. The dot color denotes the original feature value, blue a low value, and red a high value. (c) Differences in mean SHAP values between TP and TN. The asterisks (*) denote the significance level after a two-sided Mann-Whitney U test.

Figure 6 shows representative plots for two patients, a true responder and a true non-responder, highlighting the model’s ability to explain its predictions on a per-sample basis. Both genomic and transcriptomic biomarkers have contributed to these predictions.

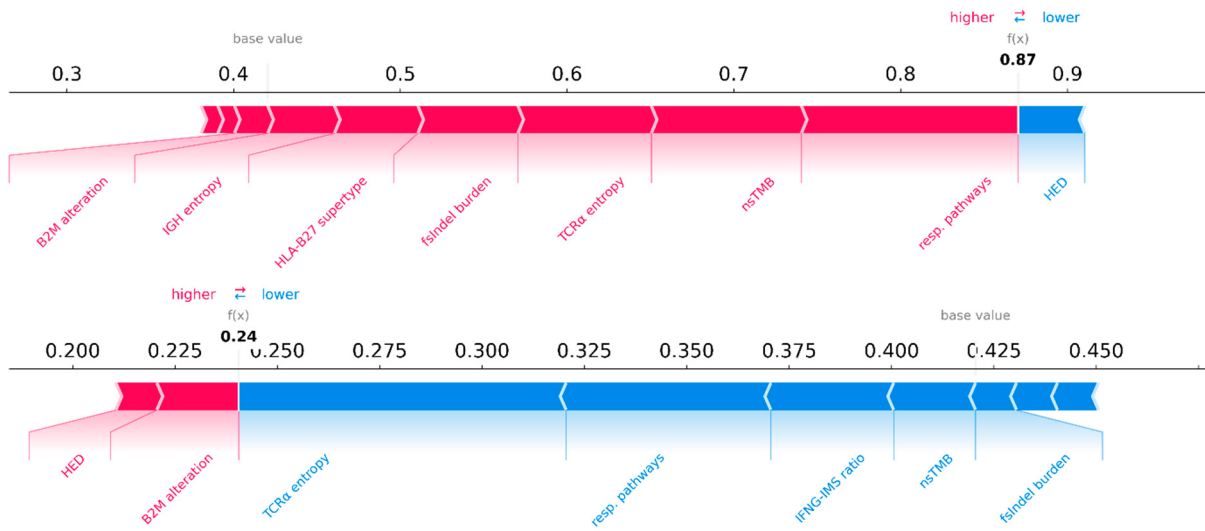


Figure 6. Two representative SHAP force plots of a TP (top) and an TN (bottom). Using these plots, we can evaluate and visualize the contribution of each feature on a specific sample directly. Both genomic and transcriptomic biomarkers contributed to the correctly identified responder and non-responder.

Discussion

Taken together, our data showed that AI models using multiple DNA- and RNA-seq-based biomarkers improved the prediction of anti-PD1 outcomes compared to the current gold standard, nsTMB. All trained models, using only genomic or transcriptomic biomarkers, as well as the multi-omic model, outperformed nsTMB on the training and the test set in terms of the ROC AUC,

suggesting that incorporating additional features can provide a more accurate prediction of anti-PD1 responses if sufficient training data are available. However, the availability of homogeneous training data is still unsatisfactory, as even after combining multiple studies into a meta-cohort, the number of samples with both WES and RNA-seq remained limited. Therefore, we expect that larger training datasets will further improve the potential of AI models.

As anti-PD1 treatment is frequently considered the last treatment option, patients frequently receive ICIs, regardless of the outcome of tumor genome diagnostics. As a result, the identification of discriminative features between responders (true positives) and non-responders (true negatives) is of particular interest to increase the explainability of the results, potentially leading to improved diagnostic interpretation. We achieved a patient-specific explainability of the outcome by using SHAP values (Figure 5(c)). These provide clinicians with a detailed breakdown of the biomarkers that contributed most to the classification of a patient. Patients showing strong indicators of non-respondents could receive another type of therapy, such as *BRAF* inhibitors, or could be spared of the side effects of the treatment.

Surprisingly, the pure RNA model (without DNA-based features) achieved a better overall performance than the well-established gold standard nsTMB alone in both the test and training datasets, although it did not include nsTMB or any other TMB-related metric. This emphasizes the potential of RNA-based features in modeling ICI responses, mostly based on features reflecting immune cell infiltration (quantified as TCR composition) and interferon signatures. Integrating DNA- and RNA-based biomarkers into a multi-omics model resulted in higher overall performance, despite the substantially reduced training dataset for which both WES and RNA-seq were available. Feature selection for the multi-omics model indicated that RNA-derived biomarkers were mostly independent of DNA-derived biomarkers, contributing equally to the top four most influential features. Indeed, we did not detect strong correlations between the genomic and transcriptomic biomarkers in the training dataset (Figure S2). However, we observed correlations between RNA-seq-based biomarkers and tumor purity. Hence, models that include RNA biomarkers may underperform in low-purity samples.

The features preferentially selected by the DNA and RNA models were highly biologically relevant (Figure 3(a)–(b)). The DNA model favored, for example, in addition to nsTMB, the related metric fsIndel burden. Frameshift mutations are a rich source of highly immunogenic neoepitopes that are recognized by the immune system[33]. In addition, the DNA model benefitted strongly from information about genetic variants in response pathways and resistance mechanisms, with changes in response pathways (*CTNNB1* alterations, *JAK1/JAK2* alterations, and defects at chr6p21.3) being the most important feature. Alterations in *B2M* are the most influential feature representing mutated resistance mechanisms. *B2M* alterations can lead to a dysfunctional MHC complex. This mechanism is crucial for antigen presentation by cancer cells[34]. Notably, no biomarker related to the predicted neoantigens exceeded the feature permutation importance threshold of 0.0055. Therefore, we assume that nsTMB is a comprehensive proxy for most neoantigen-related biomarkers. In contrast, the RNA model favored immune-related features, such as TCR α chain entropy, IFNG-IMS ratio, and IGH chain entropy, with TCR α chain entropy being by far the most important biomarker. These features capture many aspects of the tumor immune microenvironment[35]. TCR α and IGH chain entropy reflect information regarding immune cell infiltration and diversity in tumors of both T cells and B cells. With IFNG-IMS, the RNA model also selected an important marker of immune balance because it combines transcriptomic signals of immune response and immunosuppression[36]. All features selected by the DNA and RNA model retained non-zero coefficients in the multi-omics model. However, the permutation feature importance scores differed for some features in the multi-omics models compared with those in the DNA and RNA models. For example, tumor purity had a moderate feature importance score in the DNA model but was almost zero in the multi-omics model. This can be attributed to the correlation between tumor purity and TCR α chain entropy (Figure S2). Tumors with low purity often have a higher immune cell infiltration, leading to higher TCR α chain entropy.

The SHAP analysis of true responders and true non-responders in the multi-omics model confirmed the importance of the DNA and RNA features that were originally selected by feature permutation importance (Figure 5(a)-(c)). Remarkably, the SHAP methodology allowed us to directly determine the direction in which biomarkers influenced the predictions. For example, it became evident that a high TCR α chain entropy or a high nsTMB and fsInDel burden contributed to positive predictions as responders. The SHAP analysis confirmed that patients with mutated response pathways were more likely to respond to immunotherapy, whereas patients with defects in *B2M* were less likely to be classified as responders. More importantly, the application of SHAP values demonstrated the ability of our model to explain the biomarker contribution on a single-sample basis, i.e. the classification of a specific patient. In Figure 6, we showed two representative SHAP examples of single-sample predictions from the multi-omics model, providing further evidence for the per-patient interpretability of our predictions. This is an important step towards the clinical usability of machine learning models because the selection of a patient for anti-PD1 therapy requires individual explainability and assessment of the classification.

Despite showing promising results, the overall improvements in treatment response prediction of the three models over nsTMB are still unsatisfactory. One reason could be the exclusion of non-NGS patient data in our models, such as data on the gut microbiome, comorbidities, age, alcohol consumption, or serum albumin. Some of these factors, such as the gut microbiome[37], or serum albumin level[38], have been shown to correlate with ICI outcomes in various cancer types. Some studies have used AI models to investigate the interplay between clinical non-NGS features and ICI[39]. However, we did not find sufficient data in available melanoma studies capturing a comprehensive selection of such biomarkers together with genomic and transcriptomic data in melanoma. Future studies should aim to collect a broader range of clinical non-NGS biomarkers, in combination with NGS data. Another limitation of our study is that it combined data from different treatment centers into a meta-cohort. Although we applied several strategies to exclude batch effects, we cannot rule them out. The limited size of the meta-cohort, especially the number of samples with both WES and RNA-seq, was the main constraint. Therefore, using a larger cohort would likely improve model training and evaluation. It should also be noted that RECIST, as an outcome measure, is a commonly used but not universally robust metric[40]. Patients with stable disease may have survival benefits, whereas partial responders may progress to a later stage. However, response measures other than the RECIST are often unavailable for public studies. Collecting more comprehensive response data could be a starting point for future trials to address this issue. Despite these limitations, we demonstrated the general usability and interpretability of multi-omics LASSO models for ICI outcome prediction in cancer.

Conclusions

In this study, we combined genomic and transcriptomic biomarkers of a large, melanoma-specific meta-cohort to train a multi-omics LASSO model to predict the response to ICI. Multiple biomarkers derived from both whole exome sequencing and RNA-seq were selected as influential features of the multi-omics model. They showed promising improvements in predicting anti-PD1 outcomes compared with nsTMB alone. The final feature selection revealed that biomarkers describing the tumor immune microenvironment should be considered together with tumor cell-intrinsic mutation-based metrics. By applying the SHAP technique, the model predictions in this study were easier to interpret and explainable on a per-patient basis, making them more applicable in clinical decision-making. Although our study demonstrated the general applicability and explainability of multi-omics LASSO models for ICI outcome prediction, the predictions are not yet fully satisfactory. The main limitations of our study were the limited size of our meta-cohort and limited homogeneity. Future research should therefore aim at collecting large homogeneous ICI cohorts analyzed by WES and RNA-seq for AI model training.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Author Contributions: AG coordinated the project, developed the ML models and visualization, performed data curation and analysis, and wrote the manuscript. SO designed the project together with AG, helped with model development, data analysis and biomarker selection, and revised the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: There are no known conflicts of interest.

Ethics Statement: This study was approved by the independent research ethics committee of the University Hospital Tübingen (project ID 002/2024BO2), and adhered to the ethical principles of the Declaration of Helsinki.

Appendix A

DNA Biomarkers

Tumor Mutational Burden and Related Mutation Count Metrics

To calculate nsTMB, we considered only nonsynonymous variants that occurred in the overlapping target region of all studies. nsTMB was calculated as the number of nonsynonymous variants divided by the size of the target region in megabases. In addition to the classical nsTMB, we derived related burden metrics using different mutation types, normalized by the size of the common target region. We determined the different burdens using the following mutation types: i) indels, ii) frameshift indels, iii) in-frame indels, iv) splice variants, v) missense variants, vi) synonymous variants, and vii) variants that cause changes in multiple amino acids (i.e., splice and frameshift). The frameshift mutation proportion, another TMB-related biomarker, was also calculated. It is defined as the proportion of frameshift mutations among all the mutation types. Its predictive potential for anti-PD1 therapy has recently been described in a different cancer entity[41]. In addition to these TMB-related metrics, we included the ratio of all nonsynonymous mutations to all synonymous mutations, known as the dN/dS ratio[42].

Neoantigen and Neopeptide Analysis

Binding affinities to major histocompatibility complex (MHC) class I molecules were predicted using pVACseq 4.0.1 for all nonsynonymous somatic variants that lie in the common target region[43]. This pipeline was executed for all possible peptides with lengths of 8, 9, and 10, reflecting the most common MHC class I peptide binding lengths[44]. We executed pVACseq using the neopeptide binding prediction algorithms MHCflurry, MHCnuggetsI, NNAalign, NetMHC, NetMHCpan, and NetMHCpanEL. We used local instances of the IEDB MHC class I binding prediction tools 3.1.2 and MHCflurry 2.0.6 for the pVACseq predictions. We then filtered neoantigen candidates using a binding affinity cutoff of $IC_{50} < 50$ nM. After filtering, neoantigen counts were normalized to the size of the target in megabases, resulting in a TMB-like biomarker called neoantigen burden. To account for neoantigen quality, we used the binding affinity of the biomarker with the highest binding affinity as a feature. Another neoantigen metric, the differential agretopicity index (DAI), describes the difference between the predicted binding affinities of wild-type and mutated alleles[45]. Values < 0 indicate a higher binding affinity for the mutated peptide than for the wild-type peptide. We calculated the mean, median, upper decile, and maximum values of all DAIs as biomarkers, reflecting both neoantigen quantity and quality[46].

We passed any neopeptide candidate with a median binding affinity cutoff of $IC_{50} < 500$ nM to the neoantigen annotation tool NeoFox 1.1.0 to obtain further neoantigen features[47]. Three metrics

predicted by NeoFox were integrated into our model. They describe various properties of the neoepitope candidates:

- **HEX alignment score:** This score describes the similarity of tumor peptides to known viral-derived peptides[48]. High scores reflect a high similarity of the neoepitope with viral pathogens, and could be more likely recognized by the immune system. We used the maximum HEX score in the sample as a biomarker.
- **Recognition potential:** A fitness model that uses binding affinities and sequence similarity of neoantigens to known pathogens of human infectious diseases to estimate the likelihood of a neoantigen interacting with the immune system[49]. Its maximum value was used as a biomarker for our model.
- **Dissimilarity score:** This score compares the mutated peptide to the non-mutated self-proteome[50]. Neoepitopes with higher dissimilarity scores are more likely to be recognized by the immune system. Therefore, we used the maximum dissimilarity score as the biomarker.

Resistance Mechanisms

Resistance mechanisms are those by which cancer cells evade cancer treatment. In ICI, these mechanisms are typical ways in which tumors can escape immune responses. Following a literature review, we included several such biomarkers in our model. Resistance mechanisms are usually described as deleterious mutations in genes of certain pathways.

Frequent mechanisms include defects in the HLA antigen presentation pathway, which prevents T cells from recognizing cancer cells[51]. Most genes that form MHC class I molecules reside at chr6p21.3. Thus, we defined a biomarker that describes deleterious events at this locus: any deletion or loss of heterozygosity (LOH) event at chr6:29941259-33314212. This region contains all HLA class I genes and genes necessary for antigen processing, such as the transporter proteins *TAP1/TAP2* and *TAPBP*. Due to the highly polymorphic nature of this locus, we did not include potentially deleterious SNVs and small indels[52]. However, we included deleterious SNVs and indels in *B2M* as independent biomarker. We assumed a dysfunctional *B2M* protein if we detected LOH events, heterozygous deletions, mutations with a high VEP effect annotation, or variants with an annotation of the COSMIC CMC tiers 1,2, and 3. In addition to these two HLA-related mechanisms, a complete list of all the resistance mechanisms integrated into our models can be found in Table A1.

Some of the mechanisms affected only a small proportion of our training dataset. Thus, we merged the mechanisms affecting less than 20% of the training dataset into one biomarker. This resulted in the union of defects at the HLA locus, CTNNB1 alterations and JAK/JAK2 alterations as one biomarker. An analysis of odds ratio revealed that these biomarkers were more enriched in responders than in non-responders. That is why we call the resulting merged biomarker “response pathway”. Further details are provided in Supplementary Table S1.

Table A1. Resistance mechanisms. Each mechanism represents a binary biomarker (True or False). If one of the described alterations was detected, the biomarker was set as true. If none of the described alterations were found, the corresponding biomarker was set as false. Mechanisms marked by an asterisk (*) affected less than 20% of the training cohort and were merged into one biomarker.

Mechanism	Description	Sources
Defects at the HLA locus*	Any deletion or LOH event that overlaps chr6p21.3, chr6:29941259-33314212. This locus contains the majority of HLA class I-related genes.	[51]
Defects in <i>B2M</i>	<i>B2M</i> is an integral part of both MHC class I and II complexes. Dysfunctional <i>B2M</i> proteins lead to the formation of dysfunctional MHC complexes. However, the loss of both functional alleles is uncommon since such cancer clones are eradicated by NK cells. Hence, we counted heterozygous deletions, LOH events, mutations with a high VEP effect, or COSMIC tiers 1, 2, and 3 as defects in <i>B2M</i> .	[53]
<i>JAK1/JAK2</i> alteration*	Homozygous deletions, VEP-high or CMC tier 1,2, and 3 somatic variants of <i>JAK1</i> and <i>JAK2</i> . Other genes of the JAK-STAT pathway were not included because <i>JAK3</i>	[54]

	is not expressed in solid tissue and alterations in most STAT genes are difficult to interpret because STATs are at the same time both TSG and oncogenes.	
CTNNB1 pathway alteration*	CTNNB1 is an oncogene. Its pathway includes the negative regulators APC, AXIN1, and HNF1A, which are present in our intersecting target region. Thus, we counted the COSMIC tier 1, 2, and 3 mutations of CTNNB1. For its negative regulators, which act as TSG, we considered any somatic mutation with a high VEP effect or COSMIC tiers 1,2, and 3 as well as homozygous deletions.	[55]
TP53 alteration	TP53 is frequently mutated in many tumors. Defects in this TSG have manifold effects on cancer cells, including immunosuppression and evasion. Thus, we considered any deletion, LOH, mutation with a high VEP effect or COSMIC tiers 1,2, and 3 in TP53.	[56]
PTEN alteration	Deleterious events in the TSG PTEN were described as a resistance mechanism previously. Thus, we considered any deletion, LOH, mutations with a high VEP effect or COSMIC tiers 1, 2, and 3 in PTEN as deleterious.	[57,58]
STK11 alteration	STK11 is a tumor suppressor gene. It was predictive of anti-PD1 resistance in a study on another cancer entity. Hence, we counted any deletions, mutations with a high VEP effect or COSMIC tiers 1,2, and 3 in STK11.	[59]
KRAS alteration	KRAS is an oncogene. The predictive potential of KRAS mutations in anti-PD1 treatment has been previously described in another cancer entity. Thus, we considered any amplification ≥ 4 and mutations with COSMIC tiers 1, 2, and 3 in KRAS deleterious.	[60]
MDM2 alteration	MDM2 is an oncogene. Activating mutations led to the hyperprogression of melanoma in a previous study. Thus, we counted any amplification ≥ 4 and COSMIC tiers 1, 2, and 3 in MDM2 as resistance mechanisms.	[61]
MDM4 alteration	MDM4 is an oncogene. Activating mutations led to the hyperprogression of melanoma in a previous study. Thus, we counted any amplification ≥ 4 and COSMIC tiers 1, 2, and 3 in MDM4 as resistance mechanisms.	[61]
EGFR alteration	EGFR is an oncogene. Activating mutations led to the hyperprogression of melanoma in a previous study. Thus, we counted any amplification ≥ 4 and COSMIC tiers 1, 2, and 3 in EGFR as resistance mechanisms.	[61]

HLA-related Biomarkers (Germline)

Carriers of homozygous HLA alleles are thought to be at a considerable disadvantage because their MHC complexes bind to a significantly smaller neopeptide repertoire than that of heterozygous individuals. Indeed, some studies have shown that patients homozygous for *HLA-B* and *HLA-C* are less likely to respond to ICI[62]. Therefore, we included both *HLA-B* and *HLA-C* homozygosity as binary biomarkers.

HLA evolutionary divergence (HED) is a measure of the evolutionary protein distance between HLA alleles in an individual[63]. It has been described as a predictive biomarker for anti-PD1 therapy outcomes in patients with melanoma[64]. In this study, we used HED, quantified as the Grantham distance between both alleles. We included the mean HED of all three classical HLA loci (*HLA-A*, *HLA-B*, and *HLA-C*) as germline biomarkers in the model. We used an in-house tool to calculate HED (<https://github.com/axelgschwind/hlaDivergence>).

HLA alleles can be grouped into several supertypes based on their similar physicochemical properties. The three *HLA-B* supertypes, HLA-B27, HLA-B44, and HLA-B62, have been described as predictors of ICI outcomes in melanoma, and are thus included as HLA-related binary biomarkers[62,65].

Tumor Purity and Heterogeneity

Tumor purity was defined as the percentage of cancer cells in a tumor sample. It has been described as a predictor of ICI outcomes in various types of cancer [66]. In this study, we defined

tumor purity as the highest value of tumor clonality determined by ClinCNV. No CNVs were detected in very few samples. In these cases, we used the median VAF of the 15 SNVs with the highest VAF as an estimate of purity.

CNV Burden

We defined the CNV burden as the proportion of the genome that carries a CNV to the total size of the genome. We only considered CNVs that passed quality control and autosomal CNVs. The deletion burden is a related biomarker and is defined as the proportion of the genome that carries a heterozygous or homozygous deletion.

RNA Biomarkers

TCR and BCR Repertoires

We used TRUST4 to reconstruct TCR and BCR clones from RNA-seq data[67]. We obtained clones of the TCR α , TCR β , and BCR IGH chains. Clones with fewer than two reads were excluded from analysis. We then calculated the Shannon entropy based on the counts of distinct clones for each chain in each sample. This metric integrates both richness (number of distinct clones) and evenness (diversity index and distribution of clone counts) into a single, continuous biomarker.

Immune Cell Infiltration

We estimated the composition of immune cell types from RNA-seq expression data using the deconvolution algorithm quanTIseq, as implemented by the wrapper pipeline immunedeconv 2.1.0[68,69]. Using this approach, we can directly interpret the results as cell proportions in the RNA sample. In our work, we used the sum of the original quanTIseq cell types “B cell”, “T cell CD4+ (non-regulatory)”, “T cell CD8+”, “T cell regulatory (Tregs)”, and “NK cell” as a rough estimate for the lymphocyte infiltration of the sample.

IFNG-IMS Ratio

We used the ratio of two immune-related gene expression signatures as a biomarker. The impact of this ratio has previously been described as a predictor of ICI outcomes in melanoma[70]. The first signature contained interferon- γ -related genes. The second signature described patterns of immunosuppression, originally obtained by comparing ICI responders with non-responders. Using the ratio of the signatures, both inflammatory and immunosuppressive elements were combined to predict the ICI outcomes more precisely. Following [70] and [71], the IFNG signature was calculated as the arithmetic mean of the read counts of *IFNG*, *STAT1*, *IDO1*, *CXCL9*, *CXCL10*, and *HLA-DRA*. The immunosuppression signature was defined as the mean expression values of *ADAM12*, *AXL*, *BCAT1*, *CCL13*, *CCL2*, *CCL8*, *CD163*, *COL6A3*, *FAP*, *IL10*, *INHBA*, *ISG15*, *OLFML2B*, *PDGFRB*, *SIGLEC1*, *STC1*, *TWIST2*, and *VCAN*[70].

Gene Expression Profiles

Loss of HLA class I expression is a well-described mechanism of immune evasion and affects anti-PD1 treatment in melanoma[72]. Thus, we calculated the mean gene expression values (in TPM) of the three classical HLA class I loci, *HLA-A*, *HLA-B*, and *HLA-C*, as GEP biomarker. *B2M* is also crucial for the formation of MHC complexes and loss of its expression has been associated with anti-PD1 outcomes in a recent study of another cancer entity[73]. For our model, we used the gene expression of *B2M*, which was quantified as TPM. Additionally, the expression of the *PD1* ligand *PDL1* in terms of TPM was utilized as another GEP biomarker. It was included because it serves as an “official” indication for anti-PD1 treatment in other cancer entities.

References

- Hodi, F.S.; O'Day, S.J.; McDermott, D.F.; Weber, R.W.; Sosman, J.A.; Haanen, J.B.; Gonzalez, R.; Robert, C.; Schadendorf, D.; Hassel, J.C.; et al. Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. *New England Journal of Medicine* **2010**, *363*, 711-723, doi:10.1056/NEJMoa1003466.
- Barone, A.; Hazarika, M.; Theoret, M.R.; Mishra-Kalyani, P.; Chen, H.; He, K.; Sridhara, R.; Subramaniam, S.; Pfuma, E.; Wang, Y.; et al. FDA Approval Summary: Pembrolizumab for the Treatment of Patients with Unresectable or Metastatic Melanoma. *Clinical Cancer Research* **2017**, *23*, 5661-5665, doi:10.1158/1078-0432.Ccr-16-0664.
- Hazarika, M.; Chuk, M.K.; Theoret, M.R.; Mushti, S.; He, K.; Weis, S.L.; Putman, A.H.; Helms, W.S.; Cao, X.; Li, H.; et al. U.S. FDA Approval Summary: Nivolumab for Treatment of Unresectable or Metastatic Melanoma Following Progression on Ipilimumab. *Clinical Cancer Research* **2017**, *23*, 3484-3488, doi:10.1158/1078-0432.Ccr-16-0712.
- Valero, C.; Lee, M.; Hoen, D.; Zehir, A.; Berger, M.F.; Seshan, V.E.; Chan, T.A.; Morris, L.G.T. Response Rates to Anti-PD-1 Immunotherapy in Microsatellite-Stable Solid Tumors With 10 or More Mutations per Megabase. *JAMA Oncology* **2021**, *7*, 739-743, doi:10.1001/jamaoncol.2020.7684.
- Marcus, L.; Fashoyin-Aje, L.A.; Donoghue, M.; Yuan, M.; Rodriguez, L.; Gallagher, P.S.; Philip, R.; Ghosh, S.; Theoret, M.R.; Beaver, J.A.; et al. FDA Approval Summary: Pembrolizumab for the Treatment of Tumor Mutational Burden–High Solid Tumors. *Clinical Cancer Research* **2021**, *27*, 4685-4689, doi:10.1158/1078-0432.Ccr-21-0327.
- Marabelle, A.; Fakih, M.; Lopez, J.; Shah, M.; Shapira-Frommer, R.; Nakagawa, K.; Chung, H.C.; Kindler, H.L.; Lopez-Martin, J.A.; Miller, W.H.; et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *The Lancet Oncology* **2020**, *21*, 1353-1365, doi:10.1016/S1470-2045(20)30445-9.
- Martins, F.; Sofiya, L.; Sykietis, G.P.; Lamine, F.; Maillard, M.; Fraga, M.; Shabafrouz, K.; Ribi, C.; Cairolis, A.; Guex-Crosier, Y.; et al. Adverse effects of immune-checkpoint inhibitors: epidemiology, management and surveillance. *Nature Reviews Clinical Oncology* **2019**, *16*, 563-580, doi:10.1038/s41571-019-0218-0.
- Baltussen, J.C.; Welters, M.J.P.; Verdegaal, E.M.E.; Kapiteijn, E.; Schrader, A.M.R.; Slingerland, M.; Liefers, G.-J.; van der Burg, S.H.; Portielje, J.E.A.; de Glas, N.A. Predictive Biomarkers for Outcomes of Immune Checkpoint Inhibitors (ICIs) in Melanoma: A Systematic Review. *Cancers* **2021**, *13*, 6366, doi:10.3390/cancers13246366.
- Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods* **2008**, *5*, 621-628, doi:10.1038/nmeth.1226.
- Ng, S.B.; Turner, E.H.; Robertson, P.D.; Flygare, S.D.; Bigham, A.W.; Lee, C.; Shaffer, T.; Wong, M.; Bhattacharjee, A.; Eichler, E.E. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **2009**, *461*, 272-276, doi:10.1038/nature08250.
- Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **1996**, *58*, 267-288, doi:10.1111/j.2517-6161.1996.tb02080.x.
- Eisenhauer, E.A.; Therasse, P.; Bogaerts, J.; Schwartz, L.H.; Sargent, D.; Ford, R.; Dancey, J.; Arbuck, S.; Gwyther, S.; Mooney, M.; et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European Journal of Cancer* **2009**, *45*, 228-247, doi:10.1016/j.ejca.2008.10.026.
- Katz, K.; Shutov, O.; Lapoint, R.; Kimelman, M.; Brister, J.R.; O'Sullivan, C. The Sequence Read Archive: a decade more of explosive growth. *Nucleic Acids Research* **2021**, *50*, D387-D390, doi:10.1093/nar/gkab1053.
- Tryka, K.A.; Hao, L.; Sturcke, A.; Jin, Y.; Wang, Z.Y.; Ziyabari, L.; Lee, M.; Popova, N.; Sharopova, N.; Kimura, M. NCBI's Database of Genotypes and Phenotypes: dbGaP. *Nucleic acids research* **2014**, *42*, D975-D979.
- Amato, C.M.; Hintzsche, J.D.; Wells, K.; Applegate, A.; Gorden, N.T.; Vorwald, V.M.; Tobin, R.P.; Nassar, K.; Shellman, Y.G.; Kim, J.; et al. Pre-Treatment Mutational and Transcriptomic Landscape of Responding Metastatic Melanoma Patients to Anti-PD1 Immunotherapy. *Cancers* **2020**, *12*, 1943-1943, doi:10.3390/cancers12071943.

16. Cristescu, R.; Mogg, R.; Ayers, M.; Albright, A.; Murphy, E.; Yearley, J.; Sher, X.; Liu, X.Q.; Lu, H.; Nebozhyn, M.; et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* **2018**, *362*, doi:10.1126/science.aar3593.
17. Gide, T.N.; Quek, C.; Menzies, A.M.; Tasker, A.T.; Shang, P.; Holst, J.; Madore, J.; Lim, S.Y.; Velickovic, R.; Wongchenko, M.; et al. Distinct Immune Cell Populations Define Response to Anti-PD-1 Monotherapy and Anti-PD-1/Anti-CTLA-4 Combined Therapy. *Cancer Cell* **2019**, *35*, 238-255.e236, doi:10.1016/j.ccell.2019.01.003.
18. Hugo, W.; Zaretsky, J.M.; Sun, L.; Song, C.; Moreno, B.H.; Hu-Lieskovan, S.; Berent-Maoz, B.; Pang, J.; Chmielowski, B.; Cherry, G.; et al. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* **2016**, *165*, 35-44, doi:10.1016/j.cell.2016.02.065.
19. Liu, D.; Schilling, B.; Liu, D.; Sucker, A.; Livingstone, E.; Jerby-Arnon, L.; Zimmer, L.; Gutzmer, R.; Satzger, I.; Loquai, C.; et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nature Medicine* **2019**, *25*, 1916-1927, doi:10.1038/s41591-019-0654-5.
20. Pyke, R.M.; Mellacheruvu, D.; Dea, S.; Abbott, C.W.; McDaniel, L.; Bhawe, D.P.; Zhang, S.V.; Levy, E.; Bartha, G.; West, J.; et al. A machine learning algorithm with subclonal sensitivity reveals widespread pan-cancer human leukocyte antigen loss of heterozygosity. *Nature Communications* **2022**, *13*, 1925-1925, doi:10.1038/s41467-022-29203-w.
21. Riaz, N.; Havel, J.J.; Makarov, V.; Desrichard, A.; Urba, W.J.; Sims, J.S.; Hodi, F.S.; Martín-Algarra, S.; Mandal, R.; Sharfman, W.H.; et al. Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* **2017**, *171*, 934-949.e916, doi:10.1016/j.cell.2017.09.028.
22. Wolchok, J.D.; Chiarion-Sileni, V.; Gonzalez, R.; Rutkowski, P.; Grob, J.J.; Cowey, C.L.; Lao, C.D.; Wagstaff, J.; Schadendorf, D.; Ferrucci, P.F.; et al. Overall Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma. *New England Journal of Medicine* **2017**, *377*, 1345-1356, doi:10.1056/NEJMoa1709684.
23. Breiman, L. Random Forests. *Machine Learning* **2001**, *45*, 5-32, doi:10.1023/A:1010933404324.
24. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825-2830.
25. Martínez-Camblor, P.; Pardo-Fernández, J.C. The Youden Index in the Generalized Receiver Operating Characteristic Curve Context. *The International Journal of Biostatistics* **2019**, *15*, doi:doi:10.1515/ijb-2018-0060.
26. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems* **2017**, *30*.
27. Kim, S.; Scheffler, K.; Halpern, A.L.; Bekritsky, M.A.; Noh, E.; Källberg, M.; Chen, X.; Kim, Y.; Beyter, D.; Krusche, P.; et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nature Methods* **2018**, *15*, 591-594, doi:10.1038/s41592-018-0051-x.
28. Demidov, G.; Ossowski, S. ClinCNV: novel method for allele-specific somatic copy-number alterations detection. *bioRxiv* **2019**, 837971, doi:10.1101/837971.
29. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Research* **2018**, *47*, D941-D947, doi:10.1093/nar/gky1015.
30. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biology* **2016**, *17*, 122, doi:10.1186/s13059-016-0974-4.
31. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923-930, doi:10.1093/bioinformatics/btt656.
32. Farrell, J.J. The prediction of HLA genotypes from next generation sequencing and genome scan data. Dissertation, Boston University, Boston, 2014.
33. Florou, V.; Floudas, C.S.; Maoz, A.; Naqash, A.R.; Norton, C.; Tan, A.C.; Sokol, E.S.; Frampton, G.; Soares, H.P.; Puri, S.; et al. Real-world pan-cancer landscape of frameshift mutations and their role in predicting responses to immune checkpoint inhibitors in cancers with low tumor mutational burden. *Journal for ImmunoTherapy of Cancer* **2023**, *11*, e007440, doi:10.1136/jitc-2023-007440.

34. Cai, L.; Michelakos, T.; Yamada, T.; Fan, S.; Wang, X.; Schwab, J.H.; Ferrone, C.R.; Ferrone, S. Defective HLA class I antigen processing machinery in cancer. *Cancer Immunology, Immunotherapy* **2018**, *67*, 999-1009, doi:10.1007/s00262-018-2131-2.
35. Binnewies, M.; Roberts, E.W.; Kersten, K.; Chan, V.; Fearon, D.F.; Merad, M.; Coussens, L.M.; Gabrilovich, D.I.; Ostrand-Rosenberg, S.; Hedrick, C.C.; et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nature Medicine* **2018**, *24*, 541-550, doi:10.1038/s41591-018-0014-x.
36. Munn, D.H.; Bronte, V. Immune suppressive mechanisms in the tumor microenvironment. *Current Opinion in Immunology* **2016**, *39*, 1-6, doi:10.1016/j.coi.2015.10.009.
37. Lee, K.A.; Thomas, A.M.; Bolte, L.A.; Björk, J.R.; de Ruijter, L.K.; Armanini, F.; Asnicar, F.; Blanco-Miguez, A.; Board, R.; Calbet-Llopart, N.; et al. Cross-cohort gut microbiome associations with immune checkpoint inhibitor response in advanced melanoma. *Nature Medicine* **2022**, *28*, 535-544, doi:10.1038/s41591-022-01695-5.
38. Saal, J.; Ellinger, J.; Ritter, M.; Brossart, P.; Hölzel, M.; Klümper, N.; Bald, T. Pretreatment albumin is a prognostic and predictive biomarker for response to atezolizumab across solid tumors. *Clinical & Translational Immunology* **2023**, *12*, e1472, doi:doi.org/10.1002/cti2.1472.
39. Abuhelwa, A.Y.; Kichenadasse, G.; McKinnon, R.A.; Rowland, A.; Hopkins, A.M.; Sorich, M.J. Machine Learning for Prediction of Survival Outcomes with Immune-Checkpoint Inhibitors in Urothelial Cancer. *Cancers* **2021**, *13*, 2001, doi:10.3390/cancers13092001.
40. Kilickap, S.; Demirci, U.; Karadurmus, N.; Dogan, M.; Akinci, B.; Sendur, M.A.N. Endpoints in oncology clinical trials. *J buon* **2018**, *23*, 1-6.
41. Sena, L.A.; Fountain, J.; Isaacsson Velho, P.; Lim, S.J.; Wang, H.; Nizialek, E.; Rath, N.; Nussenzweig, R.; Maughan, B.L.; Velez, M.G.; et al. Tumor Frameshift Mutation Proportion Predicts Response to Immunotherapy in Mismatch Repair-Deficient Prostate Cancer. *The Oncologist* **2020**, *26*, e270-e278, doi:10.1002/onco.13601.
42. Martincorena, I.; Raine, K.M.; Gerstung, M.; Dawson, K.J.; Haase, K.; Van Loo, P.; Davies, H.; Stratton, M.R.; Campbell, P.J. Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **2017**, *171*, 1029-1041.e1021, doi:10.1016/j.cell.2017.09.042.
43. Hundal, J.; Kiwala, S.; McMichael, J.; Miller, C.A.; Xia, H.; Wollam, A.T.; Liu, C.J.; Zhao, S.; Feng, Y.-Y.; Graubert, A.P.; et al. pVACtools: A Computational Toolkit to Identify and Visualize Cancer Neoantigens. *Cancer Immunology Research* **2020**, *8*, 409-420, doi:10.1158/2326-6066.CIR-19-0401.
44. Trolle, T.; McMurtrey, C.P.; Sidney, J.; Bardet, W.; Osborn, S.C.; Kaeffer, T.; Sette, A.; Hildebrand, W.H.; Nielsen, M.; Peters, B. The Length Distribution of Class I-Restricted T Cell Epitopes Is Determined by Both Peptide Supply and MHC Allele-Specific Binding Preference. *The Journal of Immunology* **2016**, *196*, 1480-1487, doi:10.4049/jimmunol.1501721.
45. Duan, F.; Duitama, J.; Al Seesi, S.; Ayres, C.M.; Corcelli, S.A.; Pawashe, A.P.; Blanchard, T.; McMahon, D.; Sidney, J.; Sette, A.; et al. Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *Journal of Experimental Medicine* **2014**, *211*, 2231-2248, doi:10.1084/jem.20141308.
46. Ghorani, E.; Rosenthal, R.; McGranahan, N.; Reading, J.L.; Lynch, M.; Peggs, K.S.; Swanton, C.; Quezada, S.A. Differential binding affinity of mutated peptides for MHC class I is a predictor of survival in advanced lung cancer and melanoma. *Annals of Oncology* **2018**, *29*, 271-279, doi:10.1093/annonc/mdx687.
47. Lang, F.; Riesgo-Ferreiro, P.; Löwer, M.; Sahin, U.; Schrörs, B. NeoFox: annotating neoantigen candidates with neoantigen features. *Bioinformatics* **2021**, *37*, 4246-4247, doi:10.1093/bioinformatics/btab344.
48. Chiaro, J.; Kasanen, H.H.E.; Whalley, T.; Capasso, C.; Grönholm, M.; Feola, S.; Peltonen, K.; Hamdan, F.; Hernberg, M.; Mäkelä, S.; et al. Viral Molecular Mimicry Influences the Antitumor Immune Response in Murine and Human Melanoma. *Cancer Immunology Research* **2021**, *9*, 981-993, doi:10.1158/2326-6066.Cir-20-0814.
49. Łuksza, M.; Riaz, N.; Makarov, V.; Balachandran, V.P.; Hellmann, M.D.; Solovyyov, A.; Rizvi, N.A.; Merghoub, T.; Levine, A.J.; Chan, T.A.; et al. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* **2017**, *551*, 517-520, doi:10.1038/nature24473.

50. Richman, L.P.; Vonderheide, R.H.; Rech, A.J. Neoantigen Dissimilarity to the Self-Proteome Predicts Immunogenicity and Response to Immune Checkpoint Blockade. *Cell Systems* **2019**, *9*, 375-382.e374, doi:10.1016/j.cels.2019.08.009.
51. Maggs, L.; Sadagopan, A.; Moghaddam, A.S.; Ferrone, S. HLA class I antigen processing machinery defects in antitumor immunity and immunotherapy. *Trends in Cancer* **2021**, *7*, 1089-1101, doi:10.1016/j.trecan.2021.07.006.
52. Carapito, R.; Radosavljevic, M.; Bahram, S. Next-Generation Sequencing of the HLA locus: Methods and impacts on HLA typing, population genetics and disease association studies. *Human Immunology* **2016**, *77*, 1016-1023, doi:10.1016/j.humimm.2016.04.002.
53. Wang, H.; Liu, B.; Wei, J. Beta2-microglobulin(B2M) in cancer immunotherapies: Biological function, resistance and remedy. *Cancer Letters* **2021**, *517*, 96-104, doi:10.1016/j.canlet.2021.06.008.
54. Shin, D.S.; Zaretsky, J.M.; Escuin-Ordinas, H.; Garcia-Diaz, A.; Hu-Lieskovan, S.; Kalbasi, A.; Grasso, C.S.; Hugo, W.; Sandoval, S.; Torrejon, D.Y.; et al. Primary Resistance to PD-1 Blockade Mediated by JAK1/2 Mutations. *Cancer Discov* **2017**, *7*, 188-201, doi:10.1158/2159-8290.Cd-16-1223.
55. Spranger, S.; Bao, R.; Gajewski, T.F. Melanoma-intrinsic β -catenin signalling prevents anti-tumour immunity. *Nature* **2015**, *523*, 231-235, doi:10.1038/nature14404.
56. Liu, S.; Liu, T.; Jiang, J.; Guo, H.; Yang, R. p53 mutation and deletion contribute to tumor immune evasion. *Frontiers in Genetics* **2023**, *14*, doi:10.3389/fgene.2023.1088455.
57. Piro, G.; Carbone, C.; Carbognin, L.; Pilotto, S.; Ciccicarese, C.; Iacovelli, R.; Milella, M.; Bria, E.; Tortora, G. Revising PTEN in the Era of Immunotherapy: New Perspectives for an Old Story. *Cancers* **2019**, *11*, 1525, doi:10.3390/cancers11101525.
58. Roh, W.; Chen, P.-L.; Reuben, A.; Spencer, C.N.; Prieto, P.A.; Miller, J.P.; Gopalakrishnan, V.; Wang, F.; Cooper, Z.A.; Reddy, S.M.; et al. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Science Translational Medicine* **2017**, *9*, doi:10.1126/scitranslmed.aah3560.
59. Pore, N.; Wu, S.; Standifer, N.; Jure-Kunkel, M.; de los Reyes, M.; Shrestha, Y.; Halpin, R.; Rothstein, R.; Mulgrew, K.; Blackmore, S.; et al. Resistance to Durvalumab and Durvalumab plus Tremelimumab Is Associated with Functional STK11 Mutations in Patients with Non-Small Cell Lung Cancer and Is Reversed by STAT3 Knockdown. *Cancer Discovery* **2021**, *11*, 2828-2845, doi:10.1158/2159-8290.Cd-20-1543.
60. Dong, Z.-Y.; Zhong, W.-Z.; Zhang, X.-C.; Su, J.; Xie, Z.; Liu, S.-Y.; Tu, H.-Y.; Chen, H.-J.; Sun, Y.-L.; Zhou, Q.; et al. Potential Predictive Value of TP53 and KRAS Mutation Status for Response to PD-1 Blockade Immunotherapy in Lung Adenocarcinoma. *Clinical Cancer Research* **2017**, *23*, 3012-3024, doi:10.1158/1078-0432.Ccr-16-2554.
61. Forschner, A.; Hilke, F.J.; Bonzheim, I.; Gschwind, A.; Demidov, G.; Amaral, T.; Ossowski, S.; Riess, O.; Schroeder, C.; Martus, P.; et al. MDM2, MDM4 and EGFR Amplifications and Hyperprogression in Metastatic Acral and Mucosal Melanoma. *Cancers* **2020**, *12*, 540-540, doi:10.3390/cancers12030540.
62. Chowell, D.; Morris, L.G.T.; Grigg, C.M.; Weber, J.K.; Samstein, R.M.; Makarov, V.; Kuo, F.; Kendall, S.M.; Requena, D.; Riaz, N.; et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* **2018**, *359*, 582-587, doi:10.1126/science.aao4572.
63. Pierini, F.; Lenz, T.L. Divergent Allele Advantage at Human MHC Genes: Signatures of Past and Ongoing Selection. *Molecular Biology and Evolution* **2018**, *35*, 2145-2158, doi:10.1093/molbev/msy116.
64. Chowell, D.; Krishna, C.; Pierini, F.; Makarov, V.; Rizvi, N.A.; Kuo, F.; Morris, L.G.T.; Riaz, N.; Lenz, T.L.; Chan, T.A. Evolutionary divergence of HLA class I genotype impacts efficacy of cancer immunotherapy. *Nature Medicine* **2019**, *25*, 1715-1720, doi:10.1038/s41591-019-0639-4.
65. Oey, O.; Khattak, M.A.; Abed, A.; Meniawy, T.; Reid, A.; Calapre, L.; Millward, M.; Gray, E. Patient human leukocyte antigen (HLA) genotype may predict response to anti-programmed death receptor 1 (anti-PD1) in advanced melanoma. *Journal of Clinical Oncology* **2021**, *39*, e21512-e21512, doi:10.1200/JCO.2021.39.15_suppl.e21512.
66. Deng, Y.; Song, Z.; Huang, L.; Guo, Z.; Tong, B.; Sun, M.; Zhao, J.; Zhang, H.; Zhang, Z.; Li, G. Tumor purity as a prognosis and immunotherapy relevant feature in cervical cancer. *Aging (Albany NY)* **2021**, *13*, 24768-24785, doi:10.18632/aging.203714.

67. Song, L.; Cohen, D.; Ouyang, Z.; Cao, Y.; Hu, X.; Liu, X.S. TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-seq data. *Nature Methods* **2021**, *18*, 627-630, doi:10.1038/s41592-021-01142-2.
68. Finotello, F.; Mayer, C.; Plattner, C.; Laschober, G.; Rieder, D.; Hackl, H.; Krogsdam, A.; Loncova, Z.; Posch, W.; Wilflingseder, D.; et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Medicine* **2019**, *11*, 34, doi:10.1186/s13073-019-0638-6.
69. Sturm, G.; Finotello, F.; Petitprez, F.; Zhang, J.D.; Baumbach, J.; Fridman, W.H.; List, M.; Aneichyk, T. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. *Bioinformatics* **2019**, *35*, i436-i445, doi:10.1093/bioinformatics/btz363.
70. Cui, C.; Xu, C.; Yang, W.; Chi, Z.; Sheng, X.; Si, L.; Xie, Y.; Yu, J.; Wang, S.; Yu, R.; et al. Ratio of the interferon- γ signature to the immunosuppression signature predicts anti-PD-1 therapy response in melanoma. *npj Genomic Medicine* **2021**, *6*, 7-7, doi:10.1038/s41525-021-00169-w.
71. Ayers, M.; Lunceford, J.; Nebozhyn, M.; Murphy, E.; Loboda, A.; Kaufman, D.R.; Albright, A.; Cheng, J.D.; Kang, S.P.; Shankaran, V.; et al. IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade. *The Journal of Clinical Investigation* **2017**, *127*, 2930-2940, doi:10.1172/JCI91190.
72. Shklovskaya, E.; Lee, J.H.; Lim, S.Y.; Stewart, A.; Pedersen, B.; Ferguson, P.; Saw, R.P.; Thompson, J.F.; Shivalingam, B.; Carlino, M.S.; et al. Tumor MHC Expression Guides First-Line Immunotherapy Selection in Melanoma. *Cancers* **2020**, *12*, 3374, doi:10.3390/cancers12113374.
73. Zhao, Y.; Cao, Y.; Chen, Y.; Wu, L.; Hang, H.; Jiang, C.; Zhou, X. B2M gene expression shapes the immune landscape of lung adenocarcinoma and determines the response to immunotherapy. *Immunology* **2021**, *164*, 507-523, doi:10.1111/imm.13384.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.