





Machine learning identifies clinical tumor mutation landscape pathways of resistance to checkpoint inhibitor therapy in NSCLC

Vitalay Fomin ^{1,2}, WeiQing Venus So ¹, Richard Alex Barbieri,³
Kenley Hiller-Bittrolff,³ Elina Koletou ⁴, Tiffany Tu,¹ Bruno Gomes,⁵ James Cai,¹
Jehad Charo ⁶

To cite: Fomin V, So WV, Barbieri RA, *et al.* Machine learning identifies clinical tumor mutation landscape pathways of resistance to checkpoint inhibitor therapy in NSCLC. *Journal for ImmunoTherapy of Cancer* 2025;**13**:e009092. doi:10.1136/jitc-2024-009092

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/jitc-2024-009092>).

VF and WVS contributed equally.

VF and WVS are joint first authors.

Accepted 03 February 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

For numbered affiliations see end of article.

Correspondence to

Dr Jehad Charo;
jehad.charo@roche.com

ABSTRACT

Background Immune checkpoint inhibitors (CPIs) have revolutionized cancer therapy for several tumor indications. However, a substantial fraction of patients treated with CPIs derive no benefit or have short-lived responses to CPI therapy. Identifying patients who are most likely to benefit from CPIs and deciphering resistance mechanisms is therefore essential for developing adjunct treatments that can abrogate tumor resistance.

Patients and methods In this study, we used a machine learning approach that used the US-based nationwide de-identified Flatiron Health and Foundation Medicine non-small cell lung carcinoma (NSCLC) clinico-genomic database to identify genomic markers that predict clinical responses to CPI therapy. In total, we analyzed data from 4,433 patients with NSCLC.

Results Analysis of pretreatment genomic data from 1,511 patients with NSCLC identified. Of the 36 genomic signatures identified, 33 exhibited strong predictive capacity for CPI response (n=1150) compared with chemotherapy response (n=361), while three signatures were prognostic. These 36 genetic signatures had in common a core set of four genes (*BRAF*, *BRIP1*, *FGF10*, and *FLT1*). Interestingly, we observed that some (n=19) of the genes in the signatures (eg, *TP53*, *EZH2*, *KEAP1* and *FGFR2*) had alternative mutations with contrasting clinical outcomes to CPI therapy. Finally, the genetic signatures revealed multiple biological pathways involved in CPI response, including *MAPK*, *PDGF*, *IL-6* and *EGFR* signaling.

Conclusions In summary, we found several genomic markers and pathways that provide insight into biological mechanisms affecting response to CPI therapy. The analyses identified novel targets and biomarkers that have the potential to provide candidates for combination therapies or patient enrichment strategies, which could increase response rates to CPI therapy in patients with NSCLC.

INTRODUCTION

Lung cancer is the leading cause of cancer-related mortality worldwide, and approximately 85% of these tumors are of non-small cell lung cancer (NSCLC) histology.^{1,2} The discovery and subsequent regulatory approval

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Checkpoint inhibitors (CPIs) have transformed cancer treatment, particularly for non-small cell lung cancer, by achieving prolonged responses in many patients. However, a significant portion of patients either do not respond or develop resistance to CPI therapy.

WHAT THIS STUDY ADDS

⇒ Among the 36 genomic signatures identified, 33 demonstrated strong predictive capacity for CPI response relative to chemotherapy response, while three signatures (hotspot granular signatures 3, 4, and 9) did not meet the predictive threshold. The study revealed core genes such as *BRAF*, *BRIP1*, *FGF10*, and *FLT1* that are crucial for CPI outcomes. The findings also highlight that different mutations within the same gene can lead to contrasting responses to CPI therapy. Moreover, the study uncovers multiple biological pathways, including *MAPK*, *PDGF*, *IL-6*, and *EGFR* signaling, associated with CPI response, providing new insights into the mechanisms of resistance and potential targets for combination therapies.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The identification of specific genetic signatures and pathways associated with CPI response can guide the development of new predictive biomarkers and targeted therapies, potentially improving patient stratification and treatment outcomes. These insights could lead to personalized treatment strategies and the development of combination therapies to overcome CPI resistance, ultimately enhancing clinical practice and informing policy decisions on cancer treatment protocols.

of checkpoint inhibitors (CPI) has revolutionized treatment of diverse cancer types including lung cancer, by achieving prolonged responses in many patients.^{3–7} Despite the unprecedented, prolonged response rates to

CPIs, a substantial fraction of patients become resistant to CPI therapy.⁸

There is, therefore, an urgent and unmet need to advance our understanding of CPI resistance mechanisms to promote the development of predictive biomarkers and identify potential targets for novel therapies. Extensive efforts are underway to identify predictive biomarkers using various omics, histopathologic, clinical, and computational approaches.^{9–12} Presently, the only approved biomarkers for CPI therapy are programmed death ligand-1 (PD-L1), microsatellite instability and tumor mutational burden (TMB).¹³ However, these approved biomarkers do not provide the much-needed mechanistic insight behind CPI response. To gain insight into CPI resistance mechanisms in NSCLC and discover predictive genomic biomarkers, we used an interpretable machine learning (ML) framework that uses a genetic algorithm (GA) for feature selection. The patient data was obtained from the nationwide (US-based) de-identified Flatiron Health and Foundation Medicine NSCLC clinico-genomic database (FH-FMI CGDB). Using our GA framework, we discovered tumor intrinsic genetic signatures that can predict CPI response. We discovered that different mutations within the same gene may lead to opposing clinical responses. Finally, we used several network-based algorithms to discover pathways associated with CPI response. Our findings provide insights into potential mechanisms underlying CPI response in NSCLC, which are essential for developing new therapeutics that target CPI resistance and discovering biomarkers for the prediction of CPI response.

RESULTS

Patient cohorts and characteristics

Among 7,360 patients with NSCLC in the FH-FMI CGDB, we found 1,150 advanced stage patients who had received monotherapy with anti-programmed cell death protein 1 (PD-1) or anti-PD-L1 antibodies (referred to hereafter as CPI) and had either response or resistance to CPI (figure 1A,B). The de-identified longitudinal patient data originated from approximately 280 US cancer clinics (~800 sites of care). These patients had their pretreatment tumor specimens sequenced by the Foundation Medicine panels.¹⁴ We selected patients treated with CPIs as monotherapy to eliminate the confounding effect of combination therapy on CPI response.

The patients were categorized into the two response groups using clinical outcome data^{15–16} available in CGDB. Almost all patients with at least one line of therapy had real-world progression (rwP)¹⁵ data, and about 84% had real-world response (rwR)¹⁶ data. The first response group consisted of 574 patients who had durable rwR and did not show rwP during the entire study duration of 270 days were considered to have response (referred to as “response” patients).^{12–15} The second group had 576 patients who did not have rwR and showed rwP within the study and were considered to be CPI-resistant (referred

to as “resistance” patients) (figure 1A,B). Comparison of the clinical characteristics of response and resistance patients revealed no significant difference in gender, race, or disease stage (figure 1C), but did show significant differences in age, TMB, smoking status, and histology. These differences are consistent with previous reports which show the frequency of response is greater in older patients,¹⁷ patients with higher TMB,¹⁸ different histology, as well as patients with a history of smoking.^{12–19} The majority of these patients received CPI as their first-line therapy (89% 1 L, 10% 2 L).

Mutational landscape of patients with NSCLC

We characterized the mutational landscape of our patient group (figure 2A). The five most frequent short variants (single nucleotide variants and short insertions/deletions) were in *TP53*, *KRAS*, *KMT2D*, *KEAP1* and *STK11* (figure 2B). The five most frequent amplifications were of *SOX2*, *SDHA*, *RICTOR*, *KLHL6* and *PIK3CA* (figure 2C), while the top five most frequent deletions were of *CDKN2A*, *CDKN2B*, *PTEN*, *STK11* and *TEK* (figure 2D). The five most frequent rearrangements were in *CDKN2A*, *FGFR1*, *LRP1B*, *EGFR* and *TP53* (figure 2E). A comparison of the top 12 altered genes between response and resistance patients did not reveal any obvious differences in mutation frequencies between the two cohorts, with the exceptions of *STK11* mutations, which occurred at higher frequency in resistance patients (14% vs 22%) (online supplemental figure S1C). Cumulatively, the mutational landscape analysis confirmed high-frequency mutations known in NSCLC but did not reveal gross differences between response and resistance patients.

Mutations affecting response to CPIs

The conventional approach to correlating mutations with an observable characteristic involves testing for statistical significance of the occurrence of each mutation in its corresponding group. This type of analysis is useful but does not provide sufficient insight into the complex biology of response to CPI. Moreover, different mutations within the same gene can result in distinct biological outcomes,^{20–23} which can also be influenced by tissue type, tumor microenvironment, and/or xenobiotics.²⁴ With the above considerations in mind, we categorized the mutational data using three different approaches (Methods). The first approach was the traditional classification which notes if the gene is mutated or wild-type (WT), and is referred to hereafter as binary classification/input. The second approach divides gene mutations into high-frequency mutations (HFM), and low-frequency mutations (LFM) (eg, *TP53*HFM or *TP53*LFM), and is referred to after this as HFM/LFM classification/input. The third approach further separates HFM mutations and adds the location of the amino acid change (eg, *TP53* HFM 178, *TP53* HFM 248), and is referred to after this as hotspot granular classification/input. Classifying the mutations into HFM/LFM and hotspots allowed us to further delineate the underlying biological mechanism as HFM tend

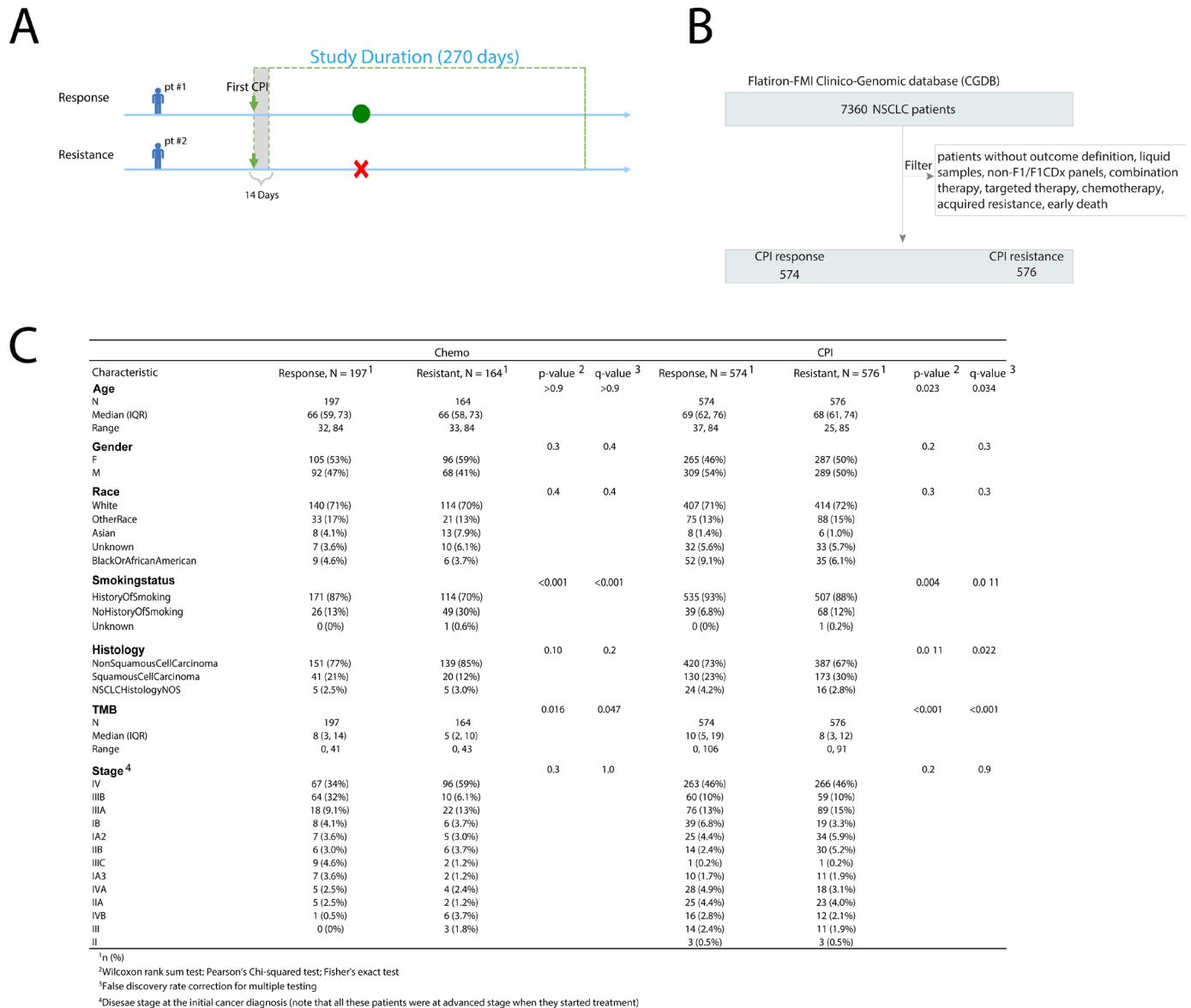


Figure 1 Definition, selection and characteristics of patient outcome groups. (A) Schematic representation of definition for response and resistance patient groups. The long blue horizontal arrows represent the patient's journey over time. The study duration (270 days), in which treatment outcomes were investigated, is marked by a green dotted open box. Green vertical arrows represent the first time a patient was treated with a CPI, and the greyed-out area represents buffer time (14 days) during which treatment outcomes were not considered because they were likely influenced by earlier treatments. The green circle represents real-world response from CPI therapy, while the red X represents disease progression. Patients in the response group exhibited durable clinical benefit without evidence of disease progression during the 270-day study period. In contrast, patients with innate resistance—defined as those who did not benefit from CPI and experienced disease progression within the same time frame—were included in the analyses. Patients with acquired resistance, characterized by an initial response followed by disease progression within the 270-day study period, were excluded. (B) Schematic representation of the number of selected patients in the CPI treatment cohort, based on the exclusion criteria depicted in the filter box in the scheme (more details in methods). (C) Clinical characteristics of the patient cohorts divided into chemotherapy (depicted as chemo) and checkpoint inhibitors (depicted as CPI). CPI, checkpoint inhibitors; NSCLC, non-small cell lung cancer; TMB, tumor mutational burden.

to be gain of function mutations and LFM tend to be loss of function mutations.²⁵ We used the Fisher's exact test within each set (ie, binary, HFM/LFM and hotspot granular) to check for mutation enrichment in response or resistance patients. The analysis revealed that in the binary classification set, *BRAF*, *BRIP1* and *ASXL1* mutations were associated with response (false discovery rate

[FDR]=0.053), while *STK11* and *CDKN2B* mutations were associated with resistance (FDR=0.053 and 0.056, respectively; figure 3A, online supplemental table S1A). Consistent with the *BRAF* effect on CPI response in NSCLC, we obtained similar data when analyzing an independent melanoma cohort (data not shown). In the HFM/LFM classification, we found that *BRAF*_HFM and *BRIP1*_LFM

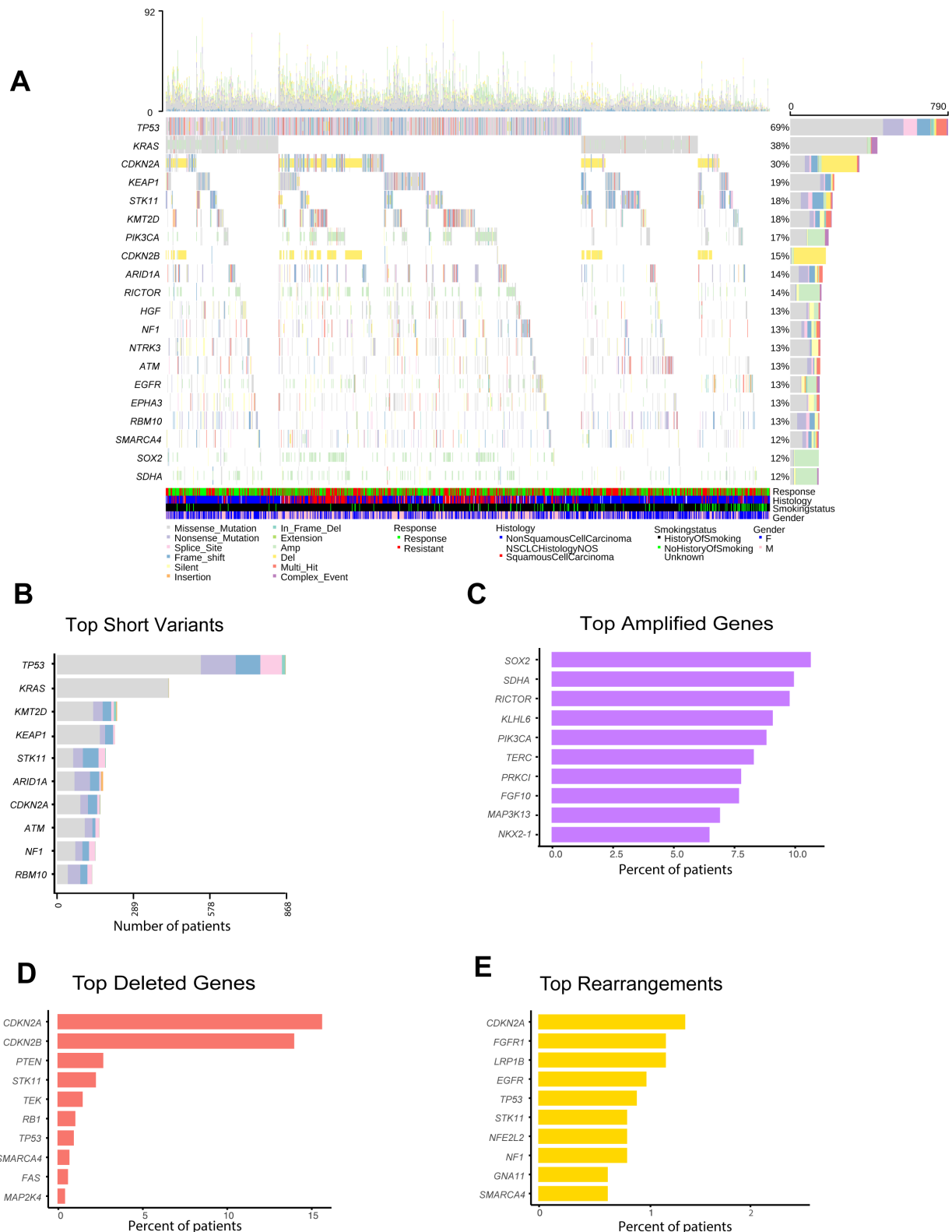


Figure 2 Mutational landscape of the CPI cohort. (A) Oncoplot of the CPI cohort (n=1,150), depicting the top 20 altered genes and some clinical characteristics. Columns represent patients, rows represent genes. The top bar plot shows mutation count and type (missense, splice site, frame shift, etc) per patient. The middle section indicates mutations in the specified gene in patients. The bottom heatmap shows clinical characteristics (outcome, histology, smoking status, and gender) with a color scheme depicted at the bottom of the figure. The right-side bar graph summarizes the total count and proportion of mutation types for each gene. (B) Bar graph of top short variants in the CPI cohort (n=1,150) with color stacking representing mutation types within each gene (same color scheme as in A). (C) Bar graph showing prevalence (%) of the top 10 amplified genes. (D) Bar graph showing prevalence (%) of the top 10 deleted genes. (E) Bar graph showing prevalence (%) of the top 10 rearrangements. CPI, checkpoint inhibitor.

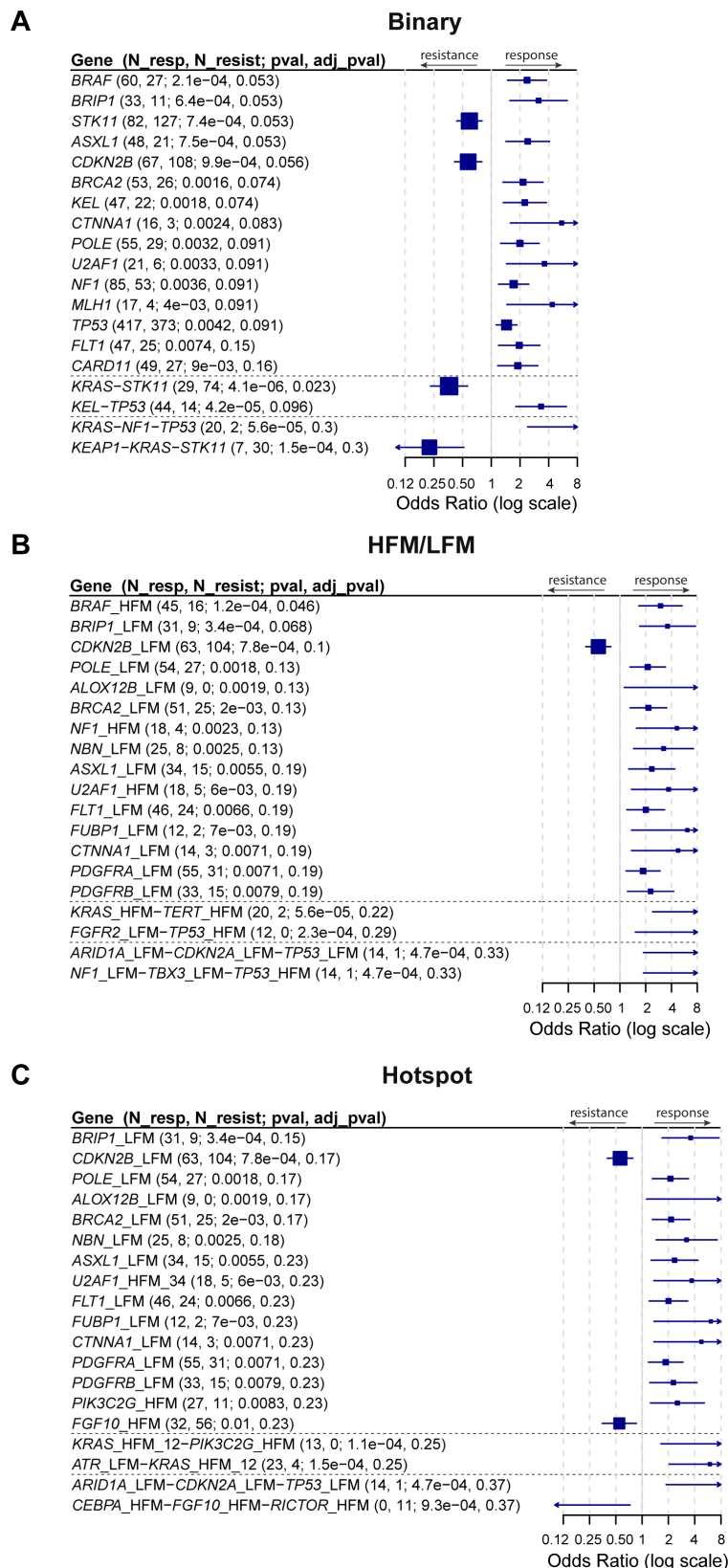


Figure 3 Mutation association with outcome in the 1,150-patient cohort. (A–C) Forest plot of the top-ranked genes of binary (A), HFM/LFM (B) and hotspot (C) mutations by Fisher's exact test comparing mutation rates between patients in response and resistance groups. Each plot shows the top 15 single gene and top 2 double and triple gene mutations co-occurrences separated by dashed horizontal lines. Each plot shows the gene name (Gene) and the number of patients who responded (N_resp) or were resistant (N_resist) having that single or combination mutation, with the x-axis representing log-scaled OR. OR > 1 indicates association of the mutation with response to CPI, and < 1 indicates association with resistance to CPI. CPI, checkpoint inhibitors; HFM, high frequency mutations; LFM, low frequency mutations.

were associated with response (FDR=0.046 and 0.068, respectively; [figure 3B](#), online supplemental table S1B), and *CDKN2B_LFM* (FDR=0.1) was associated with resistance. In the hotspot-granular set, we found no statistically significant results after multiple testing adjustments ([figure 3C](#), online supplemental table S1C). Since mutations tend to co-occur with other genomic alterations, we examined if there were any significant pairs or triples of co-occurring mutations associated with CPI response. In binary classification, we found that *STK11* and *KRAS* co-occurrence was the only significant pair (FDR=0.023) associated with resistance, and co-occurring mutations in *KEL* and *TP53* were only marginally associated with response (FDR=0.096; [figure 3A](#), online supplemental table S2A). No significant co-occurring mutations in HFM/LFM or hotspot classifications were observed, but the top result of co-occurring mutations in *KRAS_HFM* and *TERT_HFM* (FDR=0.22) and *KRAS_HFM_12* and *PIK3C2G_HFM* (FDR=0.25) were associated with response ([figure 3B,C](#), online supplemental table S2B-C). Additionally, we found no significant triplet mutations, but top results included co-occurring mutations in *KRAS-NFI-TP53* (FDR=0.3) and *ARID1A_LFM-CDKN2A_LFM-TP53_LFM* (FDR=0.33) ([figure 3A,B](#), online supplemental table S3A-B).

***BRAF*, *BRIP1*, *STK11*, *ASXL1*, *CDKN1B*, and co-occurring *STK11-KRAS* mutations affect overall survival in patients treated with CPIs**

While we found a significant association ([figure 3](#)) between several mutations and response to CPI, we wanted to investigate if mutations in the genes listed above, affect overall survival (OS) in the 1,150 patient cohort. Additionally, we wanted to explore if the mutations could be predictive for CPI-specific response or prognostic (independent of treatment). We therefore compared OS of patients treated with CPI monotherapy with 361 patients treated with chemotherapy (patient characteristics are in online supplemental figure S1A, complete analysis results are in Table S4). OS analysis revealed significantly greater OS among CPI-treated patients with mutations in *BRAF* and a trend for *BRIP1* ($p=0.09$), and significantly lower OS with mutations in *CDKN2B* and *STK11* ($p=0.02$) ([figure 4A](#)). No significant change in survival was observed in patients with *ASXL1* mutations ($p=0.5$) ([figure 4A](#)). A significantly lower OS among CPI-treated patients was also observed for patients with co-occurring mutations in *KRAS* and *STK11* ($p=0.0013$) ([figure 4B](#)). No significant change in survival was observed in chemo-treated patients with these gene mutations, indicating that the survival benefit is CPI-specific ([figure 4A,B](#)). Of note, we did observe a sharp decrease in survival among chemo-treated patients with *STK11* or *STK11* and *KRAS* co-occurring mutations ([figure 4A,B](#)), but since the survival lines crossed, it was not statistically significant, though it might be clinically relevant. Although we did not find statistically significant triplet mutations in the response data analysis ([figure 3A](#)), the top triplet mutations in

KRAS, *NFI* and *TP53* showed a significant ($p=0.00041$) increase in OS ([figure 4C](#)). The statistically non-significant result in the response data analysis might be due to the small sample sizes ($n=20$ vs 2).

We further investigated the correlation of the OS to the mutations in the genes described above, with an extended cohort of patients. This extended cohort of 4,433 patients was selected from the 7,360 patients with NSCLC in the CGDB the same way as the 1,150-patient cohort ([figure 1B](#)). The only difference was that there was no outcome definition criteria filter for the larger cohort (ie, regardless of drug response data availability) and regardless of their death status. These included 2,699 patients treated with CPI monotherapy and 1,734 chemo-treated patients. Examining the effect of the mutations analyzed in [figure 4](#), on survival in the extended cohort confirmed that CPI-treated patients with mutations in *BRAF*, *BRIP1* and *ASXL1* showed significant increases in OS when compared with WT patients ($p=0.011$, $p=0.038$, $p=0.04$, respectively) (online supplemental figure S2A). Notably, there was no significant difference in OS in chemo-treated patients with these mutations versus those without (online supplemental figure S2A). CPI-treated patients with *STK11* or *CDKN2B* mutations, and co-occurring mutations in *STK11* and *KRAS*, showed a significant reduction in OS ($p=0.0047$, $p=0.00013$ and $p=0.0015$, respectively; online supplemental figure S2A,B). We did, however, observe that chemo-treated patients with these mutations also showed significant decreases in OS ($p<0.0001$, $p=0.0068$ and $p=0.0018$, respectively), suggesting these mutations may not have a CPI-specific effect but are prognostic (online supplemental figure S2A,B). Although the top triplet mutations in *KRAS*, *NFI* and *TP53* showed a significant substantial increase in OS ($p=0.00041$) in the 1,150-patient cohort, the increase was only marginally significant ($p=0.056$) in the larger CPI-treated cohort ([figure 4C](#) and online supplemental figure S2C). The same trend was observed for co-occurring mutations in *KEL* and *TP53* ([figure 3](#) and data not shown). Cumulatively, the results of [figures 3](#) and [4](#) are in concordance, and the effect of the mutations on OR of response and resistance presented in [figure 3](#) and OS presented in [figure 4](#) followed the same pattern and directionality.

Machine learning workflow identifies genetic markers of CPI outcomes

We used a GA to evolve virtual individuals, each represented by a set of mutated genes, over multiple generations. The predictive power of each virtual individual for CPI response was assessed using Naive Bayes models, with cross-validation (CV) accuracy scores serving as the evaluation metric. The fittest individuals were selected iteratively for subsequent generations. From the top 5% of these individuals, based on their accuracy scores, we identified 12 distinct clusters using k-means clustering. Within each cluster, the set of mutated genes that appeared in more than 50% of the members was identified as the

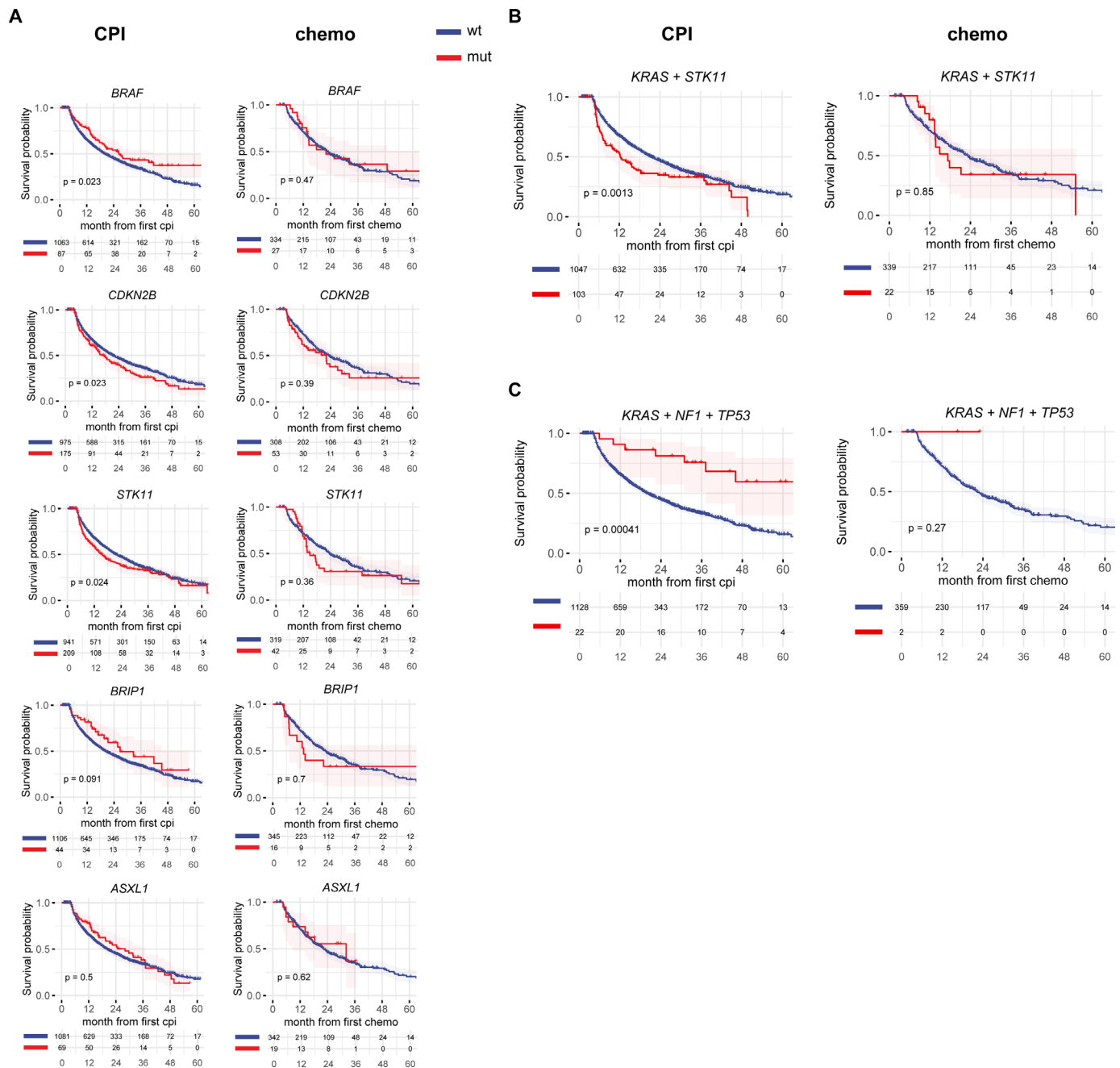


Figure 4 Overall survival of the 1,150-CPI patient cohort and 361-chemo patient cohort with mutations found to be significantly associated with CPI outcome. (A–C) Kaplan-Meier survival curves of overall survival in patients treated with CPI (left plots) or chemo (right plots) with wild-type (wt; blue lines) or mutated (mut; red lines) gene in top single (A), double (B) or triple (C) gene mutations demonstrating a trend for statistical significance in figure 3. The bottom table in each Kaplan-Meier plot shows the number of patients at risk at each time point (in month). P values indicated inside each graph were from the log-rank test comparing patient groups within each plot. CPI, checkpoint inhibitors.

genetic signature for that cluster (online supplemental figure S3, figure 5A). We identified several mutations affecting outcomes to CPI including co-occurring mutations, which suggests that these genes in tandem or as a single gene may mediate outcomes to CPI. This suggested that decoding the biological mechanism(s) behind CPI outcomes requires broadening the analysis beyond one or two genes in isolation. Such a broad analysis aims at identifying a genetic signature(s) (i.e., a collection of tumor mutations) that can be used to predict response

to CPI. TMB is an established and important biomarker that correlates with response to CPI therapy,²⁶ which we recently validated for predicting response in NSCLC.¹² We therefore used TMB as a benchmark for our subsequent results. The TMB model trained on our NSCLC cohort showed a receiver operating characteristic (ROC) with an area under the curve (AUC) score of 0.6 (figure 5B–E). Using an ML analysis workflow (described in Methods), we examined 977 out of the 1,150 patients as the initial training set. The results yielded 36 genetic signatures

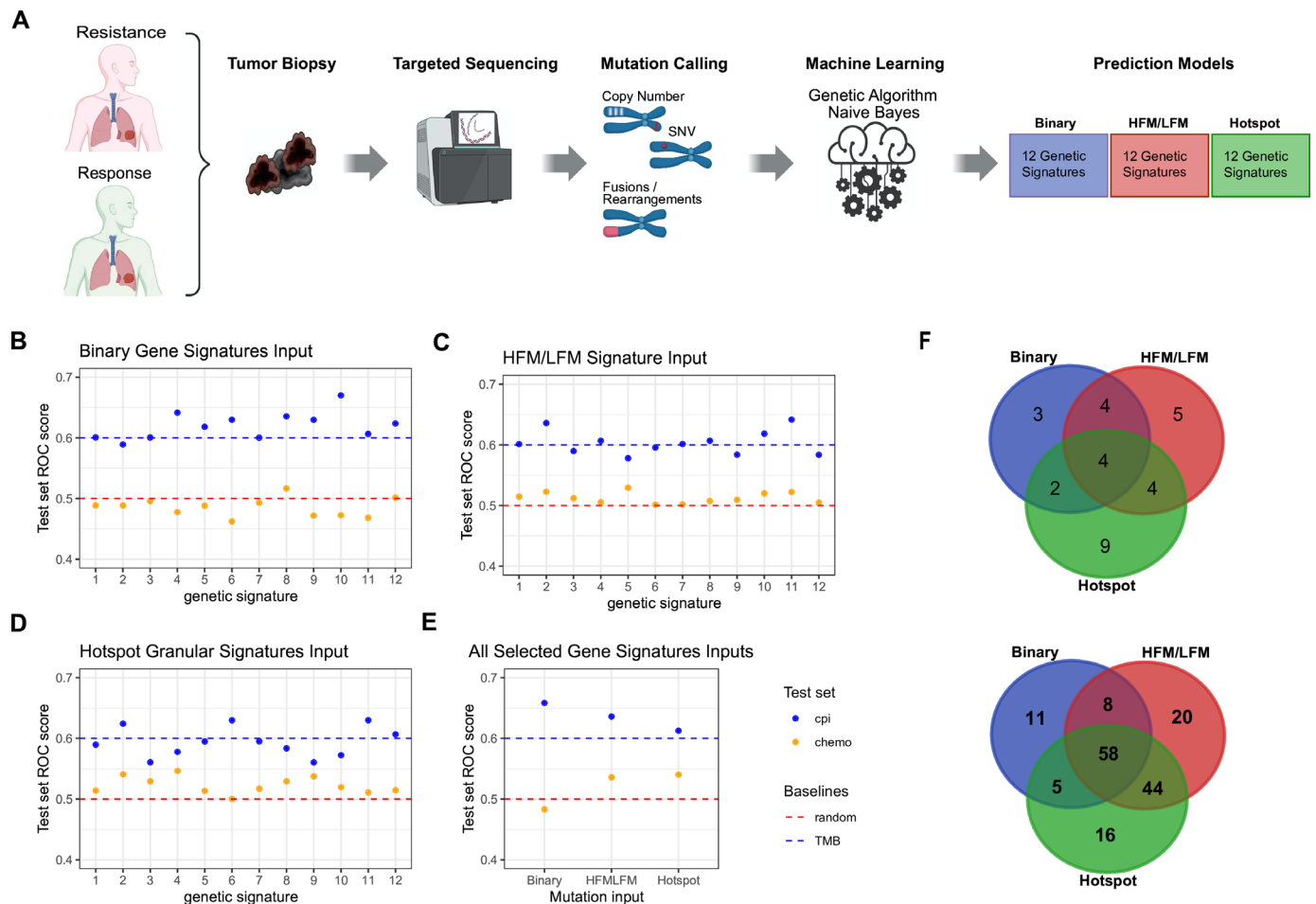


Figure 5 Machine learning algorithm identifies 36 predictive genetic signatures with a core of shared genes. (A) Schematic depiction of the machine learning algorithm. (B–E) Plots showing the area under the ROC curve (AUC) from the independent validation set ($n=173$ for CPI and $n=361$ for chemo-treated patients) using the genetic signatures from different input types. (B) Binary, (C) HFM/LFM and (D) hotspot granular and (E) AUC scores for the three inputs with all unique gene mutation features (82 for binary, 130 for HFM/LFM, and 123 for hotspot granular) from the 12 models combined into one model for each input type. For each graph, the blue dashed horizontal line represents the score for the TMB model, and the red dashed line represents the 50% AUC (depicted as random chance). (F) The top Venn diagram represents the gene overlap between the 36 individual genetic signatures across binary, HFM/LFM and hotspot granular inputs, showing that four genes are included in every single genetic signature of the 36 signatures. The bottom Venn diagram represents overlap across the unique genes in the 12 signatures of the three inputs (82 binary, 130 HFM/LFM, and 123 hotspot). Detailed information about the overlap is shown in online supplemental figure S4. CPI, checkpoint inhibitors; HFM, high frequency mutations; LFM, low frequency mutations; ROC, receiver operating characteristic; TMB, tumor mutational burden.

consisting of 12 signatures for each input: binary, HFM/LFM and hotspot (Table S5). The average CV ROC AUC scores of these signatures was 0.7 ± 0.02 (Table S6). We then used an independent test (validation) set of 173 patients to examine the association of the 36 genetic signatures to CPI response. The average resultant ROC AUC scores were 0.6 ± 0.02 compared with 0.6 with TMB (figure 5B–D, Table S6), which underscores the predictive power and generalizability of these genetic signatures. The average AUC scores for predicting response in 361 chemo-treated patients using the 36 genetic signatures was 0.5 ± 0.02 , which is consistent with random chance, confirming that the genetic signatures were CPI response-specific. These results are consistent with previously published AUC values for TMB^{27 28} and provide evidence that these genetic signatures perform as well as the established TMB

biomarker in prediction of CPI response. However, these results provide an additional, important advantage since the predictive mutations can be explored for mechanistic understanding of outcomes to CPI, which is not possible with TMB.

Characterization of 36 genetic signatures revealed a core set of genes most frequently associated with CPI response

Since each of the three ML inputs generated 12 separate predictive genetic signatures (Table S5), we first examined if each genetic signature within the inputs is diverse. Indeed, the heatmap analysis in online supplemental figure S4 shows that the 36 signatures are diverse with some overlaps between them (online supplemental figure S4). Next, we assessed the frequency of each gene mutation feature across the 36 genetic signatures that might

indicate their utility for CPI outcomes. We found that within each of the binary, HFM/LFM and hotspot inputs, there was an overlap of 13, 17, and 20 genes, respectively, within the 12 genetic signatures (online supplemental figure S4 and Table S7A). These genes were present in all 12 genetic signatures within each input, suggesting that these are necessary in generating predictive genetic signatures and play an important role in CPI outcomes. Among these overlaps, four genes (BRIP1, FLT1, FGF10 and BRAF) were shared (figure 5F top, online supplemental figure S4 and Table S7B), representing a core set of mutations most important for predicting CPI outcomes. Next, we identified 82, 130, and 123 unique gene mutation features contained in each of the binary, HFM/LFM and hotspot inputs, respectively (Table S7C). Using these gene mutation features, we built three models, one for each input type that have ROC AUC scores of 0.61–0.66 (figure 5E and Table S7D). Among the unique gene mutation features, 58 genes were overlapping, further distinguishing a core set of mutations central to predicting

CPI outcomes (figure 5F, Table S7E). Overall, this ML approach revealed a significant number of genes, many of which are previously unreported, which may play a role in outcomes to CPI therapy.

Relative contributions of individual gene mutations to overall CPI response

To interpret the biological impact of these mutations, we used the linear Naïve Bayesian model. It enabled us to: (1) quantify the relative contribution of each gene mutation feature to CPI outcomes, and (2) associate each mutation with response or resistance. Accordingly, we were able to assign a numerical rank to each gene mutation feature by their importance to CPI outcomes (figure 6A, Table S8A). In binary genetic signatures, we found that *CTNNA1* was the top contributor to response and *RARA* was the top contributor to resistance. In HFM/LFM genetic signatures, we found that *ALOX12B* LFM mutations were the top contributor to response and *MYC* LFM mutations were the top contributor to resistance. In



Figure 6 ML identifies genetic markers associated with outcomes to CPI. (A) Waterfall plot of the top 10 genes of the 36 ML-derived genetic signatures ranked by feature importance using the three different inputs (left panel for binary, middle for HFM/LFM and right for hotspot). Positive values indicate association with response (green bars), and negative values with resistance (red bars) to CPI. (B) Plot representing feature importance in genes in which HFM and LFM mutations within the same gene have divergent contribution to CPI outcome, with red associated with resistance and green with response. CPI, checkpoint inhibitors; HFM, high frequency mutations; LFM, low frequency mutations; ML, machine learning.

hotspot genetic signatures, we found that *ALOX12B* LFM was the top contributor to response and *EGFR* mutation at amino acid 746 was the top contributor to resistance. *CTNNA1* and *RARA* and other mutations that were prominent in the binary signature were also highly ranked in HFM/LFM and hotspot indicating consistency between inputs of the impact of these genes on response to CPI.

Remarkably, we observed a set of genes that had opposite effects on CPI response, depending on the type of mutation (figure 6B, Table S8B). For example, in HFM/LFM input, we found that LFM mutations in *FGFR2* were associated with response, while HFM mutations were associated with resistance (figure 6B). Furthermore, in hotspot inputs, we found that certain *TP53* HFM mutations were associated with response (*TP53_HFM_331*), while other *TP53* HFM mutations were associated with resistance (*TP53_HFM_152*) (Table S8A). These findings suggest that different mutations within the same gene can lead to opposite outcomes in CPI-treated patients.

Topology/network-based pathway enrichment analysis reveals pathways associated with CPI response

While it is important to identify individual mutations (or combinations) that affect CPI outcomes, there is a need to understand the biological processes affected by these unique gene mutations and how they affect CPI response. To explore the biological processes influenced by the genetic signatures, we investigated whether these signatures fall into meaningful biological processes. We applied a network/topology-based pathway analysis that takes the upstream and downstream gene–gene or protein–protein interactions into account. We used six different algorithms to perform topology-assisted pathway analysis (please see Methods), using the Computational Biology Methods for Drug Discovery (CBDD) R package²² (please see Methods) which facilitated the discovery of additional pathways that would otherwise not be detected with an enrichment-only approach. With binary input, the top pathways across six algorithms were associated with PDGF, oxidative stress, prolactin, thrombopoietin, AGTR1, IL-3 and YAP/TAZ and IL-6 signaling (figure 7A, Table S9A). For HFM/LFM input, the uppermost pathways were associated with androgen, IL-6, PDGF, epigenetics, oxidative stress, ErbB2 and EGFR signaling (figure 7B, Table S9B). For hotspot input, we found the uppermost pathways to be associated with PDGF, androgen, IL-6, AGTR1, HGF, EGFR and IL-2 signaling (figure 7C, Table S9C). The enrichment of PDGF and IL-6 as well as other pathways was shared across the binary, HFM/LFM and hotspots inputs, indicating consistency between these inputs to enable identification of pathways relevant for response to CPI (figure 7D, Table S9D). Since there was an overlap of 58 genes between the three ML inputs (figure 5F), we examined the pathways that were enriched in this overlap. This approach revealed pathways associated with IFN alpha/beta, PDGF, MAPK, oxidative stress, ALK, ErbB2, IL-6, leptin, prolactin and IL-2 signaling (figure 7E, Table S9E). The overall highest-ranking pathways among all

inputs (figure 7A–C) were associated with PDGF, oxidative stress, androgen, MAPK, prolactin, AGTR1, IL-6, EGFR and YAP/TAZ signaling (figure 7E, Table S9E). In summary, our pathway analysis reveals several tumor intrinsic pathways like oxidative stress and PDGF involved in outcomes to CPI as well as several previously reported immune response-related pathways like interleukin (IL)-6 and IL-2 to be involved in outcomes to CPI.

Higher IL-6 serum levels are associated with progressive disease in NSCLC

Since IL-6 signaling appeared in the results of analyses in multiple pathways, we investigated its importance in a cohort of patients with NSCLC treated with atezolizumab. We found that serum levels of IL-6 were elevated in both stable disease (SD) and progressive disease (PD) patients when compared with partial response (PR) patients using either RECIST (Response Evaluation Criteria in Solid Tumors) (figure 8A) or irRC (immune-related response criteria) (data not shown). Furthermore, OS analysis of patients with NSCLC treated with atezolizumab revealed a significantly longer survival in patients with lower IL-6 serum levels, which is consistent with previous and more recently published reports^{29–31} (figure 8B). The involvement of IL-6 in resistance to atezolizumab validates our ML algorithm, which allows greater delineation of the biological pathways and mechanism(s) behind resistance or response to CPI.

DISCUSSION

In this study, we report a comprehensive analysis of tumor genomic determinants that affect outcomes of patients with NSCLC to CPI.

We classified patients by outcomes over an extended period, which enabled a different perspective compared with objective response rate, with the extra length of time in our analyses better representing real-world outcomes of CPI therapy. We focused on innate resistance rather than acquired resistance. By targeting innate resistance, we aimed to identify genomic mechanisms intrinsic to tumors that drive resistance to CPI therapy. This distinction is crucial for understanding pre-existing actionable resistance pathways. Finally, the exclusion of early death events in our analyses increases the focus on drug response and excludes patients who had extremely advanced disease and morbidities, allowing us to focus on predictive rather than prognostic effects.¹² This approach enables more focused data analysis for biomarkers of response. The biomarkers identified from this analysis included those (eg, *STK11* and *STK11-KRAS*), which have been previously reported to be associated with response to CPI.^{32,33}

Using a traditional statistical test comparing response and resistance, we found that mutations in single genes (*BRAF*, *BRIP1*, *ASXL1*, and *CDKN2B*) as well as co-occurring mutations (*TP53-KEL* and *TP53-KRAS-NF1*) correlate with CPI response, including OS. *CDKN2B* deletion

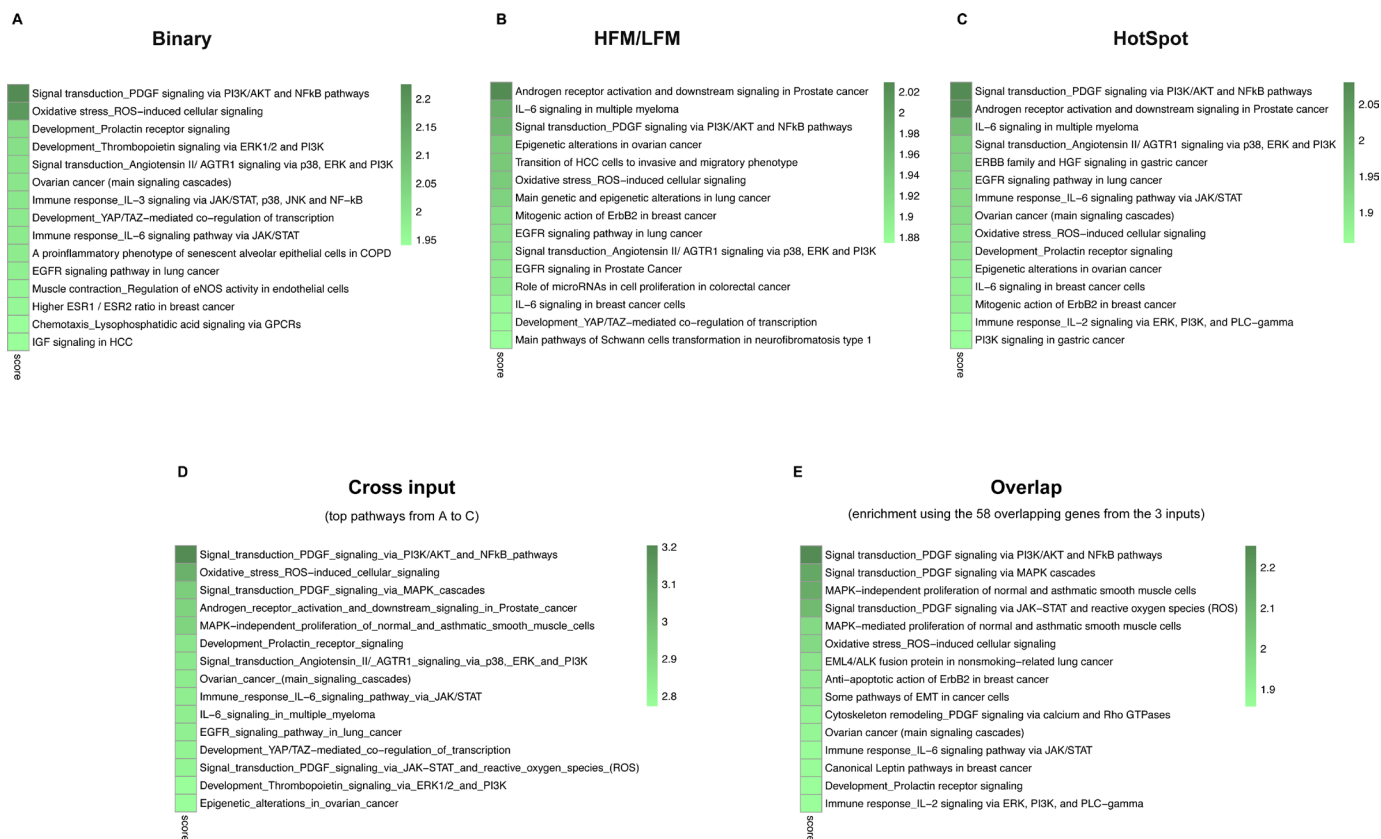


Figure 7 Pathway analysis of predictive genetic signatures associated with CPI outcomes. This figure illustrates the immune response and other biological pathways associated with CPI outcomes, identified through pathway analysis of predictive genetic signatures. The top 15 pathways were derived using the Computational Biology Methods for Drug Discovery pathway analysis, which used six different network-based algorithms. Pathways are ranked by the lowest p values (derived from the hypergeometric test) across the algorithms and are sorted according to the composite score (see Methods for details). (A–C) The top pathways identified from the (A) binary, (B) HFM/LFM, and (C) hotspot inputs. (D) The top pathways from A to C (cross input pathways). (E) Pathways enriched using the 58 overlapping genes among the genetic signatures from the three inputs, as shown in the bottom Venn diagram of [figure 5F](#) (overlap pathways). Cross input pathways (D) refer to the pathways that are consistently identified across all three input types (binary, HFM/LFM, and hotspot), indicating their broad relevance. Overlap pathways (E) refer to the pathways identified from the subset of genes that are common across all three genetic signature inputs, highlighting specific pathways that are crucial due to the shared gene set. CPI, checkpoint inhibitors; HFM, high frequency mutations; LFM, low frequency mutations.

correlated with decreased OS in CPI-treated patients, which is in keeping with a previous report in advanced urothelial carcinoma.^{34 35} However, *CDKN2B* deletion, in our studies with the larger cohort of patients, correlated with a decrease in OS in chemo-treated patients as well, suggesting this to be a prognostic rather than predictive marker for CPI outcomes. We did not observe any significant OS differences between patients with mutations in *TP53* alone but found a significant difference if *TP53* mutations co-occurred with *KRAS* and *NFI*. This data from the co-occurring mutations suggests that specific combinations of mutations may affect outcomes, and provides insight beyond conventional single-gene analyses.

We classified the mutations as HFM and LFM, where HFM and LFM tend to have gain of function and loss of function characteristics, respectively.²⁵ We found that not all HFM and LFM or hotspots mutations within the same gene lead to similar clinical outcomes with CPI, including *TP53*³⁶ and *KEAP1*.^{32 33 36} *KEAP1* was

previously reported to be associated with resistance to CPI therapy.^{32 33} However, our analysis showed that not all mutations in *KEAP1* lead to resistance, and while *KEAP1* HFM 320 was associated with resistance, LFM were associated with response. These divergent therapy outcomes to different mutations within the same gene were also found in other genes. The number of genes indicates that this observation is not anecdotal, and that binary classification (mutated vs WT) of genes is insufficient for exploring the role of mutations in outcomes to CPI, as two different mutations in the same gene may lead to opposite clinical consequences. The interpretation of how different mutations, in the same gene, lead to opposite outcomes requires experimental investigation as it can be due to different active mechanisms or non-causal factors. These divergent outcomes from different mutations in the same gene also suggest that it is necessary to assess each individual mutation and its effect on CPI outcome, but this is presently not feasible

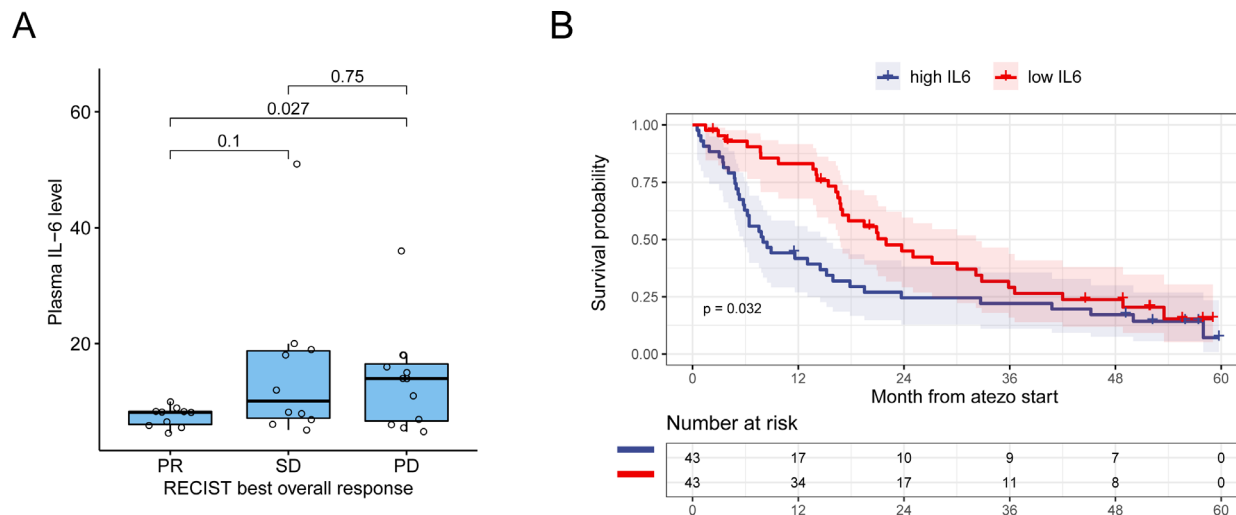


Figure 8 IL-6 validation in the atezolizumab study shows high serum levels correlate with disease progression. (A) Boxplot depicting baseline (cycle 1 day one pre-dose) serum IL-6 levels in patients whose response to atezolizumab was assessed by the RECIST criteria (CR, PR, SD or PD). Unpaired two-sided t-test p values were shown above the boxes. (B) Kaplan-Meier curves of overall survival of patients treated with atezolizumab (trial PCD4989G) comparing patients with high (blue) and low (red) IL-6 serum levels. The number of patients at risk is given below the Kaplan-Meier plot. P value indicated inside the graph was from the log-rank test comparing the two patient groups. CR, complete response; IL, interleukin; PD, progressive disease; PR, partial response; RECIST, Response Evaluation Criteria in Solid Tumors; SD, stable disease.

due to the high number of these variants that results in limitations in statistical power.

These signatures, as well as a shared signature identified from the common gene set, could inform a patient enrichment strategy for patients with NSCLC for treatment with CPI. The signature genes were analyzed with network algorithms to identify the biological pathways influencing CPI response. This identified several pathways significantly associated with CPI outcomes. Notably, some of these pathways were previously reported to be involved with CPI outcomes like IFN-alpha/beta,^{37 38} EGFR^{39 40} and IL-6,^{29 41 42} further credence to our analysis. The association of these pathways with CPI outcomes may have therapeutic implications, as there are already several Food and Drug Administration-approved drugs that can target these pathways being tested in combination with CPIs to increase response rate.⁴³ Additionally, using data available from a phase I clinical study, we confirmed IL-6 involvement in CPI resistance and showed that higher baseline serum levels of IL-6 in CPI-treated patients were associated with disease progression and poor OS. This is in line with the reports on higher IL-6 serum levels leading to worse outcomes in metastatic urothelial bladder cancer and in metastatic renal cell carcinoma treated with anti-PD-L1,^{29 44 45} implying that IL-6 signaling might be an important target in several cancers and is not limited to NSCLC. Furthermore, our pathway analysis showed that WNT/Beta-catenin, PTEN and LKB1/STK11-related pathways may affect CPI outcomes, consistent with previous reports,^{32 44 46 47} further sustaining our observations. Additionally, the BRAF/MAPK pathway has been previously reported in preclinical and clinical settings to be of central importance in inhibiting T cell-dependent antitumor activity, and inhibiting BRAF/MEK

concurrently with CPI has been demonstrated to have synergistic effects in bringing about tumor regression.^{48–50}

In addition to these previously reported pathways, we identified several resistance pathways that were not previously highlighted to affect CPI outcomes, with the most significant pathways among these being related to PDGF, ESRI, YAP/TAZ, EGFR, IGF, and AGTR1 signaling. Our pathway analyses can potentially facilitate biomarker selection. There are only a limited number of biomarkers presently available for CPI (eg, PD-L1, TMB and mismatch repair deficiency) all sharing the characteristic of being related to TMB/immunogenicity.² Identification of novel biomarkers can improve patient selection and guide better outcomes. This includes potential novel biomarkers in pathways such as PDGF, Leptin, IGF, Estrogen, and HGF that can be measured in either the tumor samples or potentially using non-invasive serum samples.

Our analyses provide a tumor-genome intrinsic view, as we used the baseline (pre-CPI treatment) tumor genetic profile to predict outcomes of CPI therapy. The pathway analyses inform on the processes that are altered in the tumor microenvironment, but an extended analysis of the processes can be obtained from RNA sequencing analysis that measures the messenger RNA expression in heterogeneous (immune, fibroblasts, endothelial and tumor) cell populations. There are several other aspects to consider when interpreting the data from the current analysis. This includes the limited data set of 286 genes, variations in tissue source (primary tumor vs metastatic site) and lines of therapy as well as the possibility that these mutations could have been selected at least in part due to their increased immunogenicity as targets and not because of their functional biological contribution.⁵¹

This latter possibility, however, is rather low, because it would require that such mutations create a *de novo* epitope presented and recognized by the immune system in all affected patients. Furthermore, the genetic data used in pathway analysis cannot reveal whether the pathway is inhibited or activated, which requires further experimental studies or previously reported differentiation in a relevant context to interpret the direction of the impact.⁴² Notwithstanding these caveats, it is encouraging that with these limited sets of genes, the models have an efficiency similar to TMB and could account for up to 65% of clinical outcomes. As more whole exome/genome sequenced patients become available, it might be possible to further refine our predictive model and deepen the biological understanding of CPI outcomes with this approach.

Of note, the adjusted FDR values from multiple Fisher's exact tests may have underestimated the significance of some mutations or co-occurring mutations, as they assume independence. It is accepted, however, that mutations are not independent of each other,⁵² and as we demonstrated above, some combinations of mutations can affect outcomes.

Patients can be stratified based on mutations identified in the signatures, and these signatures could potentially be combined with TMB and PD-L1 to create a model for patient stratification. While testing these approaches is beyond the scope of this paper, four alternative strategies with increased complexity can be used to facilitate the evaluation and use of the identified signatures in clinical trial design. First, patients can be stratified based on the four genes (BRIP1, FLT1, FGF10, and BRAF) identified in the signatures (online supplemental figure S4). Second, patients can be stratified based on one or a combination of the pathways identified, such as IL-6, or a combination of pathways (figure 7). Third, a predictive model based on the 36 predictive signatures (figure 5) can be used to stratify the patients. Fourth, our earlier data demonstrated that combining PD-L1 and TMB¹² can increase the likelihood of response. Accordingly, the identified mutations, pathways, and predictive signatures could potentially be combined with TMB and PD-L1 to create a more advanced predictive model for patient stratification.

There are presently no approved therapies to circumvent or reverse resistance to CPI, which is in part due to the lack of understanding the mechanism behind this resistance. In this report, we expand our understanding of CPI outcomes by using a GA framework, which resulted in the identification of multiple predictive genes, mutations and biological pathways with several models that can predict a patient's outcome to anti PD-(L)1 therapy using pretreatment data. The identification of mutations, pathways and potential biomarkers also provide testable hypotheses for the mechanism behind CPI resistance. Our identification of these novel targets can be used to explore predictive biomarkers, treatment selection, and patient enrichment strategies, all possibly leading to better outcomes in CPI-treated patients.

METHODS

Research design and patient cohorts

We used integrated FH-FMI CGDB which consists of longitudinal clinical data derived from electronic health records curated via technology-enabled abstraction and were linked to genomic data from the same patient who underwent FMI comprehensive genomic profiling tests by de-identified, deterministic matching.⁵³ The study included 7,360 patients diagnosed with NSCLC from January 2003 to May 2021 who underwent FoundationOne (F1) or FoundationOne CDx (F1CDx) tests (ie, only patients with T7 or DX1 baitset data) and had treatment data. We used rWR data¹⁵ (CR, PR, SD, PD or not evaluable (NE) on a specified date) and rwP data (progressed or not progressed at a specified date)¹⁶ for defining the clinical outcome groups. To study the clinical benefit of CPI therapy and to have a more balanced number of patients in each of these outcome group, CR, PR and SD from the rwR data were considered as having tumor response. Progressive disease and rwP, death or a change to a non-CPI treatment line within the study duration was considered to be disease progression. A patient was considered to have a "durable clinical benefit", hereafter referred to as "response", if there was tumor response and no disease progression starting 14 days after initiation of CPI to the end of the study duration (270 days). A patient was considered to have "innate resistance", hereafter simply referred to as "resistance" if there was disease progression without any tumor response during the study duration. Acquired resistant patients, defined as having response followed by disease progression within the 270-day study duration, were not included. Using these definitions, 574 patients who had rwP were assigned to the response group and 576 patients with disease progression, as defined above, were assigned to the resistance group. Genetic variations in these two outcome groups were then investigated by ML methods to identify key variables and pathways that might explain the drug response. No explicit cohort balancing of the two outcome groups was performed. However, baseline covariates (age, race, sex, histology, smoking history, initial disease stage) were examined and found to have no correlation with the outcomes after FDR correction. For determining the study duration, sensitivity analysis using periods ranging from 120 to 365 days, in ~2–3 months increments, was performed. The study duration of 270 days resulted in an optimal balance in patient number in each outcome group and is clinically relevant. Disease progression within the first 14 days after CPI treatment was not counted, since it may not reflect the effect of the current treatment (recommendation by Flatiron Health).

To better distinguish CPI-specific predictive effects of biomarkers from prognostic effects (independent of treatment type), we also analyzed standard-of-care chemotherapy cohorts. We also removed patients with early deaths (patients who died in the first 18 weeks after treatment start) to better delineate the two effects. This is because it has been reported that the CPI and

chemotherapy survival curves did not differentiate well until after about 18 weeks.^{5 12} To maximize the number of patients with NSCLC with FMI (Foundation Medicine) targeted DNA sequencing data, we selected patients with data from the FoundationOne (F1; T7 baitset) or FoundationOne CDx (F1CDx; DX1 baitset) panels for solid tumors. We then focused on the 288 genes profiled on both panels.

Treatment data

CPI analysis included monotherapies nivolumab (~31%), pembrolizumab (~62%), atezolizumab (~5%), durvalumab (~2%). Chemotherapy patients included patients with cytotoxic monotherapy or combinations with other cytotoxics or targeted therapies who did not have “immunotherapy” in the patient’s history in the database. For patients who had multiple lines of CPI or chemotherapy, their first CPI or chemotherapy records were used for analysis.

Tumor mutation burden

TMB (number of mutations per Mb) calculated from targeted DNA sequencing using F1 or F1CDx panels⁵³ on tumor biopsies from all analyzed patients as provided by FMI. TMB data from the most recent specimens collected before the start of treatment was used. Research Use Only calculations based on FMI’s research algorithm used at the time of collection were analyzed.⁵⁴ The TMB-only model was trained using logistic regression on the same train/validation patient split used for the mutation signature models.

IL-6 analysis

Plasma IL-6 protein levels measured by Luminex multiplex assay were from the atezolizumab phase I trial (NCT01375842).⁵⁵ Samples from cycle 1 day one predose time point from patients with NSCLC were analyzed. The survival curves were performed using the Kaplan-Meier method.

Mutation data preparation

No explicit functional filters were applied to mutations during the ML analysis, allowing the inclusion of all mutation types (eg, short variants, copy number variations, and rearrangements) to comprehensively identify genomic signatures associated with CPI resistance or response. To ensure data quality, technical filters were implemented. These included restricting the analysis to mutations identified using validated bait sets (T7 or DX1) from Foundation Medicine and excluding synonymous mutations to focus on variants with potential functional implications. This approach enabled robust identification of clinically relevant genomic features while minimizing the inclusion of technical or biologically irrelevant artifacts. After associating a CPI-resistance label to each patient, the mutations for each CPI patient were filtered to remove synonymous mutations and then used as three different sets of inputs in ML with different aggregations of mutations: (1) binary input (all mutations in a gene are aggregated into

one feature per gene, ie, whether a gene is mutated or not), (2) HFM or LFM (high/low frequency mutations; see the next paragraph for definition) inputs (all HFM in a gene are aggregated into one feature and LFM into another; 1–2 features per gene), and (3) hotspot granular input (the HFM are further separated into individual mutations). In the 1,150 patients (March 2015 to May 2021) used for analysis in this study, there were 288 binary inputs, 460 HFM/LFM inputs, and 528 hotspot granular inputs. The input data set was randomly split into training and validation subsets, stratified by CPI-resistance label, leaving 977 training patients and 173 validation patients.

HFM/LFM input categorization was performed as follows: (1) we calculated the total number of amino acids mutated per gene as X (2) we calculated the frequency of mutations in each gene (ie, how many patients had any mutation in that gene) as Y. (3) From 1 to 2 we calculated the average amino acid mutation rate:

$$\text{average aa mutation rate} = \frac{\% \text{ patients with any mutation in the gene } (Y)}{\text{total number of aa mutated in the gene } (X)}$$

HFM was assigned to any mutation that had 2× the average mutations per that specific amino acid and had more than/equal to nine mutations in that gene. LFM was assigned to any mutation that had lower than 2× the average mutations per that specific amino acid and/or had less than nine mutations. Amplifications were considered as HFM and deletions as LFM; the rationale behind this was that LFM tend to be loss of function and high-frequency mutations tend to be gain of function. Hotspot granular classification was the same as in HFM/LFM, but the amino acid mutation location to any HFM was added (eg, TP53_178, meant HFM in TP53 in amino acid 178). Any HFM that lacked the information about amino acid location was considered an amplification mutation. To assign HFM, LFM and hotspot mutations 4,433 patients with NSCLC from the CGDB were used.

Pathway analysis using CBDD

The predictive genes in the genetic signatures from the ML output were used to identify extended biological network member genes using six different network-based algorithms (CBDD R package,⁵⁶ which used the Metabase network and pathway data (Thomson Reuters)). The algorithms used were Network Propagation,⁵⁷ Interconnectivity,⁵⁸ Overconnectivity⁵⁹ Hidden nodes⁶⁰ GeneMania,⁶¹ and Causal Reasoning.^{62 63} The top 100 genes resulting from each algorithm were then used to run a pathway enrichment analysis on the Metabase pathways. Each pathway then got an enrichment p value from each of the six network algorithms.

A composite score was then calculated for each pathway using the six enrichment p values. A common way of using p values for scoring is the “ $-\log_{10}P$ ” transformation, in which higher value indicates more significance. However, the relationship between statistical and biological significance is not linear. For example, a p value of $1e-50$ ($-\log_{10}p=50$) is not 10 times more significant than

$1e-5$ ($-\log_{10}p=5$). They both mean a high level of significance. To more closely capture the biological and statistical significance reflected by p values, we performed a transformation on $-\log_{10}P$ (Equation 1) such that the final score better reflected the level of significance, which increased and then plateaued as $-\log_{10}P$ increases (Table S10 and see the interactive transformation curve on this page).

$$\text{score} = 10\sqrt{\frac{x}{12+x}} \quad x = -\log_{10}P \quad (1)$$

We then scaled the scores from each algorithm and calculated the composite score per pathway using the harmonic sum of scores from the different algorithms (Equation 2), where k is the rank of the algorithm score per pathway.

$$\sum_{k=1}^{\infty} \frac{1}{k} \quad (2)$$

Genetic algorithm feature selection

Our ML approach applied Naive Bayes models with a Bernoulli prior as the model architecture, embedded in a GA for feature selection to reveal predictive genetic signatures.^{64,65} We used GA because of the large set of available features (binary:288, HFM/LFM:460, hotspot:528) relative to the number of patients ($n=977$ in the training set), as it can efficiently find the most predictive feature combinations (out of binary: 2^{288} , HFM/LFM: 2^{460} , hotspot: 2^{528}) while including multifeature interactions.^{64,65} Naive Bayes models were chosen for the simplicity of their internal state, resistance to overtraining, and interpretability.⁶⁶ Random Forest and other ensemble-based methods were attempted but found to require heavy hyperparameter tuning to avoid overtraining during the GA search. The Bernoulli prior is appropriate for binary input data and includes a penalty term for the feature absent of mutation, differentiating it from a multinomial prior. The details are outlined in the subsequent paragraphs, accompanied by a schematic diagram in online supplemental figure S3. Python scripts (online supplemental file 3) for implementation and performance evaluation are also provided.

In this study, we defined a GA individual as a subset of the available input features of the data set, represented as a binary mask (mutated or not) over all features. The fitness of each individual is calculated based on the predictive power of a model which uses only the features contained in the subset and is trained to predict the response to CPI category, “response” or “resistance”.

For each GA individual, the fitness was the CV accuracy score of a Naive Bayes model on the training set. The accuracy score was class-balanced to avoid favoring “response” over “resistance,” or vice versa. The CV uses stratification to keep the same fraction of “response” and “resistance” patients in each fold. The number of folds used was 5, a compromise to keep the number of patients in each fold high while keeping the number of folds high enough to be confident in the result. The training data

was shuffled for each individual before CV to avoid overfitting on CV folds during GA optimization.

In each generation of the GA, individuals were selected for mutation or crossover using fitness proportionate selection,⁶⁷ which samples individuals probabilistically based on their fitness in order to maintain diversity in the GA breeding pool. The breeding pool of each generation is supplemented by a set of the highest fitness individuals from all previous generations. Mutation occurs by randomly removing or adding features to the subset while conserving the average number of features. The average number of features is conserved during mutation by partitioning the probability of mutation between adding features and removing features in order to retain (on an average) the same number of features removed and added. Without this partition, mutation would tend to increase the size of models with less than 50% of the total features used and decrease the size of models with more than 50% of the features used, regardless of the fitness of the result. Crossover occurs by randomly selecting each feature flag of the binary mask from two previous individuals and does not include further correction: crossover on average will produce offspring with the number of features midway between the parents.

The GA procedure was run with the following parameters set. The GA was run for 200 generations, each with a population of 1,000 individuals. Larger populations and larger numbers of generations were not found to produce different results, as the GA was able to find an optimum within this time. The first generation was generated randomly such that on average, 10% of the features were included in each individual. This fraction was chosen to correspond roughly to the number of features at the end of the GA. During mutation, on average, 1% of features were removed or added to an individual. The top 200 individuals over all generations were added to the breeding pool for each generation for a total breeding pool size of 1,200 in each generation after the first. To generate each successive generation, 600 individuals were formed by mutating an individual from the breeding pool while the remaining 400 of each generation were formed by crossover of two previous individuals, a relative fraction chosen to slightly favor mutation in order to increase diversity in the population.

The GA was run 10 times for each mutational input categorization (binary per-gene, gain or loss of function, hotspot). Since each run of the GA tended to find separate local optima, these 10 runs along with the clustering technique described below were used to identify multiple local optima that were too distant for a single GA run to identify.

After 10 runs of 200 generations with a population size of 1,000 there were 2,000,000 GA individuals which are available. The top 5% (100,000) of all individuals, based on their CV score, were selected and then clustered according to the similarity of the features they contain. The clustering was done using a simple k-means clustering with 12 clusters to account for the expected 10

separate local optima (one from each run of the GA) plus some leeway for outliers.

In each cluster, the set of features that appeared in more than 50% of cluster members was considered the characteristic set of features for that cluster. A final model was then trained on each of the 12 characteristic sets and evaluated on the validation set.

Feature importance calculation

In order to visualize the importance of individual mutations on the prediction outcomes, the internal state of the 36 Bernoulli Naive Bayes models were converted into their equivalent linear (logistic regression) coefficients⁶⁸. This conversion for each gene mutation feature is outlined below.

$$\begin{aligned} \text{coef}(\text{resp}) &= \ln \left(\frac{P(\text{mut}|\text{resp})}{1-P(\text{mut}|\text{resp})} \right) \\ \text{coef}(\text{resist}) &= \ln \left(\frac{P(\text{mut}|\text{resist})}{1-P(\text{mut}|\text{resist})} \right) \\ \text{coef} &= \text{coef}(\text{resp}) - \text{coef}(\text{resist}) \end{aligned}$$

Where “P(mut|resp)” is the probability of mutation for the response group, “P(mut|resist)” is the probability of mutation for the resistance group. “Coef(resp)” is the linear coefficient for the feature in the response group, while “coef(resist)” is the linear coefficient for the feature in the resistance group. “coef” is the final linear coefficient of the feature. “coef” > 0 indicates association of the feature with response while “coef” < 0 indicates association with resistance. The sum of coefficients of all the features determines the prediction of response or resistance for a patient.

When $\sum_{\text{mut}} \text{coef} > \text{threshold}$, response is predicted over resistance.

Where the threshold is the intercept of the model, which is a trained parameter of the model in addition to the coefficients.

Mutation co-occurrence analysis

To assess potential relationships between co-occurrence of mutation pairs/triples and CPI response, Fisher’s exact tests were performed, unless there was zero count(s) in the 2×2 contingency table, then the Mantel-Haenszel method with continuity correction of 0.5 was used. Multiple testing adjustment of p values used the false discovery rate method. The contingency table structure for the number of patients with mutation pairs/triples is as below:

	Response	Resistance
All gene mutation features mutated		
Not all mutated (any or none mutated)		

Each row/column sum of the above table had to be at least 5 to remove very rare mutational combinations.

Author affiliations

¹Roche Pharmaceutical Research and Early Development, Data & Analytics, Roche Innovation Center New York, Little Falls, New Jersey, USA

²Numenos, New York, NY, USA

³Data Science, Capgemini Engineering, Paris, France

⁴Roche Pharmaceutical Research and Early Development, Data and Analytics, Roche Innovation Center Basel, Basel, Switzerland

⁵Roche Pharmaceutical Research and Early Development Oncology, Roche Innovation Center Basel, Basel, Switzerland

⁶Roche Pharmaceutical Research and Early Development Oncology, Roche Innovation Center Zurich, Schlieren, Switzerland

X Elina Koletou @mynameiselina

Acknowledgements The authors gratefully acknowledge the editorial assistance provided by A Raja Choudhury, Ericka Ebot (Foundation Medicine) and Hannah Gilham (Flatiron Health) for reviewing the manuscript. The authors thank Marina Bacac, Tai-Hsien Ou Yang, Juergen Hammer and the entire Roche team for their support.

Contributors WVS and VF processed all the data and performed all the analyses in the study. TT, RAB, KH-B and EK performed parts of the analyses and reviewed the manuscript. WVS is the guarantor for this study. JCh, JCa and BG are responsible for the conception and design of the work.

Funding Authors are employees and stockholders at Roche. The work is funded by Roche.

Competing interests None declared.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval All patients accrued in clinical trials had provided informed consent and were treated in accordance with the Declaration of Helsinki. The use of historical in-house clinical study data was approved by study teams. Electronic health records (EHRs) were subscribed by Roche. EHR data from Flatiron Health did not require consent. The proper use of data contracts were signed and training taken.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data may be obtained from a third party and are not publicly available.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Vitalay Fomin <http://orcid.org/0000-0002-8382-3445>

WeiQing Venus So <http://orcid.org/0000-0002-8416-1455>

Elina Koletou <http://orcid.org/0000-0003-4539-5117>

Jehad Charo <http://orcid.org/0000-0002-5409-9160>

REFERENCES

- 1 Inamura K. Lung Cancer: Understanding Its Molecular Pathology and the 2015 WHO Classification. *Front Oncol* 2017;7:193.
- 2 Rotow J, Bivona TG. Understanding and targeting resistance mechanisms in NSCLC. *Nat Rev Cancer* 2017;17:637–58.
- 3 Antonia SJ, Borghaei H, Ramalingam SS, et al. Four-year survival with nivolumab in patients with previously treated advanced non-small-cell lung cancer: a pooled analysis. *Lancet Oncol* 2019;20:1395–408.

- 4 Bernard-Tessier A, Baldini C, Martin P, *et al.* Outcomes of long-term responders to anti-programmed death 1 and anti-programmed death ligand 1 when being rechallenged with the same anti-programmed death 1 and anti-programmed death ligand 1 at progression. *Eur J Cancer* 2018;101:160–4.
- 5 Fehrenbacher L, von Pawel J, Park K, *et al.* Updated Efficacy Analysis Including Secondary Population Results for OAK: A Randomized Phase III Study of Atezolizumab versus Docetaxel in Patients with Previously Treated Advanced Non-Small Cell Lung Cancer. *J Thorac Oncol* 2018;13:1156–70.
- 6 Mazein A, Watterson S, Hsieh W-Y, *et al.* A comprehensive machine-readable view of the mammalian cholesterol biosynthesis pathway. *Biochem Pharmacol* 2013;86:56–66.
- 7 Schadendorf D, Wolchok JD, Hodi FS, *et al.* Efficacy and Safety Outcomes in Patients With Advanced Melanoma Who Discontinued Treatment With Nivolumab and Ipilimumab Because of Adverse Events: A Pooled Analysis of Randomized Phase II and III Trials. *J Clin Oncol* 2017;35:3807–14.
- 8 Schoenfeld AJ, Hellmann MD. Acquired Resistance to Immune Checkpoint Inhibitors. *Cancer Cell* 2020;37:443–55.
- 9 Lapuente-Santana Ó, van Genderen M, Hilbers PAJ, *et al.* Interpretable systems biomarkers predict response to immune-checkpoint inhibitors. *Patterns (N Y)* 2021;2:100293.
- 10 Litchfield K, Reading JL, Puttick C, *et al.* Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* 2021;184:596–614.
- 11 Lu S, Stein JE, Rimm DL, *et al.* Comparison of Biomarker Modalities for Predicting Response to PD-1/PD-L1 Checkpoint Blockade: A Systematic Review and Meta-analysis. *JAMA Oncol* 2019;5:1195–204.
- 12 So WV, Dejardin D, Rossmann E, *et al.* Predictive biomarkers for PD-1/PD-L1 checkpoint inhibitor response in NSCLC: an analysis of clinical trial and real-world data. *J Immunother Cancer* 2023;11:e006464.
- 13 Wang Y, Tong Z, Zhang W, *et al.* FDA-Approved and Emerging Next Generation Predictive Biomarkers for Immune Checkpoint Inhibitors in Cancer Patients. *Front Oncol* 2021;11.
- 14 Frampton GM, Fichtenholtz A, Otto GA, *et al.* Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol* 2013;31:1023–31.
- 15 Ma X, Bellomo L, Magee K, *et al.* Characterization of a Real-World Response Variable and Comparison with RECIST-Based Response Rates from Clinical Trials in Advanced NSCLC. *Adv Ther* 2021;38:1843–59.
- 16 Griffith SD, Tucker M, Bowser B, *et al.* Generating Real-World Tumor Burden Endpoints from Electronic Health Record Data: Comparison of RECIST, Radiology-Anchored, and Clinician-Anchored Approaches for Abstracting Real-World Progression in Non-Small Cell Lung Cancer. *Adv Ther* 2019;36:2122–36.
- 17 Kugel CH 3rd, Douglass SM, Webster MR, *et al.* Age Correlates with Response to Anti-PD1, Reflecting Age-Related Differences in Intratumoral Effector and Regulatory T-Cell Populations. *Clin Cancer Res* 2018;24:5347–56.
- 18 Rizvi NA, Hellmann MD, Snyder A, *et al.* Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348:124–8.
- 19 Norum J, Nieder C. Tobacco smoking and cessation and PD-L1 inhibitors in non-small cell lung cancer (NSCLC): a review of the literature. *ESMO Open* 2018;3:e000406.
- 20 Brosh R, Rotter V. When mutants gain new powers: news from the mutant p53 field. *Nat Rev Cancer* 2009;9:701–13.
- 21 De Roock W, Jonker DJ, Di Nicolantonio F, *et al.* Association of KRAS p.G13D mutation with outcome in patients with chemotherapy-refractory metastatic colorectal cancer treated with cetuximab. *JAMA* 2010;304:1812–20.
- 22 Summers MG, Smith CG, Maughan TS, *et al.* BRAF and NRAS locus-specific variants have different outcomes on survival to colorectal cancer. *Clin Cancer Res* 2017.
- 23 Xu J, Qian J, Hu Y, *et al.* Heterogeneity of Li-Fraumeni Syndrome links to unequal gain-of-function effects of p53 mutations. *Sci Rep* 2014;4.
- 24 Kadosh E, Snir-Alkalay I, Venkatachalam A, *et al.* The gut microbiome switches mutant p53 from tumour-suppressive to oncogenic. *Nature New Biol* 2020;586:133–8.
- 25 Vogelstein B, Papadopoulos N, Velculescu VE, *et al.* Cancer genome landscapes. *Science* 2013;339:1546–58.
- 26 Goodman AM, Kato S, Bazhenova L, *et al.* Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers. *Mol Cancer Ther* 2017;16:2598–608.
- 27 Van Allen EM, Miao D, Schilling B, *et al.* Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 2015;350:207–11.
- 28 Miao D, Margolis CA, Vokes NI, *et al.* Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat Genet* 2018;50:1271–81.
- 29 Huseni MA, Wang L, Klementowicz JE, *et al.* CD8⁺ T cell-intrinsic IL-6 signaling promotes resistance to anti-PD-L1 immunotherapy. *Cell Rep Med* 2023;4:100878.
- 30 Keegan A, Ricciuti B, Garden P, *et al.* Plasma IL-6 changes correlate to PD-1 inhibitor responses in NSCLC. *J Immunother Cancer* 2020;8:e000678.
- 31 Powles T, Nickles D, Van Allen E, *et al.* Immune biomarkers associated with clinical benefit from atezolizumab (MPDL3280a; anti-PD-L1) in advanced urothelial bladder cancer (UBC). *J Immunotherapy Cancer* 2015;3.
- 32 Skoulidis F, Goldberg ME, Greenawalt DM, *et al.* STK11/LKB1 mutations and PD-1 inhibitor resistance in KRAS-mutant lung adenocarcinoma. *Cancer Discov* 2018.
- 33 Skoulidis F, Byers LA, Diao L, *et al.* Co-occurring genomic alterations define major subsets of KRAS-mutant lung adenocarcinoma with distinct biology, immune profiles, and therapeutic vulnerabilities. *Cancer Discov* 2015;5:860–77.
- 34 Adib E, Nassar A, El Zarif T, *et al.* Dual CDKN2A/MTAP loss compared to CDKN2A loss alone and response to immune-checkpoint inhibitors (ICI) in advanced solid tumors. *JCO* 2022;40:2622.
- 35 Adib E, Nassar AH, Akl EW, *et al.* CDKN2A Alterations and Response to Immunotherapy in Solid Tumors. *Clin Cancer Res* 2021;27:4025–35.
- 36 Cormedi MCV, Van Allen EM, Colli LM. Predicting immunotherapy response through genomics. *Curr Opin Genet Dev* 2021;66:1–9.
- 37 Liang Y, Tang H, Guo J, *et al.* Targeting IFN α to tumor by anti-PD-L1 creates feedforward antitumor responses to overcome checkpoint blockade resistance. *Nat Commun* 2018;9.
- 38 Torrejon DY, Abril-Rodriguez G, Champhekar AS, *et al.* Overcoming Genetically Based Resistance Mechanisms to PD-1 Blockade. *Cancer Discov* 2020;10:1140–57.
- 39 Jiang L, Guo F, Liu X, *et al.* Continuous targeted kinase inhibitors treatment induces upregulation of PD-L1 in resistant NSCLC. *Sci Rep* 2019;9:3705.
- 40 Gao SP, Mark KG, Leslie K, *et al.* Mutations in the EGFR kinase domain mediate STAT3 activation via IL-6 production in human lung adenocarcinomas. *J Clin Invest* 2007;117:3846–56.
- 41 Liu H, Shen J, Lu K. IL-6 and PD-L1 blockade combination inhibits hepatocellular carcinoma cancer development in mouse model. *Biochem Biophys Res Commun* 2017;486:239–44.
- 42 Bialkowski L, Van der Jeught K, Bevers S, *et al.* Immune checkpoint blockade combined with IL-6 and TGF- β inhibition improves the therapeutic outcome of mRNA-based immunotherapy. *Int J Cancer* 2018;143:686–98.
- 43 Reisländer T, Groelly FJ, Tarsounas M. DNA Damage and Cancer Immunotherapy: A STING in the Tale. *Mol Cell* 2020;80:21–8.
- 44 Yuen KC, Gupta V, Li C, *et al.* Abstract 2676: Associations of peripheral biomarkers to outcomes to anti-PD-L1 immune checkpoint blockade in metastatic urothelial cancer. *Cancer Res* 2019;79:2676.
- 45 Çelik A, Yaman H, Turan S, *et al.* Angiogenic and Immune-Related Biomarkers and Outcomes Following Axitinib/Pembrolizumab Treatment in Patients with Advanced Renal Cell Carcinoma. *J Mater Process Technol* 2018;1:1–8.
- 46 Zhao J, Chen AX, Gartrell RD, *et al.* Immune and genomic correlates of response to anti-PD-1 immunotherapy in glioblastoma. *Nat Med* 2019;25:462–9.
- 47 Luke JJ, Bao R, Sweis RF, *et al.* WNT/ β -catenin Pathway Activation Correlates with Immune Exclusion across Human Cancers. *Clin Cancer Res* 2019;25:3074–83.
- 48 Ebert PJR, Cheung J, Yang Y, *et al.* MAP Kinase Inhibition Promotes T Cell and Anti-tumor Activity in Combination with PD-L1 Checkpoint Blockade. *Immunity* 2016;44:609–21.
- 49 Ribas A, Lawrence D, Atkinson V, *et al.* Combined BRAF and MEK inhibition with PD-1 blockade immunotherapy in BRAF-mutant melanoma. *Nat Med* 2019;25:936–40.
- 50 Tian J, Chen JH, Chao SX, *et al.* Combined PD-1, BRAF and MEK inhibition in BRAF^{V600E} colorectal cancer: a phase 2 trial. *Nat Med* 2023;29:458–66.
- 51 Hartmaier RJ, Charo J, Fabrizio D, *et al.* Genomic analysis of 63,220 tumors reveals insights into tumor uniqueness and targeted cancer immunotherapy strategies. *Genome Med* 2017;9:16.

- 52 Levine AJ, Jenkins NA, Copeland NG. The Roles of Initiating Truncal Mutations in Human Cancers: The Order of Mutations and Tumor Cell Type Matters. *Cancer Cell* 2019;35:10–5.
- 53 Singal G, Miller PG, Agarwala V, et al. Association of Patient Characteristics and Tumor Genomics With Clinical Outcomes Among Patients With Non-Small Cell Lung Cancer Using a Clinicogenomic Database. *JAMA* 2019;321:1391–9.
- 54 Chalmers ZR, Connelly CF, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med* 2017;9:34.
- 55 Herbst RS, Soria J-C, Kowanetz M, et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature New Biol* 2014;515:563–7.
- 56 Emig D, Ivliev A, Pustovalova O, et al. Drug target prediction and repositioning using an integrated network-based approach. *PLoS One* 2013;8:e60618.
- 57 Vanunu O, Magger O, Ruppin E, et al. Associating genes and protein complexes with disease via network propagation. *PLoS Comput Biol* 2010;6:e1000641.
- 58 Hsu CL, Huang YH, Hsu CT, et al. Prioritizing disease candidate genes by a gene interconnectedness-based approach. *BMC Genomics* 2011;12 Suppl 3:S25.
- 59 Nikolsky Y, Sviridov E, Yao J, et al. Genome-wide functional synergy between amplified and mutated genes in human breast cancer. *Cancer Res* 2008;68:9532–40.
- 60 Dezso Z, Nikolsky Y, Nikolskaya T, et al. Identifying disease-specific genes based on their topological significance in protein networks. *BMC Syst Biol* 2009;3:36.
- 61 Mostafavi S, Ray D, Warde-Farley D, et al. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol* 2008;9 Suppl 1:S4.
- 62 Chindelevitch L, Ziemek D, Enayetallah A, et al. Causal reasoning on biological networks: interpreting transcriptional changes. *Bioinformatics* 2012;28:1114–21.
- 63 Pollard J Jr, Butte AJ, Hoberman S, et al. A Computational Model to Define the Molecular Causes of Type 2 Diabetes Mellitus. *Diabetes Technol Ther* 2005;7:323–36.
- 64 Frohlich H, Chapelle O, Scholkopf B. Feature selection for support vector machines by means of genetic algorithm. 15th IEEE International Conference on Tools with Artificial Intelligence; Sacramento, CA, USA, 2003:142–8.
- 65 Goldberg DE. *Genetic Algorithms*. Pearson Education India, 2013.
- 66 Rish I. An empirical study of the naive Bayes classifier. *IJCAI (U S)* 2001;10.
- 67 Blickle T, Thiele L. A Comparison of Selection Schemes Used in Evolutionary Algorithms. *Evol Comput* 1996;4:361–94.
- 68 Ng AY, Jordan MI. On discriminative vs. generative classifiers: a comparison of logistic regression and naive bayes. In: *Advances in Neural Information Processing Systems*. 2002: 841–8.