Journal of Translational Medicine

RESEARCH

Open Access

ProgModule: A novel computational framework to identify mutation driver modules for predicting cancer prognosis and immunotherapy response

Xiangmei Li[1†], Bingyue Pan[1†], Xilong Zhao[1†], Yinchun Su[2], Jiyin Lai[1], Siyuan Li[1], Yalan He[1], Jiashuo Wu[1] and Junwei Han[1*]

Abstract

Background Cancer originates from dysregulated cell proliferation driven by driver gene mutations. Despite numerous algorithms developed to identify genomic mutational signatures, they often suffer from high computational complexity and limited clinical applicability.

Methods Here, we presented ProgModule, an advanced computational framework designed to identify mutation driver modules for cancer prognosis and immunotherapy response prediction. In ProgModule, we introduced the Prognosis-Related Mutually Exclusive Mutation (PRMEM) score, which optimizes the balance between exclusive mutation coverage and the incorporation of mutation combination mechanisms critical for cancer prognosis.

Results Applying to BLCA and HNSC cohorts, ProgModule successfully identified driver modules that stratify patients into distinct prognostic subgroups, and the combination of these modules could serve as an effective prognostic biomarker. Extending our method to diverse cancers, ProgModule presented robust prognostic performance and stability across model parameters, including stopping criteria and network topology. Moreover, our analysis suggested that driver modules can predict immunotherapeutic benefit more effectively than existing signatures. Further analyses based on published CRISPR data indicated that genes within these modules may serve as potential therapeutic targets.

Conclusions Altogether, ProgModule emerges as a powerful tool for identifying mutation driver modules as prognostic and immunotherapy response biomarkers, and genes within these modules may be used as potential therapeutic targets for cancer, offering new insights into precision oncology.

Keywords Mutually exclusive mutation, Driver module, Prognostic biomarkers, Immune checkpoint inhibitors, Therapeutic targets

†Xiangmei Li, Bingyue Pan and Xilong Zhao contributed equally to this work.

*Correspondence:
Junwei Han
hanjunwei@ems.hrbmu.edu.cn

[1]College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China
[2]Department of Neurobiology, Harbin Medical University, Harbin 150081, China

Background

Cancer emerges through the accrual of somatic mutations, with driver mutations conferring a selective advantage that enables cells to circumvent normal growth regulation and evade apoptotic signals, thereby initiating tumorigenesis [1, 2]. Given the exponential growth in cancer genomics data, a principal objective in oncology research is to precisely differentiate cancer driver mutations and passenger mutations, deepening the comprehension of the molecular mechanisms governing tumorigenesis and enabling personalized therapy strategies tailed to specific mutational patterns.

Increasing studies suggest that driver mutations accumulate within a constrained set of critical biological pathways, particularly those governing DNA repair, cell cycle regulation, and apoptosis [3, 4]. Therefore, several network-based algorithms were developed. For example, HotNet2, developed by Leiserson et al., specifically designed to identify mutated subnetworks, thereby overcoming limitations associated with conventional single-gene, pathway, and network-centric methodologies [5]. Mutex, another approach, searches for mutually exclusive gene sets by seeding with an altered gene and sequentially expanding the set using a comprehensive directed gene network [6]. According to prior studies, driver modules typically possess two essential characteristics: high coverage and high exclusivity [7, 8]. High coverage indicates that the majority of patients possess at least one mutated gene within the module, while high exclusivity implies that most patients harbor only a single mutated gene within the same module. Therefore, many algorithms have also been proposed to detect gene modules exhibiting these properties. For example, the MEMO algorithm leverages external biological data by analyzing gene pairs and constructing graph cliques to identify mutually exclusive gene sets [9]. The RME algorithm quantifies exclusivity weight by computing the proportion of patients with exactly one mutation within a specific gene set [10]. Similarly, Dendrix, introduced by Vandin et al., measures variation in coverage and overlap of coverage within a gene set, using a greedy algorithm to identify driver pathways with high Dendrix weight [11]. However, these methods are constrained by significant high computational complexity, elevated false-positive rates, and limited clinical applicability.

Moreover, these methods are generally limited to uncovering relevant mutational patterns based on mutation frequency, coverage, and exclusivity, without systematically analyzing the mutation combination-mediated mechanisms in cancer prognosis. It is well established that cell transformation into cancer occur through multiple routes, not only via mutations targeting the gene itself but also via indirect activation or silencing of genes by the influence of other genes [12–14]. To advance our understanding of cancer initiation mechanisms, there is an urgent need for algorithms designed to identify mutation-driven modules that incorporate their prognostic significance while balancing mutation coverage and exclusivity. These algorithms may reveal complex interactions between genetic mutations and patients' survival, offering critical insights for improving cancer diagnosis and therapeutic strategies.

In our study, we presented ProgModule, an advanced computational framework designed to identify mutation driver modules for cancer prognosis and immunotherapy response prediction. In ProgModule, we introduced a PRMEM score to optimizes the balance between exclusive mutation coverage and the incorporation of mutation combination mechanism critical for cancer prognosis. Then we used the greedy algorithm to identify the candidate driver module with the maximal PRMEM score to predict patient prognosis and immunotherapy response. Applying ProgModule to BLCA and HNSC cohorts, the identified modules could be used to stratify patients into subgroups with distinct prognoses, and the combination of these modules could also serve as an effective prognostic biomarker. Extending our method to other cancer types, we found that driver modules exhibited superior prognostic performance compared to individual genes and state-of-the-art methods. Additionally, ProgModule was applied to two immunotherapy cohorts where the candidate modules effectively predict the clinical outcomes of patients and outperformed existing signatures. Furthermore, analyses based on published CRISPR data showed that the genes within these driver modules may be used as potential therapeutic targets of cancers. Finally, ProgModule is available as an R-based tool at https://CRAN.R-project.org/package=ProgModule, providing a powerful resource for identifying prognostic and immunotherapy biomarkers, as well as potential therapeutic target in cancer.

Methods

Data collection and preprocessing

We collected somatic mutation data for 9309 samples from the Cancer Genome Atlas Program (TCGA, https://portal.gdc.cancer.gov/). Somatic mutation data and clinical information were also obtained from the International Cancer Genome Consortium (ICGC, https://dcc.icgc.org/) database to further validate the robustness of our method. Patients lacking clinical survival or somatic mutation data were excluded, resulting in the retention of 8395 TCGA samples and 15,494 ICGC samples, covering 24 different cancer types. Detailed cohort information is provided in Supplementary Table S1.

Besides, we retrieved 591 driver genes from the Network of Cancer Genes & Healthy Drivers (NCG) database (http://network-cancer-genes.org/). To investigate

the role of driver genes in cancer, we also acquired three protein-protein interaction (PPI) networks from HotNet2, namely HINT + HI2012, iRefIndex, and Multinet [5]. To ensure the robustness of protein-protein interactions, we integrated the three PPI networks by retaining interactions presented in all three networks. The resulting network consists of total 5698 nodes and 14,745 edges, which provided the foundation for identifying local subnetworks during subsequent analyses.

Search local subnetworks

Somatic mutations were first extracted from Mutation Annotation Format (MAF) files, retaining only non-silent mutations. In this study, 24 MAF files were included, each of which corresponding to a specific cancer type and was sourced from the TCGA database. Subsequently, these mutations were mapped onto genes in the integrated PPI networks specific to each cancer type. Using a breadth-first search (BFS) algorithm, local subnetworks were identified by initiating the search from each driver gene (seed node) obtained from NCG database. The BFS then iteratively explored neighboring mutated genes, ensuring that only directly or indirectly connected genes were included, thus preserving the subnetwork's structural integrity and maximal connectivity. The process stops upon reaching the predefined limit of 500 genes, yielding fully connected subnetworks for subsequent analysis.

Identify prognosis-related mutually exclusive driver modules

Previous studies have shown that driver modules are characterized by high coverage and high exclusivity [7, 8]. High coverage implies that the majority of samples have at least one mutation within the module, while high exclusivity suggests that the majority of samples have only one mutation in that module. In biology, it is desired to be able to detect driver modules by harmonizing coverage with the exclusivity of module mutations. In our study, we introduced the PRMEM score to balance the exclusive coverage of mutations and the incorporation of mutation combination mechanism critical for cancer prognosis. The PRMEM score for a specific module is defined as follows:

$$PRMEM\ Score = MI\ (M) * Ex\ (M) * MutRatio\ (M) \quad (1)$$

Here, *MI(M)* represents the mutual information between the mutation states of module M and patient survival states, evaluating the influence of module mutations on prognosis. For each patient, the module M is considered as mutated if it contains at least one mutated gene in the patient; otherwise, it is classified as non-mutated. The formula for *MI(M)* is defined as:

$$MI\,(M) = H\,(M) + H\,(S) - H\,(MS) \quad (2)$$

where *H(M)* and *H(S)* respectively represent the entropy of mutation states of module M and patient survival states, and *H(MS)* is their combination entropy. *Ex(M)* denotes the exclusivity score of module M, defined as:

$$Ex\,(M) = \frac{\sum_{i \in M} \frac{EP_i}{P_i}}{N_M} \quad (3)$$

Where EP_i represents the number of samples in which gene i is exclusively mutated, P_i is the number of samples in which gene i is mutated, and N_M is the number of genes in M; *MutRatio (M)* is the mutation ratio of module M, defined as *MutRatio (M)* = n/N, where N is the total number of samples and n is the number of samples with mutations in M. The PRMEM score, to some extent, not only harmonizes coverage with the exclusivity of module mutations but also reflects the relationship between module mutations and survival states of patients.

Based on the PRMEM score, an iterative greedy algorithm was applied to search modules within local subnetworks, where the PRMEM scores reached local maxima. In each local subnetwork, candidate modules start with a single seed node (driver genes obtained from the NCG database) and expanded iteratively. In our study, the seed gene was not randomly selected from the NCG database. Instead, we systematically use each driver gene from the NCG database as a seed node to identify the optimal mutation module associated with that specific driver gene. During each iteration, the search process considers adding a gene from the set of neighboring genes within the current modules. The addition that has the same survival impact (that is if the seed node is a risk factor, hazard ratio HR > 1, then the addition should be a risk factor, and vice versa) and yields the maximal score increase is adopted. If no addition raises the score above a given improvement rate, r, the search ends. To avoid overfitting and reduce search time, the improvement rate r was chosen as 0.05. Moreover, we required the search stops when the resulting module reached a maximum size of 200 to keep the search local and prevent overly broad functional modules.

We then performed permutation analysis to determine the statistical significance of the resulting modules. Specifically, for each perturbation, we first randomly generated a sample mutation matrix with the same mutation rate as the original dataset and then recalculated the PRMEM score for each module. Subsequently, the perturbation p-value of the module was calculated as: p-value = n/N, in which n is the count of randomly generated PRMEM scores exceeding the original PRMEM score, and N denoted the total number of perturbations (N was set at 1000 in the study). To minimize the impact

of multiple comparisons, p-values were adjusted using the Bonferroni correction method [15], and modules with adjusted p-values less than 0.01 were identified as candidate driver modules. Moreover, for an individual patient, the module was deemed as mutated if it involves one or more mutated genes. The Kaplan-Meier survival was utilized to evaluate whether the mutation states of candidate modules could predict patient prognosis.

Results

Application of ProgModule on BLCA cohort

We presented ProgModule, an advanced computational framework to detect prognosis-related mutation driver modules by incorporating the exclusive coverage of mutations and patient clinical information. The schematic diagram of ProgModule is depicted in Fig. 1. ProgModule innovatively considers the mutation combination-mediated mechanisms for cancer prognosis in identifying mutually exclusive driver modules, which may discover something new modules for prognosis and immunotherapy response prediction. To explain the effectiveness of our method, we employed it on the TCGA-BLCA cohort and discovered 19 statistically significant mutation driver modules containing 138 genes (Table 1), of which 49 are cancer driver genes. We found that certain genes appear in more than one module, suggesting that some genes serve multifunctional roles, participating in multiple biological processes or pathways, making them critical regulators or hubs within interconnected pathways. These modules exhibited a diversified landscape, ranging from the smallest module with 4 genes (M2) to the largest module with 16 genes (M5) (Fig. 2A). Notably, the smallest module M2 contained AXIN1, APC, LRP5, and CTNNB1, of which 3 were well-established cancer driver genes. The functional enrichment analysis identified a strong association of M2 with pathways involved in carcinogenesis, particularly Wnt and Hippo signaling pathways (Supplementary Figure S1A). Similarly, the largest module, M5 (Fig. 2B), was significantly enriched in pathways crucial for cancer initiation and development, including the ErbB signaling pathway, chemokine signaling pathway, and focal adhesion, etc.

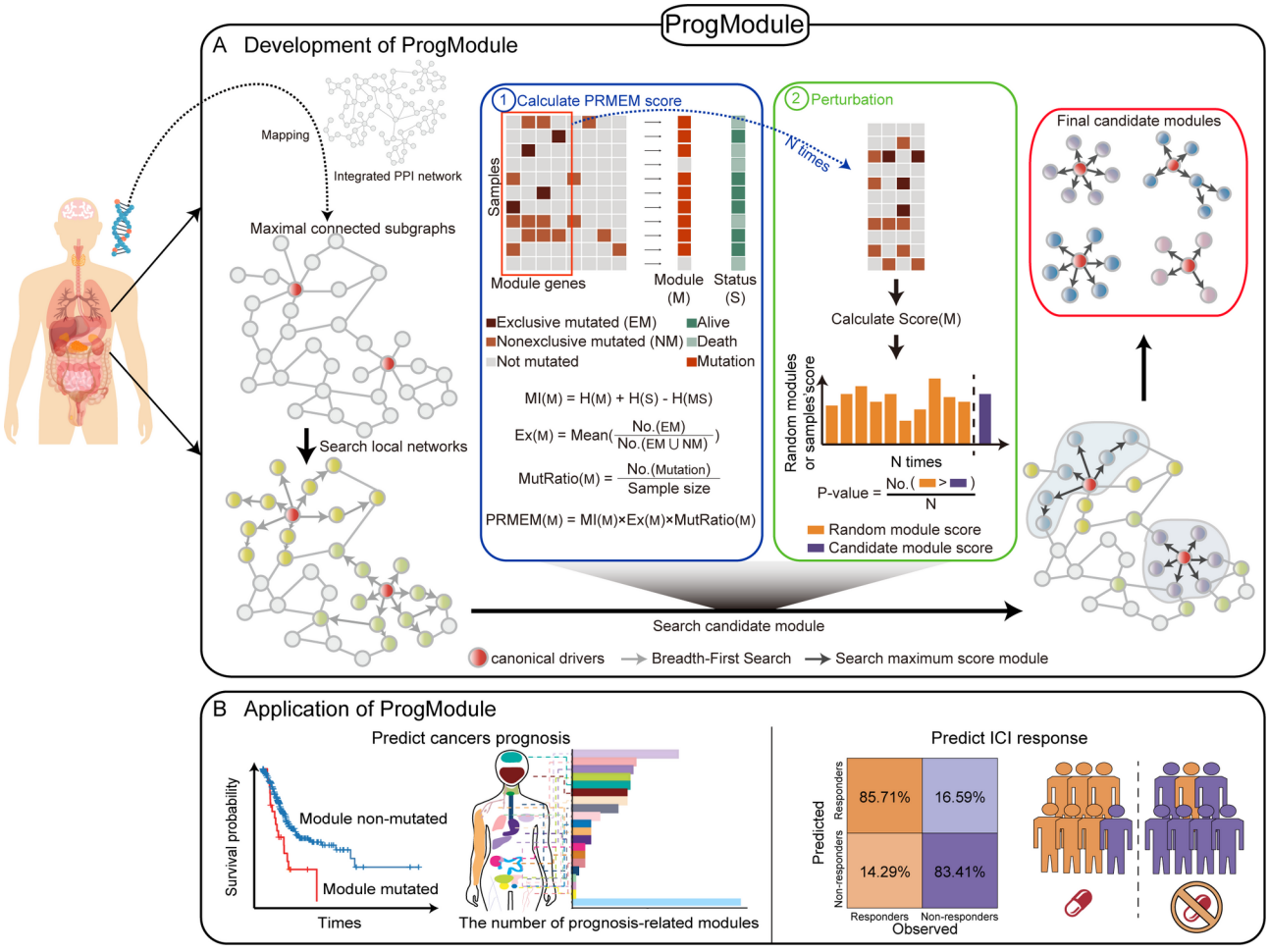

Fig. 1 The flowchart of the ProgModule method. (**A**) Elucidation of the ProgModule for identification of prognosis-related mutually exclusive modules; **(B)** Application of ProgModule, encompassing the prediction of cancer prognosis and immunotherapy response

Table 1 The detailed information of driver modules identified in BLCA cohort

Module	Genes	Coverage	*P* value	Adj. *p* value
M1	APC, SEC31A, LAMA3, NOSTRIN, FHOD1, NUP98, RANBP9, EPAS1, KRT15, NUPL2, GIGYF2, AXIN1, DLG3, CYTH2	24%	1.00E-16	6.00E-16
M2	AXIN1, APC, LRP5, CTNNB1	9%	1.00E-16	6.00E-16
M3	BIRC3, CASP3, UBC, UBE2D2, UBE2N, CASP7	3%	1.00E-16	6.00E-16
M4	CTNNB1, APC, AXIN1, PSEN1, BTRC, FBXW11, PTPRJ, CDH3, HNF1A, TRRAP, CDH1, JUP	21%	1.00E-16	6.00E-16
M5	EGFR, ERBB2, PIK3C2B, HSP90AA1, CTNND1, PLSCR1, LRIG1, GRB10, NCK1, EPN1, RGS16, STAT5B, SHC1, PTK2B, GAB1, SOS1	25%	1.00E-16	6.00E-16
M6	EP300, CREBBP, PPARG, KLF5, MYB, PPARA, HDAC1, BRCA1, STAT2, NCOA3, NAP1L1, GRIP1	35%	1.00E-16	6.00E-16
M7	ERBB2, HSP90AA1, ERBB2IP, UBC, STUB1, EGFR, SHC1	17%	1.00E-16	6.00E-16
M8	MAP3K1, MAP2K4, UBC, MAP2K1, MAP2K7, YWHAE, UBE2D4, UBE2D2, UBE2D3, MAP4K2	6%	1.00E-16	6.00E-16
M9	PDGFRB, PIK3R1, KRTAP4-12, NCK1, SHC1, UBC	5%	1.00E-16	6.00E-16
M10	PIK3R1, GAB1, PDGFRB, VAV1, IRS1, KIT, SYNJ2, UBC, ABL1, FYN, LCP2, PLCG2, LCK, PIK3CD	14%	1.00E-16	6.00E-16
M11	PLCG1, ERBB2, PDGFRB, GAB1, KDR, UBC, SHC1, LCP2, KIT, FGFR1, FLT1	20%	1.00E-16	6.00E-16
M12	RANBP2, AKAP13, TAF1, UBC, SUMO1	10%	1.00E-16	6.00E-16
M13	SMAD2, SMARCC2, EP300, SMAD3, CREBBP, HDAC1, ERBB2IP, NCOA3, SMARCC1, PIAS4, SMURF1	31%	1.00E-16	6.00E-16
M14	SMAD3, CREBBP, EP300, ERBB2IP, MECO, HOXC11, HDAC1, UBC, PIAS4, SMURF1	28%	1.00E-16	6.00E-16
M15	SMAD4, TFE3, EWSR1, SKIL, BTRC, TGFBR1, SMAD9, PIAS1, JUNB, UBE2I	9%	1.00E-16	6.00E-16
M16	SRC, RASA1, ASAP2, RXRA, BCAR1, ESRRA, STAT3, STAT1, MET, ADRB2, KHDRBS1	16%	1.00E-16	6.00E-16
M17	TNFAIP3, TBK1, TAX1BP1, UBC, RNF11, YWHAH	5%	1.00E-16	6.00E-16
M18	USP8, UBC, STAM2, CHMP1B, RNF41	5%	1.00E-16	6.00E-16
M19	ZBTB16, NCOR1, SIN3A, NCOR2, RUNX1T1, PGAM5, HDAC1, LMTK3, FCHO1, KRTAP4-12, GCSH, LDOC1, COQ6	17%	1.00E-16	6.00E-16

(Fig. 2C). Additionally, the other modules were also associated with key biological pathways (Supplementary Figure S1B), suggesting ProgModule's capability to identify modules linked to essential biological processes in cancer. Moreover, the mutational frequencies of candidate driver modules range from 4.37% (M3) to 37.14% (M6), which illustrates the candidate modules possess relatively high coverage (Supplementary Figure S1C). Moreover, the genes within each module exhibit obviously exclusivity across the patient population (Supplementary Figure S1C). For instance, in module M6, mutations in CREBBP rarely overlapped with those in EP300 (Fig. 2D). The same findings were also observed in the study by Huang et al., implying that few patients carry mutations in both the EP300 and CREBBP genes simultaneously [16]. These findings demonstrated that the candidate modules exhibit high exclusivity and coverage, consistent with the typical properties observed in driver modules, and ProgModule is capable of identifying mutated mutually exclusive gene pairs confirmed in literature.

To test if our modules could predict patient prognosis, univariate Cox regression analysis was performed on each module, respectively. For each patient, if one or more genes within a module were mutated, the module was deemed as mutated; otherwise, it was classified as wild type. According to the mutation states of modules, our findings revealed a significant association between 17 out of 19 driver modules and overall survival (OS) in patients (Cox regression $p < 0.05$, Fig. 2E). To further validate the prognostic capability of these modules, we applied these modules to the ICGC-BLCA cohort, confirming 14 modules remained predictive of patient survival (Fig. 2E). This highlighted the efficacy of ProgModule as a powerful tool for detecting modules that can predict the prognosis of bladder cancer. While cancer drivers are typically associated with cancer progression, their prognostic impact can vary depending on the tumor type, genetic context, and patient characteristics. Protective driver modules, identified in our study, reflect the diverse and context-dependent roles of driver mutations in cancer biology and prognosis [17–19].

Li et al. proposed that pathways or gene sets analysis might offer more comprehensive insights into dysregulated pathways in cancers compared to the analysis of individual genes [20]. We thus further explored whether the prognostic accuracy of driver modules surpassed that of individual genes within the module. Kaplan-Meier survival analysis revealed that the driver modules generally outperformed individual genes in predicting patient survival, except for modules M15 and M16 (Fig. 3A). Notably, within module M15, patients harboring mutations exhibited significantly worse prognosis compared to those without mutation (log-rank $P = 0.0074$, Fig. 3B). Among the genes within module M15, only TFE3 demonstrated relatively high predictive ability compared to the module itself (log-rank $P < 0.0001$, Fig. 3C). However, the clinical utility of TFE3 mutations as prognostic markers in bladder cancer was limited by their low mutation

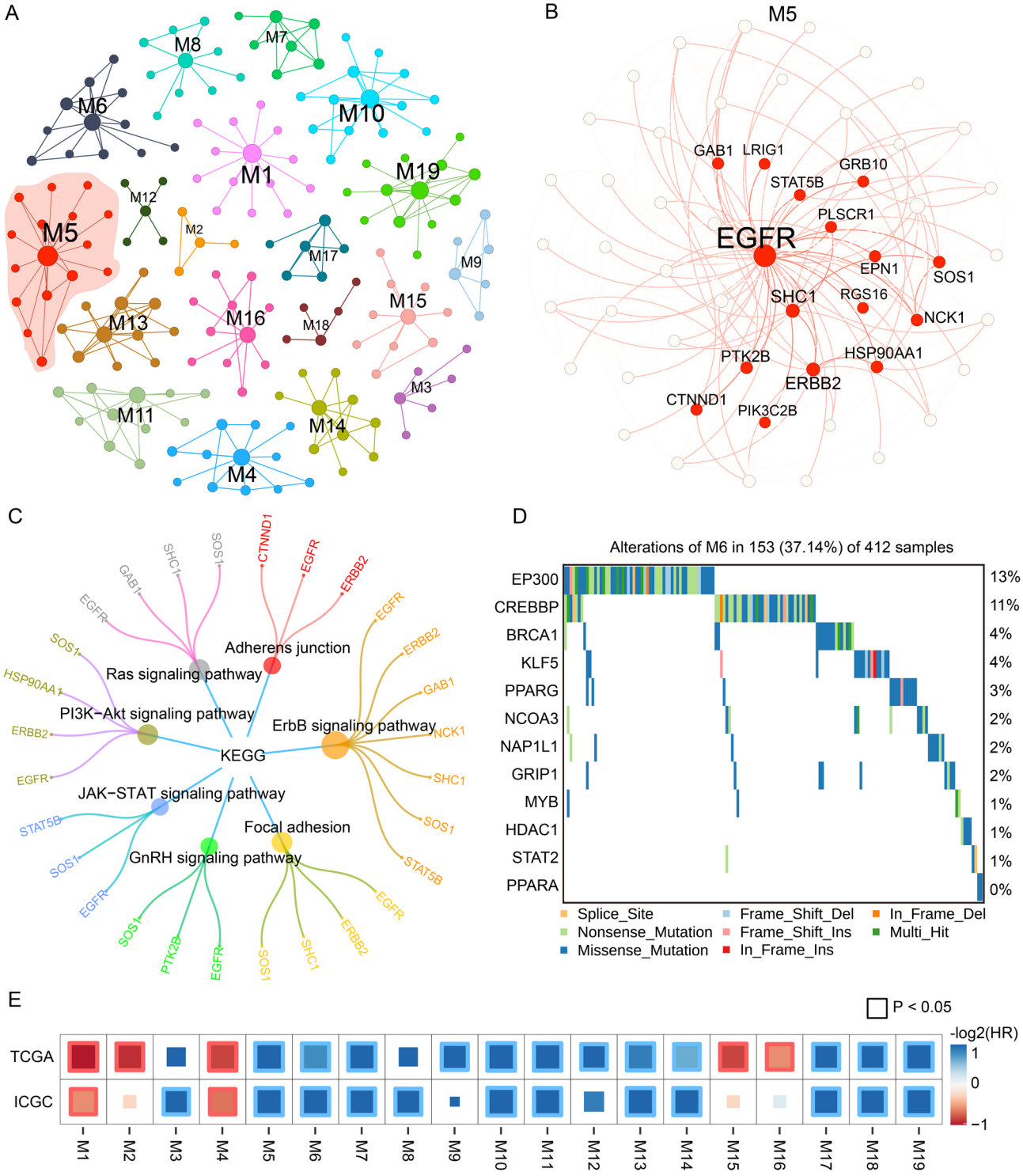

Fig. 2 Applying ProgModule to the BLCA cohort. (**A**) The landscape of candidate modules for BLCA (simplified schematic to highlight the internal PPI structure of each module). **(B)** The detailed information of module M5; **(C)** The enriched KEGG pathways of genes involved in module M5; **(D)** The waterfall plot of the module M6; **(E)** The heatmap for the Cox regression P values of candidate modules

occurrence, with only a small portion of patients (5/373) experiencing TFE3 mutations, whereas a larger proportion (33/373) harbored mutation in module M15. Similar results were observed in module M16 (Fig. 3D and E), where gene BCAR1 exhibited high predictive accuracy but limited clinical utility due to its rare mutation occurrence, rendering it impractical for a large portion of the patient population. These findings underscored the

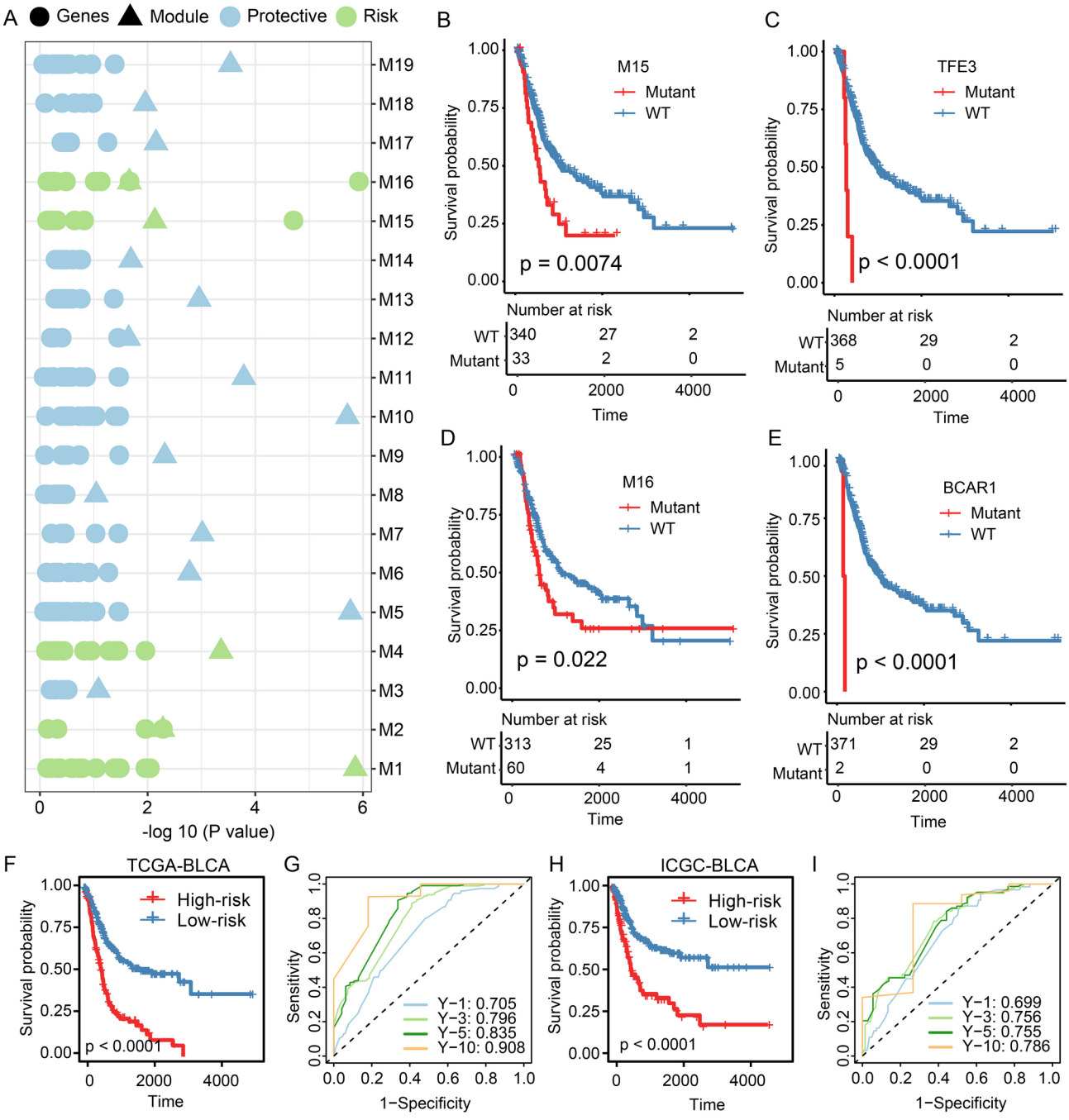

Fig. 3 Comparing the prognostic ability of modules versus individual genes. (**A**) Comparing the prognostic ability of candidate modules versus individual genes in BLCA; **(B)** Kaplan-Meier survival analysis of OS comparing the M15-Mutant and WT groups; **(C)** Kaplan-Meier survival analysis of OS comparing the TFE3-Mutant and WT groups; **(D)** Kaplan-Meier survival analysis of OS comparing the M16-Mutant and WT groups; **(E)** Kaplan-Meier survival analysis of OS comparing the BCAR1-Mutant and WT groups; **(F)** Kaplan-Meier survival curve of risk model in TCGA-BLCA; **(G)** The ROC curve of risk model to predict patient OS in TCGA-BLCA; **(H)** The Kaplan-Meier survival curve of risk model in ICGC-BLCA; **(I)** The ROC curve of risk model to predict patient OS in ICGC-BLCA

superior clinical utility of driver modules compared to single-gene markers.

Furthermore, our analysis revealed significant co-occurrence of mutations across candidate modules, indicating potential interactions among mutations between these modules (Supplementary Figure S2). To eliminate the confounding effects of module-specific mutations, multivariable Cox regression analysis was conducted. This analysis identified nine modules as independent prognostic factors for cancer overall survival

(Supplementary Table S2). To investigate the joint predictive ability of these nine modules, we developed a risk score model based on a formula that incorporates the mutation status of these modules, weighted by their respective multivariate Cox regression coefficient (Supplementary Table S2). Patients were then categorized into low-risk and high-risk groups using the median risk scores as a cutoff. Notably, patients in the high-risk group presented significantly worse prognosis compared to those in the low-risk group (median OS, 15.45 months vs. 18.07 months, log-rank $P < 0.0001$, Fig. 3F), and the performance of this risk score model in prognostic prediction was satisfactory (area under the receiver operating characteristic curve [AUC] at 3 years: 0.796; AUC at 5 years: 0.835; AUC at 10 years: 0.908, Fig. 3G). To verify the prognostic value of our risk score model, we applied it to the ICGC-BLCA cohort and found the model also presented a good predictability of clinical outcomes in bladder cancer patients (log-rank $P < 0.0001$, AUC at 3 years: 0.756; AUC at 5 years: 0.755; AUC at 10 years: 0.786, Fig. 3H and I). These findings indicated that our module-based risk model exhibits excellent prognostic value in both the TCGA-BLCA and ICGC-BLCA cohorts.

Application of ProgModule on HNSC cohort

To further prove the effectiveness of ProgModule method, we applied it to the TCGA-HNSC cohort and identified 18 statistically significant driver modules (Fig. 4A), including a total of 40 driver genes. Functional enrichment analysis highlighted that these modules were remarkably enriched in critical biological pathways, such as the TGF-beta signaling pathway, cell cycle, JAK-STAT signaling pathway, and Wnt signaling pathway, etc. (Fig. 4B-C and Supplementary Figure S3A). Additionally, the mutation landscape of driver modules revealed that the mutation genes within the same module presented significantly mutually exclusive and high coverage within the patient cohort (Fig. 4D and Supplementary Figure S3B).

Moreover, 14 out of 18 driver modules showed a significant association with overall survival in HNSC patients (Fig. 4E), in which the mutation states of these driver modules were positively or negatively influenced on patient survival. For instance, patients with mutations in the M10 module had significantly shorter OS compared to those without mutations (median OS, 18.20 months versus 21.50 months, $P < 0.0001$, Fig. 4F). Enrichment analysis unveiled those genes within M10 played critical roles in fundamental biological functions. Dysregulation of these functions can lead to abnormal cell proliferation and differentiation [21, 22], thereby influencing cancer progresses and prognosis, which partially elucidates why mutations in M10 are associated with cancer poorer survival. We then verified the prognostic value of these driver modules in the ICGC-HNSC cohort, and still 11 driver modules were remarkably associated with the survival of HNSC patients (Fig. 4E).

Next, we compared the prognostic value of driver modules with that of single genes within these modules in the TCGA-HNSC cohort. Our findings demonstrated that the prognostic values of driver modules surpassed those of individual genes within modules (Supplementary Figure S3C), suggesting that the combination of genetic mutations could provide a more comprehensive understanding of aberrant functions in cancer compared to solely single gene markers.

Finally, a risk score model based on the identified driver modules was constructed for HNSC patients to evaluate the combined predictive ability of these driver modules (Supplementary Table S3), and HNSC patients with high risk scores presented remarkably lower prognosis (median OS, 16.37 months vs. 21.10 months, log-rank $P < 0.0001$, Fig. 4G). Furthermore, the ROC curve analysis showed that model achieved good performance for predicting overall survival (AUC at 3 years: 0.669; AUC at 5 years: 0.696; AUC at 10 years: 0.787, Supplementary Figure S3D). Importantly, the risk model also exhibited good predictive performance in the ICGC-HNSC cohort (log-rank $P < 0.0001$, AUC at 10 years: 0.716, Fig. 4H and Supplementary Figure S3D). Overall, these findings indicated that ProgModule can effectively identify driver modules to predict patient prognosis, and these modules may affect cancer progresses by activating or inhibiting different anti-cancer pathways.

Comparison of molecular regulatory mechanisms across different cancer types

To further confirm the robustness of ProgModule, we also extended it to 22 additional cancer types from the TCGA database, respectively. Our analysis demonstrated that ProgModule successfully identified driver modules with significant prognostic value, and the combination of these modules could serve as powerful prognostic factors (refer to Supplementary Materials for detailed results, Supplementary Figure S4-6 and Supplementary Table S4). However, the driver modules were not identified in some cancer types (such as colon adenocarcinoma [COAD], rectum adenocarcinoma [READ], and sarcoma [SARC], etc.) due to their low mutation occurrence.

ProgModule identified driver modules for predicting patient survival not simply defined by the total number of mutations in a set of genes, but rather by considering the mutation pattern in the population and combination-mediated mechanism in cancer prognosis. Hence, we further explore whether the modules identified in specific cancer type could predict patient outcomes across different cancer types. The predictive ability was assessed using Kaplan-Meier survival analysis, where modules

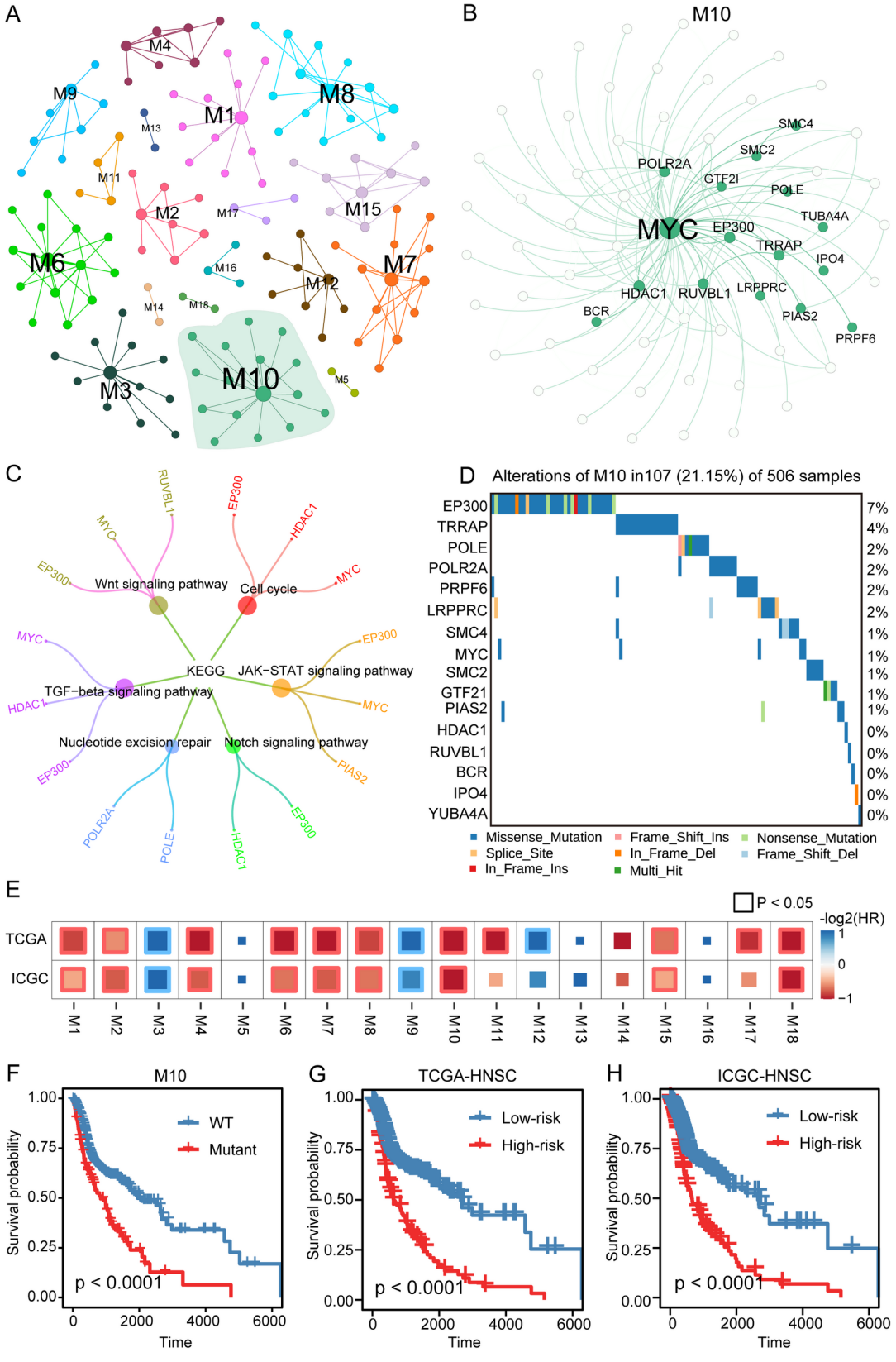

Fig. 4 Applying ProgModule to the HNSC cohort. (**A**) The landscape of candidate modules for HNSC. **(B)** The detailed information of the module M10; **(C)** The enriched KEGG pathways of genes involved in the module M10; **(D)** The waterfall plot of genes within the module M10; **(E)** The heatmap for the Cox regression P values of candidate modules; **(F)** Kaplan-Meier survival analysis of OS comparing the M10-Mutant and WT groups; **(G)** The Kaplan-Meier survival curve of risk model in TCGA-HNSC; **(H)** The Kaplan-Meier survival curve of risk model in ICGC-HNSC

achieving a significant log-rank *p*-value (< 0.05) were considered prognostic. Remarkably, we observed that modules identified for each cancer type can be generalized in predicting prognosis across multiple cancer types (Fig. 5A and Supplementary Figure S7), except for pancreatic adenocarcinoma (PAAD) and prostate adenocarcinoma (PARD). For instance, modules identified in lung adenocarcinoma can be used to predict prognosis across all the cancer types, while driver modules identified in breast cancer could serve as predictive markers for patient survival across all cancer types except skin cutaneous melanoma (SKCM). Altogether, these findings underscored ProgModule's predictive power in identifying prognostic driver modules across diverse cancer types.

To understand the molecular regulatory mechanism underlying cancer genetics, we then investigated the conservativeness of driver modules across 19 cancer types. Initially, we calculated the Jaccard index to measure the similarity of module pairs across all cancer types, identifying 15 conserved module pairs with a Jaccard index > 0.4. These conserved module pairs originated from 23 modules (11% of the total candidate modules) across 19 cancer types (Fig. 5B), highlighting that most driver modules were cancer specific. And the variation in the number of conserved modules within each cancer type reflected the diversity of shared genetic signatures across different cancer types. To further explore the underlying module preservation patterns, we then compared genes within driver modules across different cancer types by counting the overlapped genes between any two cancer types (Fig. 5C). We found that only a small fraction of selected genes overlapped between any two cancer types except BRCA and lung cancers, consistent with the limited conservation of candidate driver modules (only 11% of the total candidate modules were conserved). The high number of overlapping genes between BRCA and lung cancers (LUAD and LUSC) may reflect shared biological functions, such as immune evasion, inflammatory responses, and angiogenesis. And treatments for breast cancer, including radiotherapy and chemotherapy, are known risk factors for lung cancer, contributing to its occurrence as a secondary malignancy [23, 24]. Additionally, metastatic interactions, as

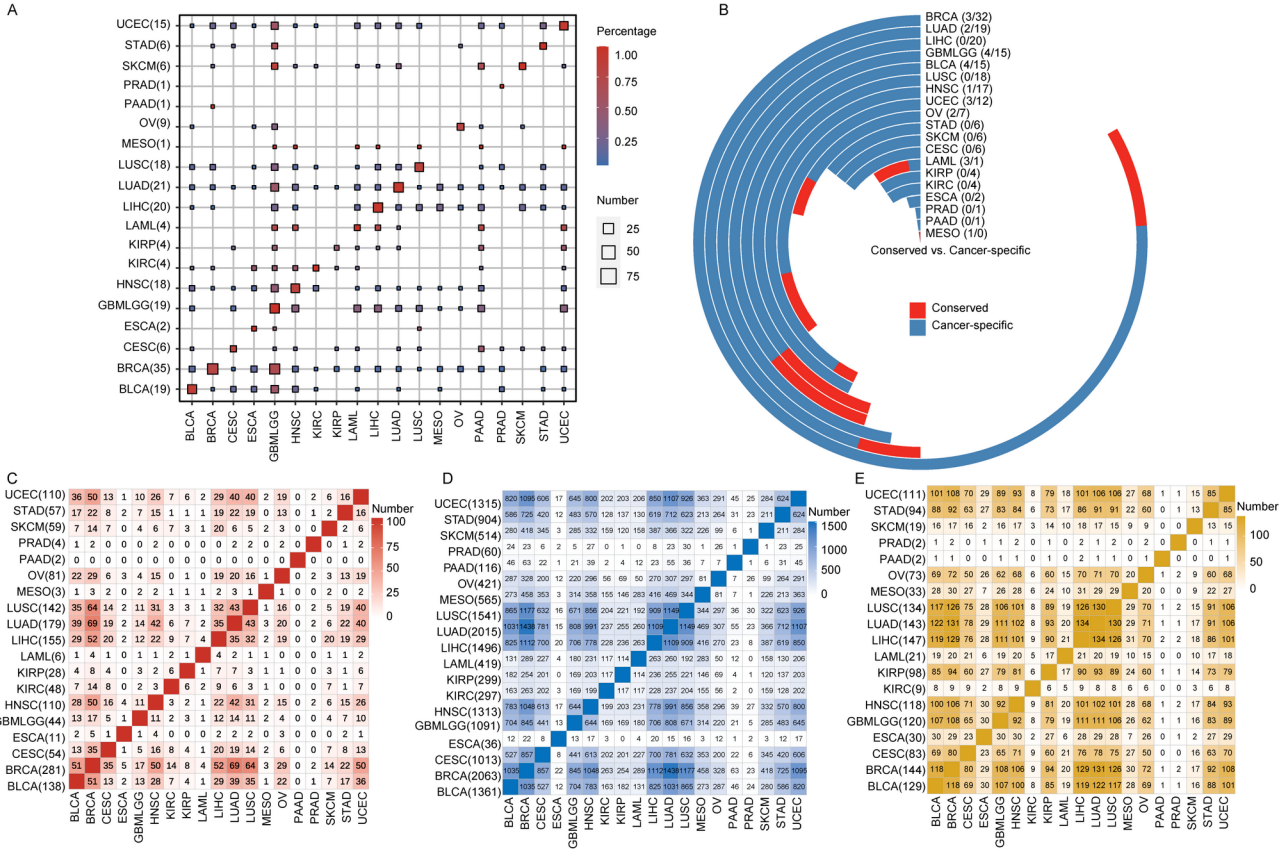

Fig. 5 The modules identified by ProgModule were cancer-specific. (**A**) Applying the cancer-specific model to other cancer type datasets, each pixel represents the number of modules identified in one cancer type that can predict patient prognosis in another cancer type. **(B)** The number of conserved modules across all cancer types. **(C)** Heatmap showing the number of overlapping Gene Ontology (GO) terms between cancer types. The color scale represents the number of shared GO terms, with darker colors indicating a higher number of overlaps. **(F)** The number of overlapped KEGG pathways between two cancer types

demonstrated by recent studies, indicate that breast cancer cells can reprogram the stromal microenvironment to facilitate pulmonary metastasis [25]. Surprisingly, gene ontology (GO) and KEGG function enrichment analyses revealed considerable overlap of GO terms (Fig. 5D) and KEGG pathways (Fig. 5E) between any two cancer types, implying shared regulatory mechanism in cancer development. For example, although no gene overlap was observed between STAD and MESO, they shared numerous GO terms and pathways (Supplementary Table S5), including cell cycle G1/S phase transition [26, 27], p53 binding [21], and the Wnt signaling pathway [28]. The dysregulation of the cell cycle is a common feature in many cancers, often caused by factors such as aberrant expression of CDK and mutations in p53, which may lead to excessive activation of the G1/S phase transition, thereby promoting tumor development [29]. Furthermore, analyze of the top ten significantly enriched biological functions in different cancers revealed that most cancers were enriched in the same biological functions, including cell cycle, DNA binding, and p53 binding, which highlights the shared underlying mechanism in cancer pathogenesis and partly explains why modules identified in specific cancers can predict the prognosis of patients with other cancer types (Supplementary Figure S8). These findings revealed that although module genes were cancer-specific, they may participate in common or similar biological processes, suggesting a single biological pathway or process may be dysregulated through different routes.

Robustness analysis of ProgModule

To further assess the robustness of ProgModule, we explored its performance with respect to the model parameters, including stopping criteria, initial seed nodes, network topology. Specifically, the greedy algorithm used for module identification initially applied an improvement rate of $r = 0.05$. To evaluate the impact of this parameter, we applied ProgModule to each cancer dataset with r values ranging from 0.01 to 0.10 at an increment of 0.01, and subsequently compared the identified modules at different values of r with those obtained at $r = 0.05$. The results showed that the identified modules remained relatively consistent across all r values (Supplementary Figure S9), demonstrating the negligible effect of the improvement rate on the results. Additionally, we also explored the effect of network topology by successively removing 2%, 4%, 6%, 8%, and 10% of the edges from the PPI network, while keeping other parameters unchanged. For each removal, we compared the identified modules in each instance of removal with original modules. The results revealed that relatively congruent modules were identified even after removing 10% of the network edges (Supplementary Figure S10), attesting to the robustness of ProgModule with respect to network topology. Altogether, these results demonstrate that the ProgModule is robust to the model parameters such as stopping criteria, network topology.

The modules identified by ProgModule could predict the immunotherapy response in cancer

Patient-specific neoantigens arising from tumor-specific mutations have been proposed as predictive biomarkers for response to immunotherapy [30, 31]. To investigate whether the mutated driver modules can predict immunotherapy response, we applied ProgModule to the Van Allen cohort comprising 105 melanoma patients [32]. Utilizing 591 driver genes sourced from the NCG database as seed nodes, we ultimately pinpointed 5 statistically significant driver modules. Through functional enrichment analysis, these modules were found to be participated in important biological processes (Fig. 6A). Besides, all modules were markedly linked to survival in melanoma patients, with three modules (M1, M2, and M4) were identified as risk factors and two modules (M3 and M5) as protective factors for OS of patients (Fig. 6B). Moreover, patients with mutations in modules M1, M2, and M4 exhibited remarkable higher immunotherapy response rates, whereas those with mutations in modules M3 and M5 displayed lower immunotherapy response rates compared to those without mutations in these modules (Supplementary Figure S11A-E).

Furthermore, the risk modules (M1, M2, and M4) showed obviously co-occurrences between each other, which were mutually exclusive with the protective modules (M3 and M5) (Fig. 6B). According to these modules, a risk score model was constructed using the multivariable Cox proportional hazards regression. Patients with low risk scores exhibited longer overall survival (median OS, 26.60 months vs. 6.63 months, log-rank $P < 0.0001$, Fig. 6C) and higher objective response rates (ORR) to immunotherapy (31.1% vs. 5%, Fisher's exact test $P = 0.0004$, Fig. 6D) compared to those with high risk scores. To evaluate the effectiveness of the risk score model, we applied it to another independent immunotherapy cohort (Miao cohort) [33]. With the risk score, patients were stratified into subgroups with remarkably distinct prognosis (median OS, 24.93 months vs. 9.47 months, log-rank test $P < 0.0001$, Fig. 6E) and varying benefits from immunotherapy (35.3% vs. 15.2%, Fisher's exact test $P = 0.0067$, Fig. 6F). Finally, we assessed the predictive ability of our module-based risk model in comparison with several published immunotherapy biomarkers, including cancer-associated fibroblasts (CAF), PD-L1 expression levels, CTL.flag, M2 type of tumor-associated macrophages (TAM.M2), tumor immune dysfunction and exclusion (TIDE), TMB, as well as signatures proposed by Wang et al. [34] and Long et al. [35].

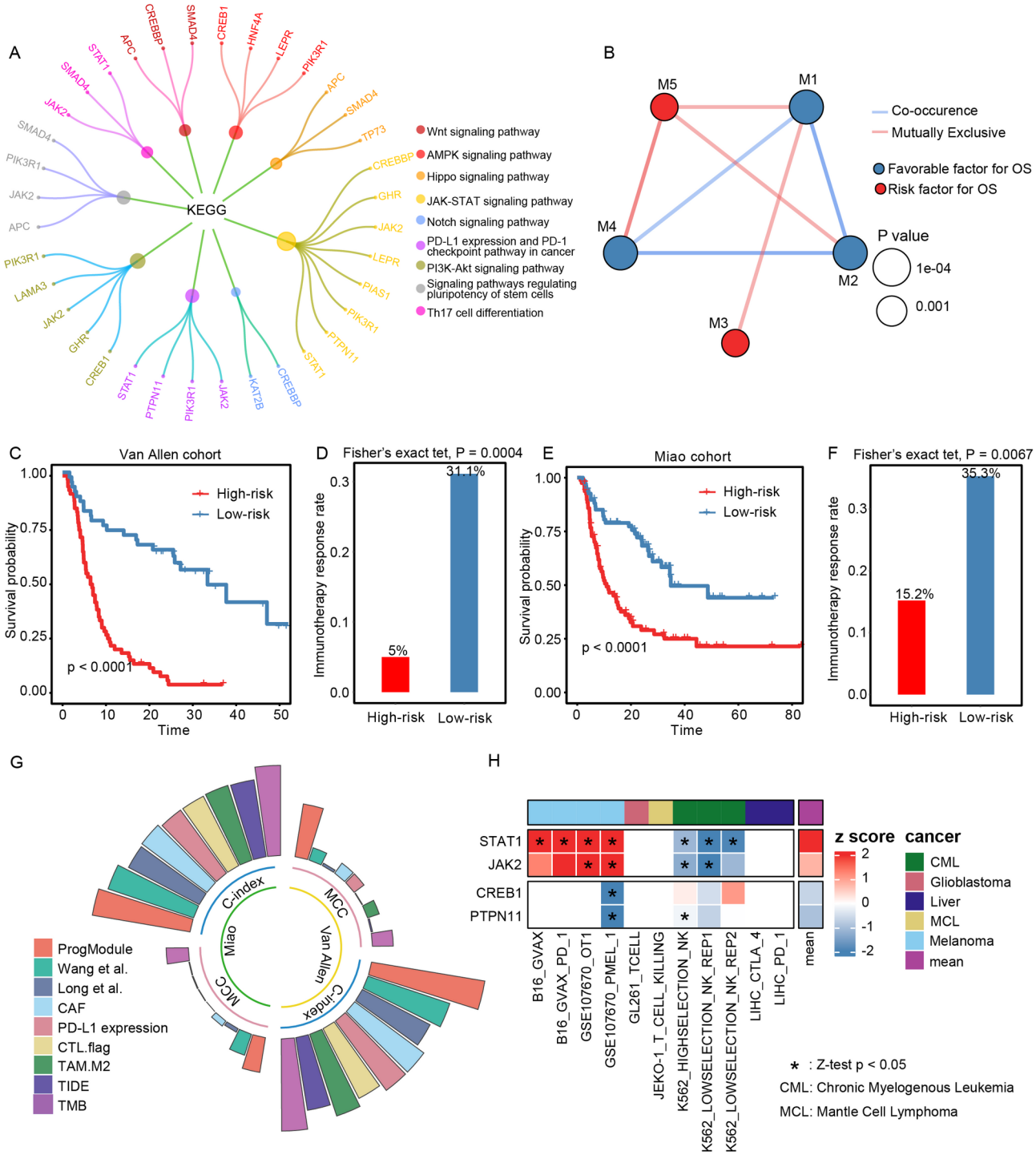

Fig. 6 ProgModule can be used to predict the clinical outcomes of patients receiving immunotherapy. (**A**) The enriched KEGG pathways of genes involved in all candidate modules in the Van Allen cohort. **(B)** Circle plot depicting the impact on melanoma overall survival of five candidate module mutations. **(C)** Kaplan-Meier survival analysis of OS comparing the High-risk and Low-risk groups from the Van Allen cohort. **(D)** Comparison of ORR between the High-risk and Low-risk groups in the Van Allen cohort. **(E)** Kaplan-Meier survival analysis of OS comparing the High-risk and Low-risk groups from the Miao cohort. **(F)** Comparison of ORR between the High-risk and Low-risk groups in the Miao cohort. **(G)** Compared the performance of our method with other published immunotherapy biomarkers based on C-index and MCC. **(H)** Heatmap depicting the Z score of seven candidate genes in the top 10% of ranked genes across different CRISPR datasets

Applying these biomarkers in the Van Allen and Miao cohorts, respectively, we observed that our module-based risk model outperformed these biomarkers in predicting immunotherapy benefit through comparing C-index and Matthews Correlation Coefficient (MCC) (Fig. 6G and Supplementary Figure S11F-M), highlighting the promising potential of driver modules identified by ProgModule as biomarkers for predicting immunotherapy efficacy.

Although immunotherapy has indeed revolutionized cancer treatment, it is accompanied by serious adverse effects [36], highlighting the urgent need for a deeper understanding of the biological processes that underpin tumor progression, which are not only for improving risk stratification but also to identify new possible therapeutic strategies. CRISPR and related tools have significantly advanced our understanding of therapeutic vulnerabilities in cancer [37]. Hence, we investigated potential therapeutic targets among driver module genes using published CRISPR screen data. Firstly, we collected 11 CRISPR datasets that assessed the individual effects of gene knockout on tumor immunity, sourced from various independent studies (Supplementary Table S6). In the original studies associated with these datasets, the Z-test was employed to quantify the phenotypic impact of gene knockout on the anti-tumor immunity. Genes with positive Z-scores, termed as immune-resistant genes, may promote anti-tumor immunity after knockout, whereas genes with negative Z-scores, termed as immune-sensitive genes, are likely to suppress anti-tumor immunity after knockout. We then filtered genes with statistically significant effects (Z-test P-value < 0.05) from each dataset and subsequently intersected them with genes in our candidate driver modules. Thus, we identified four genes within our candidate driver modules significantly associated with anti-tumor immunity: two immune-sensitive (CREB1 and PTPN11) genes and two immune-resistant genes (STAT1 and JAK2) (Fig. 6H). Notably, these genes exhibited consistent effects on anti-tumor immunity in melanoma datasets but showed variability in other cancer datasets, such as chronic myelogenous leukemia, suggesting the underlying anti-tumor mechanisms varied from cancer. Researchers have highlighted that PTPN11-encoded SHP2 plays a critical role in regulating various tumorigenesis-related signaling pathways, making it as a promising target for anti-tumor drug development [38, 39]. Additionally, the deactivation of STAT1 enhances anti-tumor responses by activating natural killer cells, macrophages, and CD8 + T cell-mediated cytolytic activity [40, 41]. Furthermore, Chan et al. demonstrated that the IL-6/JAK1 pathway drives PD-L1 Y112 phosphorylation, thereby promoting cancer immune evasion [42]. Consequently, several IL-6/JAK1 pathway-blocking antibodies or inhibitors have received FDA approval for the treatment of hepatocellular carcinoma [42, 43]. Furthermore, prospective validation in vitro and in vivo experiments has further confirmed the utility of the identified candidate driver genes as predictive biomarkers [44–47], underscoring the clinical relevance of our findings. In conclusion, these findings implied that genes within driver modules could serve as potential therapeutic targets in synergy with the immune-checkpoint blockade, which provides clinicians with novel insights into drug development.

Discussion

Cancer originates from dysregulated cell proliferation driven by critical driver gene mutations. Prior studies indicated that these driver genetic mutations tend to converge on a limited number of important biological pathways, characterized by high coverage and exclusivity [7, 8]. Despite numerous algorithms have been developed to uncover mutational signatures from somatic mutation catalogs balancing these two properties [5, 48], they often suffer from high computational complexity, high false-positive rate, and limited clinical applications. To advance our understanding of the mechanisms driving cancer initiation, we propose the development of computational approaches to detect driver modules that incorporate their prognostic significance while balancing mutation coverage and exclusivity. These algorithms may reveal complex interactions between genetic mutations and survival, offering critical insights for cancer diagnosis.

In our study, we introduced a novel algorithm, named ProgModule, designed to detect mutation driver modules for cancer prognosis and immunotherapy prediction. ProgModule innovatively defined a PRMEM score, which optimizes the balance between the exclusive coverage of mutations and the incorporation of mutation combination mechanism critical for cancer prognosis. Applying ProgModule to each of the 24 cancer type datasets, respectively, we found that mutations in identified driver modules were remarkably related to patient prognosis. Notably, we observed that the combination of these driver modules could also serve as effective prognostic biomarkers in both TCGA and ICGC cohorts. Importantly, the majority of driver modules showed superior predictive power for patient survival compared to single-gene markers. Besides, robustness analysis demonstrated that the ProgModule is robust to the model parameters such as stopping criteria, network topology. Furthermore, our analysis also revealed that ProgModule exhibited good predictive performance for predicting immunotherapy benefits and outperformed other published immunotherapy biomarkers, including CAF, PD-L1 expression levels, CTL.flag, TAM.M2, TIDE, TMB, as well as signatures proposed by Wang et al. and Long et al. Finally, ProgModule was developed as

an R-based tool freely available at https://CRAN.R-project.org/package=ProgModule.

Our approach presents several advantages. First, we innovatively defined a PRMEM score, which effectively optimizes the balance between the exclusive coverage of mutations and the incorporation of mutation combination mechanism critical for cancer prognosis, for which no approach has yet been developed to do this. Secondly, ProgModule outperformed existing methodologies in predicting survival outcomes and immunotherapy response for various cancers and exhibited superior clinical utility. However, like other computational approaches, ProgModule has its limitations. For instance, no driver module was detected in some cancer types (such as COAD, READ, and SARC, etc.) due to their low mutation occurrence, highlighting the necessity for further refining ProgModule to enhance its sensitivity in detecting driver modules across diverse cancer types in future.

Conclusions

We developed ProgModule, an advanced computational framework designed to detect mutation driver modules for cancer prognosis and immunotherapy response prediction. Compared to state-of-the-art methods, ProgModule demonstrated superior performance in identifying cancer driver genes and accurately predicting patient survival across 24 cancer types. Moreover, ProgModule was also applicable to immunotherapy cohorts and can predict immunotherapeutic benefit more effectively than existing signatures, highlighting its universality and scalability. Additionally, analyses based on published CRISPR data revealed that genes within these modules may serve as potential therapeutic targets, offering novel insights into the advancement of precision oncology.

Abbreviations

PRMEM	Prognosis-Related Mutually Exclusive Mutation
BLCA	bladder cancer
HNSC	head and neck squamous cell carcinoma
TCGA	the Cancer Genome Atlas Program
ICGC	the International Cancer Genome Consortium
NCG	the Network of Cancer Genes & Healthy Drivers
PPI	protein-protein interaction
OS	overall survival
ROC	receiver operating characteristic curve
AUC	area under the ROC
COAD	colon adenocarcinoma
READ	rectum adenocarcinoma
SARC	sarcoma
PAAD	pancreatic adenocarcinoma
PARD	prostate adenocarcinoma
SKCM	skin cutaneous melanoma
GO	gene ontology
ORR	objective response rates
DDR	DNA damage response

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-06497-0.

Supplementary Material 1

Acknowledgements

Not applicable.

Authors' contributions

JH and XL conceived and designed the study. XL, BP, and XZ developed the method. JL, SL, and MJ analyzed the data and implemented the methodology. YH and JW provided constructive discussions. JH and XL drafted the manuscript. All the authors read and agreed to the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China under Grant 62372143 and 62072145; the Natural Science Foundation of Heilongjiang Province under Grant LH2019C042.

Data availability

ProgModule is freely available as an R package at https://CRAN.R-project.org/package=ProgModule, and the data are accessible in a public, open-access repository.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All the authors read and approved the publication of the final manuscript.

Competing interests

The authors declare no conflict of interests.

Received: 23 December 2024 / Accepted: 13 April 2025
Published online: 08 May 2025

References

1. Yates LR, Campbell PJ. Evolution of the cancer genome. Nat Rev Genet. 2012;13:795–806.
2. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144:646–74.
3. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. 2004;10:789–99.
4. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA. Jr., Kinzler KW: Cancer genome landscapes. Science. 2013;339:1546–58.
5. Leiserson MD, Vandin F, Wu HT, Dobson JR, Eldridge JV, Thomas JL, Papoutsaki A, Kim Y, Niu B, McLellan M, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. Nat Genet. 2015;47:106–14.
6. Babur O, Gonen M, Aksoy BA, Schultz N, Ciriello G, Sander C, Demir E. Systematic identification of cancer driving signaling pathways based on mutual exclusivity of genomic alterations. Genome Biol. 2015;16:45.
7. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455:1061–8.
8. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. Nature. 2011;474:609–15.
9. Ciriello G, Cerami E, Sander C, Schultz N. Mutual exclusivity analysis identifies oncogenic network modules. Genome Res. 2012;22:398–406.
10. Miller CA, Settle SH, Sulman EP, Aldape KD, Milosavljevic A. Discovering functional modules by identifying recurrent and mutually exclusive mutational patterns in tumors. BMC Med Genomics. 2011;4:34.
11. Vandin F, Upfal E, Raphael BJ. De Novo discovery of mutated driver pathways in cancer. Genome Res. 2012;22:375–85.

12. Garraway LA, Lander ES. Lessons from the cancer genome. Cell. 2013;153:17–37.
13. Consortium ITP-CAWG. Pan-cancer analysis of whole genomes. Nature. 2020;578:82–93.
14. Bradner JE, Hnisz D, Young RA. Transcriptional addiction in Cancer. Cell. 2017;168:629–43.
15. Sedgwick P. Multiple hypothesis testing and Bonferroni's correction. BMJ. 2014;349:g6284.
16. Huang YH, Cai K, Xu PP, Wang L, Huang CX, Fang Y, Cheng S, Sun XJ, Liu F, Huang JY, et al. CREBBP/EP300 mutations promoted tumor progression in diffuse large B-cell lymphoma through altering tumor-associated macrophage polarization via FBXW7-NOTCH-CCL2/CSF1 axis. Signal Transduct Target Ther. 2021;6:10.
17. Yan H, Parsons DW, Jin G, McLendon R, Rasheed BA, Yuan W, Kos I, Batinic-Haberle I, Jones S, Riggins GJ, et al. IDH1 and IDH2 mutations in gliomas. N Engl J Med. 2009;360:765–73.
18. Nunes L, Li F, Wu M, Luo T, Hammarstrom K, Torell E, Ljuslinder I, Mezheyeuski A, Edqvist PH, Lofgren-Burstrom A, et al. Prognostic genome and transcriptome signatures in colorectal cancers. Nature. 2024;633:137–46.
19. Germani MM, Vetere G, Santamaria F, Intini R, Ghelardi F, Bensi M, Boccaccino A, Minelli A, Carullo M, Ciraci P, et al. Treatment of patients with BRAF(V600E)-mutated metastatic colorectal cancer after progression to Encorafenib and cetuximab: data from a real-world nationwide dataset. ESMO Open. 2024;9:102996.
20. Li X, He Y, Jiang Y, Pan B, Wu J, Zhao X, Huang J, Wang Q, Cheng L, Han J. PathwayTMB: A pathway-based tumor mutational burden analysis method for predicting the clinical outcome of cancer immunotherapy. Mol Ther Nucleic Acids. 2023;34:102026.
21. Evan GI, Vousden KH. Proliferation, cell cycle and apoptosis in cancer. Nature. 2001;411:342–8.
22. Zhang Y, Wang X. Targeting the Wnt/beta-catenin signaling pathway in cancer. J Hematol Oncol. 2020;13:165.
23. Lorigan P, Radford J, Howell A, Thatcher N. Lung cancer after treatment for Hodgkin's lymphoma: a systematic review. Lancet Oncol. 2005;6:773–9.
24. Lorigan P, Califano R, Faivre-Finn C, Howell A, Thatcher N. Lung cancer after treatment for breast cancer. Lancet Oncol. 2010;11:1184–92.
25. Cohen N, Mundhe D, Deasy SK, Adler O, Ershaid N, Shami T, Levi-Galibov O, Wassermann R, Scherz-Shouval R, Erez N. Breast Cancer-Secreted factors promote lung metastasis by signaling systemically to induce a fibrotic pre-metastatic niche. Cancer Res. 2023;83:3354–67.
26. Fuchs Y, Steller H. Programmed cell death in animal development and disease. Cell. 2011;147:742–58.
27. Fulda S. Tumor resistance to apoptosis. Int J Cancer. 2009;124:511–5.
28. Kumagai S, Koyama S, Nishikawa H. Antitumour immunity regulated by aberrant ERBB family signalling. Nat Rev Cancer. 2021;21:181–97.
29. Kern SE. p53: tumor suppression through control of the cell cycle. Gastroenterology. 1994;106:1708–11.
30. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. Science. 2015;348:69–74.
31. Sahin U, Tureci O. Personalized vaccines for cancer immunotherapy. Science. 2018;359:1355–60.
32. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Foppen MHG, Goldinger SM, et al. Genomic correlates of response to CTLA-4 Blockade in metastatic melanoma. Science. 2015;350:207–11.
33. Miao D, Margolis CA, Vokes NI, Liu D, Taylor-Weiner A, Wankowicz SM, Adeegbe D, Keliher D, Schilling B, Tracy A, et al. Genomic correlates of response to immune checkpoint Blockade in microsatellite-stable solid tumors. Nat Genet. 2018;50:1271–81.
34. Wang Z, Zhao J, Wang G, Zhang F, Zhang Z, Zhang F, Zhang Y, Dong H, Zhao X, Duan J, et al. Comutations in DNA damage response pathways serve as potential biomarkers for immune checkpoint Blockade. Cancer Res. 2018;78:6486–96.
35. Long J, Wang D, Wang A, Chen P, Lin Y, Bian J, Yang X, Zheng M, Zhang H, Zheng Y, et al. A mutation-based gene set predicts survival benefit after immunotherapy across multiple cancers and reveals the immune response landscape. Genome Med. 2022;14:20.
36. Riley RS, June CH, Langer R, Mitchell MJ. Delivery technologies for cancer immunotherapy. Nat Rev Drug Discov. 2019;18:175–96.
37. Katti A, Diaz BJ, Caragine CM, Sanjana NE, Dow LE. CRISPR in cancer biology and therapy. Nat Rev Cancer. 2022;22:259–79.
38. Liu M, Gao S, Elhassan RM, Hou X, Fang H. Strategies to overcome drug resistance using SHP2 inhibitors. Acta Pharm Sin B. 2021;11:3908–24.
39. Wang LB, Karpova A, Gritsenko MA, Kyle JE, Cao S, Li Y, Rykunov D, Colaprico A, Rothstein JH, Hong R, et al. Proteogenomic and metabolomic characterization of human glioblastoma. Cancer Cell. 2021;39:509–e528520.
40. Takeda K, Akira S. STAT family of transcription factors in cytokine-mediated biological responses. Cytokine Growth Factor Rev. 2000;11:199–207.
41. Chang HC, Han L, Goswami R, Nguyen ET, Pelloso D, Robertson MJ, Kaplan MH. Impaired development of human Th1 cells in patients with deficient expression of STAT4. Blood. 2009;113:5887–90.
42. Chan LC, Li CW, Xia W, Hsu JM, Lee HH, Cha JH, Wang HL, Yang WH, Yen EY, Chang WC, et al. IL-6/JAK1 pathway drives PD-L1 Y112 phosphorylation to promote cancer immune evasion. J Clin Invest. 2019;129:3324–38.
43. Johnson DE, O'Keefe RA, Grandis JR. Targeting the IL-6/JAK/STAT3 signalling axis in cancer. Nat Rev Clin Oncol. 2018;15:234–48.
44. Wang Z, Chen Y, Fang H, Xiao K, Wu Z, Xie X, Liu J, Chen F, He Y, Wang L, et al. Reprogramming cellular senescence in the tumor microenvironment augments cancer immunotherapy through multifunctional nanocrystals. Sci Adv. 2024;10:eadp7022.
45. Nicolet CM, Surfus JM, Hank JA, Sondel PM. Transcription factor activation in lymphokine activated killer cells and lymphocytes from patients receiving IL-2 immunotherapy. Cancer Immunol Immunother. 1998;46:327–37.
46. Nie J, Wang C, Zheng L, Liu Y, Wang C, Chang Y, Hu Y, Guo B, Pan Y, Yang Q, et al. Epigenetic agents plus anti-PD-1 reprogram the tumor microenvironment and restore antitumor efficacy in hodgkin lymphoma. Blood. 2024;144:1936–50.
47. Nguyen TT, Ramsay L, Ahanfeshar-Adams M, Lajoie M, Schadendorf D, Alain T, Watson IR. Mutations in the IFNgamma-JAK-STAT pathway causing resistance to immune checkpoint inhibitors in melanoma increase sensitivity to oncolytic virus treatment. Clin Cancer Res. 2021;27:3432–42.
48. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013;499:214–8.

Publisher's note