



OPEN

## Construction of an anoikis-associated lncRNA-miRNA-mRNA network reveals the prognostic role of $\beta$ -elemene in non-small cell lung cancer

Kai Tan<sup>1</sup>, Changhui Zhang<sup>1</sup>, Zuomei He<sup>2,3</sup> & Puhua Zeng<sup>2,3</sup>✉

$\beta$ -Elemene is the main active ingredient in *Curcuma Rhizoma* that exerts antitumour effects. Anoikis affects tumour development through various biological pathways in non-small cell lung cancer (NSCLC), but the regulation between  $\beta$ -elemene and anoikis remains to be explored. First, we explored the molecular expression patterns of anoikis-associated genes (AAGs) using consensus clustering and characterized the impact of AAGs on patient prognosis, clinical characteristics, and genomic instability. In addition, we revealed that AAG regulatory genes have rich interactions with  $\beta$ -elemene targets, and established a lncRNA-miRNA-mRNA network to explain the effect of  $\beta$ -elemene on anoikis. Finally, to reveal the prognostic effect of their correlation, the prognostic scoring model and clinical nomogram of  $\beta$ -elemene and anoikis were successfully established by least absolute shrinkage and selection operator (LASSO) and random forest algorithms. This prognostic scoring model containing noncoding RNA (ncRNA) can indicate the immunotherapy and mutational landscape, providing a novel theoretical basis and direction for the study of the antitumour mechanism of  $\beta$ -elemene in NSCLC patients.

### Abbreviations

NSCLC	Non-small cell lung cancer
TCM	Traditional Chinese medicine
CHM	Chinese herbal medicine
ZIC2	Zic family member 2
Src	Steroid receptor coactivator
FAK	Focal adhesion kinase
PTEN	Phosphatase and tensin homolog
AKT	Protein kinase B
FAM188B	Family with sequence similarity 188 member B
EGFR	Epidermal growth factor receptor
AMPK	The AMP-activated protein kinase
MAPK	Mitogen-activated protein kinases
PI3K	Phosphoinositide 3-kinase
mTOR	Mechanistic target of rapamycin
ncRNA	Noncoding RNA
lncRNA	Long noncoding RNA
miRNA	MicroRNA

<sup>1</sup>Hunan University of Chinese Medicine, Changsha 410208, Hunan, People's Republic of China. <sup>2</sup>Cancer Research Institute of Hunan Academy of Traditional Chinese Medicine, Changsha 410006, Hunan, People's Republic of China. <sup>3</sup>Hunan Academy of Traditional Chinese Medicine Affiliated Hospital, Changsha 410006, Hunan, People's Republic of China. ✉email: zph120@126.com

circRNA	Circular RNA
ceRNA	Competing endogenous RNA
TCGA	The Cancer Genome Atlas
CNV	Copy number variation
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
AAG	Anoikis-associated gene
PCA	Principal component analysis
GSEA	Gene set enrichment analysis
BAG	$\beta$ -Elemene and anoikis-associated genes
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
LASSO	Least absolute shrinkage and selection operator
ROC	Receiver operating characteristic
DCA	Decision curve analysis
TIDE	Tumor Immune Dysfunction and Exclusion
MAF	Mutation annotation format
ECM	Extracellular matrix
AUC	Area under the curves
TMB	Tumour mutation burden
PPARA	Peroxisome proliferator activated receptor alpha
CNR2	Cannabinoid receptor 2
FOXO1	Forkhead box-1
NAC	Acetyl-Cysteine
PDK1	3-Phosphoinositide-dependent protein kinase 1
KLF10	KLF transcription factor 10
LRP8	Low-density lipoprotein receptor-related protein-8
SATB2	Special AT-rich sequence-binding protein 2
TNS1	Tensin1
FIH	Hypoxia inducible factor 1 $\alpha$ inhibitor
HIF	Hypoxia inducible factor
VDR	Vitamin D receptor
VEGFC	Vascular endothelial growth factor C
PTGS1	Prostaglandin-endoperoxide synthase 1
PD-L1	Programmed cell death ligand 1
PTGS2	Prostaglandin-endoperoxide synthase 2
PGE2	Prostaglandin E2
MSCs	Mesenchymal stem cells
5-FU	5-Fluorouracil
NF- $\kappa$ B	Nuclear factor-kappaB

Lung cancer remains the leading cause of cancer death, accounting for 18% of all cancer deaths and 11.4% of all cancers diagnosed according to GLOBOCAN 2020, making it the deadliest type of cancer worldwide<sup>1</sup>. Non-small cell lung cancer (NSCLC) accounts for approximately 85% of lung cancer cases<sup>2</sup>. Due to the insidious symptoms of early-onset NSCLC, many patients are in advanced stages at the time of diagnosis, and thus systemic drug treatments such as chemotherapy, targeted therapy and immunotherapy are usually needed but do not provide the desired benefits due to drug resistance and adverse effects<sup>3,4</sup>.

Traditional Chinese medicine (TCM), especially Chinese herbal medicine (CHM), has been widely used in China and many other countries for the treatment of cancer. CHM not only relieves symptoms and improves the quality of life of cancer patients but also reduces the adverse effects and complications caused by chemotherapy, targeted therapy or immunotherapy<sup>5</sup>. *Curcumae Rhizoma* is a commonly used drug in the TCM treatment of lung cancer, and  $\beta$ -elemene (C<sub>15</sub>H<sub>24</sub>) is the main active ingredient extracted from it. Several previous studies have demonstrated that  $\beta$ -elemene plays an antitumour role in NSCLC patients not only by inhibiting NSCLC cell proliferation, invasion and migration and inducing apoptosis of NSCLC cells but also by increasing the sensitivity of drugs and other mechanisms, which are closely related to the development of NSCLC<sup>6–8</sup>.

Anoikis is a specific form of programmed cell death induced by the loss of cellular exposure to the extracellular matrix, which plays a key role in the maintenance of tissue homeostasis<sup>9</sup>. However, tumour cells have the ability to evade cell death and usually show resistance to anoikis, which leads to tumour progression and metastatic spread of cancer cells<sup>10</sup>. An increasing number of studies have confirmed the involvement of anoikis in NSCLC biological processes: Liu et al.<sup>11</sup> showed that silencing *Zic* family member 2 (*ZIC2*) could downregulate the migration, invasion and anoikis resistance ability of NSCLC cells by inhibiting steroid receptor coactivator/focal adhesion kinase (Src/FAK) signalling; McCarroll et al.<sup>12</sup> found that  $\beta$ III-tubulin induced NSCLC development and anoikis resistance through the phosphatase and tensin homolog/protein kinase B (PTEN/AKT) signalling axis; Jang et al.<sup>13</sup> demonstrated that knockdown of family with sequence similarity 188 member B (*FAM188B*) downregulated the activity of various signalling pathways involved in anti-anoikis downstream of epidermal growth factor receptor (EGFR), sensitizing NSCLC cells to anoikis and inhibiting tumour metastasis. It is evident that anoikis resistance is regulated by multiple signalling pathways in NSCLC cells. Several studies have shown that the main pathways of  $\beta$ -elemene anti-NSCLC include the AMP-activated protein kinase/

mitogen-activated protein kinases (AMPK/MAPK), phosphoinositide 3-kinase (PI3K)/AKT/mechanistic target of rapamycin (mTOR) and FAK-Src pathways<sup>6,8,14</sup>. It is thus evident that  $\beta$ -elemene may exert anti-NSCLC effects by participating in the regulation of anoikis-related pathways and that the targets of  $\beta$ -elemene may play key roles in the anoikis process.

Noncoding RNAs (ncRNAs) are unique RNA transcripts that are widely found in eukaryotes, and a variety of ncRNAs, including long noncoding RNAs (lncRNAs), microRNAs (miRNAs), and circular RNAs (circRNAs), are oncogenic drivers and tumour suppressors of major tumours<sup>15</sup>. Extensive interactions also exist between ncRNAs, with lncRNAs usually acting as specific competing endogenous RNAs (ceRNAs), competing for complementary miRNA binding sites to influence and regulate the expression of cancer target genes<sup>16,17</sup>. A variety of ncRNAs play key roles in NSCLC and can influence NSCLC development through various mechanisms<sup>18–20</sup>. In particular, anoikis-associated ncRNAs have been shown to be key markers for tumour metastasis and progression, including breast cancer<sup>21</sup>, hepatocellular carcinoma<sup>22</sup>, and prostate cancer<sup>23</sup>. The latest research shows that lncRNA-miRNA interactions are successfully predicted based on multiple network algorithms, providing novel and valuable insights into ncRNA prediction of prognosis of NSCLC patients<sup>24,25</sup>. However, studies involving the regulatory relationship of anoikis-related ncRNAs in NSCLC have been less frequently reported. Additionally, it is not yet known whether the target of  $\beta$ -elemene interacts with anoikis-related ncRNAs. Thus, elucidating their roles in NSCLC may improve our understanding of the mechanism of action of  $\beta$ -elemene in anoikis as well as new therapeutic strategies against NSCLC.

In this study, we explored the molecular expression pattern of anoikis prognostic factors in NSCLC patients by mining The Cancer Genome Atlas (TCGA) database, and investigated the biological function and prognostic significance of these molecular clusters. In addition, by constructing a lncRNA-miRNA-mRNA network of anoikis and  $\beta$ -elemene targets, the regulatory relationship of anoikis-associated ncRNAs on  $\beta$ -elemene targets was clarified. Finally, potential targets were obtained by constructing a prognostic regression model, and binding stability was evaluated for targets and  $\beta$ -elemene through molecular docking, providing a theoretical basis and new possibilities for the diagnosis and treatment of NSCLC.

## Methods and materials

### Data acquisition

First, the transcriptional expression profiles, clinical information, survival information, somatic mutation data, copy number variation (CNV) data, and miRNA expression profiles (isoform expression quantification) of the lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) cohorts were obtained from the TCGA database (<https://portal.gdc.cancer.gov/>). The TCGA-LUAD cohort included data for 59 paracancerous tissue samples and 541 tumour tissue samples, and the TCGA-LUSC cohort included data for 51 paracancerous tissue samples and 502 tumour tissue samples. Using the ComBat function of the "sva" package, the gene expression profiles of the two cohorts were combined, batch effects were eliminated, and the merged matrix was normalized using the log<sub>2</sub> function. Only 970 NSCLC tumour samples with full transcriptome, miRNA expression profiles, and survival information were included for the construction of prognostic models. Second, 501 anoikis-associated genes (AAGs) with relevance scores greater than 0.4 were obtained from the GeneCards database<sup>26</sup> (<https://www.genecards.org/>). In addition, 139 AAGs were obtained from the Harmonizome database<sup>27</sup> (<http://amp.pharm.mssm.edu/Harmonizome>). A total of 638 AAGs were obtained after deduplicating. In addition, the PubChem database<sup>28</sup> (<https://pubchem.ncbi.nlm.nih.gov/>) was explored to obtain canonical SMILES for  $\beta$ -elemene (PubChem CID: 6,918,391). Potential targets of  $\beta$ -elemene were predicted with the SwissTarget-Prediction database<sup>29</sup> (<http://swisstargetprediction.ch/>). Next, the UniProt IDs of the 23 predicted targets were converted to gene symbols using the UniProt database<sup>30</sup> (<https://www.uniprot.org/>). Finally, 26 potential target genes of  $\beta$ -elemene were obtained.

### Molecular characterization of AAGs

First, the expression of AAGs in paracancerous and tumour tissues of NSCLC patients was compared to obtain AAGs that were significantly differentially expressed in tumour tissues ( $|\log_{2}FC| > 1$  and  $FDR < 0.05$ ). Then, univariate Cox regression analysis was performed on the differential AAGs to obtain the AAGs significantly associated with prognosis of NSCLC ( $P < 0.05$ ). To further explore the genomic activity of potentially prognostic AAGs, we assessed the frequency of CNV occurrence in these AAGs and visualized the location of CNVs. In addition, to explore the interactions of prognostically relevant AAGs, we analysed the interactions of proteins encoded by AAGs using the String database<sup>31</sup> (<https://string-db.org/>) (confidence > 0.4).

### Consensus clustering of AAGs

To explore the molecular expression patterns of prognostic AAGs, consensus clustering was performed based on the expression profiles of AAGs using the k-means algorithm of the "ConsensusClusterPlus" package<sup>32</sup>. The maximum number of clusters was set to 9. A resampling scheme was used to sample 80% of the sample and resampled 50 times to find stable and reliable subgroup classification. Principal component analysis (PCA) and survival analysis were applied to test the differences in clustering and to compare the expression levels of these AAGs between clusters. In addition, the correlation between AAG clusters and the clinical information of NSCLC patients was explored, and differences in biological functions between AAG clusters were compared by gene set enrichment analysis (GSEA).

### Construction and biological function analysis of the ceRNA network

To investigate the mechanism of action of  $\beta$ -elemene in NSCLC and its correlation with the anoikis process, we constructed an anoikis- $\beta$ -elemene ceRNA network. First, anoikis-associated differential mRNAs, lncRNAs, and

miRNAs ( $|\log FC| > 0.585$  and  $P < 0.05$ ) were obtained by differential significance analysis among AAG clusters for all genes. Then, the targeting relationships of these mRNAs and lncRNAs were predicted by the DIANA-LncBase tool<sup>33</sup> (<http://www.microrna.gr/LncBase>). The targets of differential miRNAs were also predicted by the TargetScan 8.0 database ([www.targetscan.org](http://www.targetscan.org)) and compared with anoikis-related differential mRNAs and  $\beta$ -elemene targets. In addition, by obtaining the targeting relationships between anoikis-related miRNAs and  $\beta$ -elemene targets and by performing correlation analyses between anoikis-related mRNAs and  $\beta$ -elemene targets, anoikis- $\beta$ -elemene gene relationship pairs were established for the  $\beta$ -elemene targets that are both miRNA targets and coexpressed genes of anoikis-related mRNAs. Finally, only mRNAs with relevance to  $\beta$ -elemene targets were retained, and a lncRNA-miRNA-mRNA network was established based on ncRNA-target relationship pairs. The  $\beta$ -elemene-regulated anoikis genes used to construct the ceRNA network were defined as  $\beta$ -elemene and anoikis-associated genes (BAGs). To explore the biological functions of BAGs, we analysed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment of BAGs constructing ceRNA networks ( $P < 0.05$  and  $Q < 0.1$ ).

### Prognostic modelling of feature BAGs

To further explore the impact of BAGs on patient survival, we used the random number method to divide NSCLC patients into training and validation groups at a ratio of 1:1 and then used least absolute shrinkage and selection operator (LASSO) and random forest algorithms to screen for prognostically relevant feature BAGs. First, in the training group, we used the "randomForest" package to construct a random forest model of OS.status predictions from BAGs by randomly cycling through all possible random numbers of the variables. The number of decision trees (ntree) contained in the random forest was 500, and the number of variables used in the nodes for binary trees (mtry) was 8. The Gini coefficient method was used to determine the importance of BAGs for OS.status prediction in the random forest model, and the top 10 BAGs were taken as feature genes. In addition, a LASSO regression model was constructed using the "glmnet" package with BAGs as the independent variable and OS.status as the dependent variable in the training group. The LASSO model was used for binary discrete dependent variables (family = "binomial"), and the cross-validated loss function was expressed in terms of the mean squared error (type.measure = "deviance") with a tenfold cross-test (nfolds = 10). The number of variables corresponding to the smallest mean square error was optimal, and these BAGs were screened to be feature genes. Finally, the intersection of the feature BAGs obtained by the two algorithms was taken and the intersecting genes were used to establish a multivariate Cox regression model for predicting the prognosis of NSCLC patients. The patient's prognostic risk score was calculated using the sum of model genes and coefficient products and denoted as the BAG\_Score. The training, validation, and entire groups were categorized into high and low BAG\_Score groups based on the median BAG\_Score of the training group. Finally, the accuracy of the prognostic model was assessed using receiver operating characteristic (ROC) curves and Kaplan–Meier survival analyses, and multivariate Cox regression analysis was performed to assess the independent prognostic power of the BAG\_Score with clinical factors.

### Construction of a clinical nomogram

A nomogram is built on the basis of multifactor regression analysis, the integration of multiple predictive indicators, and then the use of scaled line segments, according to a certain scale plotted on the same plane, to be used to express the interrelationships between the variables in the predictive model. To explore the clinical application value of the BAG\_Score, we combined the BAG\_Score with clinical factors (age, gender, pathologic M, pathologic N, pathologic T, tumour stage, tobacco smoking history and histological type) to construct a nomogram, assign scores according to each variable of the patient, and predict the 1-, 3-, and 5-year survival probability of the patient based on the total points. In addition, the accuracy of the nomogram for predicting OS for patient prognosis was assessed using calibration curves, cumulative risk curves, and decision curve analysis (DCA).

### Immune correlation analysis of the BAG\_Score

To investigate the correlation between the BAG\_Score and immune cells, we used seven algorithms, TIMER, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, MCPOUNTER, XCELL, and EPIC<sup>34</sup>, to assess the abundance of immune cells in patients, and the Wilcoxon test was used to compare the degree of immune infiltration with high and low BAG\_Scores. In addition, patients' levels of immune escape and immune response were assessed using the Tumor Immune Dysfunction and Exclusion (TIDE) database<sup>35</sup> (<http://tide.dfci.harvard.edu>), comparing the degree of sensitivity of the immune response in patients in the high and low BAG\_Score groups.

### Mutational landscape of different BAG\_Score groups

To identify the mutation profiles of patients with different PPG\_scores, the "maftools" package was used to create the mutation annotation format (MAF) for the LUAD and LUSC cohorts and visualize the mutational landscapes of high and low BAG\_Scores. In addition, the tumour mutational (mutation frequency per million bases) was calculated and compared between the two groups.

### Molecular docking

We obtained the protein structures of potential targets associated with  $\beta$ -elemene involved in building the BAG\_Score model from the PDB database (<https://www.rcsb.org/>). Then, the 2D structure of  $\beta$ -elemene was obtained from the PubChem database and converted into a 3D structure using Chem3D software. Finally, the small-molecule ligand  $\beta$ -elemene and the potential receptor protein were molecularly docked using AutoDock Vina 1.1.2<sup>36</sup> and PyMOL 2.5.5<sup>37</sup>, and the free energy of molecular binding was calculated to evaluate the stability of the binding.

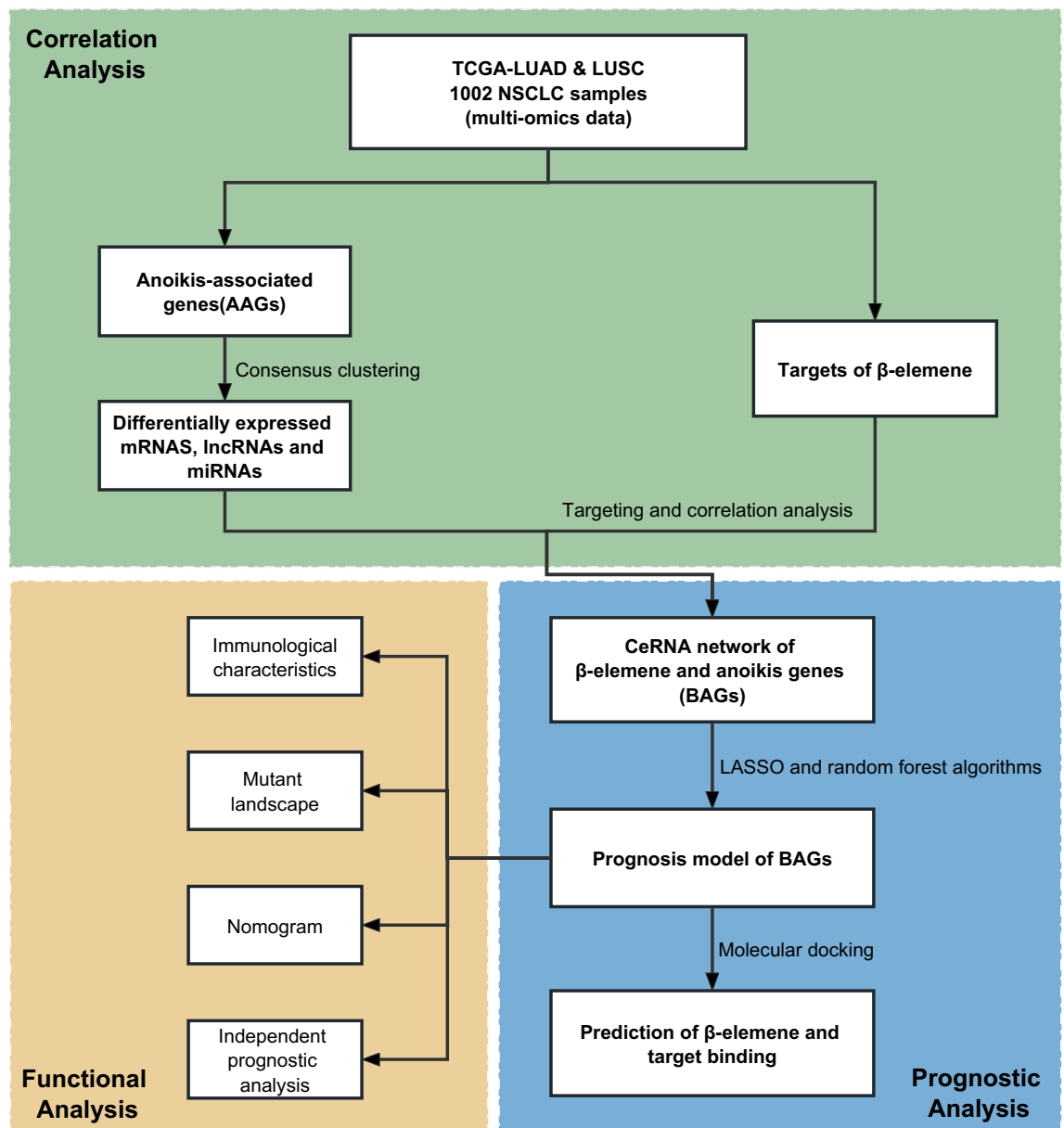
## Statistical analysis

R software (version 4.2.1) and corresponding packages were used for statistical analysis. Cytoscape v3.9.1 was used to visualize the lncRNA-miRNA-mRNA network. Multivariate Cox regression was used to construct a prognostic model. The Kaplan–Meier method and log-rank test were used to assess prognosis. Correlation analysis was performed using Pearson's and Spearman's methods. The Wilcoxon test was used to compare the differences between the two groups. All tests with  $P < 0.05$  indicate statistical significance.

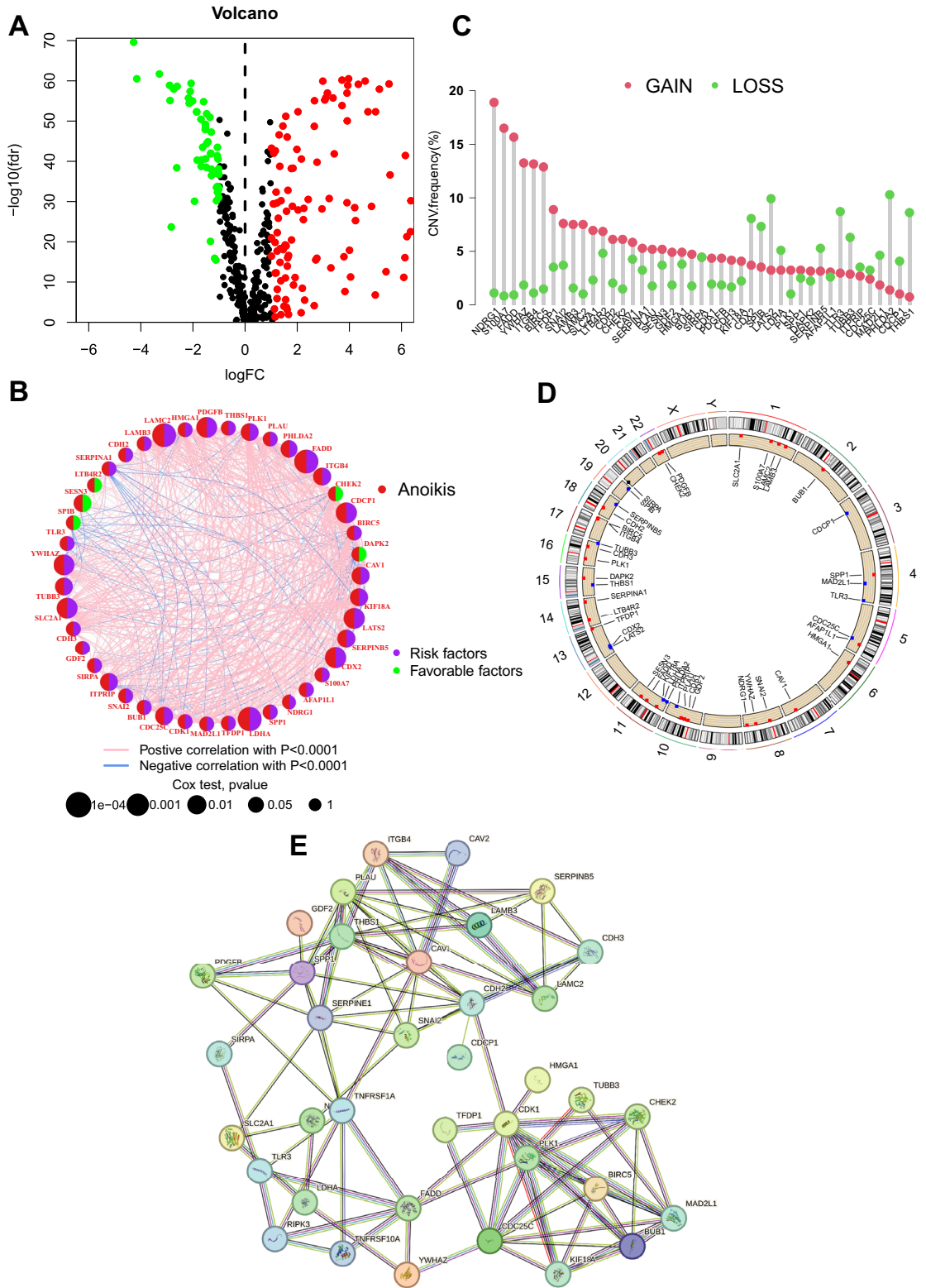
## Results

### Molecular characterization of AAGs

The workflow diagram of the study is displayed in Fig. 1. We performed a differential significance analysis of 638 AAGs and obtained 162 AAGs that were differentially expressed in paraneoplastic and tumour samples of NSCLC patients (Fig. 2A). Univariate Cox regression analysis of these 162 differential AAGs yielded 43 AAGs that were significantly associated with prognosis (Fig. 2B). Among them, 38 AAGs were risk prognostic factors and most of them were positively correlated with each other (Fig. 2B). Then, by analysing the frequency of CNVs in prognosis-associated AAGs, we found that *NDRG1*, *S100A7*, and *FADD* had the highest frequency of copy number gain, whereas *THBS1*, *LATS2*, and *PHLDA2* had the highest frequency of copy number loss, and that most of the AAGs had high-frequency CNVs (Fig. 2C), which often occurred on chromosomes 1, 10, 11, etc. (Fig. 2D). In addition, Fig. 2E shows that the 37 AAGs encode proteins that have interacting relationships with



**Figure 1.** Workflow diagram.

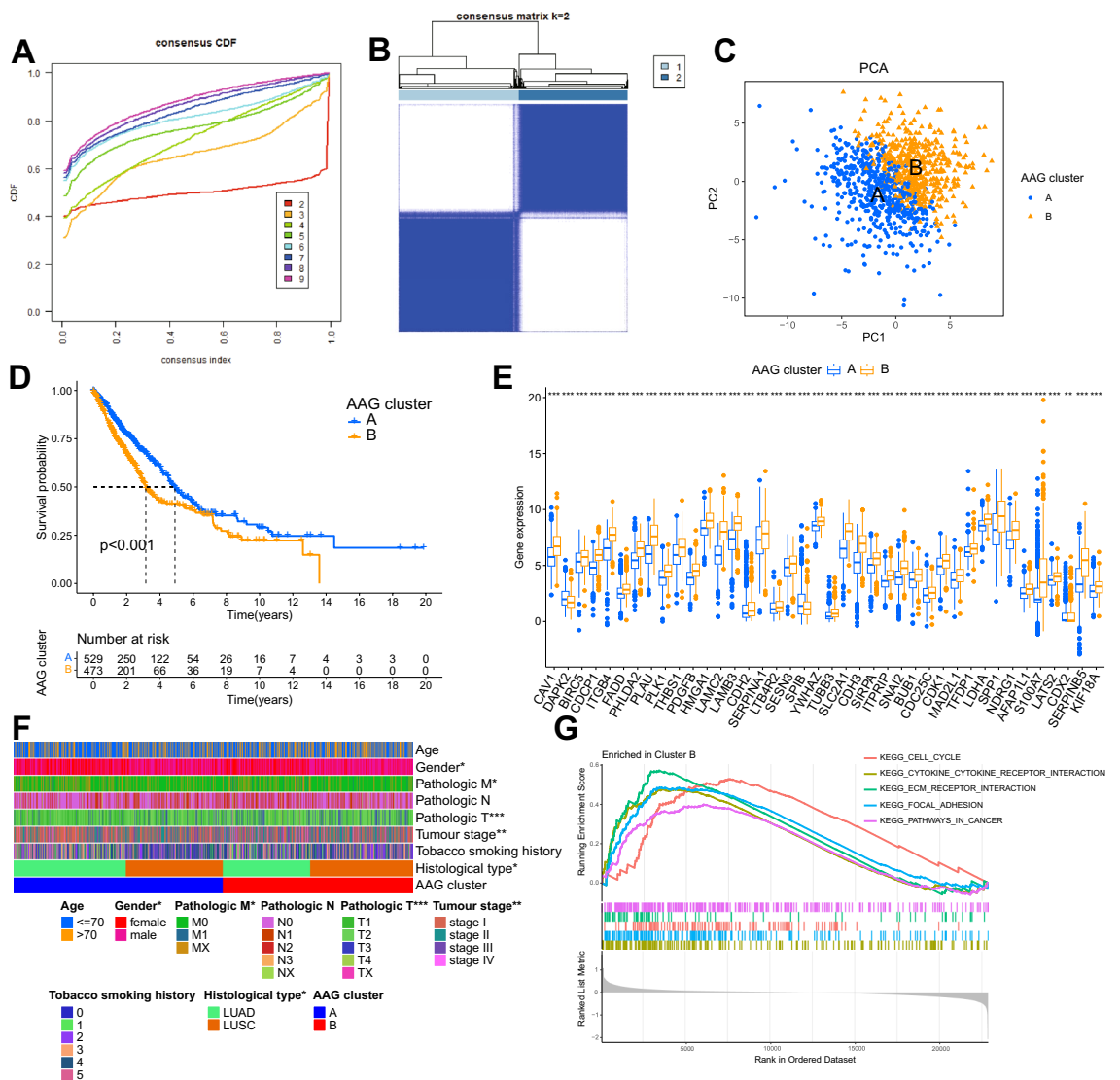


**Figure 2.** Molecular characteristics of anoikis associated genes (AAGs) (A) Significant difference analysis of AAGs between paracancerous (control group) and tumour tissues. (B) Correlation and prognosis analysis of AAGs. (C) Copy number variation (CNV) frequency of AAGs. (D) Location of CNV occurrence of AAGs. (E) Interaction of proteins encoded by AAGs.

each other. All these results indicate that AAGs are aberrantly expressed and genomically unstable in NSCLC tumour tissues and may contribute to the poor prognosis of patients through synergistic effects.

### Generation and functional characterization of AAG clusters

To explore the molecular expression patterns of 43 prognostically relevant AAGs, we used consensus clustering to characterize the expression of AAGs in NSCLC. Based on the plot of the cumulative distribution function, it was determined that the clustering had optimal stability at  $k=2$  (Fig. 3A,B). PCA demonstrated the differential distribution of AAG cluster A and B (Fig. 3C). Kaplan–Meier analysis illustrated that cluster A had a higher survival probability than cluster B (Fig. 3D,  $P < 0.001$ ). Most AAGs were significantly highly expressed in cluster B compared to cluster A (Fig. 3E,  $P < 0.05$ ). According to Fig. 3F, the clinical characteristics of gender, pathologic M, pathologic N, pathologic T, tumour stage and histological type were significantly different between the two clusters (Fig. 3F,  $P < 0.05$ ). To explore the molecular functions of the differences between the two AAG clusters, we performed GSEA, which showed that compared with cluster A, cell cycle, cytokine receptor interaction, extracellular matrix (ECM)-receptor interaction, focal adhesion, and pathways in cancer were significantly enriched in cluster B (Fig. 3G). All of these results reveal that NSCLC can be clustered into two clusters with different prognostic features, clinical traits, and molecular functions based on the expression of AAGs and that AAGs may be related to the cell cycle and the pathway of tumour development in NSCLC.



**Figure 3.** Molecular classification and function of AAGs. (A) Cumulative distribution function for  $k=2-9$ . (B) Consensus clustering heatmap with  $k=2$ . (C) PCA between two AAG clusters. (D) Kaplan–Meier survival analysis between two AAG clusters. (E) Differential expression of AAGs between two AAG clusters. (F) Difference analysis of clinical characteristics between AAG clusters. (G) GSEA to compare two clusters.

### Construction of the ceRNA network of anoikis and $\beta$ -elemene

First, we performed a significant difference analysis of all the genes of AAG cluster A and B using  $|\log_{2}FC| > 0.585$  and  $P < 0.05$  as thresholds, respectively, and obtained 731 mRNAs, 42 lncRNAs and 10 miRNAs significantly related to AAGs. Then, the DIANA-LncBase tool was used to compare the targeting relationship between the differential lncRNAs and miRNAs, and the relationship pairs of 9 miRNAs and 7 lncRNAs were obtained. The target genes of 9 miRNAs were predicted by TargetScan 8.0: (i) target genes of 9 miRNAs were compared with 731 differential mRNAs, and 208 mRNAs related to AAGs were obtained; (ii) 15 genes were the target genes of AAG-associated miRNAs by comparison with 26  $\beta$ -elemene target genes. In addition, 15  $\beta$ -elemene target genes were subjected to Spearman correlation analysis with 208 AAG-related mRNAs, and 47 genes related to  $\beta$ -elemene and anoikis were obtained, including 40 differential mRNAs, 8  $\beta$ -elemene targets, and the intersecting gene *PTGS2* (Fig. 4A,B). Finally, based on the lncRNA-miRNA and miRNA-mRNA targeting relationships and the correlation relationship between  $\beta$ -elemene and anoikis, a ceRNA network consisting of 7 lncRNAs, 9 miRNAs, and 47 mRNAs was successfully constructed (Fig. 4C, Table 1). A total of 63 BAGs were obtained based on  $\beta$ -elemene and anoikis regulatory relationships in the network, which will help us further understand the role of  $\beta$ -elemene in regulating anoikis in the prognosis of NSCLC patients.

### Functional enrichment analysis of BAGs

To explore the molecular functions of  $\beta$ -elemene and anoikis in NSCLC, we performed GO and KEGG enrichment analysis of 63 BAGs ( $P < 0.05$  and  $Q < 0.1$ ). These genes were significantly enriched in functions and pathways such as regulation of cellular response to growth factor stimulus (GO:0,090,287), epithelium migration (GO:0,090,132), tissue migration (GO:0,090,130), focal adhesion (hsa04510), ECM-receptor interaction (hsa04512), and chemical carcinogenesis (hsa05204) (Fig. 5A,B). Among them, epithelial and tissue migration are important characteristics of tumour cells; focal adhesion is the key to signal transduction; the extracellular matrix (ECM) is closely related to the immune microenvironment; and chemical carcinogens are a common driver of tumours. This indicates that BAGs may be involved in tumour characteristics, carcinogenic factors, and immune function-related functions, thus affecting the disease process of NSCLC patients.

### Prognostic BAG\_Score model construction

To explore the prognostic role of the  $\beta$ -elemene- and anoikis-regulated ceRNA network, we screened and modelled 63 BAGs. First, 970 patients were randomized into the training group (485 patients) and the validation group (485 patients). The random forest algorithm was used for the training group to construct a classifier for predicting OS.status, and according to the model error plot, the error tended to remain stable at a decision tree of 300 (Fig. 6A). The top 10 genes were obtained as feature BAGs based on the Gini coefficient method with the MeanDecreaseGini index as the importance score (Fig. 6B). Then, with OS.status as the response variable and the normalized expression matrix of the 63 BAGs as the independent variables in the training group, a tenfold cross-test was performed to obtain the smallest mean squared error (binomial deviation) of the model for a variable number of 24 (Fig. 6C). The intersection of the feature genes obtained by the two algorithms was taken to obtain seven feature BAGs for constructing the BAG\_Score model (Fig. 6D). The risk score was calculated by multivariate Cox regression analysis as follows:

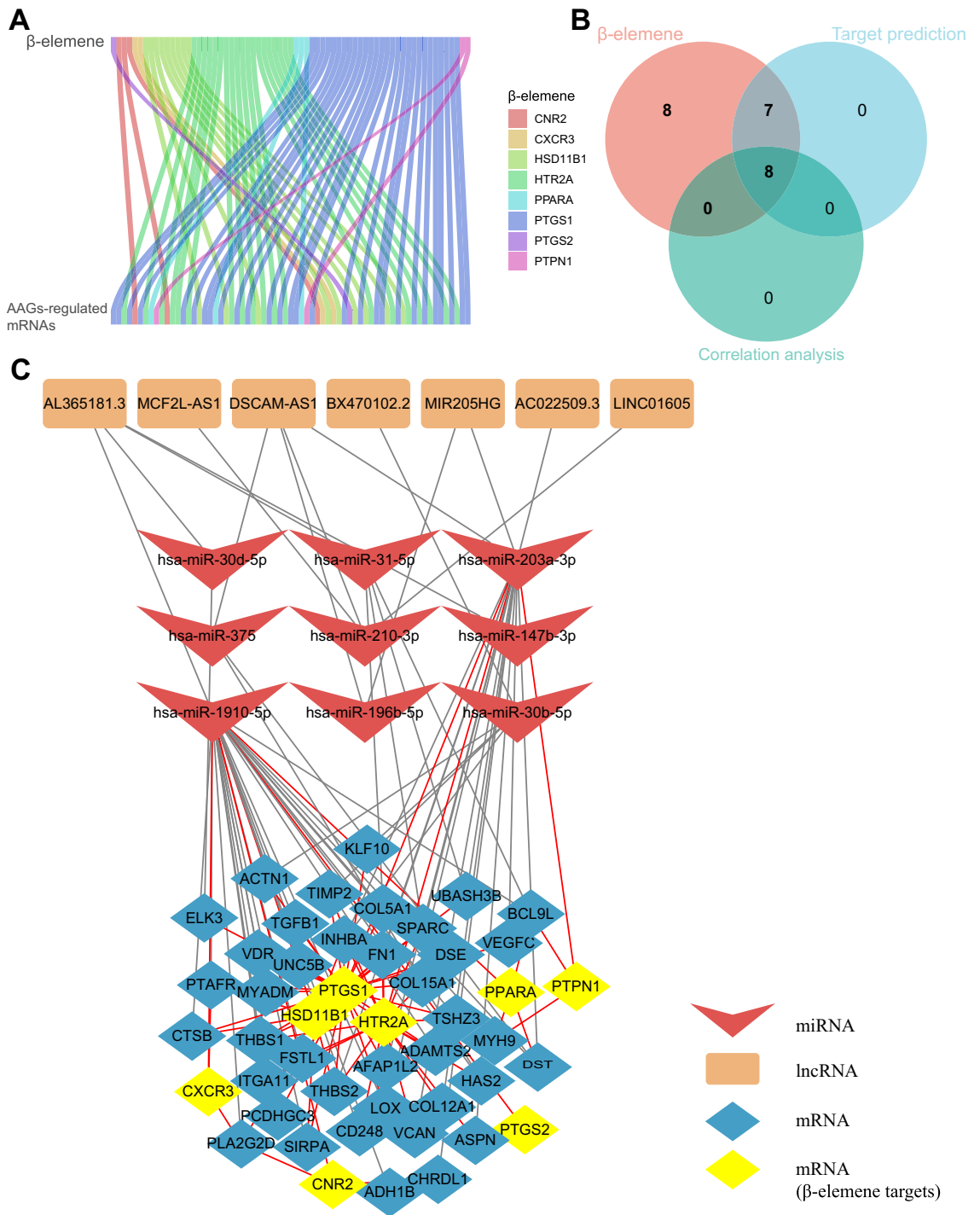
$$\begin{aligned} \text{BAG\_Score} = & (0.015) * \text{Exp}(hsa - miR - 30b - 5p) \\ & + (-0.243) * \text{Exp}(PPARA) + (-0.22) * \text{Exp}(CNR2) \\ & + (0.078) * \text{Exp}(AL365181.3) + (0.294) * \text{Exp}(KLF10) \\ & + (-0.25) * \text{Exp}(VDR) + (0.247) * \text{Exp}(VEGFC). \end{aligned}$$

According to the BAG\_Score model in Fig. 6E, *PPARA* and *VDR* were prognostic protective factors for NSCLC ( $HR < 1$  and  $P < 0.05$ ), while *KLF10* and *VEGFC* were prognostic risk factors ( $HR > 1$  and  $P < 0.05$ ). According to the risk heatmap, *CNR2*, *PPARA*, and *VDR* were relatively highly expressed in the low BAG\_Score group, whereas *AL365181.3*, *KLF10*, and *VEGFC* were relatively highly expressed in the high BAG\_Score group (Fig. 6F). These four genes not only are the key genes by which  $\beta$ -elemene regulates anoikis but also serve as potential prognostic molecular markers for NSCLC patients, providing a research basis for the study of effective targets for  $\beta$ -elemene.

### Assessment of prognostic BAG\_Score models

Multivariate regression analysis of clinical characteristics and the BAG\_Score showed that age, T3, and the BAG\_Score were independent prognostic factors in NSCLC patients (Fig. 7A,  $HR > 1$ ,  $P < 0.05$ ). ROC analysis showed that the area under the curves (AUCs) for 1-, 3-, and 5-year survival probability were all greater than 0.57 in the training group, greater than 0.53 in the test group, and greater than 0.55 in the entire group (Fig. 7B,C,D). In particular, the AUC for 1- and 3-year survival in the entire group was greater than 0.61 (Fig. 7D), suggesting that the model has some prognostic predictive ability and is more accurate in predicting early-stage disease. Kaplan–Meier analysis demonstrated that patients in the high and low BAG\_Score groups had significantly different OS. Patients in the low BAG\_Score group had a higher survival probability than those in the high BAG\_Score group (Fig. 7E,F,G, log-rank test:  $P < 0.05$ ). By evaluating the BAG\_Score model, we found that the BAG\_Score was an independent predictor and significantly correlated with NSCLC patient prognosis.





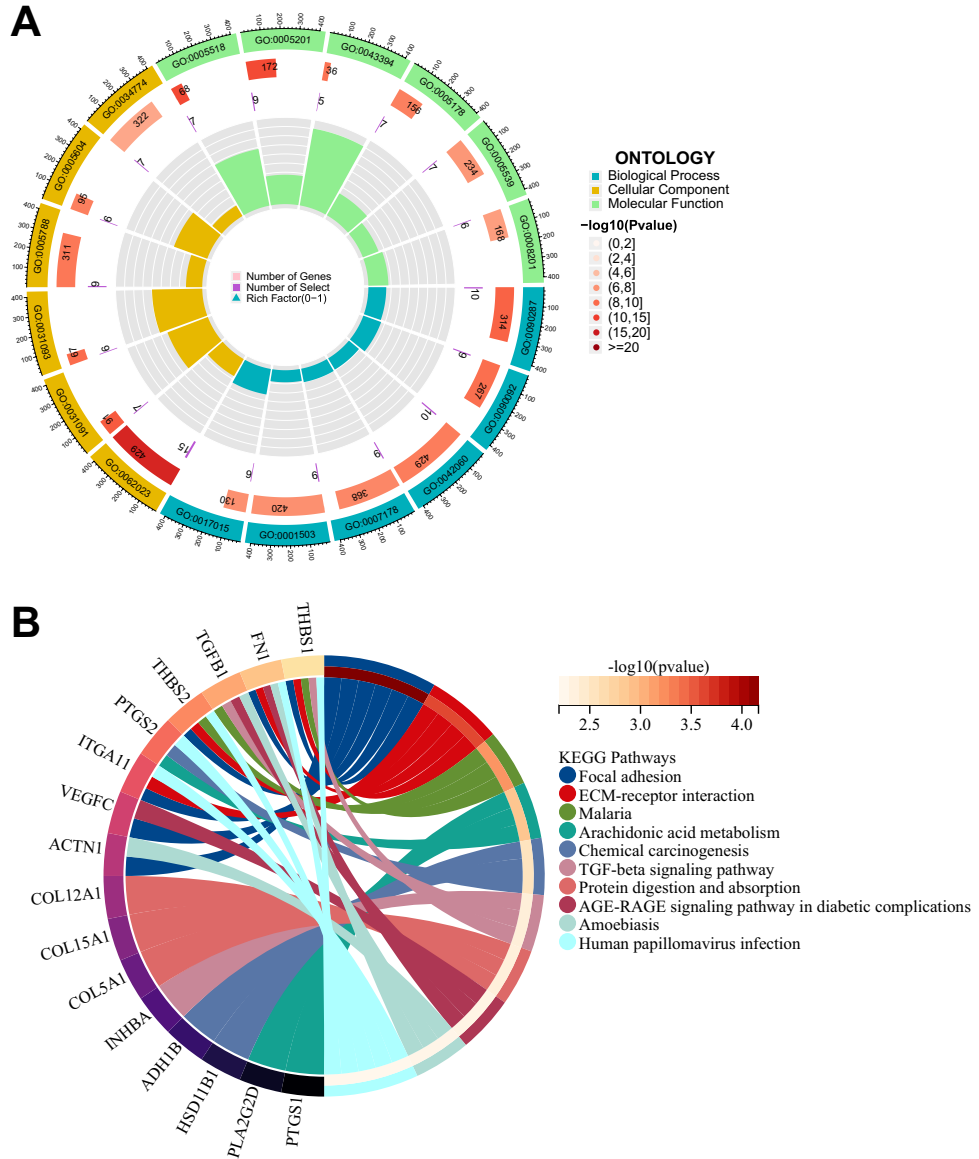
**Figure 4.** Construction of the lncRNA-miRNA-mRNA network. (A) Correlation analysis of  $\beta$ -elemene targets and AAG-regulated mRNAs. (B) Screening the gene sets related to anoikis among the  $\beta$ -elemene targets. (C) LncRNA-miRNA-mRNA network of the interaction between  $\beta$ -elemene and anoikis.

### Construction of the BAG\_Score combined nomogram

To explore the value of applying the BAG\_Score with clinical factors, we constructed a clinical nomogram of the combined BAG\_Score, aiming at the prediction of patient OS at 1, 3, and 5 years. In the nomogram, age, BAG\_Score, and pathologic T independently predicted patient prognosis (Fig. 8A). The calibration curves show that the OS predictions for the 1-, 3-, and 5-year nomograms are extremely close to the actual observed values, demonstrating the excellent accuracy of the model (Fig. 8B). The cumulative hazard analysis showed that the cumulative hazard was higher in patients with high-nomoRisk than in the low-nomoRisk group (Fig. 8C). In addition, the 1-, 3-, and 5-year DCA curves showed that the prognostic predictive power of the BAG\_Score and

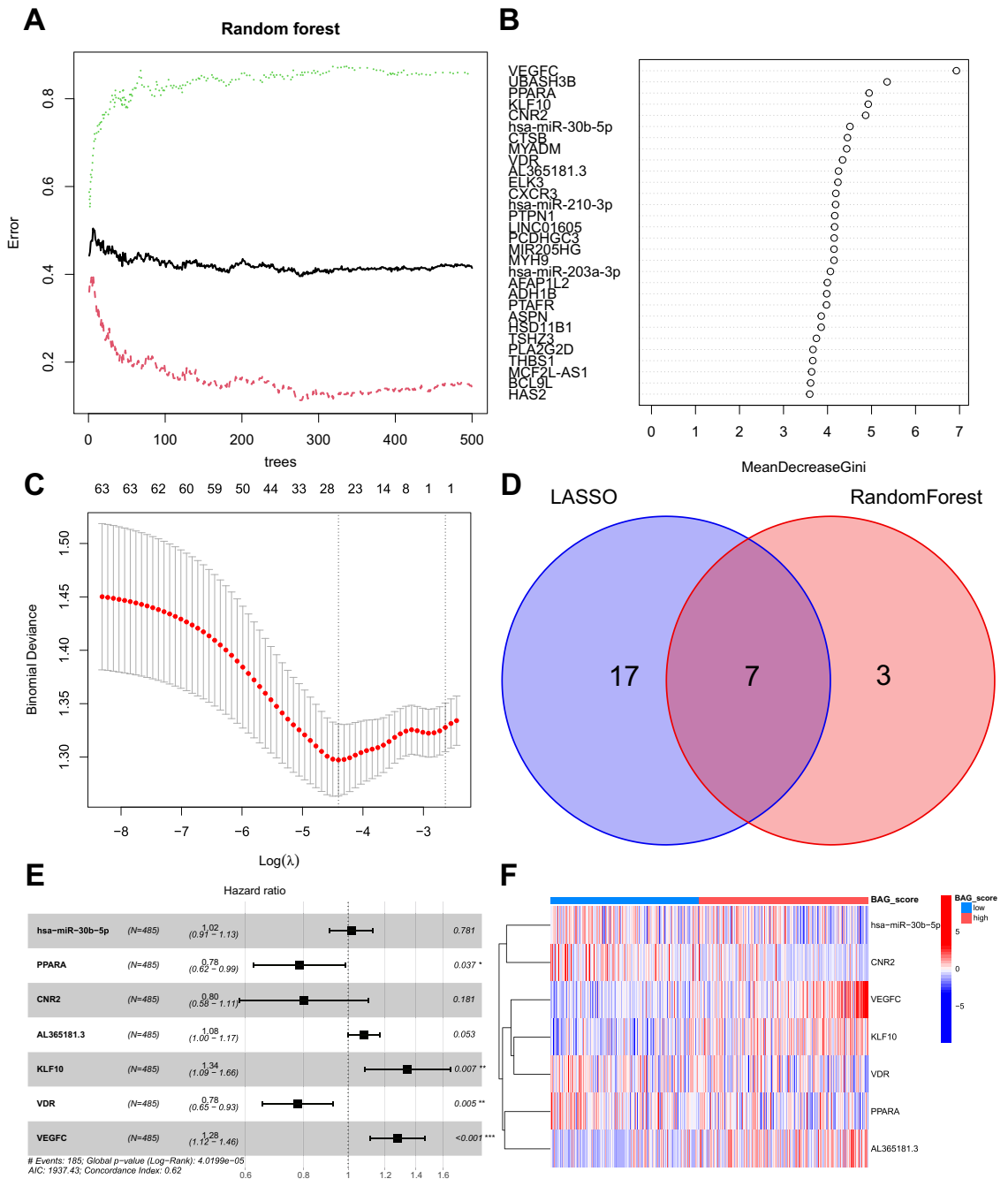
Biotype	Genes
lncRNA	AL365181.3, DSCAM-AS1, MIR205HG, AC022509.3, LINC01605, MCF2L-AS1, BX470102.2
miRNA	hsa-miR-31-5p, hsa-miR-1910-5p, hsa-miR-375, hsa-miR-30b-5p, hsa-miR-30d-5p, hsa-miR-210-3p, hsa-miR-196b-5p, hsa-miR-147b-3p, hsa-miR-203a-3p
mRNA	PTPN1, PPARA, PTGS1, HSD11B1, CXCR3, HTR2A, CNR2, PTGS2, MYH9, BCL9L, DST, ACTN1, ELK3, DSE, AFAP1L2, CTSB, KLF10, COL5A1, INHBA, TGFB1, UNC5B, THBS2, LOX, TSHZ3, TIMP2, ADAMTS2, VDR, FN1, UBASH3B, PCDHGC3, SIRPA, SPARC, VEGFC, COL15A1, MYADM, PTAFR, FSTL1, THBS1, CHRDL1, ADH1B, PLA2G2D, COL12A1, CD248, VCAN, ITGA11, HAS2, ASPN

**Table 1.** ceRNA network constructed by interacting genes of  $\beta$ -elemene and anoikis.



**Figure 5.** Functional enrichment of the lncRNA-miRNA-mRNA network. (A) GO and (B) KEGG analysis of interacting genes of  $\beta$ -elemene and anoikis in the network.

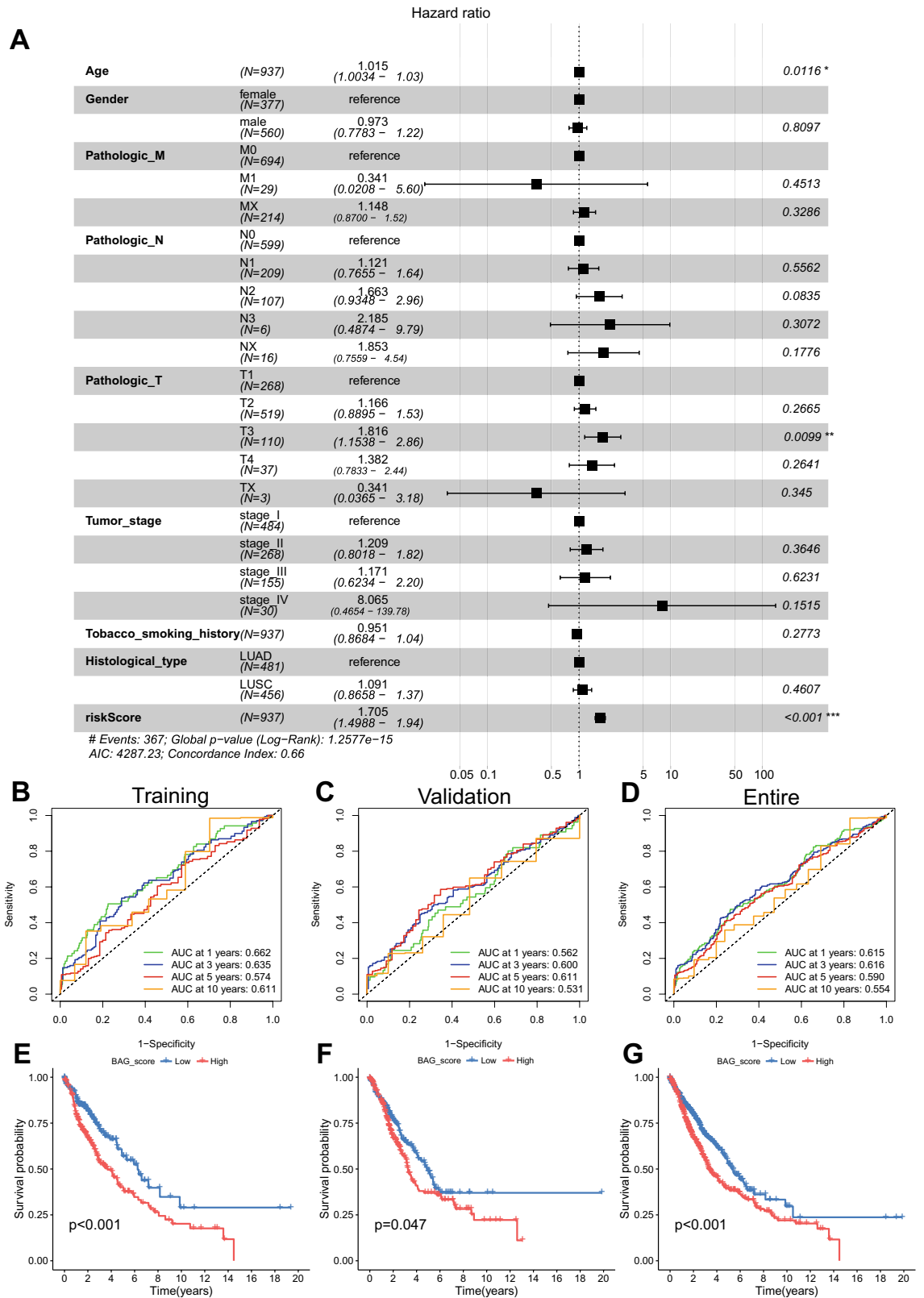
nomogram was superior to that of other clinical factors, resulting in a higher net clinical benefit for patients (Fig. 8D,E,F). The clinical nomogram with the BAG\_Score has better predictive validity than single clinical factors and can provide an effective reference for the clinical prediction of OS in NSCLC patients.



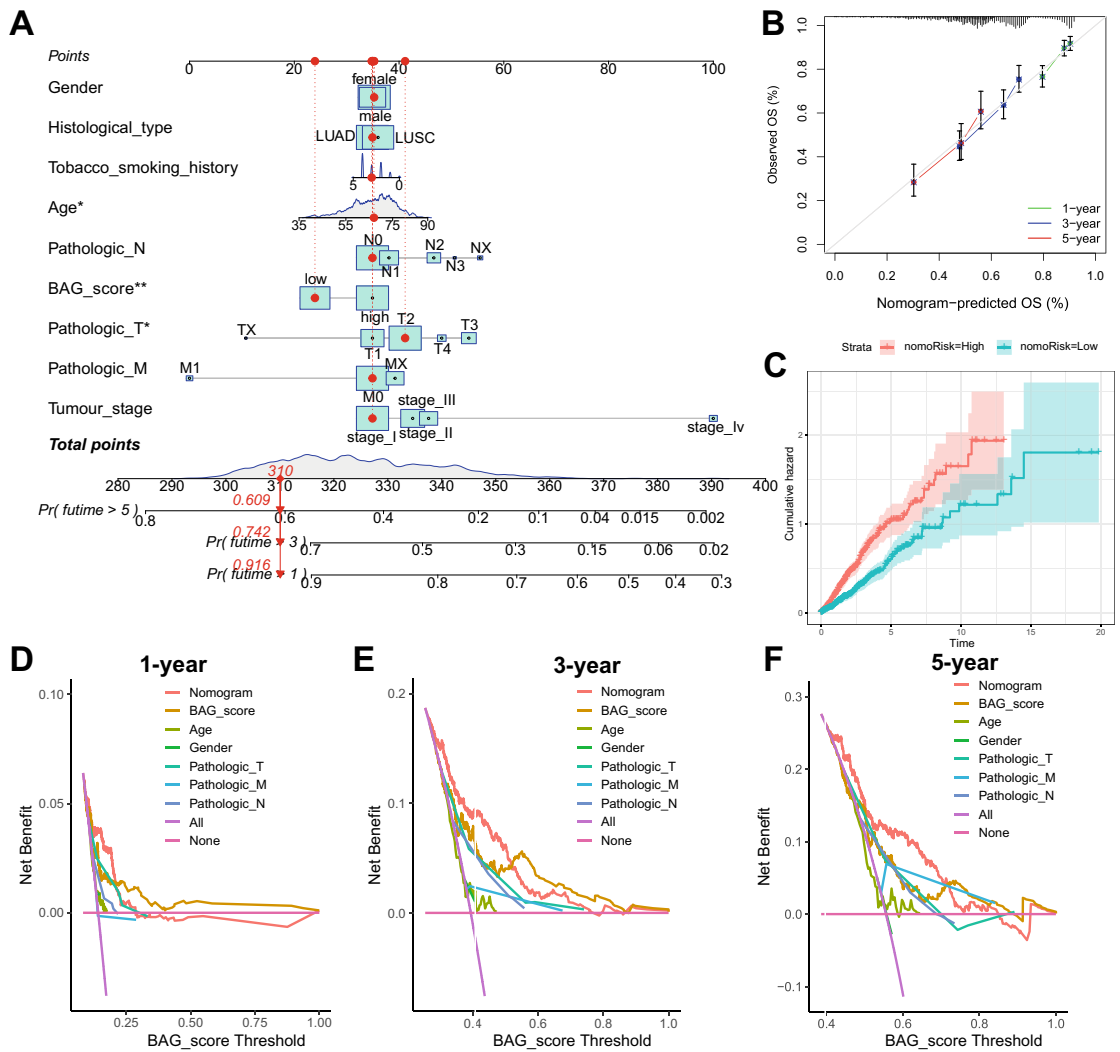
**Figure 6.** Construction of the prognostic BAG\_Score model. (A) Model error plot for random forest. (B) Gene importance scoring with the random forest algorithm. (C) Optimal number of variables for LASSO regression analysis. (D) Intersecting genes of the two algorithms. (E) Multivariate Cox regression to construct the BAG\_Score model. (F) Risk heatmap for BAG\_Score modelling genes.

### Immunization and mutational landscapes for different BAG\_Scores

Using seven immune cell assessment algorithms to correlate the intensity of immune cell infiltration and the BAG\_Score in NSCLC patients, the abundances of a total of 65 kinds of immune cells were correlated with the BAG\_Score. Most of the immune cell levels were high in the low BAG\_Score group, such as CD4+ T cell central memory, CD4+ T cell effector memory, and CD8+ T cell (Fig. 9A) levels, suggesting that the immune cells were more active in patients with low BAG\_Scores. The low BAG\_Scores group had significantly lower TIDE scores than the high BAG\_Score group, suggesting that the low BAG\_Score group was less prone to immune escape (Fig. 9B). Moreover, the incidence of immune response reactions was significantly higher in the low BAG\_Score group than in the high BAG\_Score group (Fig. 9C). Among the modelling genes, *CNR2* and *PPARA* were significantly correlated with more immune cells, CD8 T cells and follicular helper T cells were negatively correlated with the modelling genes, and neutrophils and M0 macrophages were positively correlated with the



**Figure 7.** Evaluation of the BAG\_Score model. (A) Independent prognostic analysis of BAG\_Score and clinical variables. (B–D) ROC curves of the BAG\_Score in the training (B), validation (C) and entire (D) groups. (E–G) Kaplan–Meier survival analysis of the BAG\_Score in the training (E), validation (F) and entire (G) groups.



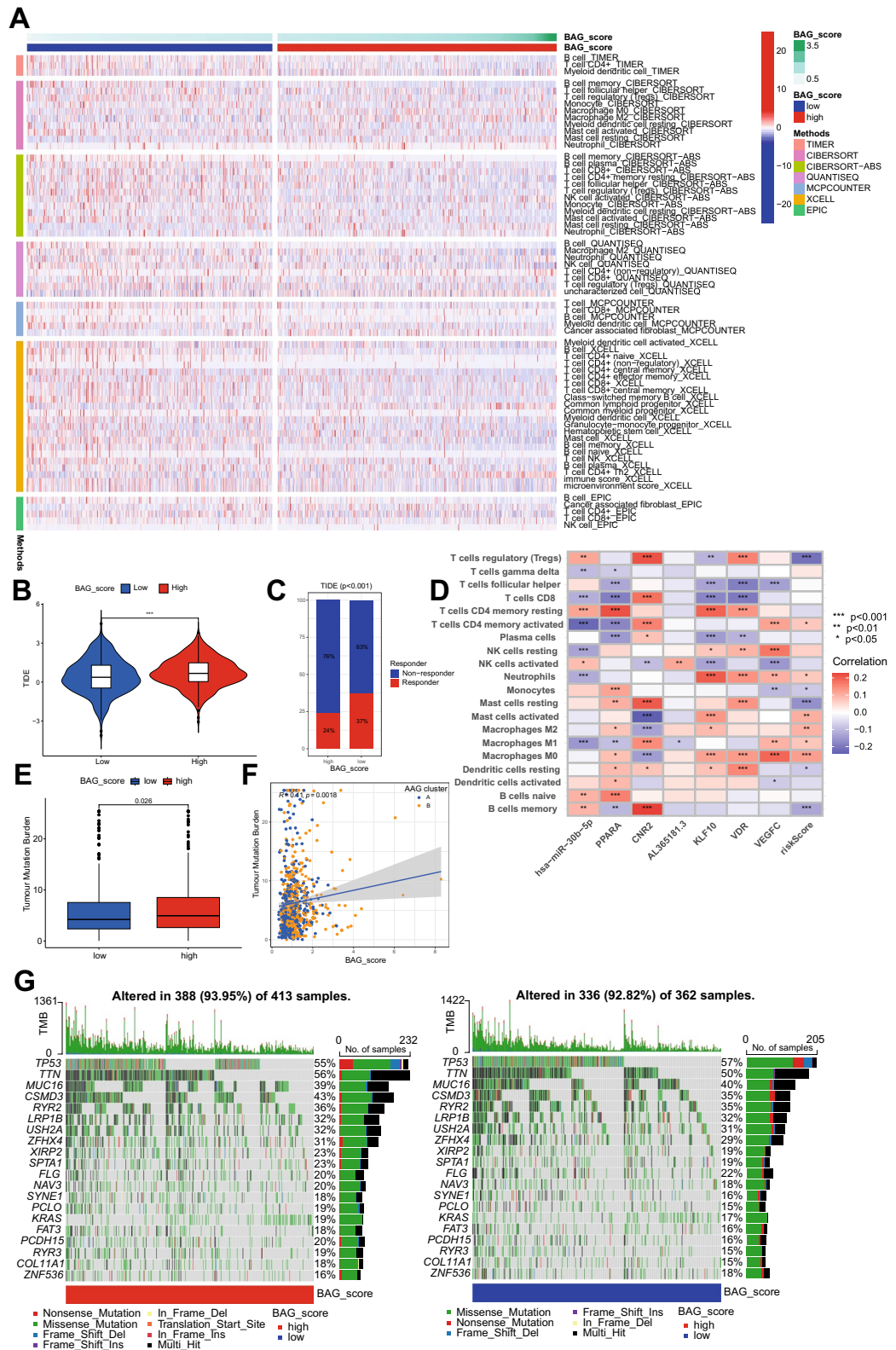
**Figure 8.** Construction of the clinical nomogram. (A) Construction of the nomogram for the BAG\_Score and clinical characteristics. (B–F) Calibration curves (B), cumulative hazard curves (C) and 1-, 3-, and 5-year DCA curves (D–F) to assess the performance of the nomogram.

modelling genes (Fig. 9D). This indicates that the BAG\_Score is closely related to the immune level of NSCLC patients and that patients with a low BAG\_Score have stronger immune cell activity and a higher success rate of immune response occurrence.

To explore the relationship between the BAG\_Score and genomic instability, we assessed and compared the differences in tumour mutation burden (TMB) between the high and low BAG\_Score groups. The outcomes revealed that TMB was significantly higher in the high BAG\_Score group than in the low BAG\_Score group (Fig. 9E,  $P < 0.05$ ). Moreover, the BAG\_Score was significantly positively correlated with TMB ( $COR = 0.11$ ,  $P = 0.0018$ ), and AAG cluster A had a lower TMB and BAG\_Score (Fig. 9F). Furthermore, visualizing the immune landscape of the high and low BAG\_Score groups in Fig. 9G, gene alterations occurred more frequently in the high BAG\_Score group (93.95%) than in the low BAG\_Score group (92.82%). The most frequent mutation type in both groups was missense mutation. The most mutation-prone genes in both groups were *TP53*, *TTN*, *MUC16*, *CSMD3* and *RYR2*. Except for *TP53*, all of these genes were mutated more frequently in the high BAG\_Score group than in the low BAG\_Score group (Fig. 9G). In summary, it can be concluded that patients in the high BAG\_Score group have a higher probability of mutations and greater genomic instability, which may lead to faster tumour progression and worse prognosis.

### Molecular docking of $\beta$ -elemene

According to the BAG\_Score model, *CNR2* and *PPARA* are potential targets of  $\beta$ -elemene; *KLF10*, *VDR* and *VEGFC* all have a coexpression relationship with the  $\beta$ -elemene target *PTGS1*; *PTGS2* and *PTGS1* belong to the same gene family and are also potential targets of  $\beta$ -elemene; and in particular, *PTGS2* is also a gene related to anoikis. Therefore, we selected *CNR2*, *PPARA*, *PTGS1* and *PTGS2*, four key genes that have direct or indirect effects on prognosis in NSCLC, and conducted molecular docking with  $\beta$ -elemene. The molecular



**Figure 9.** Comparison of immunization and mutational landscapes between different BAG\_Score groups. (A) Correlation between the BAG\_Score and immune cell abundance. (B) Differences in TIDE scores between the high and low BAG\_Score groups. (C) Differences in the incidence of immune response between the high and low BAG\_Score groups. (D) Correlation of BAG\_Score-modelled genes and immune cell content. (E) Differences in TMB between the high and low BAG\_Score groups. (F) Correlation between the BAG\_Score and TMB. (G) Immune landscape of the high and low BAG\_Score groups.

docking effect of these four macromolecule receptors and the small-molecule ligand  $\beta$ -elemene was evaluated through AutoDock Vina 1.1.2. The results showed that the four proteins all have the ability to bind to  $\beta$ -elemene (Fig. 10A,B,C,D). Among them, the combination of *PPARA* and  $\beta$ -elemene was the most stable (binding free energy  $-6.0$  kcal/mol), followed by *PTGS2* ( $-5.8$  kcal/mol) and *PTGS1* ( $-5.2$  kcal/mol), and *CNR2* was the least stable ( $-4.4$  kcal/mol). This suggests that the four proteins all have strong affinity for  $\beta$ -elemene and may be potential prognostic targets for  $\beta$ -elemene in NSCLC patients.

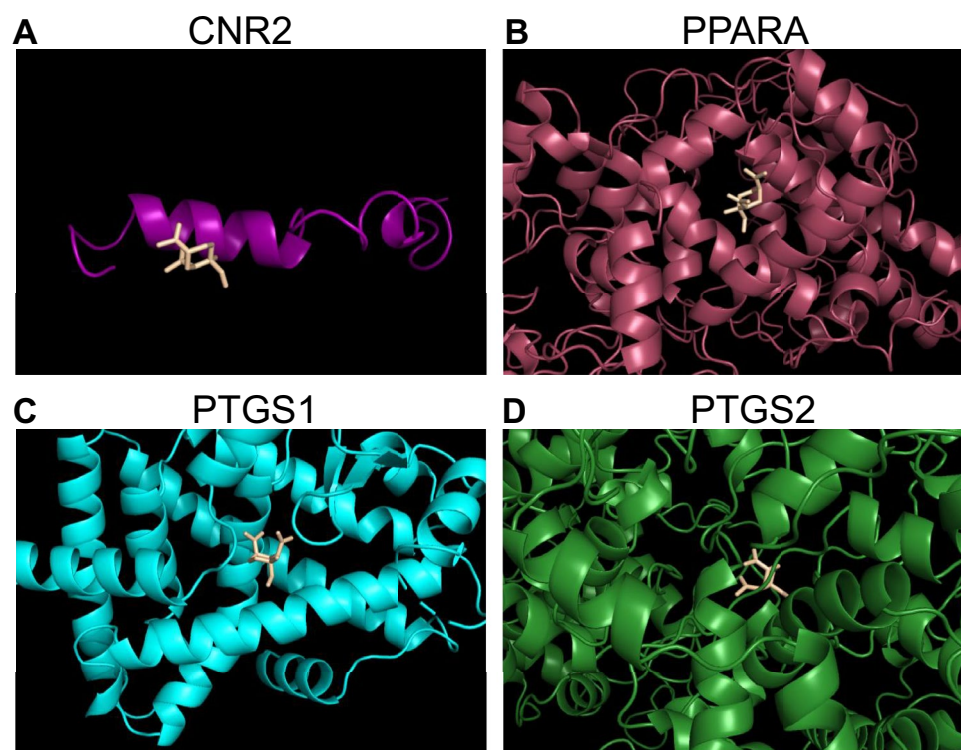
## Discussion

NSCLC is a serious global public health problem, and despite ongoing advances in diagnostic and therapeutic technologies, NSCLC remains an incurable disease for most patients<sup>38</sup>. Curcuma Rhizoma has a long history of use in TCM clinical NSCLC treatment, and modern studies have demonstrated that Curcuma Rhizoma exerts anti-NSCLC effects via multiple pathways through a variety of active constituents, among which  $\beta$ -elemene is the most important sesquiterpene compound isolated from Curcuma Rhizoma and is rich in prominent antitumour activity<sup>39,40</sup>. Therefore, the continuous elucidation of the  $\beta$ -elemene antitumour mechanism is of increasing significance for exploring the innovative path of antitumour drug development.

In this study, we characterized the genetic and epigenetic profiles of anoikis in NSCLC patients by gene expression and clinical information from the LUSC and LUAD cohorts of the TCGA database. Anoikis genes have significant prognostic differences and genomic instability, suggesting that AAGs are closely associated with tumour development. To explore the molecular mechanisms regulating anoikis in NSCLC patients, we used consensus clustering to analyse the molecular expression patterns of patients. According to the AAG expression level, patients can be classified into two different AAG molecular subtypes, and NSCLC patients with different clusters have different prognoses, clinical features and immune functions.

Then, to further explore the molecular regulatory mechanism of AAGs on tumour cells, we successfully constructed a lncRNA-miRNA-mRNA regulatory network of  $\beta$ -elemene interaction with anoikis by using the targeting relationship of AAG-related lncRNAs, miRNAs and mRNAs as well as through the interactions between  $\beta$ -elemene and AAG-related genes. This ceRNA network demonstrated the rich interactions between  $\beta$ -elemene targets and related genes of anoikis in NSCLC, which provided a new theoretical basis and direction to further explore the antitumour mechanism of  $\beta$ -elemene in NSCLC.

In addition, to further explore the prognostic role of  $\beta$ -elemene in NSCLC, we screened and constructed a prognostic score model for all the genes in the ceRNA network using LASSO regression and a random forest algorithm. The scoring model can be used as an independent prognostic factor for the prediction of prognosis in NSCLC patients and in conjunction with clinical factors to construct a nomogram that provides the best net benefit for the prediction of survival in clinical patients. Molecular docking of  $\beta$ -elemene and model-related targets is beneficial for elucidating the mechanism of  $\beta$ -elemene in NSCLC patients.



**Figure 10.** Molecular docking. Purple (A), red (B), blue (C) and green (D) represent the protein structures of *CNR2*, *PPARA*, *PTGS1* and *PTGS2*, respectively. The wheat-coloured structure represents  $\beta$ -elemene.

In the prognostic scoring model, we identified Peroxisome proliferator activated receptor alpha (*PPARA*) and Cannabinoid receptor 2 (*CNR2*) as potential target genes of  $\beta$ -elemene, which we hypothesized could be regulated by upstream *hsa-miR-1910-5p*. *PPARA* was the first PPAR isoform identified, and several studies have demonstrated the oncogenic effects of *PPARA* in NSCLC: *PPARA* activation inhibits NSCLC growth and angiogenesis and reduces metastasis<sup>41</sup>; the *PPARA* agonist fenofibrate promotes NSCLC resistance to gefitinib by modulating in a *PPARA*-dependent manner the AMPK/AKT/forkhead box-1 (FOXO1) pathway to promote gefitinib-induced apoptosis, thereby alleviating NSCLC resistance to gefitinib<sup>42</sup>; N-Acetyl-Cysteine (NAC) in combination with *PPARA* induces *p53* and reduces p65 protein expression to inhibit 3-phosphoinositide-dependent protein kinase 1 (PDK1) activity, leading to inhibition of NSCLC cell growth<sup>43</sup>. *CNR2* is widely present in NSCLC tissues, and its expression influences NSCLC development. Xu et al.<sup>44</sup> found that deletion of *CNR2* inhibited the progression of NSCLC cells, suggesting that *CNR2* has a pro-carcinogenic role in NSCLC. Sarsembayeva et al.<sup>45</sup> demonstrated that *CNR2* in the tumour microenvironment hinders the antitumour activity of CD8 T cells and NK cells, thereby promoting NSCLC growth. It has been shown that *CNR2* expression in early-stage NSCLC is associated with prolonged survival and fewer lymph node metastases<sup>46</sup>. Both *PPARA* and *CNR2*, potential target genes of  $\beta$ -elemene, have been shown to be involved in the tumour development process of NSCLC. Among them, *PPARA* is an oncogene, which is consistent with our observation, so we speculate that  $\beta$ -elemene exerts tumour suppression in NSCLC through the oncogene *PPARA*.

Both KLF transcription factor 10 (*KLF10*) and *hsa-miR-30b-5p* were modelled, and interestingly, *KLF10* is a potential target gene for *hsa-miR-30b-5p*. We hypothesized that *hsa-miR-30b-5p* might improve the prognosis of patients by targeting *KLF10*. Studies have shown that *miR-30b-5p* is closely associated with the survival and prognosis of lung cancer patients<sup>47,48</sup>. Qiu et al.<sup>49</sup> found that *miR-30b-5p* plays an oncogenic role in lung cancer and sensitizes lung cancer cells to cisplatin by targeting low-density lipoprotein receptor-related protein-8 (LRP8). It has been demonstrated that *miR-30b* can induce anoikis resistance by downregulating Caspase 3 expression<sup>50</sup>.

We found that the anoikis-associated lncRNA *AL365181.3* may be regulated by *hsa-miR-31-5p* to exert pro-oncogenic effects. In previous studies, *AL365181.3* was identified as a neutrophil extracellular trap-associated lncRNA and has potential prognostic value for LUAD patients<sup>51</sup>. In addition, *AL365181.3* has been identified as an iron death-associated lncRNA and has been verified to be significantly upregulated in a variety of NSCLC cell<sup>52</sup>. Yu et al.<sup>53</sup> found that *miR-31-5p* was heavily enriched in the exosomes of hypoxic LUAD cells and promoted LUAD invasion and migration by decreasing special AT-rich sequence-binding protein 2 (SATB2) expression and activating MEK/ERK signalling pathway transduction. Zhu et al.<sup>54</sup> demonstrated that *miR-31-5p* is highly expressed in LUAD and determined that it promotes LUAD progression through the tensin1 (*TNSI*)/*p53* axis. Zhu et al.<sup>55</sup> demonstrated that *miR-31-5p* targeting the hypoxia inducible factor 1 $\alpha$  inhibitor/hypoxia inducible factor (FIH/HIF) mechanism enhances the Warburg effect, induces glycolysis and promotes NSCLC cell proliferation. Thus, both *hsa-miR-31-5p* and *AL365181.3* are upregulated in NSCLC, leading to anoikis inhibition and exerting a potential pro-oncogenic role.

In addition, both vitamin D receptor (*VDR*) and vascular endothelial growth factor C (*VEGFC*) were predictive of NSCLC prognosis, and both were coexpressed with the target gene  $\beta$ -elemene and the target gene prostaglandin-endoperoxide synthase 1 (*PTGS1*). A high expression level of *VDR* is associated with elevated LUAD survival and with anti-proliferative and G1 arrest<sup>56</sup>. Several studies have shown that *VDR* polymorphisms can significantly reduce NSCLC risk<sup>57,58</sup>. High expression of *VEGFC* is significantly associated with poor prognosis in NSCLC. Qin et al.<sup>59</sup> found that coexpression of *VEGFC* and programmed cell death ligand 1 (PD-L1) was an indicator of high risk of recurrence and poor prognosis in LUAD. Regan et al.<sup>60</sup> demonstrated that *VEGFC* is a major driver of tumour lymphangiogenesis in NSCLC. *PTGS1* is a coexpressed gene of *VDR* and *VEGFC* in NSCLC, while homologous prostaglandin-endoperoxide synthase 2 (*PTGS2*) is one of the target genes of  $\beta$ -elemene. *PTGS1* and *PTGS2* are important anticancer targets, and studies have shown that *PTGS1* and *PTGS2* are highly associated with lung tumorigenesis<sup>61,62</sup>. Several studies have demonstrated the association of *PTGS2* with anoikis: *PTGS2*-mediated prostaglandin E2 (PGE2) synthesis renders three-dimensionally cultured mesenchymal stem cells (MSCs) resistant to anoikis<sup>63</sup>; *PTGS2* provides hepatocyte growth factor-mediated resistance to anoikis and promotes human head and neck squamous cell carcinoma growth<sup>64</sup>; and a *PTGS2* inhibitor (celecoxib) enhances the effect of cisplatin and induces anoikis in osteosarcoma through the PI3K/Akt pathway<sup>65</sup>. For  $\beta$ -elemene, Su et al.<sup>66</sup> found that  $\beta$ -elemene combined with 5-fluorouracil (5-FU) inhibited triple-negative breast cancer growth by interfering with the nuclear factor-kappaB (NF- $\kappa$ B)/*PTGS2* pathway. Cai et al.<sup>67</sup> found that  $\beta$ -elemene inhibited *PTGS2* expression in NSCLC. From this, we hypothesize that the active substance  $\beta$ -elemene regulates the expression of the potential molecular markers *VDR* and *VEGFC* by acting on *PTGS1* and *PTGS2* to affect anoikis in NSCLC cells.

In summary, by studying the interaction relationship between  $\beta$ -elemene targets and anoikis-related genes, this study obtained multiple gene pairs related to the prognosis of NSCLC patients, which provides new research ideas and a theoretical basis for the antitumour mechanism and clinical treatment of  $\beta$ -elemene-acting anoikis.

This study also has some limitations. External validation of the prognostic scoring model with other cohorts was not performed because of limited miRNA-seq data. Moreover, our data were obtained from public clinical databases, and relevant molecular and animal experiments are yet to be carried out to validate the interaction between  $\beta$ -elemene targets and genes related to loss of anoikis. At the same time, this article has certain shortcomings in methodology. The prognostic prediction feature used in this article is the expression level of gene pairs, and the existing research on the interaction between ncRNAs is very mature<sup>24,25</sup>. Single-cell multiomic data and related algorithms have also been widely used in gene/protein association analysis<sup>68,69</sup>; in addition, the potential association between metabolites and diseases can be directly predicted based on deep learning models<sup>70-72</sup>, but this article has not yet conducted an in-depth study of diseases based on metabolites. In the future, this research should be based on the theoretical model of the gene/protein signalling network<sup>73-75</sup>, and use deep learning



model theory to connect with clinical practice to gain an in-depth understanding of the regulatory mechanism of the disease and find potential therapeutic targets.

## Conclusion

We comprehensively analysed the genetic and clinical characteristics and prognosis of AAGs and, for the first time, discovered a rich interaction relationship between  $\beta$ -elemene and AAG-regulated genes in NSCLC patients. Furthermore, we observed the effect of interacting gene pairs on prognosis, which provided us with enlightening significance and a new perspective to further explore the mechanism of action of  $\beta$ -elemene in NSCLC patients.

## Data availability

All of the data used in this investigation was obtained from public clinical databases. The data set numbers are mentioned in the publication. All dataset numbers are mentioned in the article and are publicly available. All data analysis in this study is based on the R software (version 4.2.1). The methods are all from the published R package, and the specific methods can be found in the text and references. Further inquiries can be directed to the corresponding author.

Received: 13 September 2023; Accepted: 1 November 2023

Published online: 18 November 2023

## References

- Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249. <https://doi.org/10.3322/caac.21660> (2021).
- Wang, M., Herbst, R. S. & Boshoff, C. Toward personalized treatment approaches for non-small-cell lung cancer. *Nat. Med.* **27**, 1345–1356. <https://doi.org/10.1038/s41591-021-01450-2> (2021).
- Tian, X., Gu, T., Lee, M. H. & Dong, Z. Challenge and countermeasures for EGFR targeted therapy in non-small cell lung cancer. *Biochim. Biophys. Acta Rev. Cancer* **1877**, 188645. <https://doi.org/10.1016/j.bbcan.2021.188645> (2022).
- Passaro, A., Brahmer, J., Antonia, S., Mok, T. & Peters, S. Managing resistance to immune checkpoint inhibitors in lung cancer: Treatment and novel strategies. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **40**, 598–610. <https://doi.org/10.1200/jco.21.01845> (2022).
- Zhang, X., Qiu, H., Li, C., Cai, P. & Qi, F. The positive role of traditional Chinese medicine as an adjunctive therapy for cancer. *Biosci. Trends* **15**, 283–298. <https://doi.org/10.5582/bst.2021.01318> (2021).
- Zhang, H. *et al.*  $\beta$ -elemene inhibits non-small cell lung cancer cell migration and invasion by inactivating the FAK-Src pathway. *Exp. Ther. Med.* **22**, 1095. <https://doi.org/10.3892/etm.2021.10529> (2021).
- Song, G. Q. *et al.* Elemene induces cell apoptosis via inhibiting glutathione synthesis in lung adenocarcinoma. *J. Ethnopharmacol.* **311**, 116409. <https://doi.org/10.1016/j.jep.2023.116409> (2023).
- Wang, J. *et al.*  $\beta$ -elemene enhances the antitumor activity of erlotinib by inducing apoptosis through AMPK and MAPK pathways in TKI-resistant H1975 lung cancer cells. *J. Cancer* **12**, 2285–2294. <https://doi.org/10.7150/jca.53382> (2021).
- Frisch, S. M. & Francis, H. Disruption of epithelial cell-matrix interactions induces apoptosis. *J. Cell Biol.* **124**, 619–626. <https://doi.org/10.1083/jcb.124.4.619> (1994).
- Khan, S. U., Fatima, K. & Malik, F. Understanding the cell survival mechanism of anoikis-resistant cancer cells during different steps of metastasis. *Clin. Exp. Metastasis* **39**, 715–726. <https://doi.org/10.1007/s10585-022-10172-9> (2022).
- Liu, A. *et al.* Silencing ZIC2 abrogates tumorigenesis and anoikis resistance of non-small cell lung cancer cells by inhibiting Src/FAK signaling. *Mol. Therap. Oncolytics* **22**, 195–208. <https://doi.org/10.1016/j.omto.2021.05.008> (2021).
- McCarroll, J. A. *et al.* TUBB3/ $\beta$ III-tubulin acts through the PTEN/AKT signaling axis to promote tumorigenesis and anoikis resistance in non-small cell lung cancer. *Cancer Res.* **75**, 415–425. <https://doi.org/10.1158/0008-5472.Can-14-2740> (2015).
- Jang, E. J. *et al.* FAM188B downregulation sensitizes lung cancer cells to anoikis via EGFR downregulation and inhibits tumor metastasis in vivo. *Cancers* **13**, 247. <https://doi.org/10.3390/cancers13020247> (2021).
- Cheng, G. *et al.*  $\beta$ -elemene suppresses tumor metabolism and stem cell-like properties of non-small cell lung cancer cells by regulating PI3K/AKT/mTOR signaling. *Am. J. Cancer Res.* **12**, 1535–1555 (2022).
- Anastasiadou, E., Jacob, L. S. & Slack, F. J. Non-coding RNA networks in cancer. *Nat. Rev. Cancer* **18**, 5–18. <https://doi.org/10.1038/nrc.2017.99> (2018).
- Salmena, L., Poliseno, L., Tay, Y., Kats, L. & Pandolfi, P. P. A ceRNA hypothesis: The rosetta stone of a hidden RNA language?. *Cell* **146**, 353–358. <https://doi.org/10.1016/j.cell.2011.07.014> (2011).
- Chan, J. J. & Tay, Y. Noncoding RNA: RNA regulatory networks in cancer. *Int. J. Mol. Sci.* **19**, 1310. <https://doi.org/10.3390/ijms19051310> (2018).
- Chen, Z. *et al.* Integrative analysis of NSCLC identifies LINC01234 as an oncogenic lncRNA that interacts with HNRNPA2B1 and regulates miR-106b biogenesis. *Mol. Therap. J. Am. Soc. Gene Therapy* **28**, 1479–1493. <https://doi.org/10.1016/j.jymthe.2020.03.010> (2020).
- Bing, Z., Han, J., Zheng, Z. & Liang, N. FOXO3-induced oncogenic lncRNA CASC9 enhances gefitinib resistance of non-small-cell lung cancer through feedback loop. *Life Sci.* **287**, 120012. <https://doi.org/10.1016/j.lfs.2021.120012> (2021).
- Liang, G. *et al.* miR-196b-5p-mediated downregulation of TSPAN12 and GATA6 promotes tumor progression in non-small cell lung cancer. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 4347–4357. <https://doi.org/10.1073/pnas.1917531117> (2020).
- Lu, Q. *et al.* lncRNA APOC1P1-3 promoting anoikis-resistance of breast cancer cells. *Cancer Cell Int.* **21**, 232. <https://doi.org/10.1186/s12935-021-01916-w> (2021).
- Fu, X. T. *et al.* MicroRNA-30a suppresses autophagy-mediated anoikis resistance and metastasis in hepatocellular carcinoma. *Cancer Letters* **412**, 108–117. <https://doi.org/10.1016/j.canlet.2017.10.012> (2018).
- Yu, Y. *et al.* CircCEMIP promotes anoikis-resistance by enhancing protective autophagy in prostate cancer cells. *J. Exp. Clin. Cancer Res.* **CR 41**, 188. <https://doi.org/10.1186/s13046-022-02381-7> (2022).
- Wang, W., Zhang, L., Sun, J., Zhao, Q. & Shuai, J. Predicting the potential human lncRNA-miRNA interactions based on graph convolution network with conditional random field. *Brief. Bioinform.* **23**, bbac463. <https://doi.org/10.1093/bib/bbac463> (2022).
- Zhang, L., Yang, P., Feng, H., Zhao, Q. & Liu, H. Using network distance analysis to predict lncRNA-miRNA interactions. *Interdiscip. Sci. Comput. Life Sci.* **13**, 535–545. <https://doi.org/10.1007/s12539-021-00458-z> (2021).
- Safran, M. *et al.* GeneCards Version 3: The human gene integrator. *Database J. Biol. Databases Curation* **2010**, 020. <https://doi.org/10.1093/database/baq020> (2010).
- Rouillard, A. D. *et al.* The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database J. Biol. Databases Curation* **2016**, 100. <https://doi.org/10.1093/database/baw100> (2016).
- Kim, S. *et al.* PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res.* **49**, D1388–d1395. <https://doi.org/10.1093/nar/gkaa971> (2021).

29. Daina, A., Michielin, O. & Zoete, V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res.* **47**, W357–w364. <https://doi.org/10.1093/nar/gkz382> (2019).
30. Pundir, S., Martin, M. J. & O'Donovan, C. UniProt Tools. *Curr. Prot. Bioinform.* **53**, 21–15. <https://doi.org/10.1002/0471250953.bi0129s53> (2016).
31. Szklarczyk, D. *et al.* The STRING database in 2023: Protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **51**, D638–d646. <https://doi.org/10.1093/nar/gkac1000> (2023).
32. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinform. Oxf. Eng.* **26**, 1572–1573. <https://doi.org/10.1093/bioinformatics/btq170> (2010).
33. Karagkouni, D. *et al.* DIANA-LncBase v3: Indexing experimentally supported miRNA targets on non-coding transcripts. *Nucleic Acids Res.* **48**, D101–d110. <https://doi.org/10.1093/nar/gkz1036> (2020).
34. Yi, Q., Pu, Y., Chao, F., Bian, P. & Lv, L. ACAP1 deficiency predicts inferior immunotherapy response in solid tumors. *Cancers* **14**, 5951. <https://doi.org/10.3390/cancers14235951> (2022).
35. Fu, J. *et al.* Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med.* **12**, 21. <https://doi.org/10.1186/s13073-020-0721-z> (2020).
36. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461. <https://doi.org/10.1002/jcc.21334> (2010).
37. Seeliger, D. & de Groot, B. L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J. Comput. Aided Mol. Design* **24**, 417–422. <https://doi.org/10.1007/s10822-010-9352-6> (2010).
38. Yang, S. R. *et al.* Precision medicine in non-small cell lung cancer: Current applications and future directions. *Semin. Cancer Biol.* **84**, 184–198. <https://doi.org/10.1016/j.semcancer.2020.07.009> (2022).
39. Chen, Y. *et al.* Terpenoids from *Curcuma* Rhizoma: Their anticancer effects and clinical uses on combination and versus drug therapies. *Biomed. Pharmacother.* **138**, 111350. <https://doi.org/10.1016/j.biopha.2021.111350> (2021).
40. Zhai, B. *et al.* Molecular targets of  $\beta$ -elemene, a herbal extract used in traditional Chinese medicine, and its potential role in cancer therapy: A review. *Biomed. Pharmacotherap.* **114**, 108812. <https://doi.org/10.1016/j.biopha.2019.108812> (2019).
41. Skrypnik, N. *et al.* PPAR $\alpha$  activation can help prevent and treat non-small cell lung cancer. *Cancer Res.* **74**, 621–631. <https://doi.org/10.1158/0008-5472.Can-13-1928> (2014).
42. Wang, M. S. *et al.* PPAR $\alpha$  agonist fenofibrate relieves acquired resistance to gefitinib in non-small cell lung cancer by promoting apoptosis via PPAR $\alpha$ /AMPK/AKT/FoxO1 pathway. *Acta Pharmacol. Sin.* **43**, 167–176. <https://doi.org/10.1038/s41401-021-00638-z> (2022).
43. Hann, S. S., Zheng, F. & Zhao, S. Targeting 3-phosphoinositide-dependent protein kinase 1 by N-acetyl-cysteine through activation of peroxisome proliferators activated receptor alpha in human lung cancer cells, the role of p53 and p65. *J. Exp. Clin. Cancer Res.* **CR 32**, 43. <https://doi.org/10.1186/1756-9966-32-43> (2013).
44. Xu, S., Ma, H., Bo, Y. & Shao, M. The oncogenic role of CB2 in the progression of non-small-cell lung cancer. *Biomed. Pharmacother.* **117**, 109080. <https://doi.org/10.1016/j.biopha.2019.109080> (2019).
45. Sarsembayeva, A. *et al.* Cannabinoid receptor 2 plays a pro-tumorigenic role in non-small cell lung cancer by limiting anti-tumor activity of CD8(+) T and NK cells. *Front. Immunol.* **13**, 997115. <https://doi.org/10.3389/fimmu.2022.997115> (2022).
46. Vidlarova, M. *et al.* Cannabinoid receptor 2 expression in early-stage non-small cell lung cancers identifies patients with good prognosis and longer survival. *Transl. Lung Cancer Res.* **11**, 2040–2050. <https://doi.org/10.21037/tlcr-22-247> (2022).
47. Cheng, Y. *et al.* Molecular characterization of lung cancer: A two-miRNA prognostic signature based on cancer stem-like cells related genes. *J. Cell. Biochem.* **121**, 2889–2900. <https://doi.org/10.1002/jcb.29525> (2020).
48. MacDonagh, L. *et al.* MicroRNA expression profiling and biomarker validation in treatment-naïve and drug resistant non-small cell lung cancer. *Transl. Lung Cancer Res.* **10**, 1773–1791. <https://doi.org/10.21037/tlcr-20-959> (2021).
49. Qiu, H. *et al.* miR-30b-5p inhibits cancer progression and enhances cisplatin sensitivity in lung cancer through targeting LRP8. *Apoptosis Int. J. Programm. Death* **26**, 261–276. <https://doi.org/10.1007/s10495-021-01665-1> (2021).
50. Moreno-Mateos, M. A. *et al.* Novel small RNA expression libraries uncover hsa-miR-30b and hsa-miR-30c as important factors in anoikis resistance. *RNA N.Y.* **19**, 1711–1725. <https://doi.org/10.1261/rna.039461.113> (2013).
51. Ding, W., Li, B., Zhang, Y., He, L. & Su, J. A neutrophil extracellular traps-associated lncRNA signature predicts the clinical outcomes in patients with lung adenocarcinoma. *Front. Genet.* **13**, 1047231. <https://doi.org/10.3389/fgene.2022.1047231> (2022).
52. Yi, W. W., Guo, X. Q., Xu, Y., Liang, B. & Song, P. A prognostic model based on ferroptosis-related long non-coding RNA signatures and immunotherapy responses for non-small cell lung cancer. *Eur. Rev. Med. Pharmacol. Sci.* **27**, 2591–2604. [https://doi.org/10.26355/eurev\\_202303\\_31796](https://doi.org/10.26355/eurev_202303_31796) (2023).
53. Yu, F. *et al.* Hypoxic tumor-derived exosomal miR-31-5p promotes lung adenocarcinoma metastasis by negatively regulating SATB2-reversed EMT and activating MEK/ERK signaling. *J. Exp. Clin. Cancer Res.* **CR 40**, 179. <https://doi.org/10.1186/s13046-021-01979-7> (2021).
54. Zhu, C. *et al.* miR-31-5p modulates cell progression in lung adenocarcinoma through TNSI/p53 axis. *Strahlenther. Onkol.* **198**, 304–314. <https://doi.org/10.1007/s00066-021-01895-x> (2022).
55. Zhu, B. *et al.* MicroRNA-31-5p enhances the Warburg effect via targeting FIH. *FASEB J. Off. Pub. Fed. Am. Soc. Exp. Biol.* **33**, 545–556. <https://doi.org/10.1096/fj.20180803R> (2019).
56. Kim, S. H. *et al.* Characterization of vitamin D receptor (VDR) in lung adenocarcinoma. *Lung Cancer Amst. Neth.* **77**, 265–271. <https://doi.org/10.1016/j.lungcan.2012.04.010> (2012).
57. Pineda Lancheros, L. E. *et al.* Effect of single nucleotide polymorphisms in the vitamin D metabolic pathway on susceptibility to non-small-cell lung cancer. *Nutrients* **14**, 4668. <https://doi.org/10.3390/nu14214668> (2022).
58. Wu, X., Cheng, J. & Yang, K. Vitamin D-related gene polymorphisms, plasma 25-hydroxy-vitamin D, cigarette smoke and non-small cell lung cancer (NSCLC) risk. *Int. J. Mol. Sci.* **17**, 1597. <https://doi.org/10.3390/ijms17101597> (2016).
59. Qin, T. *et al.* Clinical importance of VEGFC and PD-L1 co-expression in lung adenocarcinoma patients. *Thorac. Cancer* **11**, 1139–1148. <https://doi.org/10.1111/1759-7714.13354> (2020).
60. Regan, E. *et al.* Identification of gene expression differences between lymphangiogenic and non-lymphangiogenic non-small cell lung cancer cell lines. *PLoS one* **11**, e0150963. <https://doi.org/10.1371/journal.pone.0150963> (2016).
61. Wang, L. *et al.* Plasma RNA sequencing of extracellular RNAs reveals potential biomarkers for non-small cell lung cancer. *Clin. Biochem.* **83**, 65–73. <https://doi.org/10.1016/j.clinbiochem.2020.06.004> (2020).
62. Wang, X. *et al.* Combined histone deacetylase and cyclooxygenase inhibition achieves enhanced antiangiogenic effects in lung cancer cells. *Mol. Carcinog.* **52**, 218–228. <https://doi.org/10.1002/mc.21846> (2013).
63. Komatsu, N. *et al.* Cox2-mediated PGE2 production via p38/JNK-c-fos signaling inhibits cell apoptosis in 3D floating culture clumps of mesenchymal stem cell/extracellular matrix complexes. *Biochem. Biophys. Res. Commun.* **530**, 448–454. <https://doi.org/10.1016/j.bbrc.2020.05.100> (2020).
64. Zeng, Q., McCauley, L. K. & Wang, C. Y. Hepatocyte growth factor inhibits anoikis by induction of activator protein 1-dependent cyclooxygenase-2 Implication in head and neck squamous cell carcinoma progression. *J. Biol. Chem.* **277**, 50137–50142. <https://doi.org/10.1074/jbc.M208952200> (2002).
65. Liu, B., Yan, S., Qu, L. & Zhu, J. Celecoxib enhances anticancer effect of cisplatin and induces anoikis in osteosarcoma via PI3K/Akt pathway. *Cancer Cell Int.* **17**, 1. <https://doi.org/10.1186/s12935-016-0378-2> (2017).

66. Su, P., Ahmad, B., Zou, K. & Zou, L.  $\beta$ -elemene enhances the chemotherapeutic effect of 5-fluorouracil in triple-negative breast cancer via PI3K/AKT, RAF-MEK-Erk, and NF- $\kappa$ B signaling pathways. *Onco Therap* **13**, 5207–5222. <https://doi.org/10.2147/ott.S242820> (2020).
67. Cai, H., Ren, L., Wang, Y. & Zhang, Y. Beta-elemene reduces the malignancy of non-small cell lung cancer by enhancing C3orf21 expression. *Front. Oncol.* **11**, 571476. <https://doi.org/10.3389/fonc.2021.571476> (2021).
68. Hu, H. *et al.* Gene function and cell surface protein association analysis based on single-cell multiomics data. *Comput. Biol. Med.* **157**, 106733. <https://doi.org/10.1016/j.compbiomed.2023.106733> (2023).
69. Meng, R., Yin, S., Sun, J., Hu, H. & Zhao, Q. scAAGA: Single cell data analysis framework using asymmetric autoencoder with gene attention. *Comput. Biol. Med.* **165**, 107414. <https://doi.org/10.1016/j.compbiomed.2023.107414> (2023).
70. Gao, H. *et al.* Predicting metabolite-disease associations based on auto-encoder and non-negative matrix factorization. *Brief. Bioinform.* **24**, bbad259. <https://doi.org/10.1093/bib/bbad259> (2023).
71. Sun, F., Sun, J. & Zhao, Q. A deep learning method for predicting metabolite-disease associations via graph neural network. *Brief. Bioinform.* **23**, 266. <https://doi.org/10.1093/bib/bbac266> (2022).
72. Wang, T., Sun, J. & Zhao, Q. Investigating cardiotoxicity related with hERG channel blockers using molecular fingerprints and graph attention mechanism. *Comput. Biol. Med.* **153**, 106464. <https://doi.org/10.1016/j.compbiomed.2022.106464> (2023).
73. Li, X. *et al.* Caspase-1 and gasdermin D afford the optimal targets with distinct switching strategies in NLRP1b inflammasome-induced cell death. *Res. Wash. D. C.* **2022**, 9838341. <https://doi.org/10.34133/2022/9838341> (2022).
74. Li, X. *et al.* RIP1-dependent linear and nonlinear recruitments of caspase-8 and RIP3 respectively to necrosome specify distinct cell death outcomes. *Protein Cell* **12**, 858–876. <https://doi.org/10.1007/s13238-020-00810-x> (2021).
75. Xu, F. *et al.* Specificity and competition of mRNAs dominate droplet pattern in protein phase separation. *Phys. Rev. Res.* **5**, 023159. <https://doi.org/10.1103/PhysRevResearch.5.023159> (2023).

## Acknowledgements

We would like to thank the staff members of the TCGA database (<https://portal.gdc.cancer.gov/>), the GenCards database (<https://www.genecards.org/>), the Harmonizome database (<http://amp.pharm.mssm.edu/Harmonizome>), the PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>), the SwissTargetPrediction database (<http://swisstargetprediction.ch/>), the Uniprot database (<https://www.uniprot.org>), the String database (<https://string-db.org/>), the DIANA-LncBase tool (<http://www.microrna.gr/LncBase>), the TIDE database (<http://tide.dfc.harvard.edu>).

## Author contributions

K.T. and P.Z. conceived and designed the study. K.T. contributed to data collection and curation. K.T., C.Z. and Z.H. analyzed the datasets and interpreted the results. All authors interpreted the data, contributed to the writing and review of the manuscript, had full access to all the data in the study, and had final consent for publication. The authors read and approved the final manuscript.

## Funding

This work was supported by National Natural Science Foundation of China (82074425); National Natural Science Foundation of China Youth Science Foundation Project (82205227); Young Qihuang Scholars Talent Project of National Administration of Traditional Chinese Medicine; Top Technology Leading Talents Project of Hunan Province; Key R&D projects in Hunan Province (2021SK2006); Natural Science Foundation of Hunan Province (2021JJ30417); Natural Science Foundation of Hunan Province Youth Fund (2021JJ40310); Hunan Provincial Engineering Research Center of Anti-tumour Chinese Medicine Creation Technology Project; Hunan Provincial Health Commission Traditional Chinese Medicine Shennong Leading Talent Project; Youth Fund of Hunan Provincial Administration of Traditional Chinese Medicine (2021176).

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to P.Z.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023