RESEARCH

Open Access

An integrative analysis reveals cancer risk associated with artificial sweeteners

Jumin Xie[1]* , Ying Zhu[1†], Zixuan Yang[1†], Zhang Yu[1†], Mingzhi Yang[1†] and Qingzhi Wang[2]*

Abstract

Background Artificial sweeteners (AS) have been widely utilized in the food, beverage, and pharmaceutical industries for decades. While numerous publications have suggested a potential link between AS and diseases, particularly cancer, controversy still surrounds this issue. This study aims to investigate the association between AS consumption and cancer risk.

Methods Targets associated with commonly used AS were screened and validated using databases such as CTD, STITCH, Super-PRED, Swiss Target Prediction, SEA, PharmMapper, and GalaxySagittarius. Cancer-related targets were sourced from GeneCards, OMIM, and TTD databases. AS-cancer targets were identified through the intersection of these datasets. A network visualization ('AS-targets-cancer') was constructed using Cytoscape 3.9.0. Protein–protein interaction analysis was conducted using the STRING database to identify significant AS-cancer targets. GO and KEGG enrichment analyses were performed using the DAVID database. Core targets were identified from significant targets and genes involved in the 'Pathways in cancer' (map05200). Molecular docking and dynamics simulations were employed to verify interactions between AS and target proteins. Pan-cancer and univariate Cox regression analyses of core targets across 33 cancer types were conducted using GEPIA 2 and SangerBox, respectively. Gene chip datasets (GSE53757 for KIRC, GSE21354 for LGG, GSE42568 for BRCA, and GSE46602 for PRAD) were retrieved from the GEO database, while transcriptome and overall survival data were obtained from TCGA. Data normalization and identification of differentially expressed genes (DEGs) were performed on these datasets using R (version 4.3.2). Gene Set Enrichment Analysis (GSEA) was employed to identify critical pathways in the gene expression profiles between normal and cancer groups. A cancer risk prognostic model was constructed for key targets to further elucidate their significance in cancer initiation and progression. Finally, the HPA database was utilized to investigate variations in the expression of key AS-cancer target proteins across KIRC, LGG, BRCA, PRAD, and normal tissues.

Results Seven commonly used AS (Aspartame, Acesulfame, Sucralose, NHDC, Cyclamate, Neotame, and Saccharin) were selected for study. A total of 368 AS-cancer intersection targets were identified, with 48 notable AS-cancer targets, including TP53, EGFR, SRC, PIK3R1, and EP300, retrieved. GO biological process analysis indicated that these targets are involved in the regulation of apoptosis, gene expression, and cell proliferation. Thirty-five core targets were identified from the intersection of the 48 significant AS-cancer targets and genes in the 'Pathways in cancer' (map05200). KEGG enrichment analysis of these core targets revealed associations with several cancer types and the PI3K-Akt signaling pathway. Molecular docking and dynamics simulations confirmed interactions

---

†Ying Zhu, Zixuan Yang, Zhang Yu and Mingzhi Yang have contributed equally to this work.

*Correspondence:
Jumin Xie
xiejumin@hbpu.edu.cn; xiejm922@163.com
Qingzhi Wang
wangqz1001@163.com

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 2 of 31

between AS and these core targets. HSP90AA1 was found to be highly expressed across the 33 cancer types, while EGF showed the opposite trend. Univariate Cox regression analysis demonstrated strong associations of core targets with KIRC, LGG, BRCA, and PRAD. DEGs of AS-cancer core targets across these four cancers were analyzed. GSEA revealed upregulated and downregulated pathways enriched in KIRC, LGG, BRCA, and PRAD. Cancer risk prognostic models were constructed to elucidate the significant roles of key targets in cancer initiation and progression. Finally, the HPA database confirmed the crucial function of these targets in KIRC, LGG, BRCA, and PRAD.

Conclusion This study integrated data mining, machine learning, network toxicology, molecular docking, molecular dynamics simulations, and clinical sample analysis to demonstrate that AS increases the risk of kidney cancer, low-grade glioma, breast cancer, and prostate cancer through multiple targets and signaling pathways. This paper provides a valuable reference for the safety assessment and cancer risk evaluation of food additives. It urges food safety regulatory agencies to strengthen oversight and encourages the public to reduce consumption of foods and beverages containing artificial sweeteners and other additives.

Keywords Artificial sweetener, Cancer, Core targets, Molecular dynamics simulation, Machine learning, Network toxicology, Pathways in cancer

Introduction

Artificial sweeteners (AS) refer to a group of low-calorie, calorie-free, or non-nutritive additives employed as sugar substitutes in the food, beverage, and pharmaceutical industries, experiencing a global rise in consumption annually [1]. Prolonged use of these non-sugar sweeteners is strongly linked to obesity, type 2 diabetes (T2D), cancer, cardiovascular diseases (CVD), and mortality [2]. Currently, the World Health Organization (WHO) does not explicitly enumerate the types and species of AS approved for use in its latest guidelines [3]. However, commonly recognized and widely accepted AS deemed safe include aspartame, saccharin, sucralose, acesulfame K, neotame, cyclamate, and stevia [1, 3]. The types of AS approved and commonly used vary between countries, a phenomenon influenced by factors such as the authorized types of artificial sweeteners in each country, obesity rates, and dietary preferences.

The use of artificial sweeteners (AS) spans nearly a century, owing to their no or low caloric and intense sweetness (ranging from 200 to 13,000 times that of sucrose), making them highly sought after by dieters and diabetes patients [4–6]. Artificial sweeteners were initially introduced into the food industry in the 1950s. However, since the 2000s, there has been an exponential rise in their consumption [7]. As of August 2020, approximately 250,000 tons of AS had been consumed globally, with China, Japan, the United States, Germany, France, and the United Kingdom ranking among the largest consumers. Forecasts indicate that the global AS market will expand from $7.2 billion in 2022 to $9.7 billion by 2028, representing a compound annual growth rate of 5.05% [8].

In Europe, there are 19 approved artificial sweeteners (AS) [9]. In Spain, the most commonly consumed AS include acesulfame K, aspartame, cyclamate, and sucralose [10]. In the United States, six high-intensity sweeteners are FDA-approved as food additives: saccharin, aspartame, acesulfame potassium (Ace-K), sucralose, neotame, and advantame [11]. The Food and Drug Administration (FDA) has received Generally Recognized as Safe (GRAS) notices for two additional types of high-intensity sweeteners: certain steviol glycosides obtained from the leaves of the stevia plant (Stevia rebaudiana) and extracts from Siraitia grosvenorii Swingle fruit, also known as Luo Han Guo or monk fruit [11].

Currently, China has approved 20 types of sweeteners for use, many of which are also authorized for international and regional use, including aspartame, acesulfame, saccharin, and cyclamate [2, 12]. The Joint Expert Committee on Food Additives (JECFA), under the United Nations Food and Agriculture Organization (FAO) and the WHO, is the foremost authority globally for evaluating the safety of food additives. JECFA conducts comprehensive toxicity assessments, encompassing acute, subchronic, mutagenicity, carcinogenicity, reproductive toxicity, chronic toxicity, and metabolic pathways of each sweetener under consideration for approval. These rigorous evaluations lead to the establishment of Acceptable Daily Intake (ADI) values, which factor in variables such as race, gender, and age. JECFA asserts that normal consumption of sweeteners within ADI limits poses no safety risks. Adhering to relevant regulatory standards ensures that sweeteners do not pose health hazards when used appropriately [3, 12]. Several food and beverage products have been found to contain excessive levels of AS, and researches indicate that the consumption of these sweeteners may lead to various health issues [13–15].

Obesity and weight gain have emerged as significant global health concerns [2, 16]. The WHO recommends limiting sugar intake to less than 10% of daily caloric intake to manage body weight and prevent obesity [17]. However, large cohort studies indicated that high

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 3 of 31

consumption of AS is associated with increased risks of all-cause mortality, cardiovascular diseases, coronary artery disease, cerebrovascular events, and cancer [18, 19]. The growing body of research suggests a significant correlation between prolonged AS consumption and the development of various diseases. Furthermore, aside from the heightened risk of cardiovascular events, AS have been shown to elevate the likelihood of diabetes mellitus [5, 6, 13, 20], inflammatory bowel disease [21–23], glucose intolerance [24, 25], hypertension [26, 27], stroke [28, 29], neurotoxicity [30, 31], obesity [26, 32, 33], and cancers [14, 34, 35].

Recent publications have linked artificial sweetener (AS) consumption with increased cancer risks [14]. The first report, published in the 1980s, revealed an association between AS and bladder cancer [35–37]. Subsequently, several other cancers have also been reported to correlate with AS consumption. An animal study using Sprague–Dawley rats tested the carcinogenic effects of aspartame at doses close to the acceptable daily intake for humans. The results demonstrate that lifelong exposure to aspartame, beginning in fetal life, increases carcinogenic effects, resulting in higher incidences of lymphomas/leukemias in both male and female rats, and mammary cancer in females [38, 39]. The large French population-based cohort, NutriNet-Santé (2009–2021), which included 102,865 adults with a median follow-up time of 7.8 years, found that compared to non-consumers, those who consumed higher amounts of artificial sweeteners had a greater risk of overall cancer (n = 3358 cases; hazard ratio = 1.13 [95% CI 1.03 to 1.25], P-trend = 0.002). Increased risks were also observed for breast cancer and obesity-related cancers [19].

A recent NHANES study reported that higher absolute saccharin intake was associated with an increased risk of all-cause mortality (hazard ratio [HR]: 1.41, 95% CI 1.05, 1.90), CVD mortality (HR: 1.93, 95% CI 1.15, 3.25), and cancer mortality (HR: 2.26, 95% CI 1.10, 4.45) among diabetic and pre-diabetic individuals. For the overweight population, elevated saccharin intake was linked to a heightened risk of cancer mortality (HR: 7.369, 95% CI 2.122, 25.592). The study concluded substantial risks associated with increased saccharin intake for all-cause mortality, CVD mortality, and cancer mortality [40]. Sucralose has been reported to increase colorectal cancer risk in a murine model, accompanied by gut microbiota dysbiosis, impaired inactivation of digestive proteases, damage to the gut barrier, and heightened inflammation [41].

Large cohort studies, epidemiological research, and animal experiments have shown a strong association between AS and the risk of various types of cancer. Currently, the debate over whether artificial sweetener (AS)

cause cancer is intensifying. However, the existing experimental methods used to evaluate the safety and carcinogenicity of AS, including epidemiological and animal studies, face several limitations. These include uncertainty in establishing causation, the presence of confounding factors and biases, differences between animal models and human biology, unreasonable experimental conditions and doses, among others [42]. This paper aims to employ data mining, machine learning, network toxicology, molecular docking, and molecular dynamics simulation techniques to comprehensively assess the carcinogenic potential of AS, providing valuable references for their safety evaluation. In this study, after rigorous screening, eight key core targets were identified and found to be strongly associated with the carcinogenic effects of AS. The flowchart for this study is illustrated in Fig. 1.

Materials and methods

Target collection of artificial sweeteners (AS)

Based on the commonly used types of AS in China, the United States, the European Union, and other nations, seven specific AS have been chosen for further investigation. These include 'Aspartame,' 'Acesulfame,' 'Sucralose,' 'NHDC,' 'Cyclamate,' 'Neotame,' and 'Saccharin' respectively. Detailed information is provided in Table 1.

Seven commonly used AS were chosen as keywords for target selection in the Comparative Toxicogenomics Database (CTD, https://ctdbase.org/) [43], STITCH (http://stitch.embl.de/) [44], and Super-PRED (https://prediction.charite.de/) [45] databases. The isomeric SMILES and sdf molecular formats of these seven AS were obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov) database. Subsequently, the targets associated with these AS were searched using Swiss Target Prediction (http://www.swisstargetprediction.ch) [46] and SEA (https://sea.bkslab.org/) [47] databases, employing the isomeric SMILES. The AS sdf molecular format was converted to mol2 format using Chem3D software, and the corresponding targets were identified via PharmMapper (https://lilab-ecust.cn/pharmmapper/index.html) [48] and GalaxySagittarius (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=SAGITTARIUS) [49] databases. The species selected for analysis was '*Homo sapiens*', and the criteria for selection were as follows: in Swiss Target Prediction, 'Probability > 0.01'; in Super-PRED, 'Probability ≥ 50%'; and in PharmMapper, 'Norm Fit ≥ 0.7'.

The targets of the seven AS, obtained from various databases, were amalgamated to ascertain the individual targets for each sweetener. Subsequently, the respective targets of these seven sweeteners were merged to derive the collective targets for AS, facilitating subsequent research. A bioinformatics platform (https://www.bioin
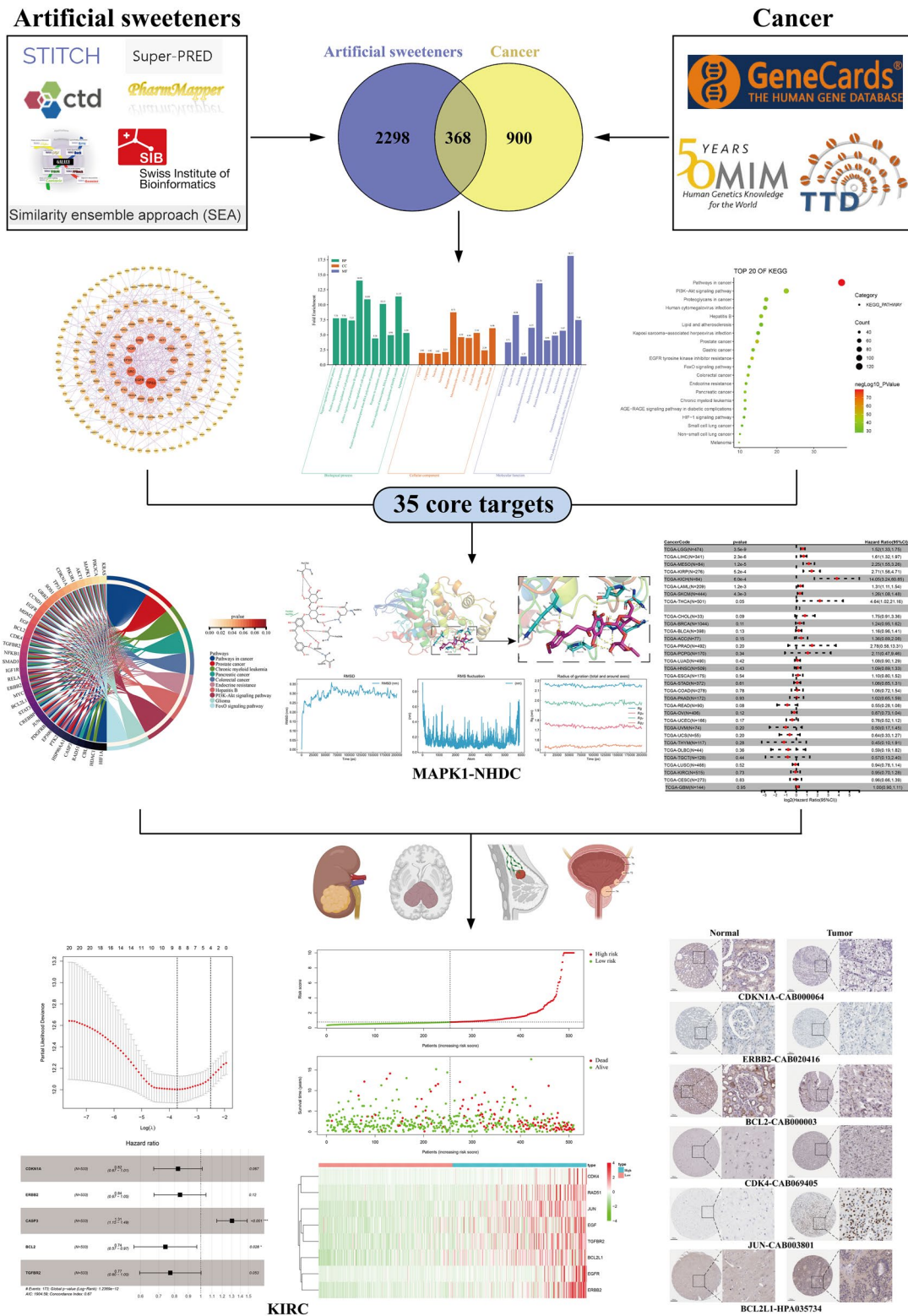
Fig. 1 Flowchart of this study

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 5 of 31

Table 1 Detailed information on commonly used artificial sweeteners

NO	Molecular Name	CAS ID	Formulas	MW(g/moL)
1	Aspartame	22,839–47-0	$C_{14}H_{18}N_2O_5$	294.3
2	Acesulfame	33,665–90-6	$C_4H_5NO_4S$	163.15
3	Sucralose	56,038–13-2	$C_{12}H_{19}Cl_3O_8$	397.6
4	NHDC	20,702–77-6	$C_{28}H_{36}O_{15}$	612.6
5	Cyclamate	100–88-9	$C_6H_{13}NO_3S$	179.24
6	Neotame	165,450–17-9	$C_{20}H_{30}N_2O_5$	378.5
7	Saccharin	81–07-2	$C_7H_5NO_3S$	183.19

formatics.com.cn/) was utilized to generate Venn diagrams depicting the overlapping targets among different AS.

Cancer-related target collection

The pertinent target associated with the keyword "cancer" was retrieved, focusing on the biological species '*Homo sapiens*,' across the GeneCards (https://www.genecards.org/) [50], OMIM (https://www.omim.org/) [51], and TTD (https://db.idrblab.net/ttd/) [52] databases.

GeneCards, the human gene database, allows researchers to efficiently explore and correlate a vast array of human genes, diseases, variants, proteins, cells, and biological pathways. OMIM is intended for use primarily by physicians and other professionals concerned with genetic disorders, by genetics researchers, and by advanced students in science and medicine. TTD is a database providing information about the known and explored therapeutic protein and nucleic acid targets, the targeted disease, pathway information and the corresponding drugs directed at each of these targets.

In the GeneCards database, targets were filtered using a criterion of 'Relevance score ≥ 15'. Cancer-related targets were then compiled by consolidating all retrieved targets from the GeneCards, OMIM, and TTD databases, utilizing the Hiplot (ORG) plotting website (https://hiplot.cn/).

Artificial sweetener-cancer target acquisition and protein–protein interaction (PPI) network construction

The intersecting targets between AS and cancer were identified using Venny 2.1 (https://bioinfogp.cnb.csic.es/tools/venny/) software. These intersection targets were then subjected to analysis of target protein interactions using the STRING database (https://cn.string-db.org/) [53], with a confidence coefficient set at ≥ 0.995 as the screening threshold. The STRING database integrates both known and predicted protein associations, including physical interactions and functional relationships.

Following the retrieval of PPI results, non-essential targets were filtered out, leading to the creation of a target protein interaction map. The resulting interaction network, provided in TSV format, was imported into Cytoscape 3.9.0 software (https://cytoscape.org/) [54] for visualization of target protein interactions. Cytoscape provides a range of network analysis tools that calculate various network attributes, such as Degree, Closeness Centrality, and Betweenness Centrality. These tools help assess the significance and relevance of network connections.

Utilizing the 'analyze network' module within Cytoscape 3.9.0, the topological attributes of target protein interactions were computed, yielding attribute data for target proteins. Selection criteria were defined as follows: Degree ≥ double the median value, Betweenness Centrality ≥ the median value, and Closeness Centrality ≥ the median value [55], thereby identifying significant targets associated with AS and cancer. The greater the network connectivity degree, the more important the node is within the network. Similarly, the higher the values of the two topological measures, "Betweenness Centrality" and "Closeness Centrality," the more significant the network node.

GO and KEGG enrichment analysis

The targets associated with AS and cancer were imported into the DAVID platform (https://david.ncifcrf.gov/) [56] for Gene Ontology (GO) [57] function and Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathway enrichment analysis, focusing on the species '*Homo sapiens*'. The DAVID platform offers a suite of powerful tools that enable researchers to explore large gene lists from multiple biological perspectives, extracting associated biological insights to the fullest extent.

The analyzed data regarding Biological Process (BP), Cellular Component (CC), and Molecular Function (MF) were obtained and visually represented using a bioinformatics platform (https://www.bioinformatics.com.cn/). The top 10 terms for BP, CC, and MF were selected based on their P-Value, arranged from lower to higher, to generate the GO bar chart. Similarly, the top 10 enriched terms from the KEGG analysis were visualized using the SangerBox platform (http://sangerbox.com/) to produce the KEGG pathway diagram.

Artificial sweetener-cancer core target screening and network construction

All genes pertaining to the 'pathways in cancer' (map05200) were acquired from the KEGG database. The core targets of AS-cancer were determined by intersecting the significant target genes of AS-cancer with the genes within the map05200 pathway. Subsequently,

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 6 of 31

an AS-Target-Cancer network was established using Cytoscape 3.9.0 software. Concurrently, the PPI of the AS-cancer core targets was examined via the STRING database, and the outcomes were visualized using Cytoscape 3.9.0 software.

Molecular docking and visualization

To further investigate potential interactions between the core targets and artificial sweeteners (AS), the core targets were retrieved from the UniProt database, filtered by "Reviewed" and "Human" entries. The corresponding target entries were then copied and searched in the RCSB PDB database (https://www.rcsb.org/) [58]. Based on *Homo sapiens*, near full-length sequences, the number of unique ligands, and low-resolution methods, the corresponding target protein crystal structures were selected as macromolecules.

The 2D structures of seven sweeteners were obtained from the PubChem database as small molecules. In the preliminary stage, Chem3D software was used to optimize the small molecule structures, which were then saved in mol2 format. For macromolecules, water molecules and extraneous residues were removed using PyMOL, and the structures were converted to pdb format. Both macromolecules and small molecules were processed using AutoDockTools 1.5.7 and converted into pdbqt format.

Subsequently, molecular docking between artificial AS molecules and the core targets was carried out using Autodock Vina (https://vina.scripps.edu/) software (Version 1.5.7) [59, 60].

The AS-core target associations were visualized using PyMOL (version 2.2) (http://www.pymol.org/2/), following the parameters previously outlined in our report [55]. Complexes with low binding energy and favorable conformations were selected for visualization in PyMOL to assess hydrogen bonds.

Additionally, 2D molecular docking interactions were conducted using the PROTEINS PLUS (https://proteins.plus/) web tools [61]. Strong interactions between the core targets and AS were observed when the affinity score was ≤ -5.0 kcal/mol [62]. The binding energy data obtained from the molecular docking results were imported into the GENESCLOUD website (https://www.genescloud.cn/chart/ChartOverview) to generate multi-axis bubble heat maps.

Molecular dynamics simulation

We conducted molecular dynamics (MD) simulations using GROMACS (Version 2020.6) to evaluate the binding patterns of the docking complexes [63]. The Amber GAFF2 force field was utilized to model the behavior of artificial sweetener (AS) molecules, while the FF14SB protein force field was employed for simulating the molecular dynamics [64, 65]. The binding complex was immersed in a $10 \times 10 \times 10$Å cubic box filled with TIP3P water molecules and neutralized with Na^+ and Cl^- ions.

The energy of the binding complex was minimized first using the steepest descent method and subsequently with the conjugate gradient method. Next, the thermodynamic temperature of the binding complex was gradually increased from 0 to 310 K under constant volume and a uniform heating rate. Subsequently, a 200 ps NVT (constant Number of particles, Volume, and Temperature) simulation was executed after the binding complex stabilized at a thermodynamic temperature of 310 K.

Upon pre-equilibration, an extended molecular dynamics simulation lasting 200 ns was performed employing the Nosé-Hoover algorithm-based Parrinello-Rahman constant-pressure barostat and V-rescale thermal bath method.

The Nosé-Hoover algorithm-based Parrinello-Rahman constant-pressure barostat is a standard method for maintaining temperature control in the NVT ensemble, making it suitable for equilibrium state simulations. The V-rescale thermostat, which uses velocity scaling with random noise terms, ensures proper canonical ensemble behavior, similar to Berendsen coupling but providing more accurate statistical properties. After molecular simulations, periodic correction of the binding complex is performed to account for potential numerical drift during the simulation.

Subsequent to the molecular simulation, the binding complex underwent periodic corrections, and the Root Mean Square Fluctuation (RMSF), Root Mean Square Deviation (RMSD), and Radius of Gyration (Rg) of the AS-target were calculated.

RMSF provides insight into the flexibility of different protein regions, with lower values indicating greater stability in those regions. RMSD measures the overall structural deviation of the protein over time, with lower values signifying a more stable structure. A smaller Rg suggests a more compact protein conformation, which may correlate with increased stability and is potentially linked to the binding affinity and specificity of AS to its target [66].

DuIvyTools was used to generate the visualization (https://duivytools.readthedocs.io/en/latest/DIT.html).

Pan-cancer analysis of AS-cancer core targets

Retrieve the expression data of AS-cancer core targets across 33 types of cancer by importing them into the GEPIA 2 website (http://gepia2.cancer-pku.cn/#index) [67]. Next, input this expression data into the CNSknowall website (https://cnsknowall.com/#/HomePage) to generate a heatmap. The full names and abbreviations for the 33 types of cancer can be found in Table 2.

Table 2 Full names and abbreviations of cancer types

Cohort	Full name
TCGA-ACC	Adrenocortical carcinoma
TCGA-BLCA	Bladder Urothelial Carcinoma
TCGA-BRCA	Breast invasive carcinoma
TCGA-CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
TCGA-CHOL	Cholangiocarcinoma
TCGA-COAD	Colon adenocarcinoma
TCGA-DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
TCGA-ESCA	Esophageal carcinoma
TCGA-GBM	Glioblastoma multiforme
TCGA-HNSC	Head and Neck squamous cell carcinoma
TCGA-KICH	Kidney Chromophobe
TCGA-KIRC	Kidney renal clear cell carcinoma
TCGA-KIRP	Kidney renal papillary cell carcinoma
TCGA-LAML	Acute Myeloid Leukemia
TCGA-LGG	Brain Lower Grade Glioma
TCGA-LIHC	Liver hepatocellular carcinoma
TCGA-LUAD	Lung adenocarcinoma
TCGA-LUSC	Lung squamous cell carcinoma
TCGA-MESO	Mesothelioma
TCGA-OV	Ovarian serous cystadenocarcinoma
TCGA-PAAD	Pancreatic adenocarcinoma
TCGA-PCPG	Pheochromocytoma and Paraganglioma
TCGA-PRAD	Prostate adenocarcinoma
TCGA-READ	Rectum adenocarcinoma
TCGA-SARC	Sarcoma
TCGA-STAD	Stomach adenocarcinoma
TCGA-SKCM	Skin Cutaneous Melanoma
TCGA-TGCT	Testicular Germ Cell Tumors
TCGA-THCA	Thyroid carcinoma
TCGA-THYM	Thymoma
TCGA-UCEC	Uterine Corpus Endometrial Carcinoma
TCGA-UCS	Uterine Carcinosarcoma
TCGA-UVM	Uveal Melanoma

Univariate cox regression analysis of AS-cancer core targets

The AS-cancer core targets were submitted to the SangerBox website for univariate Cox regression analysis, resulting in the acquisition of differential data. Subsequently, the CNSknowall website was utilized to generate a heatmap visualizing the results of the univariate Cox regression analysis.

Univariate Cox regression analysis is used to assess the association between each feature and survival individually. The purpose of employing univariate Cox regression here is to preliminarily identify cancers associated with the core target for further investigation. In the subsequent analysis of the four selected cancer studies, we controlled for potential confounding factors using multivariate analysis, LASSO regression, and other methods.

Data mining from GEO and TCGA databases

Following the integration of enriched KEGG pathways and univariate Cox regression analysis results concerning AS-cancer core targets, we selected renal clear cell carcinoma, low-grade glioma, breast cancer, and prostate cancer for further investigation. We sequentially procured gene chip datasets for these cancers from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) [68] as follows: GSE53757 for renal clear cell carcinoma, GSE21354 for low-grade glioma, GSE42568 for breast cancer, and GSE46602 for prostate cancer. Detailed information about the datasets is listed in Table 3. Additionally, transcriptome and overall survival data for these cancers were obtained from the TCGA database (https://portal.gdc.cancer.gov/), namely TCGA-KIRC for renal clear cell carcinoma, TCGA-LGG for low-grade glioma, TCGA-BRCA for breast cancer, and TCGA-PRAD for prostate cancer. The respective patient counts were 533 for KIRC, 510 for LGG, 1081 for BRCA, and 501 for PRAD. These datasets will be utilized in constructing prognosis models.

Screening of differentially expressed genes (DEGs) in GEO datasets and construction of target-pathway network

We retrieved all targets associated with "hsa05211: Renal cell carcinoma," "hsa05214: Glioma," "hsa05224: Breast cancer," and "hsa05215: Prostate cancer" from the KEGG database. Data normalization and identification of DEGs were performed on four GEO datasets (GSE53757, GSE21354, GSE42568, and GSE46602) using the limma package in R (version 4.3.2) [69]. The screening criteria

Table 3 GEO chip datasets

GEO	Tissue	Platform	Normal	Tumor	References
GSE53757	Kidney	GPL570	72	72	Von roemeling CA et al. (2014)
GSE21354	Brain	GPL570	4	14	Liu Z et al. (2011)
GSE42568	Breast	GPL570	17	104	Clarke C et al. (2013)
GSE46602	Prostate	GPL570	14	36	Mortensen MM et al. (2015)

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 8 of 31

included $p < 0.05$ and an absolute value of fold change greater than 2 ($| Log_2$ Fold Change $| > 1$) to identify both up-regulated and down-regulated genes. A gradient volcano map was generated using an online tool (http://bioinfoapp.com:3838/tools/).

The screening criteria depended on KEGG pathway targets. We intersected cancer targets with artificial sweetener targets to obtain the intersection target termed 'AS-cancer target.' This target set was imported into the STRING database, with a confidence coefficient set to 0.90. Subsequently, free targets were removed, and the target-pathway network was constructed using Cytoscape 3.9.0.

Gene set enrichment analysis (GSEA)

GSEA is employed to evaluate whether a predefined gene set exhibits statistically significant differences between two biological phenotypes, such as a control group and a cancer group, and to pinpoint critical pathways within the overall gene expression matrix [70]. Initially, all sequenced genes are ranked based on their differential factors, from largest to smallest. Subsequently, the input gene set is ranked to assess its enrichment across various biological functions and signaling pathways. The KEGG gene set was selected for screening based on a significance threshold of $p < 0.05$. The 'limma,' 'org.Hs.eg. db,' 'clusterProfiler,' and 'enrichplot' packages in the R (version 4.3.2) are utilized to visualize four cancer datasets [69]. The GSEA results include enrichment scores (ES) for the top 5 and bottom 5 biological processes and KEGG pathways. The enrichment score (ES), normalized enrichment score (NES), P-value, and adjusted P-value are computed through 1000 permutations. A positive ES indicates activation of the gene set, while a negative ES suggests suppression.

Construction and evaluation of cancer risk prognostic model

To further ascertain the significance of AS-Cancer core targets in cancer onset and progression, sequencing data and overall survival information for renal clear cell carcinoma (KIRC), low-grade glioma (LGG), breast cancer (BRCA), and prostate cancer (PRAD) were retrieved from the TCGA database. The validation of AS-Cancer core targets was conducted across these four cancer types. Univariate Cox proportional hazards regression analysis was performed on the AS-Cancer core targets within these datasets to generate forest plots. Genes exhibiting a significance level of $p < 0.05$ in the univariate Cox regression analysis were identified as prognostically relevant genes for subsequent analyses [71].

To mitigate overfitting, LASSO regression analysis was conducted using the 'glmnet' package in the R (version 4.3.2) [72]. Subsequently, a risk model was developed through stepwise multivariate Cox regression. The risk score formula was defined as follows: Risk Score $(RS) = \sum (Coef_{gene} \times exp_{genes})$, where $Coef_{gene}$ represents the coefficient of each prognostic gene, and exp_{gene} denotes the expression level of each gene (https://intaver.com/blog-project-management-project-risk-analysis/risk-scores-2/). Utilizing this formula, the risk score for each cancer patient was computed, and patients were stratified into high-risk and low-risk groups based on the median risk score.

Kaplan–Meier survival analysis was conducted using the R packages 'Survival' and 'survminer' in R (version 4.3.2) to compare survival differences between the high- and low-risk groups [73]. A significance level of $p < 0.05$ was adopted.

Additionally, in conjunction with survival duration and status data, receiver operating characteristic (ROC) curves for 1, 3, and 5 years were plotted using the 'survivalROC' R package to evaluate the prognostic model's predictive accuracy [74]. The area under the curve (AUC) was utilized for this assessment [75]. Furthermore, the change in risk score from low to high was graphed, and the association between follow-up time, survival status, gene expression, and risk score alteration was analyzed.

In addition to the algorithms used in this study, we aim to provide other useful and interesting algorithms for data collection and analysis [76–89].

Key protein expression profile in normal and cancer tissues.

Key protein expression profile between normal and cancerous tissues were analyzed using the Human Protein Atlas (HPA) database (https://www.proteinatlas.org/) [90]. Specifically, the HPA database was employed to investigate variations in the expression of key AS-cancer target proteins across KIRC, LGG, BRCA, PRAD, and normal tissues.

Results

Artificial sweetener (AS) targets

After compiling targets obtained from seven commonly used databases, we identified 1039 targets for Aspartame, 633 for Acesulfame, 686 for Sucralose, 710 for NHDC, 708 for Cyclamate, 864 for Neotame, and 703 for Saccharin. In total, the targets of these seven types of commonly used AS were combined, resulting in 2666 AS targets (Fig. 2 A-D).

AS-cancer targets

Using 'cancer' as the search keyword, we screened 1172 targets from the GeneCards database, 101 targets from the TTD database, and 168 targets from the OMIM
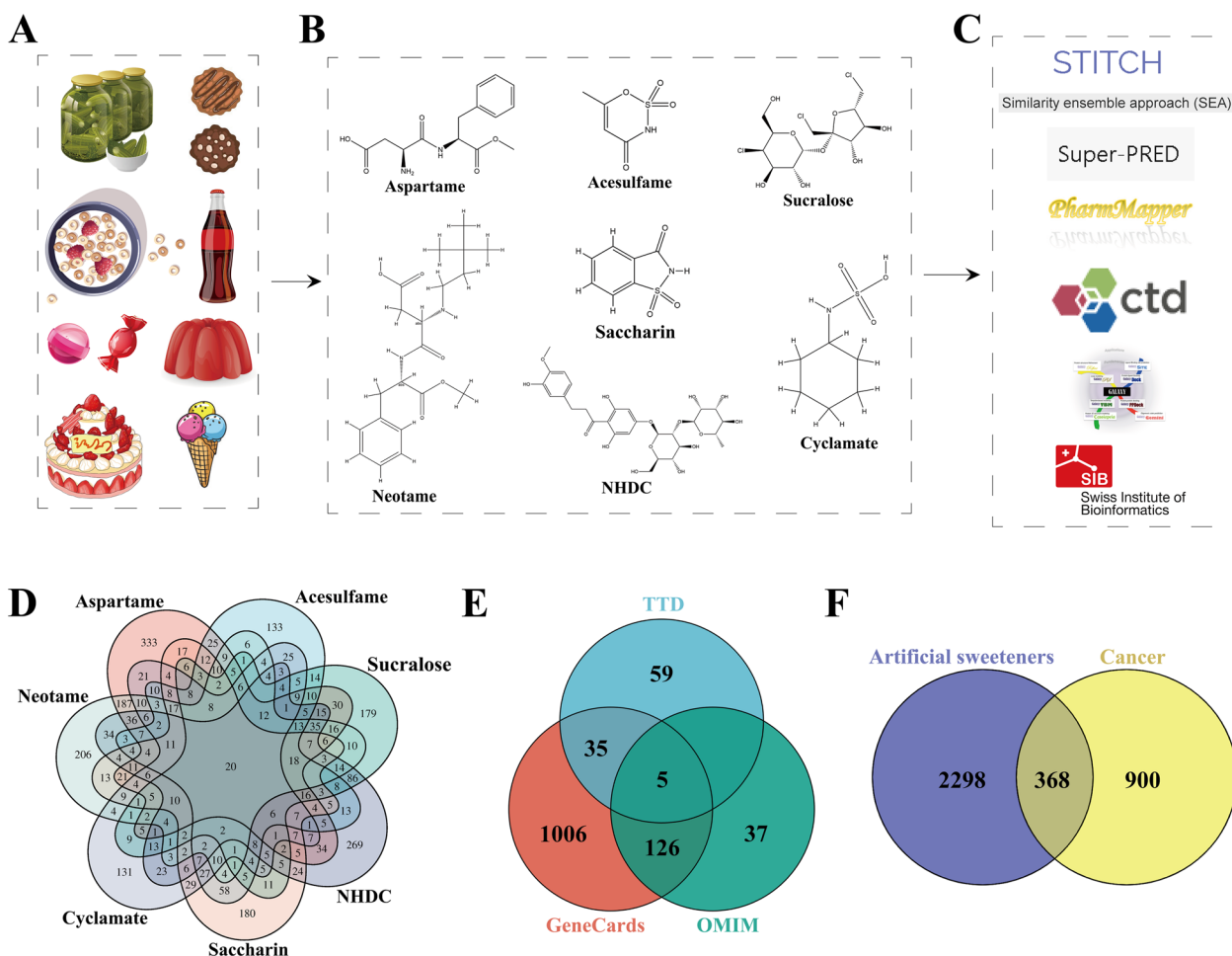
Fig. 2 Artificial sweetener-cancer targets obtain. **A** Daily consumption of foods containing artificial sweeteners (AS). **B** Molecular structures of seven commonly used AS. **C** AS targets acquisition database. **D** AS targets Venn diagram. **E** Cancer targets Venn diagram. **F** AS-cancer targets Venn diagram

database. By combining these three databases, a total of 1268 cancer targets were identified (Fig. 2E). The intersection of 2666 AS targets and 1268 cancer targets resulted in a total of 368 AS-cancer intersection targets (Fig. 2F).

Protein–protein interaction (PPI) network construction and AS-cancer significant targets selection

A total of 368 AS-cancer intersection targets were imported into the STRING database for PPI analysis with a confidence score set to 0.995. After analysis, 201 AS-cancer targets remained following the removal of free targets (Supplemental Fig. 1A).

Using Cytoscape 3.9.0, a visual PPI network consisting of 201 nodes and 433 edges was generated, excluding unconnected proteins (Supplemental Fig. 1B). These targets were ranked by Degree value, where darker colors and larger circle sizes indicate stronger interactions.

The "Analyze Network" tool was used to calculate topological attributes, from which gene target attribute data were derived. A total of 48 key AS-cancer targets were selected, including TP53, EGFR, SRC, PIK3R1, and EP300, among others, based on a degree ≥ 2 times the median value, betweenness $\geq$ median value, and closeness $\geq$ median value.

GO and KEGG enrichment analysis

The 368 targets associated AS-cancer were analyzed using the DAVID database, yielding 1362, 155, and 258 enriched GO-BP, GO-CC, and GO-MF terms, respectively. The top 10 most significant terms were selected to generate a GO bar chart (Supplemental Fig. 1C). The analysis indicated that these targets are primarily involved in negative regulation of apoptosis, positive regulation of gene expression, cell proliferation, and protein kinase B signaling in BP. In terms of cellular components,

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 10 of 31

the targets are predominantly located in the cytoplasm and nucleus. Molecular function analysis revealed that their functions mainly revolve around protein binding.

Furthermore, KEGG enrichment analysis identified 186 signaling pathways, with the top 20 most significant pathways depicted on a KEGG bubble chart (Supplemental Fig. 1D). Among these pathways, those associated with AS-cancer significant targets include Pathways in cancer, Prostate cancer, EGFR tyrosine kinase inhibitor resistance, and the PI3K-Akt signaling pathway.

AS-cancer core targets collection and network construction

Genes from the 'Pathways in cancer' (map05200, 325 genes) signaling pathway were retrieved from the KEGG database and intersected with 48 significant AS-cancer targets, resulting in the identification of 35 core AS-cancer targets.

The AS-core Target-Cancer network was constructed using Cytoscape 3.9.0 software (Fig. 3A).

PPI, GO and KEGG enrichment analysis of AS-cancer core targets

The PPI analysis of 35 AS-cancer core targets was conducted using the STRING database, and the resulting network was visualized using Cytoscape 3.9.0. This network consisted of 35 nodes and 76 edges (Fig. 3B).

The 35 AS-cancer core targets were further analyzed using the DAVID database, revealing 342 enriched GO-BP, 37 GO-CC, and 84 GO-MF terms. The top 10 most significant terms were selected to generate a GO bar chart (Fig. 4A). The analysis highlighted that AS-cancer core targets are primarily involved in negative regulation of apoptosis, positive regulation of cell proliferation, response to xenobiotic stimulus, and the epidermal growth factor receptor signaling pathway in BP. Regarding CC, these targets are predominantly located in the cytoplasm and nucleus. MF analysis indicated that their functions mainly involve enzyme binding.

Additionally, KEGG enrichment analysis identified 150 signaling pathways, with the top 10 most significant pathways depicted on a KEGG pathway chordal graph (Fig. 4B). Among these pathways, those associated with AS-cancer core targets include Pathways in cancer, Prostate cancer, Chronic myeloid leukemia, Pancreatic cancer, Colorectal cancer, Glioma, and the PI3K-Akt signaling pathway.

## Molecular docking and molecular dynamics simulation validation

The validation of 35 core protein targets in AS-cancer involved molecular docking with 7 AS molecules. A multi-axis bubble heat map was generated based on binding energy data from the molecular docking analysis (Fig. 5A). Combinations exhibiting lower binding energies for each AS were identified, such as CASP3-Sucralose, CCND1-Aspartame, CDK4-Cyclamate, KRAS-Neotame, KRAS-Saccharin, MAPK1-Acesulfame, MAPK1-NHDC, and BCL2L1-NHDC. The interactions between the AS molecules and the target proteins, including the formation of hydrogen bonds and involved amino acids, were visualized in both 2D and 3D structures.

Aspartame binds to seven amino acid residues within the CCND1 protein (MET-52, SER-56, VAL-74, ALA-76, ASN-146, LYS-147, and LYS-149), forming nine hydrogen bonds with bond lengths ranging from 2.4 to 3.3 Å (Fig. 5B). Acesulfame interacts with four amino acid residues in the MAPK1 protein (ALA-92, GLN-97, LYS-99, and ASP-100), forming seven hydrogen bonds ranging from 2.2 to 3.5 Å (Fig. 5C). Sucralose binds to five residues in the CASP3 protein (LEU-33, ASP-34, ASN-35, SER-36, and ASP-40), resulting in eight hydrogen bonds with distances from 2.2 to 3.3 Å (Fig. 5D). NHDC targets six residues in the MAPK1 protein (PRO-176, ASP-179, HIS-180, ASN-201, LYS-207, and PRO-298), establishing seven hydrogen bonds between 2.1 and 2.7 Å (Fig. 5E). Cyclamate binds to two residues in the CDK4 protein (ASN-83 and ASP-86), forming two hydrogen bonds with lengths from 2.1 to 2.8 Å (Fig. 5F). Neotame interacts with seven amino acids in the KRAS protein (GLY-13, VAL-14, GLY-15, LYS-16, SER-17, VAL-29, and ASP-30), forming nine hydrogen bonds with lengths between 2.2 and 3.4 Å (Fig. 5G). Saccharin binds to three amino acids in the KRAS protein (ASN-116, ASP-119, and ALA-146), forming three hydrogen bonds with distances from 2.0 to 2.7 Å (Fig. 5H). NHDC binds to six residues in the BCL2L1 protein (CLU-96, TYR-101, ASN-136, TRP-137, GLY-138, and ASN-198), forming seven hydrogen bonds with bond lengths ranging from 2.0 to 3.5 Å (Fig. 5I).

The binding energies for the AS complexes are as follows: CCND1-Aspartame, −6.8 kcal/mol; CDK4-Cyclamate, −6.5 kcal/mol; MAPK1-Acesulfame, −5.2 kcal/mol; KRAS-Neotame, −7.9 kcal/mol; CASP3-Sucralose, −5.5 kcal/mol; KRAS-Saccharin, − 6.9 kcal/mol; MAPK1-NHDC, − 8.6 kcal/mol; and BCL2L1-NHDC, − 8.3 kcal/mol.

The RMSD curve of the MAPK1-NHDC complex indicated that the complex achieved stability after 50 ns, reaching dynamic equilibrium (Fig. 6A). The RMSF curve showed minimal fluctuation in the MAPK1 protein, indicating stability of its domains and conserved functional regions (Fig. 6B). The Rg curve demonstrated consistent compactness throughout the simulation, highlighting the structural integrity of the complex (Fig. 6C). Molecular dynamics simulations confirmed the effective binding of MAPK1 and NHDC.
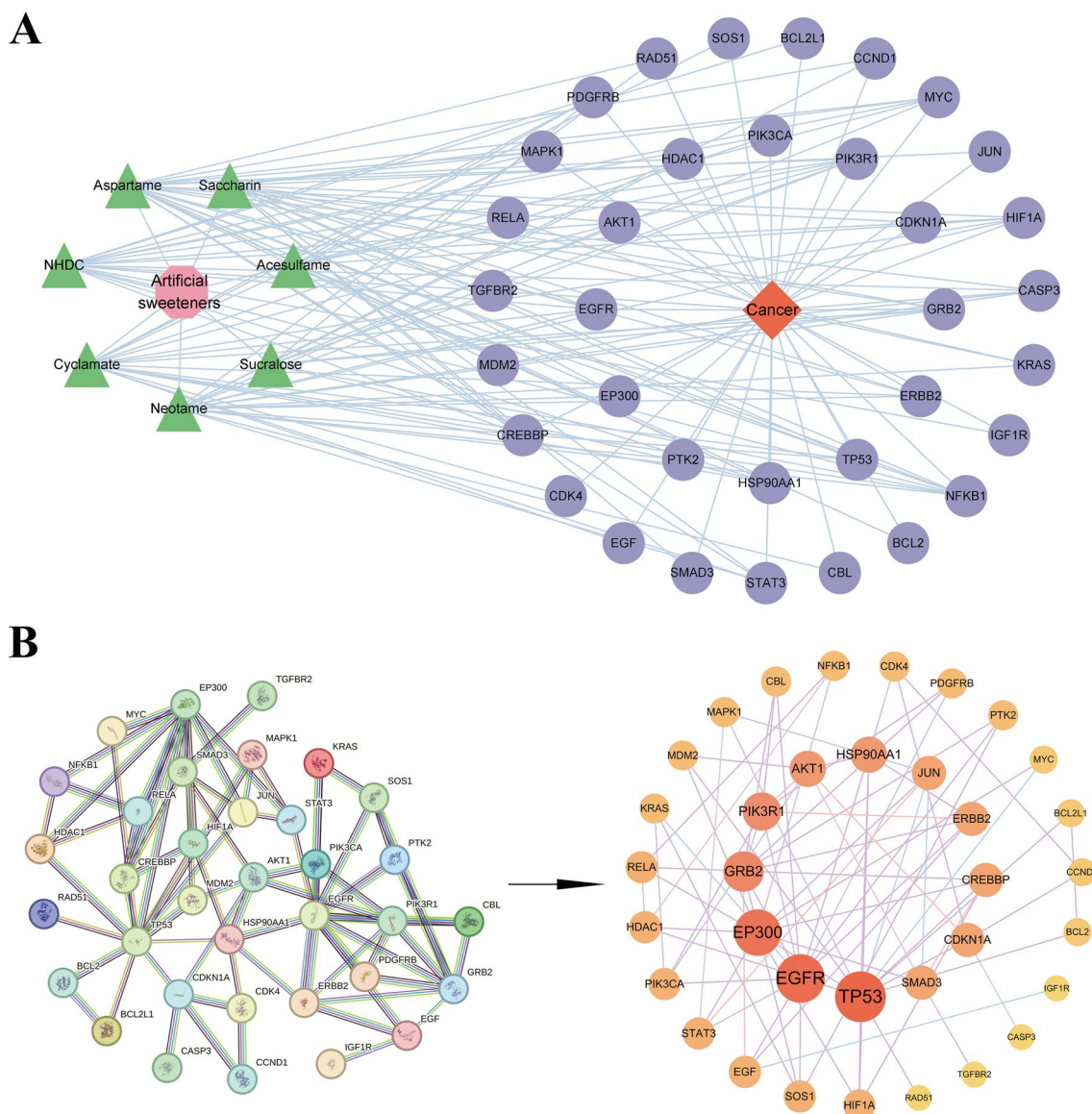
Fig. 3 AS-cancer core targets acquisition. **A** AS-target-cancer network. The pink octagon symbolizes AS, the green triangle denotes seven commonly used types of artificial sweeteners, the orange diamond represents cancer, the purple circle signifies 35 core targets, and the blue lines primarily indicate the relationships between the seven AS and the core targets. **B** PPI network diagram and visualization of AS-cancer core targets. In the PPI visualization diagram, core targets are sorted based on their degree value. A darker color and larger circle area indicate a higher Degree value, signifying stronger interactions among targets

Similarly, the RMSD curve of the CDK4-Cyclamate complex stabilized after 50 ns, confirming its structural stability akin to the MAPK1-NHDC complex (Fig. 6D). Analysis of RMSF and Rg curves further supported the structural integrity and flexibility of the CDK4-Cyclamate complex during simulation (Fig. 6E & F).

For the MAPK1-Acesulfame complex, stability was achieved after 100 ns, with RMSF and Rg curves reinforcing its structural integrity (supplemental Fig. 2A–C). The KRAS-Neotame and KRAS-Saccharin complexes also

demonstrated stability after 100 ns, as corroborated by RMSF and Rg curves (supplemental Fig. 2D-I).

However, the BCL2-Aspartame complex showed instability within 200 ns based on the RMSD curve, despite minimal fluctuation in the protein skeleton observed in the RMSF curve and stable tightness reflected in the Rg curve (supplemental Fig. 2J-L). Similarly, the CASP3-Sucralose complex exhibited instability after 200 ns according to the RMSD curve, with consistent protein skeleton fluctuation observed in the RMSF curve and
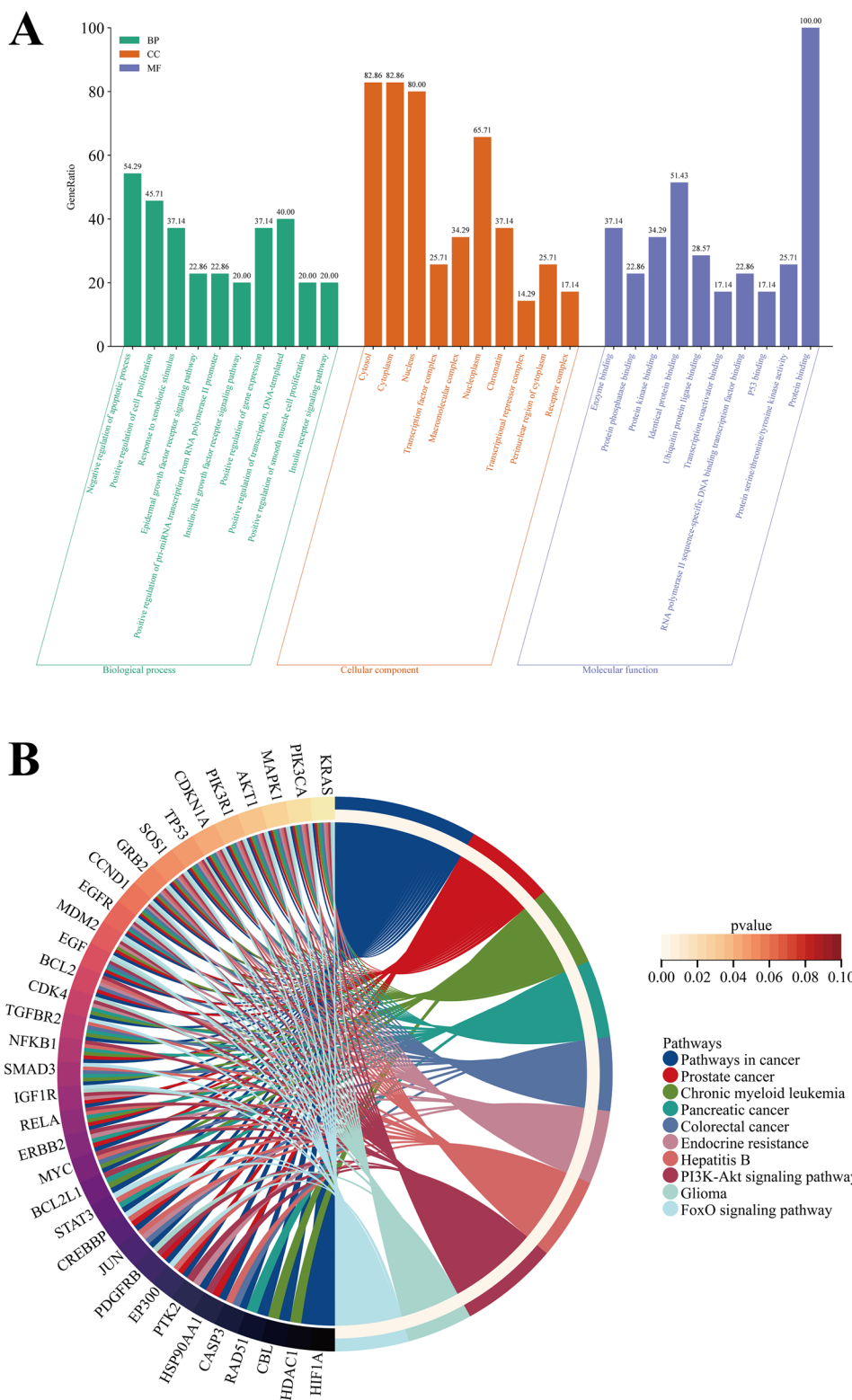
Fig. 4 GO and KEGG enrichment analysis of AS-cancer core targets. **A** Triadic histogram depicting the GO biological processes (BP), molecular functions (MF), and cellular components (CC) of AS-cancer core targets. In this visualization, green denotes the BP terms, orange represents the CC terms, and purple signifies the MF terms. Each column reflects the GeneRatio associated with its respective terms. **B** KEGG enrichment analysis of top 10 pathways for AS-cancer core targets. Different colors on the right half of the ring represent various pathways. On the left half, different colors indicate PValue for the 35 core targets. The lines within the circle illustrate relationships between pathways and targets
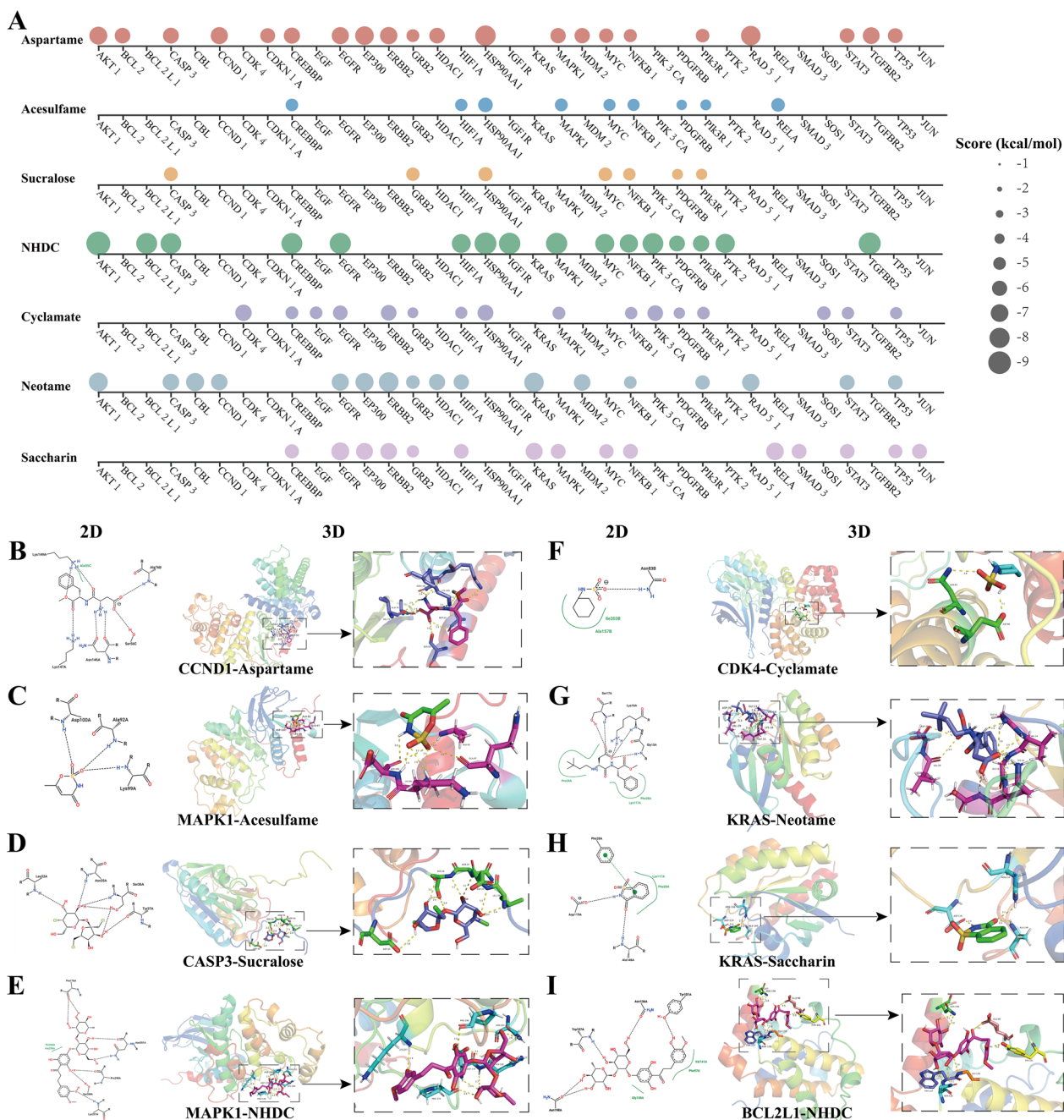
Fig. 5 Molecular docking between AS and AS-cancer core targets. **A** The molecular docking heat map. The results display seven different circles in various colors, each representing a different AS molecule. The size of each circle correlates with the value of the binding energy. **B**–**I** 2D and 3D interacted structures of AS and AS-cancer core targets. CCND1-Aspartame, MAPK1-Acesulfame, CASP3-Sucralose, MAPK1-NHDC, CDK4-Cyclamate, KRAS-Neotame, KRAS-Saccharin, and BCL2L1-NHDC

relatively stable system tightness reflected in the Rg curve (supplemental Fig. 2M–O).

Pan-cancer analysis results of AS-cancer core targets

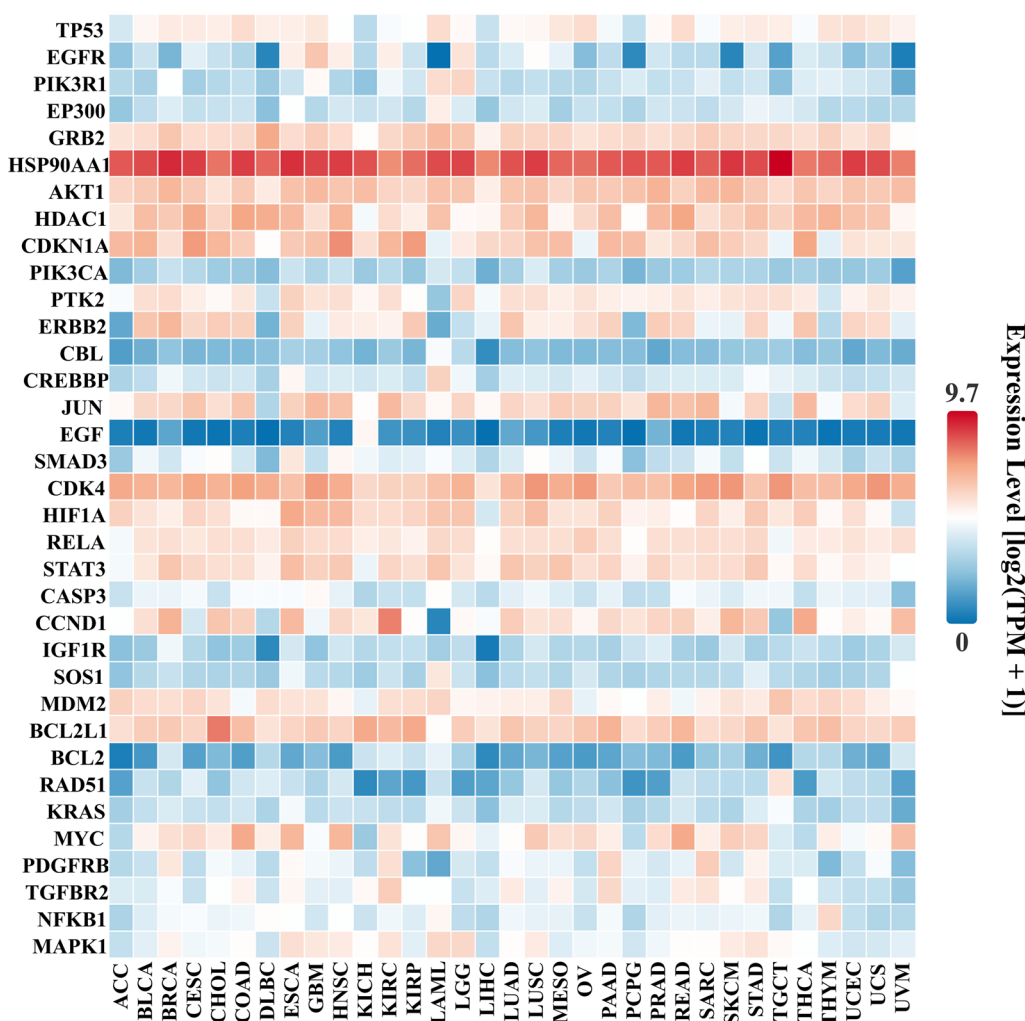The expression data for 35 AS-cancer core targets across 33 types of cancer were obtained from the GEPIA2 database. These expression profiles were then utilized on the CNSknowall website to generate a heatmap illustrating variations in expression levels among these targets across the different cancer types (Fig. 7). The analysis revealed that HSP90AA1 consistently exhibits high expression levels across all cancers, followed by CDK4,
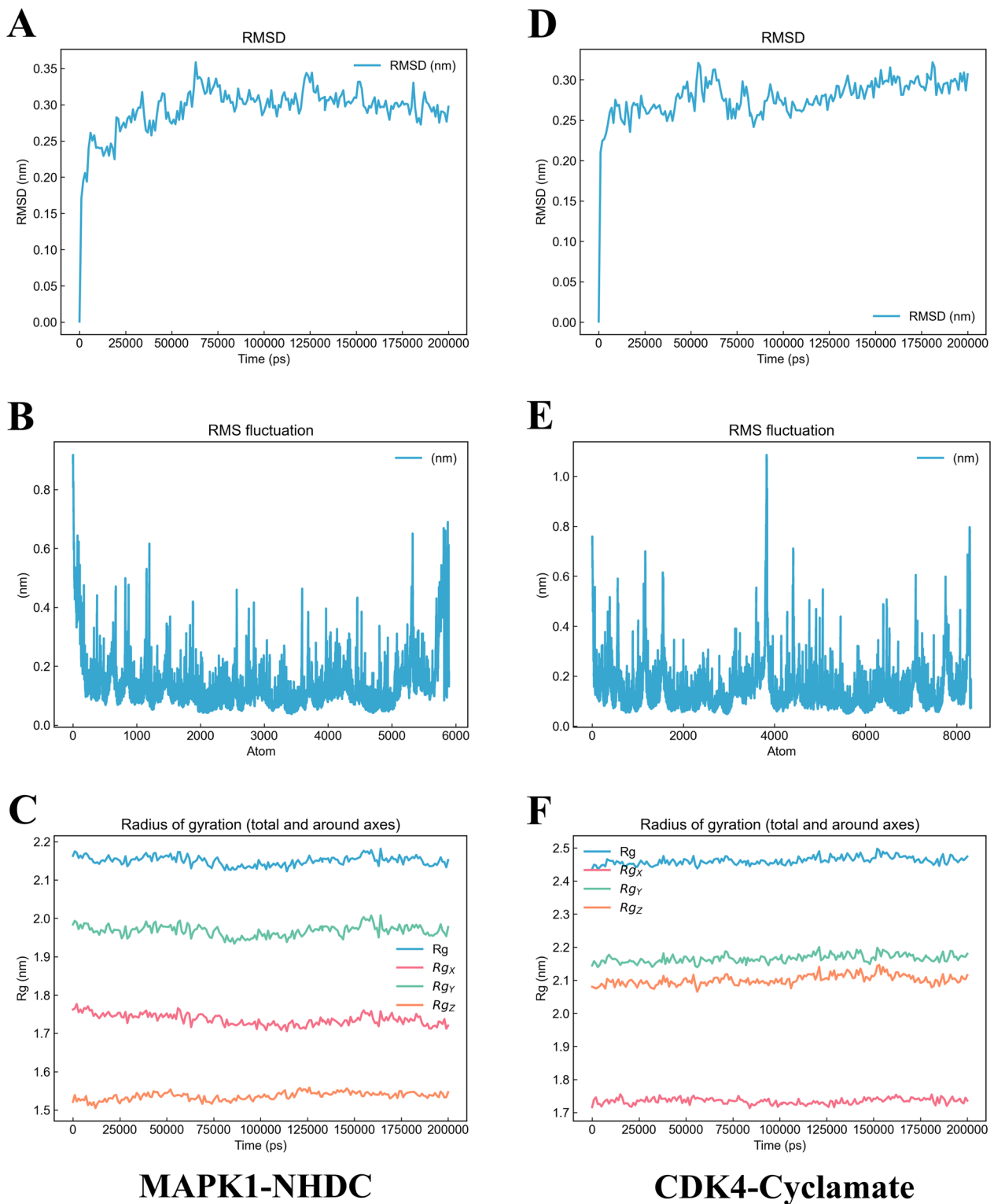
Fig. 6 Molecular dynamics simulation. **A** and **D** Root mean square deviation (RMSD) of MAPK1-NHDC and CDK4-cyclamate. **B** and **E** Root mean square fluctuation (RMSF). **C** and **F** The radius of gyration (RG) and its values along the three axes (Rgx, Rgy, Rgz)

Fig. 7 AS-cancer core targets pan-cancer analysis expression difference heat map. The expression patterns of 35 AS-cancer core targets across 33 types of cancer are visualized. Red indicates higher expression levels, while blue indicates lower expression levels

AKT1, BCL2L1, and HDAC1. Conversely, the five target proteins with the lowest expression levels across all cancers are EGF, BCL2, CBL, RAD51, and PIK3CA.
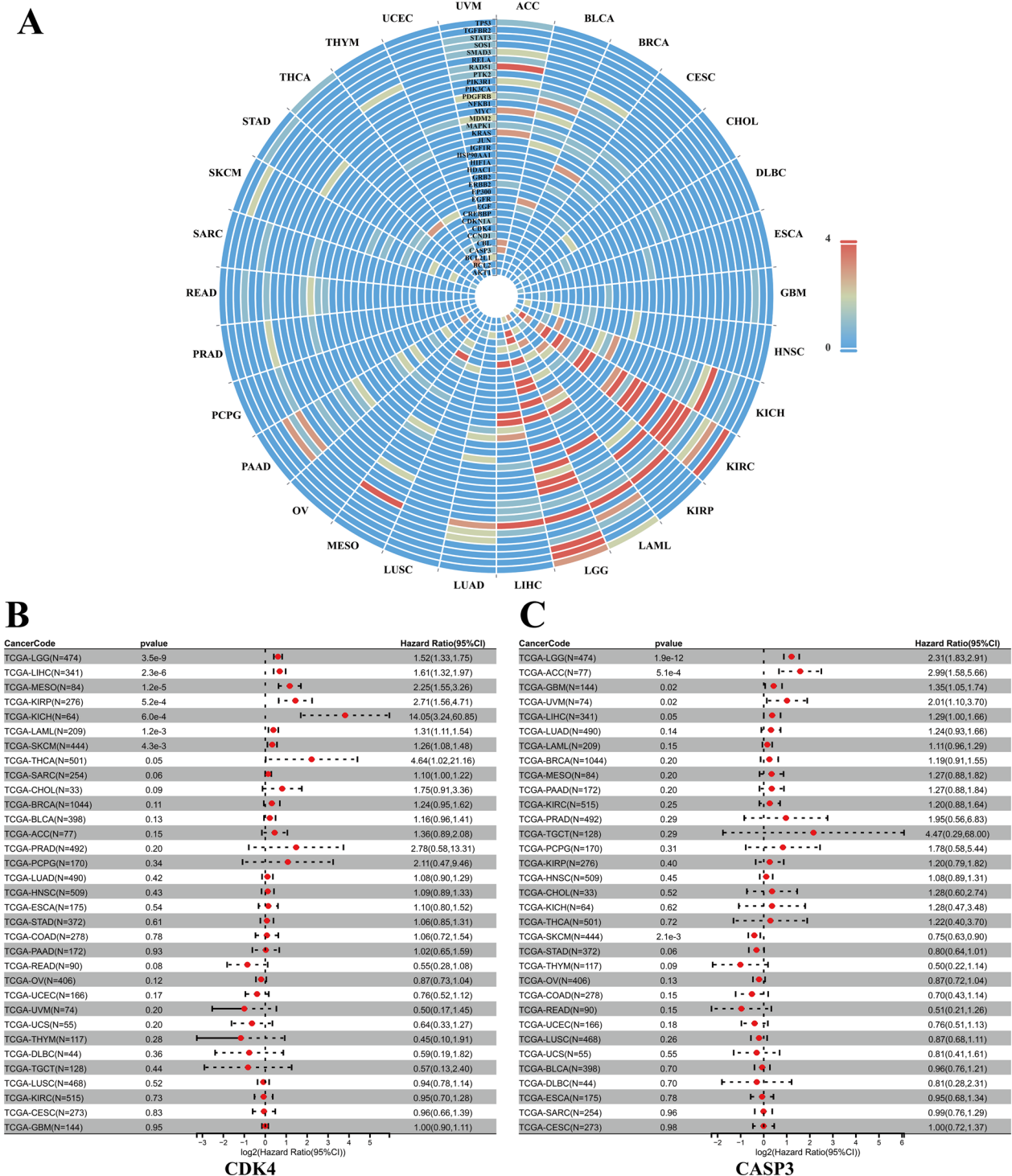
Univariate cox regression analysis results

The 35 AS-cancer core targets underwent univariate COX regression analysis using the SangerBox website, resulting in 35 sets of forest plots and difference data. These difference data were then utilized on the CNSknowall website to generate a heatmap illustrating variations identified through univariate COX regression analysis among the core targets (Fig. 8A). Significant differences in expression were observed in cancers such as 'LGG' and 'KIRC' based on these heatmaps.

The log2 (hazard ratio (95% CI)) of the 35 core targets across 33 types of cancer was calculated using the SangerBox website, and each target's association with the

cancers was visualized in a forest plot. Specifically, the log2 (hazard ratio (95% CI)) values of CDK4 and CASP3 across the 33 cancers were presented (Fig. 8B and C). A log2 (hazard ratio (95% CI)) value of 0 indicates little to no correlation between the target and cancer (hazard ratio, HR=1). A HR greater than 1 suggests a positive association between the target and cancer, while a value less than 1 indicates a negative association. CDK4 shows high correlation with LGG, LIHC, and MESO, whereas CASP3 is strongly associated with SKCM, ACC, and LGG (Fig. 8B and C).

Differentially expressed genes (DEGs) of AS-cancer core targets

The gene expression profiles of normal and cancer samples from GSE53757 (KIRC), GSE21354 (LGG), GSE42568 (BRCA), and GSE46602 (PRAD) were

Fig. 8 Univariate Cox regression analysis of AS-cancer core targets. **A** Visualization heat maps display the differential expression of 35 AS-cancer core targets across 33 types of cancer. In these maps, 'ns' denotes '0', '*' denotes '1', '**' denotes '2', '***' denotes '3', and '****' denotes '4'. **B** Forest plots of univariate Cox regression for CDK4 across 33 cancers show that an HR = 1 indicates no effect, HR < 1 indicates a favorable effect, and HR > 1 indicates an adverse effect. CDK4 demonstrates significant differences in LGG, LIHC, MESO, KIRP, KICH, SKCM, and LAML. **C** Forest maps of univariate Cox regression for CASP3 across 33 types of cancer reveal significant differences in LGG, ACC, SKCM, GBM, LIHC, and UVM

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 17 of 31

retrieved from the GEO database. The original gene expression data for KIRC, LGG, BRCA, and PRAD (Supplemental Fig. 3A, C, E and G) were normalized using the R 'limma' package to produce standardized data (Supplemental Fig. 3B, D, F, & H).

DEGs analysis of the GSE53757 dataset revealed significant upregulation of CCND1, MAPK1, MYC, CDKN1A, and PDGFRB, while EGF was notably downregulated (Fig. 9A). In the GSE21354 dataset, analysis indicated significant upregulation of RELA, BCL2L1, MYC, TP53, IGF1R, HIF1A, CCND1, and CDK4 (Fig. 9E). Analysis of the GSE42568 dataset identified HDAC1, PTK2, and IGF1R as significantly upregulated, whereas PIK3CA, TGFBR2, and JUN were downregulated (Supplemental Fig. 4A). In the GSE46602 dataset, MYC was found to be significantly upregulated, while BCL2 was significantly downregulated (Supplemental Fig. 4E).

AS-cancer targets expression analysis

In the expression analysis of Clear Cell Renal Cell Carcinoma (ccRCC or KIRC), significant upregulation was observed in MAPK1, MET, ETS1, CDKN1A, and EGLN3, while PAK6 and ARNT2 showed significant downregulation (Fig. 9B). For Lower Grade Glioma (LGG), AKT3, ARAF, SOS2, RB1, CDK4, GADD45G, RAF1, PIK3R3, IGF1R, CDKN1A, CCND1, and TP53 were notably upregulated, whereas CAMK2A, PRKCB, PRKCG, DDB2, CAMK2B, PDGFA, CALML5, SHC3, MAP2K1, PDGFB, CAMK4, CALM1, and KRAS were significantly downregulated (Fig. 9F). In Breast Invasive Carcinoma (BRCA), IGF1R, RPS6KB2, E2F3, JAG2, and FZD6 were significantly upregulated, whereas MAP2K1, JUN, FOS, FGF2, FGF10, FGF1, PIK3CA, IGF1, TCF7L2, FZD4, FZD5, and WNT11 were downregulated (Supplemental Fig. 4B). Lastly, in Prostate Adenocarcinoma (PRAD), CREB3L1 was found to be significantly upregulated, while FGFR2, IGF1, TCF7L1, BCL2, and TCF7L2 were all significantly downregulated (Supplemental Fig. 4F).
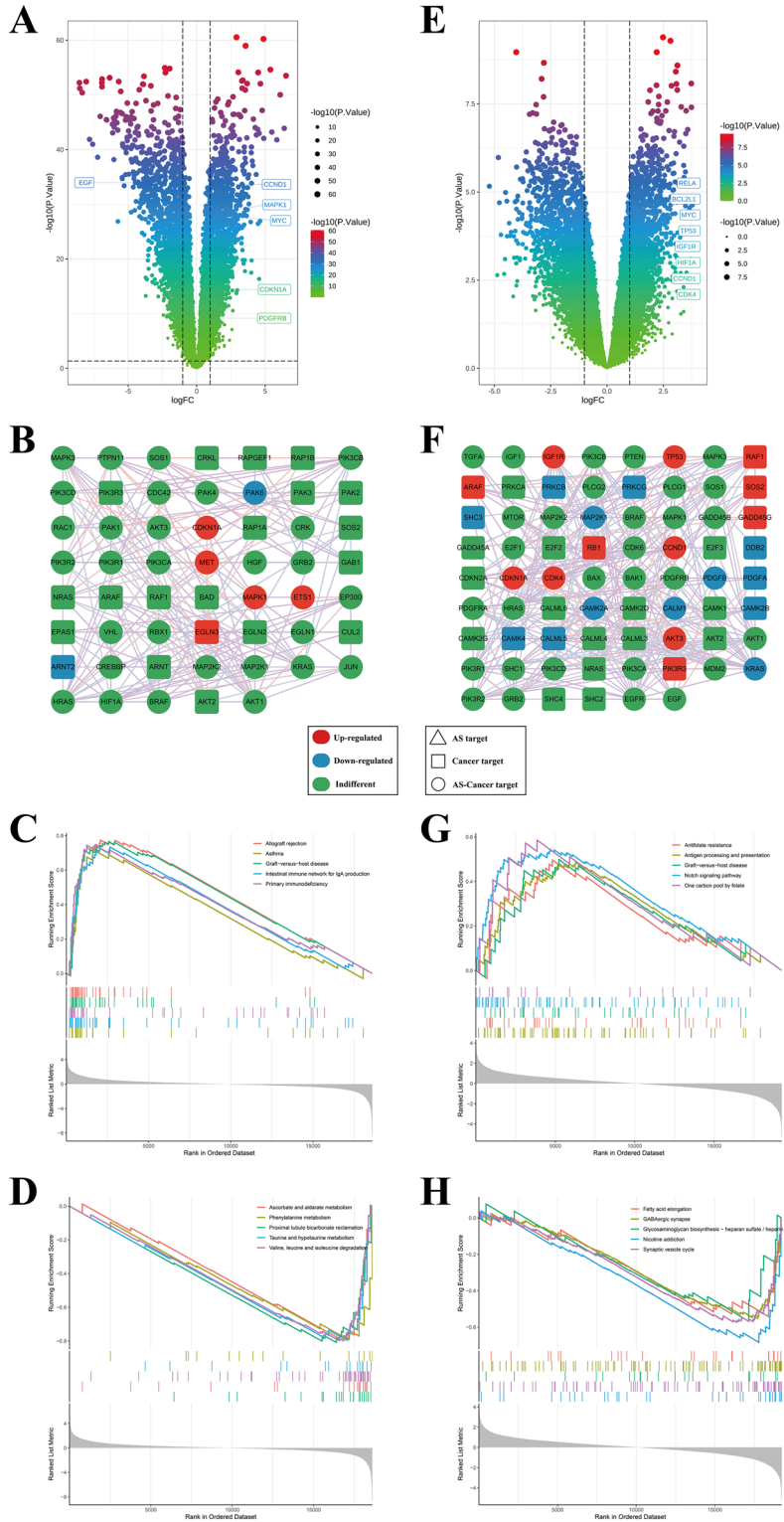
GSEA results

The GSEA analysis revealed the molecular functions and biological processes implicated in DEGs across four cancers. In the KIRC dataset GSE53757, pathways such as Allograft rejection, Asthma, Graft-versus-host disease, Intestinal immune network for IgA production, and Primary immunodeficiency were significantly up-regulated (Fig. 9C). Conversely, pathways including Ascorbate and aldarate metabolism, Phenylalanine metabolism, Proximal tubule bicarbonate reclamation, Taurine and hypotaurine metabolism, and Valine, leucine, and isoleucine degradation showed significant down-regulation (Fig. 9D). In the LGG dataset GSE21354, pathways such as Antifolate resistance, Antigen processing and presentation, Graft-versus-host disease, Notch signaling pathway, and One carbon pool by folate were markedly up-regulated (Fig. 9G). Conversely, pathways like Fatty acid elongation, GABAergic synapse, Glycosaminoglycan biosynthesis-heparan sulfate/heparin, Nicotine addiction, and Synaptic vesicle cycle were significantly down-regulated (Fig. 9H).

For the BRCA dataset GSE42568, pathways including Base excision repair, DNA replication, Fanconi anemia pathway, Pyrimidine metabolism, and Vibrio cholerae infection were notably up-regulated (Supplemental Fig. 4C). Conversely, pathways such as Citrate cycle (TCA cycle), Fatty acid degradation, PPAR signaling pathway, Propanoate metabolism, and Regulation of lipolysis in adipocytes exhibited significant down-regulation (Supplemental Fig. 4D). In the PRAD dataset GSE46602, pathways like Base excision repair, Biosynthesis of nucleotide sugars, Mannose type O-glycan biosynthesis, Nucleotide metabolism, and Pyrimidine metabolism were significantly up-regulated (Supplemental Fig. 4G). Conversely, pathways such as Chemical carcinogenes-DNA adducts, Drug metabolism-cytochrome P450, Graft-versus-host disease, Histidine metabolism, and Metabolism of xenobiotics by cytochrome P450 showed marked down-regulation (Supplemental Fig. 4H).

(See figure on next page.)

Fig. 9 Differentially expressed genes (DEGs) screening, target-pathway diagram and GSEA pathway enrichment analysis in clear cell renal cell carcinoma (GSE53757) and brain low-grade glioma (GSE21354). **A** Volcano plots depicting DEGs in a renal cancer dataset (GSE53757) reveal significant up-regulation of CCND1, MAPK1, MYC, CDKN1A, and PDGFRB genes in KIRC patients, while the EGF gene shows significant down-regulation in these patients. **B** Target-pathway diagrams in AS and KIRC. In these diagrams, red indicates genes significantly upregulated, blue denotes significantly down-regulated genes, and green represents genes with insignificant changes. Shapes correspond to different targets: triangles denote sweetener targets, squares denote disease targets, and circles denote targets common to both sweeteners and diseases. **C** The top 5 GSEA enrichment pathways up-regulated in KIRC. **D** The top 5 GSEA enrichment pathways down-regulated in KIRC. **E** Volcano plot of DEGs in a low-grade glioma dataset (GSE21354) shows significant up-regulation of RELA, BCL2L1, MYC, TP53, IGF1A, CCND1, and CDK4 in patients with LGG. **F** Target-pathway diagrams in AS and LGG. **G** The top 5 GSEA enrichment pathways up-regulated in LGG. **H** The top 5 GSEA enrichment pathways down-regulated in LGG

Fig. 9 (See legend on previous page.)

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 19 of 31

Cancer risk prognostic model

Univariate Cox regression analysis was conducted on 35 AS-cancer core target genes in TCGA-KIRC, revealing that 20 genes were significantly associated with overall KIRC survival ($p < 0.05$) (Fig. 10A). To avoid overfitting, a LASSO regression analysis was then performed on these 20 genes (Fig. 10B and C). Subsequently, multivariate Cox regression analysis was conducted on the 9 genes identified from the LASSO analysis. This analysis identified 5 genes (CDKN1A, ERBB2, CASP3, BCL2, TGFBR2) for constructing the KIRC prognostic model (Fig. 10D). The risk score (RS) for each patient was calculated as follows: RS = (0.1925 * CDKN1A) + (0.1753 * CASP3) + (0.2663 * ERBB2) + (0.2996 * TGFBR2) + (0.2579 * BCL2). RS was used to classify patients into high and low-risk groups using the median as the cutoff. Survival analysis indicated that patients in the low-risk group had significantly better survival outcomes compared to those in the high-risk group, with a highly significant statistical difference ($p < 0.001$) (Fig. 10E). The area under the ROC curve (AUC) for overall survival prediction by this model at 1, 3, and 5 years was 0.695, 0.681, and 0.697, respectively (Fig. 10F). Patients were ranked by their RS, and a RS plot was generated. This plot showed that as RS increased, the expression levels of CASP3 increased while those of CDKN1A, ERBB2, BCL2, and TGFBR2 decreased. Moreover, higher RS scores were associated with shorter patient survival times and increased mortality (Fig. 10 G).

Similarly, univariate Cox regression analysis was performed on 35 AS-cancer core target genes in TCGA-LGG, identifying 19 genes significantly associated with overall survival of LGG ($p < 0.05$) (Fig. 11A). To prevent overfitting, LASSO regression analysis was applied to these 19 genes (Fig. 11B&C). Subsequently, multivariate Cox regression analysis was conducted on the 11 genes identified from LASSO analysis. This analysis resulted in 8 genes (EGFR, ERBB2, JUN, EGF, CDK4, BCL2L1, RAD51, TGFBR2) being selected for constructing the LGG prognostic model (Fig. 11D). The risk score (RS) for each patient was calculated as: RS = (0.1500 * EGFR) + (0.3603 * ERBB2) + (0.1934 * JUN) + (0.1309 * EGF) + (0.2287 * CDK4) + (0.2350 * BCL2L1) + (0.2043 * RAD51) + (0.1495 * TGFBR2). RS was used to classify patients into high and low-risk groups using the median as the cutoff. Survival analysis showed that patients in the low-risk group had significantly better survival outcomes compared to those in the high-risk group, with a highly significant statistical difference ($p < 0.001$) (Fig. 11E). The AUC of the ROC curve for overall survival predicted by this model was 0.843, 0.795, and 0.752 at 1st, 3rd, and 5th years respectively (Fig. 11F). Patients were ranked by their RS, and a RS plot was generated. This plot showed that with increasing RS, the expression levels of each gene increased. Additionally, higher RS scores were associated with shorter patient survival times and increased mortality (Fig. 11G).

Univariate Cox regression analysis was performed on 35 AS-cancer core target genes in the TCGA-BRCA dataset, identifying 6 genes significantly associated with overall survival in breast cancer (BRCA) ($p < 0.05$) (Supplemental Fig. 5A). To minimize overfitting, these 6 genes were selected for LASSO regression analysis (Supplemental Fig. 5B&C). Subsequently, multivariate Cox regression analysis was conducted on the genes derived from the LASSO analysis, leading to the selection of 3 genes (HSP90AA1, PIK3CA, RAD51) for constructing a prognostic model for BRCA (Supplemental Fig. 5D). The risk score for each patient was calculated as follows: RS = (0.1578*HSP90AA1) + (0.1424*PIK3CA) + (0.1053*RAD51). Patients were stratified into high and low-risk groups using the median risk score as the cutoff. Survival analysis indicated that patients in the low-risk group had significantly better outcomes than those in the high-risk group ($p = 0.041$) (Supplemental Fig. 5E). The AUC of the ROC curve predicting overall survival (OS) was 0.578, 0.602, and 0.593 at one, three, and five years, respectively (Supplemental Fig. 5F). Patients were ranked by risk score from lowest to highest, and a risk score graph was generated. Along with increasing risk scores, gene expression levels rose, survival times decreased, and a higher number of patients died (Supplemental Fig. 5G).

Univariate Cox regression analysis was performed on 35 AS-cancer core target genes in the TCGA-PRAD dataset, identifying 2 genes significantly associated with overall survival in prostate cancer (PRAD) ($p < 0.05$) (Supplemental Fig. 6A). To minimize overfitting, these 2

(See figure on next page.)

Fig. 10 Renal cancer prognosis targets screening, and a cancer risk prognostic model constructing. **A** Univariate Cox regression analysis initially identified 20 prognostic genes (P < 0.05). **B** A coefficient profile plot was generated against the log(λ) values in the LASSO regression analysis model. The optimal parameter (λ) was marked with a dashed line on the left. **C** LASSO regression analysis of the 20 prognostic genes. **D** Multivariate Cox analysis identified 5 prognostic genes. **E** Kaplan–Meier survival analysis in high and low risk groups. **F** ROC curves of overall survival at 1, 3 and 5 years of the prognostic model. **G** The relationship between the survival status of renal cancer patients and the risk score of the prognostic model, along with changes in gene expression levels as the risk score increases. The dashed line delineates the boundary between the low-risk and high-risk groups
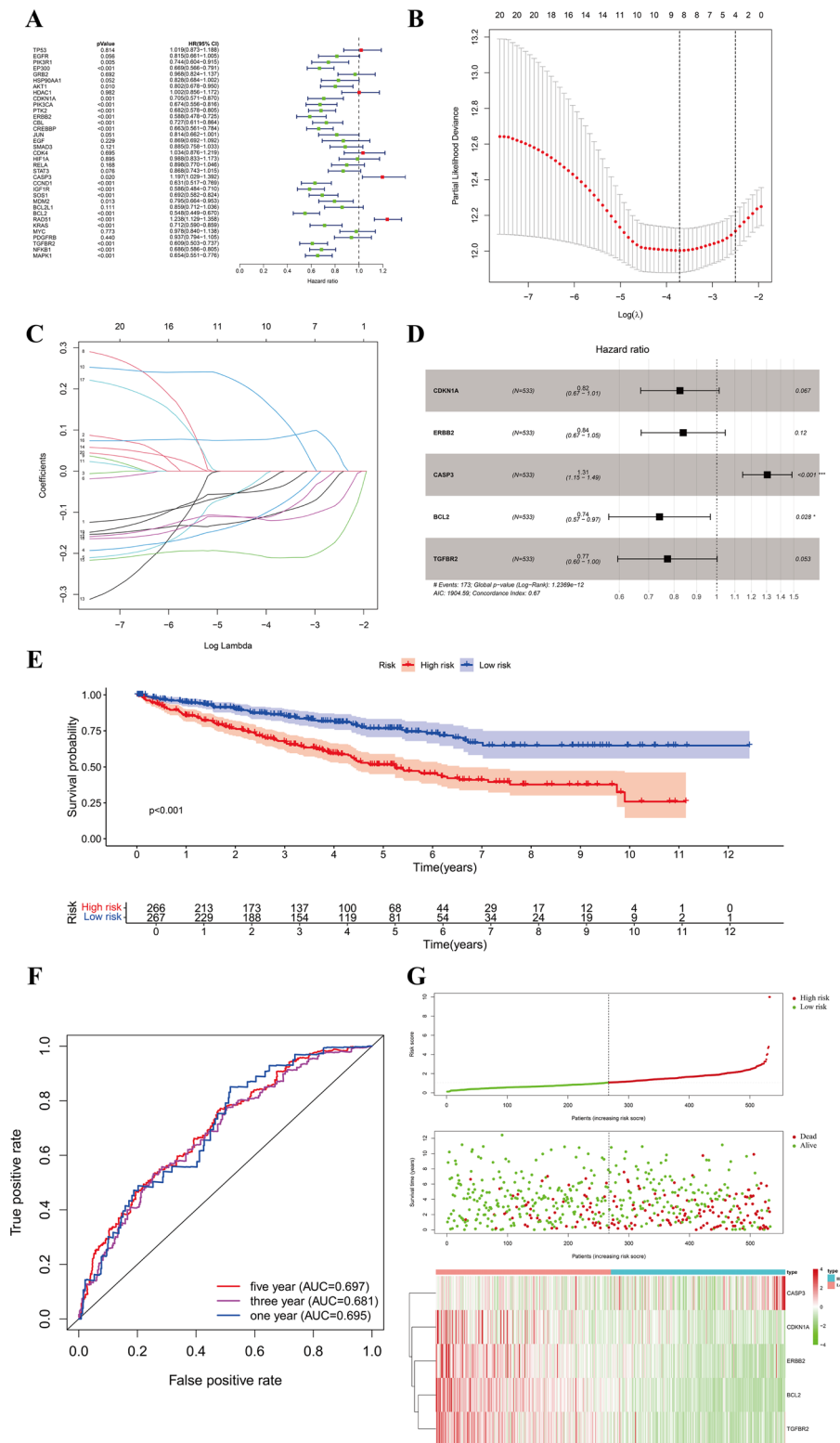
Fig. 10 (See legend on previous page.)

genes were selected for LASSO regression analysis (Supplemental Fig. 6B&C). Subsequently, multivariate Cox regression analysis was conducted on the genes derived from the LASSO analysis, leading to the selection of 2 genes (GRB2, RAD51) for constructing a prognostic model for PRAD (Supplemental Fig. 6D). The risk score for each patient was calculated as follows: RS = (0.7069* GRB2) + (0.2665*RAD51). Patients were stratified into high and low-risk groups using the median risk score as the cutoff. Survival analysis indicated that patients in the low-risk group had slightly better outcomes than those in the high-risk group (p = 0.079) (Supplemental Fig. 6E). The AUC of the ROC curve predicting overall survival (OS) was 0.990, 0.607, and 0.645 at one, three, and five years, respectively (Supplemental Fig. 6F). Patients were ranked by their RS, and a RS plot was generated. This plot showed that with increasing RS, the expression levels of each gene increased. Additionally, higher RS scores were associated with shorter patient survival times and increased mortality (Supplemental Fig. 6G).

HPA validation

Immunohistochemical results of normal kidney tissue and KIRC tissue stained with CDKN1A, ERBB2, and BCL2 corresponding antibodies CAB000064, CAB020416, and CAB000003 were obtained from the HPA database, respectively. Immunohistochemical staining results showed that, compared with normal kidney tissue, the immunohistochemical staining of KIRC tissues was light, indicating that the three proteins were significantly underexpressed in KIRC tissues (Fig. 12A–C). Immunohistochemical results of normal brain tissue and LGG tissue stained with CDK4, JUN, and BCL2L1 corresponding antibodies CAB069405, CAB003801, and HPA035734 were obtained from the HPA database, respectively. Immunohistochemical staining results showed that, compared with normal tissue, the immunohistochemical staining of LGG tissue was deeper, indicating that these three proteins were significantly overexpressed in LGG tissue (Fig. 12D–F).

Immunohistochemical results for normal breast tissue and BRCA tissue stained with antibodies CAB017804, CAB002058, and HPA039310 targeting PIK3CA,

HSP90AA1, and RAD51 were obtained from the HPA database, respectively. Immunohistochemical staining revealed that compared to normal breast tissue, BRCA tissue exhibited deeper staining, indicating high expression levels of these three proteins (Supplemental Fig. 7A–C).

Similarly, immunohistochemical results for normal prostate tissue and PRAD tissue stained with GRB2 and RAD51 antibodies CAB002589 and CAB010381 were retrieved from the HPA database, respectively. Immunohistochemical analysis of PRAD tissue demonstrated deeper staining compared to normal tissue, highlighting high expression levels of these two proteins in PRAD tissue (Supplemental Fig. 7D–E).

Discussion

Artificial sweeteners (AS) are widely added to food and beverages due to their low-calorie content and intense sweetness. By 2028, the market for AS is expected to approach nearly \$10 billion [8]. The extensive use of AS has raised concerns regarding environmental pollution and human health. In many countries, AS serve as indicators of pollution in groundwater, rivers, lakes, and seas [91–94]. Following the consumption of AS, there have been observed changes in the type and quantity of human intestinal microbiota, potentially linked to the development of various diseases [24, 25, 95–97]. AS consumption has been associated with premature births, DNA fragmentation, cancer, diabetes, obesity, and the onset of various diseases [13, 14, 29, 98–100].

The widespread use of artificial sweeteners has been linked to a rise in the prevalence of various diseases, posing a significant public health concern. As artificial sweeteners are mass-produced and broadly applied, their potential harm is unavoidable for many individuals. This issue demands urgent attention from researchers and healthcare professionals alike to mitigate associated health risks. This study aims to investigate AS-related targets and signaling pathways involved in carcinogenesis.

Based on commonly used AS in China, USA, and Europe, aspartame, acesulfame, Sucralose, NHDC, cyclamate, neotame, and saccharin, were selected for further investigation in this study. A total of 2666 AS-related

(See figure on next page.)

Fig. 11 Brain low-grade glioma prognosis targets screening, and a glioma cancer risk prognostic model constructing. **A** Univariate Cox regression analysis initially identified 19 prognostic genes (P < 0.05). **B** A coefficient profile plot was generated against the log(λ) values in the LASSO regression analysis model. The optimal parameter (λ) was marked with a dashed line on the left. **C** LASSO regression analysis of the 19 prognostic genes. **D** Multivariate Cox analysis identified 8 prognostic genes. **E** Kaplan–Meier survival analysis in high and low risk groups. **F** ROC curves of overall survival at 1, 3 and 5 years of the prognostic model. **G** The relationship between the survival status of brain low-grade glioma patients and the risk score of the prognostic model, along with changes in gene expression levels as the risk score increases. The dashed line delineates the boundary between the low-risk and high-risk groups
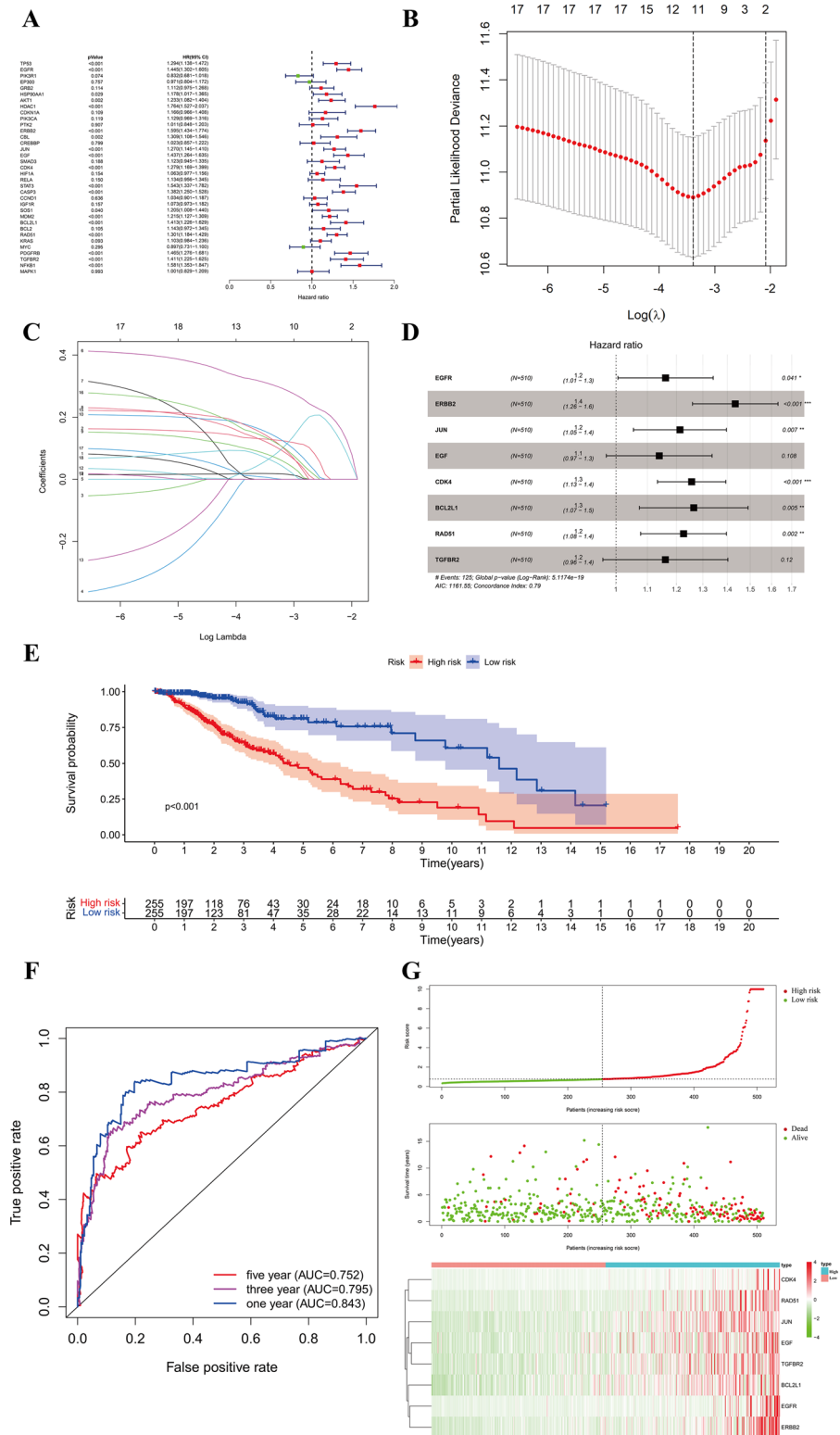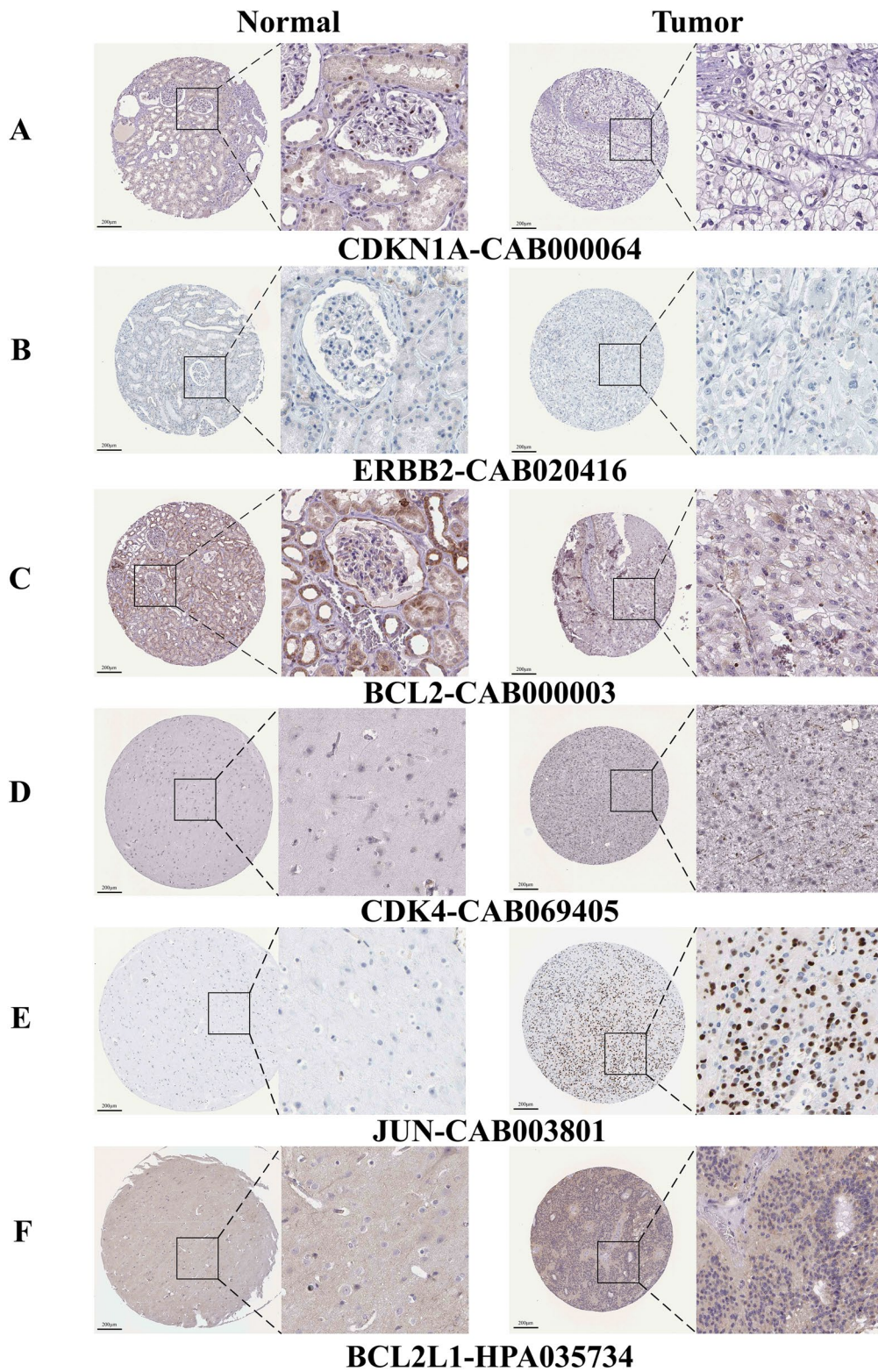
Fig. 11 (See legend on previous page.)

Fig. 12 Validation of the expression of core target proteins in KIRC and LGG, as well as in their respective normal tissues. **A**–**C** Immunohistochemical staining of CDKN1A, ERBB2, and BCL2 proteins in KIRC and normal renal tissues. **D**–**F** Immunohistochemical staining of CDK4, JUN, and BCL2L1 in LGG and normal brain tissue

targets were identified, alongside 1268 cancer-related targets, resulting in 368 AS-cancer intersection targets. GO enrichment analysis indicated strong associations of these 368 AS targets with cell proliferation, migration, apoptosis, and gene expression regulation. Furthermore, KEGG pathway analysis highlighted cancer-related pathways, which comprised half of the Top 20 enriched signaling pathways. The results highlight the critical role of AS in increasing cancer risk across various types.

After initial screening, 48 significant AS-cancer targets were identified. These targets were then intersected with those 325 genes found in the 'pathways in cancer (map05200)', resulting in 35 core AS-cancer targets. Subsequent research focused on these 35 targets, including TP53, EGFR, SRC, PIK3R1, AKT1, HSP90AA1, JUN, and CDKN1A, which are key genes in carcinogenesis. TP53 is one of the most well-known molecules in cancer biology, playing a key role in carcinogenesis, tumor angiogenesis, immune infiltration, metastasis, and cancer treatment across various cancer types [101–103]. The EGFR mutant is a key regulator in tumorigenesis and cancer treatment [104, 105]. SRC, PIK3R1, JUN, and AKT1 are key signaling molecules involved in various physiological processes in cancer cells [106–111]. HSP90AA1 has been identified as a significant target in cancer treatment [112–114].

A network illustrating the association between AS compounds and cancer targets was constructed to elucidate their relationship. GO-BP enrichment analysis indicated that these core targets are involved in cell proliferation, apoptosis, positive regulation of gene expression, and the insulin receptor signaling pathway. KEGG enrichment analysis revealed strong associations between these core targets and pathways related to various cancers, including prostate, pancreatic, colorectal cancers, glioma, and chronic myeloid leukemia.

Moreover, the 35 targets were strongly linked to pathways in cancer, endocrine resistance, and the PI3K-Akt and FoxO signaling pathways. These cancer pathways are integral to nearly every physiological process in the tumor cell lifecycle [115–117]. Endocrine resistance plays a pivotal role in the treatment of various cancers [118–120]. The PI3K-Akt signaling pathway is closely associated with tumorigenesis, cell growth and proliferation, invasion, metastasis, apoptosis, and autophagy [121–123]. FoxO collaborates with the PI3K-Akt pathway, playing crucial roles in cancer inhibition and therapy [124–126].

To further explore the interactions between the 35 core targets and the 7 commonly used AS, molecular docking was employed to validate their association. Among the seven AS, Aspartame, NHDC, and Neotame demonstrated strong affinity for the core targets, whereas Sucralose and Acesulfame showed weaker binding. Eight

molecular docking results were selected from a total of 245 docking outcomes for presentation: CCND1-Aspartame, MAPK1-Acesulfame, MAPK1-NHDC, CASP3-Sucralose, CDK4-Cyclamate, KRAS-Saccharin, KRAS-Neotame, and BCL2L1-NHDC. Analysis of both 2D and 3D docking results indicated a substantial association between the AS and the core targets.

CCND1 is implicated in disease pathogenesis, specifically through chromosomal aberration such as translocation t(11;14)(q13;q32), which may lead to B-lymphocytic malignancies like mantle-cell lymphoma, multiple myeloma, and parathyroid adenomas [127–129]. Activation of CCND1 potentially contributes to oncogenesis by modulating the G1/S transition in the cell cycle [130, 131].

MAPK1 serves as a critical component of the MAP kinase signaling transduction pathway, playing a pivotal role in the MAPK/ERK cascade that regulates cell growth, survival, adhesion, and differentiation through the control of transcription, translation, and cytoskeletal rearrangements [132, 133]. CASP3 acts as a significant effector caspase involved in the execution phase of apoptosis [134, 135]. Additionally, CASP3 cleaves and activates several proteins, including caspase-6, −7, and −9, interleukin-18, huntingtin, AKT1, and gasdermin-E, among others [135–138].

CDK4 forms complexes that phosphorylate and inhibit RB1, thereby regulating the G1/S transition in the cell cycle, sharing functional similarities with CCND1 [139]. KRAS is the most well-known oncogene, characterized by a high mutation rate across various cancers and linked to several lethal cancers and clinical syndromes [140–142]. BCL2L1 functions as a potent inhibitor of cell death by blocking activation of caspases, regulating cell death through inhibition of the voltage-dependent anion channel (VDAC), and acting as a regulator of the G2 checkpoint [143, 144].

Molecular dynamics simulation is a powerful technique for assessing molecule docking results and simulating the binding process between small molecules and proteins. Molecular dynamics simulations were conducted to confirm the association between the AS compounds and core targets. Significant parameters including RMSD, RMSF, and Rg were measured to assess the stability of these interactions. Following the simulation process, complexes such as MAPK1-NHDC, CDK4-Cyclamate, KRAS-Saccharin, MAPK1-Acesulfame, and KRAS-Neotame exhibited good stability and protein structure flexibility. However, BCL2-Aspartame and CASP3-Sucralose displayed fluctuations within the 200 ns simulation period and did not achieve stability. Molecular dynamics simulations validated the molecular docking results and confirmed the interactions between AS compounds

and the core targets. AS elevated cancer risk by binding to key targets, thereby disrupting signal transduction and inducing carcinogenesis across various cancer types. Different AS molecules bind to distinct targets and are involved in various signaling pathways. The molecular mechanisms underlying carcinogenesis also differ.

The expression profiles of 35 AS-cancer core targets were presented in a heatmap to clearly illustrate the differences in expression levels across 33 types of cancer. HSP90AA1 exhibited high expression across nearly all 33 cancers, whereas EGF showed the opposite pattern. Although the expression of HSP90AA1 and EGF significantly differed, both played crucial roles in various cancer types by participating in essential cell signaling pathways related to carcinogenesis, invasion, metastasis, and cancer treatment [145–147].

The expression differences among the 35 core targets were calculated across 33 cancer types and illustrated in a heatmap. Significant variations in the expression of these core targets were observed in low grade brain glioma (LGG) and kidney renal clear cell carcinoma (KIRC), indicating a strong association of AS with these specific cancer types.

The hazard ratio (95% CI) of the 35 core targets across 33 types of cancer was calculated. CDK4 shows a high correlation with KICH, THCA, and PRAD, while CASP3 is strongly associated with TGCT, ACC, and LGG. Previous studies have emphasized the critical roles of CDK4 in kidney cancer [148], thyroid carcinoma [149], and prostate cancer [150, 151]. CASP3 has been shown to be strongly associated with prostate cancer [152], bladder cancer [153], gastric cancer [154], lung cancer [154], and others [155]. These publications strongly support our findings.

After normalizing gene expression data, differential expression analysis of GSE53757 (KIRC), GSE21354 (LGG), GSE42568 (BRCA), and GSE46602 (PRAD) identified up- or down-regulated AS-cancer core targets. These findings elucidate the correlation between differentially expressed genes (DEGs) across various cancers and AS-cancer core targets, thereby substantiating AS's role in carcinogenesis in KIRC, LGG, BRCA, and PRAD. GSEA analysis revealed the pathways associated with DEGs in these cancers, which play pivotal roles in cancer development.

After integrating univariate Cox regression, LASSO regression, and multivariate Cox regression analyses, CDKN1A, ERBB2, CASP3, BCL2, and TGFBR2 were selected to construct the prognostic model for KIRC. CDKN1A emerged as a critical predictor of poor outcomes in patients with chromophobe renal cell carcinoma [156]. Aberrant ERBB2 expression has been shown to correlate with renal cancers [157, 158], while BCL2 is

recognized as a significant therapeutic target for renal cancer treatment [159, 160]. Regulation of TGFBR2 expression represents an important strategy in clear cell renal cell carcinoma (ccRCC) therapy [161]. As the risk score (RS) increased, CASP3 expression levels rose, while CDKN1A, ERBB2, BCL2, and TGFBR2 levels decreased. Additionally, higher RS scores correlated with shorter patient survival times and increased mortality in renal carcinoma.

Moreover, as the RS increased, the expression levels of all 8 genes (EGFR, ERBB2, JUN, EGF, CDK4, BCL2L1, RAD51, TGFBR2) also increased. Higher RS scores were linked to shorter patient survival times and increased mortality in brain lower grade glioma. As RS increased, HSP90AA1, PIK3CA, and RAD51 gene expression levels rose, leading to decreased survival times and increased mortality in breast cancer. With increasing RS, the expression levels of GRB2 and RAD51 increased. Furthermore, higher RS scores were associated with shorter patient survival times and increased mortality in prostate cancer.

To validate the significant roles of AS-cancer core targets in the progression of KIRC, LGG, BRCA, and PRAD, the HPA database was utilized to provide immunohistochemical data from clinical cancer samples. Specific antibodies targeting AS-cancer key proteins were used for staining tumor and normal tissue sections. CDKN1A, ERBB2, and BCL2 were found to play critical roles in KIRC progression. Additionally, CDK4, JUN, and BCL2L1 were significantly overexpressed in LGG tissues. High expression levels of PIK3CA, HSP90AA1, and RAD51 were detected in BRCA tissues, while elevated expression of GRB2 and RAD51 was observed in PRAD tissues.

Through an integrative analysis encompassing data mining, machine learning, network toxicology, molecular docking, molecular dynamics simulations, and HPA validation, we concluded that commonly used artificial sweetener (AS) molecules were correlated with 35 core targets. These targets regulate the PI3K-Akt, endocrine resistance, and FoxO signaling pathways, thereby increasing the risk of cancer in kidney cancer (KIRC), low-grade glioma (LGG), breast cancer (BRCA), and prostate cancer (PRAD).

However, this study has several limitations: 1. The findings are primarily derived from data mining, machine learning, bioinformatics analysis, and computational methods, lacking direct experimental evidence to substantiate the link between artificial sweeteners and increased cancer risk. 2. The results may be impacted by issues related to data quality, source variability, and algorithmic biases. 3. Potential deviations in results due to unidentified influencing factors. Although this study

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 26 of 31

does not include direct in vivo experimental results, it still provides a valuable methodological framework for assessing the safety of food additives, offering significant guidance for future research.

Based on our findings, we strongly advise the public to reduce their intake of artificial sweeteners to alleviate common health issues such as cardiovascular disease, diabetes, and irritable bowel syndrome. Additionally, we urge food manufacturers to limit the use of artificial sweeteners and other additives, prioritizing public health over financial gains. Finally, we call on government bodies and relevant agencies to conduct thorough evaluations of food additive safety and to develop policies that guide the public towards healthier dietary choices and greater health awareness.

Moving forward, our research will focus on: 1. Investigating the specific impact of commonly used artificial sweetener molecules on cancer risk. 2. Enhancing our findings with physiological and biochemical data from cellular and animal studies, pathological images, and transcriptomic and proteomic sequencing to provide direct evidence of the carcinogenic potential of artificial sweeteners. 3. Incorporating global cohort studies related to artificial sweetener usage to enrich public health data. 4. Gathering data from clinical trials, which, despite being challenging, could furnish direct experimental proof that artificial sweeteners contribute to cancer development.

Conclusion

In this study, we employed data mining, machine learning, network toxicology, molecular docking, molecular dynamics simulations, and clinical sample staining to explore the relationship between artificial sweeteners (AS) and cancer. Our findings indicate that AS may potentially increase the incidence of kidney cancer, low-grade glioma, breast cancer, and prostate cancer through multiple targets and signaling pathways.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-024-06047-0.

Supplementary material 1: Supplemental Figure 1. Artificial sweeteners-cancer targets analysis.Protein-protein interactionnetworks of AS-cancer targets. In the PPI network, light blue and dark purple lines represent known interactions, while green, red, and dark blue lines represent predicted interactions. Yellow, black, and light purple lines indicate correlations based on text mining, co-expression, and protein homology. The 3D structure within each node is either known or predicted.Visualization of the PPI network for AS-cancer targets. In the PPI visualization, the core targets are ranked by their degree values. The darker the color and the larger the circle area, the higher the degree value, indicating stronger reciprocity among gene targets.Triadic histogram depicting biological processes, molecular functions, and cellular componentsassociated with AS-cancer targets.Bubble maps illustrating the top 20 pathways from KEGG enrichment analysis of AS-cancer targets. Supplemental Figure 2. Molecular

dynamics simulation.Root mean square deviation, root mean square fluctuation, and radius of gyrationof MAPK1-Acesulfame.RMSD, RMSF, and RG of KRAS-Neotame.RMSD, RMSF, and RG of KRAS-Saccharin.RMSD, RMSF, and RG of BCL2-Aspartame.RMSD, RMSF, and RG of CASP3-Sucralose. Supplemental Figure 3. Normalization of the four cancer datasets from GEO database.KIRC, LGG, BRCA, and PRAD untreated and normalized datasets. Supplemental Figure 4. Differentially expressed genesscreening, target-pathway diagram and GSEA pathway enrichment analysis in breast invasive carcinomaand prostate adenocarcinoma.Volcano plots depicting DEGs in a breast cancer datasetreveal significant up-regulation of HDAC1, PTK2, and IGF1R genes in BRCA patients, while PIK3CA, TGFBR2, and JUN genes show significant down-regulation in these patients.Target-pathway diagrams in AS and BRCA. In these diagrams, red indicates genes significantly upregulated, blue denotes significantly down-regulated genes, and green represents genes with insignificant changes. Shapes correspond to different targets: triangles denote sweetener targets, squares denote disease targets, and circles denote targets common to both sweeteners and diseases.The top 5 GSEA enrichment pathways up-regulated in BRCA.The top 5 GSEA enrichment pathways down-regulated in BRCA. Volcano plot of DEGs in a prostate adenocarcinomashows significant up-regulation of the MYC gene in PRAD patients, while the BCL2 gene shows significant down-regulation in these patients.Target-pathway diagrams in AS and PRAD.The top 5 GSEA enrichment pathways up-regulated in PRAD.The top 5 GSEA enrichment pathways down-regulated in PRAD. Supplemental Figure 5. Breast cancer prognosis targets screening, and the cancer risk prognostic model constructing.Univariate Cox regression analysis initially identified 6 prognostic genes.A coefficient profile plot was generated against the logvalues in the LASSO regression analysis model. The optimal parameterwas marked with a dashed line on the left.LASSO regression analysis of the 6 prognostic genes.Multivariate Cox analysis identified 3 prognostic genes.Kaplan-Meier survival analysis in high and low risk groups.ROC curves of overall survival at 1, 3 and 5 years of the prognostic model.The relationship between the survival status of breast cancer patients and the risk score of the prognostic model, along with changes in gene expression levels as the risk score increases. The dashed line delineates the boundary between the low-risk and high-risk groups. Supplemental Figure 6. Prostate cancer prognosis targets screening, and the cancer risk prognostic model constructing.Univariate Cox regression analysis initially identified 2 prognostic genes.A coefficient profile plot was generated against the logvalues in the LASSO regression analysis model. The optimal parameterwas marked with a dashed line on the left. LASSO regression analysis of the 2 prognostic genes.Multivariate Cox analysis identified 2 prognostic genes.Kaplan-Meier survival analysis in high and low risk groups.ROC curves of overall survival at 1, 3 and 5 years of the prognostic model.The relationship between the survival status of prostate cancer patients and the risk score of the prognostic model, along with changes in gene expression levels as the risk score increases. The dashed line delineates the boundary between the low-risk and high-risk groups. Supplemental Figure 7. Validation of the expression of core target proteins in BRCA and PRAD, as well as in their respective normal tissues. Immunohistochemical staining of PIK3CA, HSP90AA1, and RAD51 proteins in BRCA and normal breast tissues.Immunohistochemical staining of GRB2 and RAD51 in PRAD and normal prostate tissue

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 27 of 31

Data availability

The processed and raw data are freely serviced from the first and corresponding authors.

Declarations

Ethics approval and consent to participate

The public data we used to validate our key genes were sourced from the TCGA, GEO, and HPA databases, as cited in our paper. This data was utilized solely for research purposes.

Consent for publication

The authors have all agreed to publication.

Competing interests

The authors declare no competing interests.

Author details

[1]Hubei Key Laboratory of Renal Disease Occurrence and Intervention, Medical School, Hubei Polytechnic University, Guilin North Road No 16, Huangshi 435003, Hubei, People's Republic of China. [2]Medical College of YiChun University, Xuefu Road No 576, Yichun 336000, Jiangxi, People's Republic of China.

References

1. What Are Artificial Sweeteners? [https://health.usnews.com/wellness/food/articles/a-guide-to-artificial-sweeteners-the-next-generation].
2. Kmietowicz Z. Artificial sweeteners do not aid weight loss or improve health, says WHO. BMJ. 2023;381:1112.
3. Guidelines Review Committee NaFSN. Use of non-sugar sweeteners: WHO guideline. Geneva: World Health Organization; 2023.
4. Wilk K, Korytek W, Pelczyńska M, Moszak M, Bogdański P. the effect of artificial sweeteners use on sweet taste perception and weight loss efficacy: a review. Nutrients. 2022;14:1261.
5. Debras C, Deschasaux-Tanguy M, Chazelas E, Sellem L, Druesne-Pecollo N, Esseddik Y, Szabo de Edelenyi F, Agaësse C, De Sa A, Lutchia R, et al. Artificial sweeteners and risk of type 2 diabetes in the prospective nutrinet-santé cohort. Diabetes Care. 2023;46:1681–90.
6. Daoust L. Artificial sweeteners and type 2 diabetes. Nat Food. 2023;4:739.
7. Basson AR, Rodriguez-Palacios A, Cominelli F. Artificial sweeteners: history and new concepts on inflammation. Front Nutr. 2021;8: 746247.
8. Martkets RA. Artificial and natural sweetener market forecasts from 2023 to 2028. Res Martkets. 2023;2023:207.
9. Samaniego-Vaesken ML, Ruiz E, Partearroyo T, Aranceta-Bartrina J, Gil Á, González-Gross M, Ortega RM, Serra-Majem L, Varela-Moreiras G. Added sugars and low- and no-calorie sweeteners in a representative sample of food products consumed by the Spanish ANIBES study population. Nutrients. 2018;10:4.
10. Redruello-Requejo M, González-Rodríguez M, Samaniego-Vaesken Mª DL, Montero-Bravo A, Partearroyo T, Varela-Moreiras G. Low- and no-calorie sweetener (LNCS) consumption patterns amongst the Spanish adult population. Nutrients. 1845;2021:13.
11. High-Intensity Sweeteners [https://www.fda.gov/food/food-additives-petitions/high-intensity-sweeteners]
12. Rios-Leyvraz MMJ. Health effects of the use of non-sugar sweeteners: a systematic review and meta-analysis. Geneva: World Health Organization; 2022.
13. Zhang Y, Tang Z, Shi Y, Li L. Associations between artificial sweetener intake from cereals, coffee, and tea and the risk of type 2 diabetes mellitus: a genetic correlation, mediation, and mendelian randomization analysis. PLoS ONE. 2024;19: e0287496.
14. Pan H, Feng C, Zhou Z, Huang J, Deng J, Zhou Y, Wang Y, Mu X, Wang Q, Wang K, Lu Z. The causal association between artificial sweeteners and the risk of cancer: a Mendelian randomization study. Food Funct. 2024;15:4527–37.
15. Palomar-Cros A, Straif K, Romaguera D, Aragonés N, Castaño-Vinyals G, Martin V, Moreno V, Gómez-Acebo I, Guevara M, Aizpurua A, et al. Consumption of aspartame and other artificial sweeteners and risk of cancer in the Spanish multicase-control study (MCC-Spain). Int J Cancer. 2023;153:979–93.
16. Alharthi SAS, Alaisayi KHA, Alalawi LYS, Alamri ROS, Abu-Elfotuh K, Alenazi TS, Amirthalingam P, Albariqi HAH, Mohammed AA, Alsubayti N, et al. The consumption pattern and perception of using artificial sweeteners among the public in Tabuk region, Saudi Arabia. Front Public Health. 2023;11:1166868.
17. WH Organization. Guideline: sugars intake for adults and children. Geneva: World Health Organization; 2015.
18. Debras C, Chazelas E, Sellem L, Porcher R, Druesne-Pecollo N, Esseddik Y, de Edelenyi FS, Agaësse C, De Sa A, Lutchia R, et al. Artificial sweeteners and risk of cardiovascular diseases: results from the prospective NutriNet-Santé cohort. BMJ. 2022;378: e071204.
19. Debras C, Chazelas E, Srour B, Druesne-Pecollo N, Esseddik Y, Szabo de Edelenyi F, Agaësse C, De Sa A, Lutchia R, Gigandet S, et al. Artificial sweeteners and cancer risk: Results from the NutriNet-Santé population-based cohort study. PLoS Med. 2022;19:e1003950.
20. Iizuka K. Is the use of artificial sweeteners beneficial for patients with diabetes mellitus? The advantages and disadvantages of artificial sweeteners. Nutrients. 2022;14:40.
21. Kränkel N, Rauch-Kroehnert U. Artificial sweetener sucralose: a possible modulator of autoimmune diseases. Signal Transduct Target Ther. 2023;8:377.
22. Basson AR, Katz J, Singh S, Celio F, Cominelli F, Rodriguez-Palacios A. Sweets and inflammatory bowel disease: patients favor artificial sweeteners and diet foods/drinks over table sugar and consume less fruits/vegetables. Inflamm Bowel Dis. 2023;29:1751–9.
23. Qin X. The possible link between artificial sweeteners such as saccharin and sucralose and inflammatory bowel disease deserves further study. Inflamm Bowel Dis. 2016;22:E17.
24. Suez J, Korem T, Zeevi D, Zilberman-Schapira G, Thaiss CA, Maza O, Israeli D, Zmora N, Gilad S, Weinberger A, et al. Artificial sweeteners induce glucose intolerance by altering the gut microbiota. Nature. 2014;514:181–6.
25. Greenhill C. Gut microbiota: not so sweet–artificial sweeteners can cause glucose intolerance by affecting the gut microbiota. Nat Rev Endocrinol. 2014;10:637.
26. Cai S, Xie N, Zheng L, Li Q, Zhang S, Huang Q, Luo W, Wu M, Wang Y, Du Y, et al. Sweet taste receptors play roles in artificial sweetener-induced enhanced urine output in mice. NPJ Sci Food. 2024;8:2.
27. Shil A, Zhang J, Chichger H. Investigating the use and awareness of artificial sweeteners among diabetic patients in Bangladesh. PLoS ONE. 2023;18: e0295272.
28. Girigosavi KB, Etta I, Kambham S, Panjiyar BK. Sweet surprises: an in-depth systematic review of artificial sweeteners and their association with cerebrovascular accidents. Curr Nutr Rep. 2024;10:1.
29. Witkowski M, Nemet I, Alamri H, Wilcox J, Gupta N, Nimer N, Haghikia A, Li XS, Wu Y, Saha PP, et al. The artificial sweetener erythritol and cardiovascular event risk. Nat Med. 2023;29:710–8.
30. Wiklund AE, Guo X, Gorokhova E. Cardiotoxic and neurobehavioral effects of sucralose and acesulfame in Daphnia: toward understanding ecological impacts of artificial sweeteners. Comp Biochem Physiol C Toxicol Pharmacol. 2023;273: 109733.
31. Tagorti G, Yalçın B, Güneş M, Burgazlı AY, Kaya B. Comparative evaluation of natural and artificial sweeteners from DNA damage, oxidative stress, apoptosis, to development using Drosophila melanogaster. Drug Chem Toxicol. 2023;47:1–12.
32. Patil S, Jalal RAS, Albar DH, Bansal SJ, Patil S, Nagaral S, Finch J, Bernard CA, Baeshen HA, Awan KH. Intake of artificial sweeteners by children: boon or bane? J Contemp Dent Pract. 2023;24:137–45.
33. Pearlman M, Obert J, Casey L. The association between artificial sweeteners and obesity. Curr Gastroenterol Rep. 2017;19:64.
34. Han S, Yang J, Park JE, Kim JH. Artificial sweeteners and pancreatic cancer: is aspartame a culprit or a coincidence? Pak J Med Sci. 2024;40:790.

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 28 of 31

35. Balint IB, Erdodi BT. Is there a promoting role for artificial sweeteners in the evolution of bladder cancer? A meta-analysis of current literature. Minerva Surg. 2024;79:92–9.

36. Hoover RN, Strasser PH. Artificial sweeteners and human bladder cancer. Preliminary results. Lancet. 1980;1:837–40.

37. Ozonoff D. Artificial sweeteners and bladder cancer. N Engl J Med. 1980;303:341–2.

38. Soffritti M, Belpoggi F, Tibaldi E, Esposti DD, Lauriola M. Life-span exposure to low doses of aspartame beginning during prenatal life increases cancer effects in rats. Environ Health Perspect. 2007;115:1293–7.

39. Soffritti M, Belpoggi F, Degli Esposti D, Lambertini L, Tibaldi E, Rigano A. First experimental demonstration of the multipotential carcinogenic effects of aspartame administered in the feed to Sprague-Dawley rats. Environ Health Perspect. 2006;114:379–85.

40. Gao Y, Yin L, Zhang Y, Li X, Liu L. Associations of saccharin intake with all-cause, cardiovascular and cancer mortality risk in USA adults. Br J Nutr. 2024;132:1–9.

41. Li X, Liu Y, Wang Y, Li X, Liu X, Guo M, Tan Y, Qin X, Wang X, Jiang M. Sucralose promotes colitis-associated colorectal cancer risk in a murine model along with changes in microbiota. Front Oncol. 2020;10:710.

42. Mela DJ, McLaughlin J, Rogers PJ. Perspective: standards for research and reporting on low-energy ("artificial") sweeteners. Adv Nutr. 2020;11:484–91.

43. Davis AP, Wiegers TC, Wiegers J, Wyatt B, Johnson RJ, Sciaky D, Barkalow F, Strong M, Planchart A, Mattingly CJ. CTD tetramers: a new online tool that computationally links curated chemicals, genes, phenotypes, and diseases to inform molecular mechanisms for environmental health. Toxicol Sci. 2023;195:155–68.

44. Szklarczyk D, Santos A, von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. Nucleic Acids Res. 2016;44:D380-384.

45. Gallo K, Goede A, Preissner R, Gohlke BO. SuperPred 3.0: drug classification and target prediction-a machine learning approach. Nucleic Acids Res. 2022;50:W726-w731.

46. Daina A, Michielin O, Zoete V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. Nucleic Acids Res. 2019;47:W357-w364.

47. Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. Nat Biotechnol. 2007;25:197–206.

48. Wang X, Shen Y, Wang S, Li S, Zhang W, Liu X, Lai L, Pei J, Li H. PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. Nucleic Acids Res. 2017;45:W356-w360.

49. Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res. 2012;40:W294-297.

50. Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Stein TI, Nudel R, Lieder I, Mazor Y, et al. The GeneCards suite: from gene data mining to disease genome sequence analyses. Curr Protoc Bioinf. 2016;54:1.

51. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. 2005;33:D514-517.

52. Zhou Y, Zhang Y, Lian X, Li F, Wang C, Zhu F, Qiu J, Chen Y. Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents. Nucleic Acids Res. 2021. https://doi.org/10.1093/nar/gkab953.

53. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021;49:D605-d612.

54. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.

55. Xie J, Chen R, Wang Q, Mao H. Exploration and validation of Taraxacum mongolicum anti-cancer effect. Comput Biol Med. 2022;148: 105819.

56. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC, Lempicki RA. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res. 2007;35:W169-175.

57. Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuermann M, Gaudet P, Harris NL, Hill DP, et al. The gene ontology knowledgebase in 2023. Genetics. 2023;224:1.

58. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic Acids Res. 2000;28:235–42.

59. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455–61.

60. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem. 2009;30:2785–91.

61. Schöning-Stierand K, Diedrich K, Fährrolfes R, Flachsenberg F, Meyder A, Nittinger E, Steinegger R, Rarey M. ProteinsPlus: interactive analysis of protein-ligand binding interfaces. Nucleic Acids Res. 2020;48:W48-w53.

62. Shamsol Azman ANS, Tan JJ, Abdullah MNH, Bahari H, Lim V, Yong YK. Network pharmacology and molecular docking analysis of active compounds in Tualang Honey against atherosclerosis. Foods. 2023;12:35.

63. Arshia AH, Shadravan S, Solhjoo A, Sakhteman A, Sami A. De novo design of novel protease inhibitor candidates in the treatment of SARS-CoV-2 using deep learning, docking, and molecular dynamic simulations. Comput Biol Med. 2021;139: 104967.

64. He X, Man VH, Yang W, Lee TS, Wang J. A fast and high-quality charge model for the next generation general AMBER force field. J Chem Phys. 2020;153: 114502.

65. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput. 2015;11:3696–713.

66. Fu T, Zheng G, Tu G, Yang F, Chen Y, Yao X, Li X, Xue W, Zhu F. Exploring the binding mechanism of metabotropic glutamate receptor 5 negative allosteric modulators in clinical trials by molecular dynamics simulations. ACS Chem Neurosci. 2018;9:1492–502.

67. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45:W98-w102.

68. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, et al. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41:D991-995.

69. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43: e47.

70. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102:15545–50.

71. Tay JK, Narasimhan B, Hastie T. Elastic net regularization paths for all generalized linear models. J Stat Softw. 2023;106:36.

72. Wang H, Lengerich BJ, Aragam B, Xing EP. Precision Lasso: accounting for correlations and linear dependencies in high-dimensional genomic data. Bioinformatics. 2019;35:1181–7.

73. May WL. Kaplan-Meier survival analysis. In: Schwab M, editor. Encyclopedia of cancer. Berlin: Springer, Berlin Heidelberg; 2009. p. 1590–3.

74. Lorent M, Giral M, Foucher Y. Net time-dependent ROC curves: a solution for evaluating the accuracy of a marker to predict disease-related mortality. Stat Med. 2014;33:2379–89.

75. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol. 2022;75:25–36.

76. Abualigah L, AlNajdawi S, Ikotun MA, Forestiero A, Gul F, Ezugwu AE, Jia H, Zare M, Mahajan S, Alshinwan M. 11—quantum approximate optimization algorithm: a review study and problems. In: Abualigah L, editor. Metaheuristic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 147–65.

77. Abualigah L, Al Turk AA, Ikotun AM, Zitar RA, Alsoud AR, Khodadadi N, Hussien AG, Jia H. 2—social spider optimization algorithm: survey and new applications. In: Abualigah L, editor. Metaheuristic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 15–31.

78. Abualigah L, Alshatti SM, Ikotun AM, Zitar RA, Alsoud AR, Khodadadi N, Ezugwu AE, Hanandeh ES, Jia H, Zare M. 9—spider monkey optimiza-tions: application review and results. In: Abualigah L, editor. Metaheuris-tic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 117–31.

79. Abualigah L, Odah S, Ikotun MA, Alsoud AR, Forestiero A, Ezugwu AE, Hanandeh ES, Jia H, Zare M. 10—marine predator's algorithm: a survey of recent applications. In: Abualigah L, editor. Metaheuristic optimiza-tion algorithms. Burlington: Morgan Kaufmann; 2024. p. 133–45.

80. Abualigah L, Ra A, Ikotun AM, Zitar RA, Alsoud AR, Khodadadi N, Ezugwu AE, Hanandeh ES, Jia H. 8 - Whale optimization algorithm: analysis and full survey. In: Abualigah L, editor. Metaheuristic Optimiza-tion Algorithms. Burlington: Morgan Kaufmann; 2024. p. 105–15.

81. Abualigah L, Abusaleem A, Ikotun AM, Zitar RA, Alsoud AR, Khoda-dadi N, Ezugwu AE, Hanandeh ES, Jia H. 6—arithmetic optimization algorithm: a review and analysis. In: Abualigah L, editor. Metaheuristic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 73–87.

82. Abualigah L, Al-Abadi L, Ikotun AM, Al-Saqqar F, Izci D, Zhang P, Şahin CB, El-Bashir M, Sumari P. 15—a review of mothflame optimization algorithm: analysis and applications. In: Abualigah L, editor. Metaheuris-tic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 205–19.

83. Abualigah L, Al-Hilo G, Raza A, Ezugwu AE, Nasar MRA, Mughaid A, AlZu'bi S, Aldiabat K, Al-diabat M. 13—a review of Henry gas solubil-ity optimization algorithm: a robust optimizer and applications. In: Abualigah L, editor. Metaheuristic Optimization Algorithms. Burlington: Morgan Kaufmann; 2024. p. 177–92.

84. Abualigah L, Sheikhan A, Ikotun MA, Zitar RA, Alsoud AR, Al-Shourbaji I, Hussien AG, Jia H. 1—particle swarm optimization algorithm: review and applications. In: Abualigah L, editor. Metaheuristic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 1–14.

85. Abualigah L, Ababneh A, Ikotun AM, Zitar RA, Alsoud AR, Khodadadi N, Ezugwu AE, Hanandeh ES, Jia H. 4—a survey of cuckoo search algorithm: optimizer and new applications. In: Abualigah L, editor. Metaheuristic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 45–57.

86. Abualigah L, Sbenaty B, Ikotun AM, Zitar RA, Alsoud AR, Khodadadi N, Ezugwu AE, Hanandeh ES, Jia H. 7—aquila optimizer: review, results and applications. In: Abualigah L, editor. Metaheuristic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 89–103.

87. Abualigah L, Ahmad EN, Ikotun AM, Zitar RA, Alsoud AR, Khodadadi N, Ezugwu AE, Jia H. 3—animal migration optimization algorithm: novel optimizer, analysis, and applications. In: Abualigah L, editor. Metaheuris-tic Optimization Algorithms. Burlington: Morgan Kaufmann; 2024. p. 33–43.

88. Abualigah L, Hawamdeh W, Zitar RA, AlZu'bi S, Mughaid A, Hanandeh ES, Alsoud AR, El-kenawy E-SM. 18—Salp swarm algorithm: survey, analysis, and new applications. In: Abualigah L, editor. Metaheuristic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 241–58.

89. Abualigah L, Faweer S, Raza A, Gul F, Ezugwu AE, Alshinwan M, Nasar MRA, Mughaid A, AlZu'bi S. 12—crow search algorithm: a survey of novel optimizer and its recent applications. In: Abualigah L, editor. Metaheuristic optimization algorithms. Burlington: Morgan Kaufmann; 2024. p. 167–76.

90. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. Proteomics tissue-based map of the human proteome. Science. 2015;347:1260419.

91. Yu X, Yu F, Li Z, Shi T, Xia Z, Li G. Occurrence, distribution, and ecological risk assessment of artificial sweeteners in surface and ground waters of the middle and lower reaches of the Yellow River (Henan section, China). Environ Sci Pollut Res Int. 2023;30:52609–23.

92. Li D, Zheng Q, Thomas KV, Dang AK, Binh VN, Anh NTK, Thai PK. Use of artificial sweeteners and caffeine in a population of Hanoi: an assess-ment by wastewater-based epidemiology. Sci Total Environ. 2023;868: 161515.

93. Sultana T, Metcalfe CD. Calibration and field validation of POCIS passive samplers for tracking artificial sweeteners as indicators of municipal wastewater contamination in surface waters. Environ Monit Assess. 2022;194:564.

94. Tran NH, Hu J, Li J, Ong SL. Suitability of artificial sweeteners as indica-tors of raw wastewater contamination in surface water and groundwa-ter. Water Res. 2014;48:443–56.

95. Shil A, Chichger H. Artificial sweeteners negatively regulate pathogenic characteristics of two model gut bacteria, E. coli and E. faecalis. Int J Mol Sci. 2021;22:5228.

96. Yu Z, Henderson IR, Guo J. Non-caloric artificial sweeteners modulate conjugative transfer of multi-drug resistance plasmid in the gut micro-biota. Gut Microbes. 2023;15:2157698.

97. Hosseini A, Barlow GM, Leite G, Rashid M, Parodi G, Wang J, Morales W, Weitsman S, Rezaie A, Pimentel M, Mathur R. Consuming artificial sweeteners may alter the structure and function of duodenal microbial communities. iScience. 2023;26:108530.

98. Zhu C, Ji D, Ma J, Da M. Association between artificial sweeteners-aspartame consumption and colorectal cancer risk: evidence-based strategies. Oncology. 2023. https://doi.org/10.1159/000534812.

99. Ye X, Zhang Y, He Y, Sheng M, Huang J, Lou W. Association between consumption of artificial sweeteners and breast cancer risk: a system-atic review and meta-analysis of observational studies. Nutr Cancer. 2023;75:795–804.

100. van Eyk AD. The effect of five artificial sweeteners on Caco-2, HT-29 and HEK-293 cells. Drug Chem Toxicol. 2015;38:318–27.

101. Yoon AR, Lee S, Kim JH, Park Y, Koo T, Yun CO. CRISPR-mediated ablation of TP53 and EGFR mutations enhances gefitinib sensitivity and anti-tumor efficacy in lung cancer. Mol Ther. 2024;32:3618–28.

102. Yoo YA, Quan S, Yang W, Guo Q, Rodríguez Y, Chalmers ZR, Dufficy MF, Lackie B, Sagar V, Unno K, et al. Asparagine dependency is a targetable metabolic vulnerability in TP53-altered castration-resistant prostate cancer. Cancer Res. 2024;84:3004–22.

103. Cederquist GY, Boe LA, Walsh M, Freedman GM, Maxwell KN, Taunk N, Braunstein LZ. Radiation-associated secondary cancer in patients with breast cancer harboring TP53 germline variants. JAMA Oncol. 2024;10:1588–90.

104. Wang S, Lai JC, Li Y, Tang C, Lu J, Han M, Ye X, Jia L, Cui W, Yang J, et al. Loss of CDKN2A enhances the efficacy of immunotherapy in EGFR mutant non-small cell lung cancer. Cancer Res. 2024. https://doi.org/10.1158/0008-5472.CAN-24-1817.

105. Voronkova MA, Johnson B, Gandhi N, Koomen JM, Patrick M, Shanthi Bhupathi S, Wu VM, Elliott A, Vanderwalde A, Halmos B, Hazlehurst LA. ERO1A levels are a prognostic indicator in EGFR mutated non small cell lung cancer. NPJ Precis Oncol. 2024;8:250.

106. Dent RA, Kim SB, Oliveira M, Barrios C, O'Shaughnessy J, Isakoff SJ, Saji S, Freitas-Junior R, Philco M, Bondarenko I, et al. Ipatasertib plus paclitaxel for patients with PIK3CA/AKT1/PTEN-altered locally advanced unresect-able or metastatic triple-negative breast cancer in the IPATunity130 phase III trial. Clin Cancer Res. 2024;30:4329–38.

107. Huang TT, Chiang CY, Nair JR, Wilson KM, Cheng K, Lee JM. AKT1 inter-acts with DHX9 to mitigate R loop-induced replication stress in ovarian cancer. Cancer Res. 2024;84:887–904.

108. Egeland EV, Seip K, Skourti E, Øy GF, Pettersen SJ, Pandya AD, Dahle MA, Haugen MH, Kristian A, Nakken S, et al. The SRC-family serves as a thera-peutic target in triple negative breast cancer with acquired resistance to chemotherapy. Br J Cancer. 2024;131:1656–67.

109. Gupta I, Gaykalova DA. Unveiling the role of PIK3R1 in cancer: a com-prehensive review of regulatory signaling and therapeutic implications. Semin Cancer Biol. 2024;106–107:58–86.

110. Li H, Lin R, Zhang Y, Zhu Y, Huang S, Lan J, Lu N, Xie C, He S, Zhang W. N6-methyladenosine-modified circPLPP4 sustains cisplatin resistance in ovarian cancer cells via PIK3R1 upregulation. Mol Cancer. 2024;23:5.

111. Huang Z, Li Y, Qian Y, Zhai E, Zhao Z, Zhang T, Liu Y, Ye L, Wei R, Zhao R, et al. Tumor-secreted LCN2 impairs gastric cancer progression via autocrine inhibition of the 24p3R/JNK/c-Jun/SPARC axis. Cell Death Dis. 2024;15:756.

112. Song Q, Zhou Z, Bai J, Liu N. A comprehensive analysis of immunogenic cell death and its key gene HSP90AA1 in bladder cancer. Clin Transl Oncol. 2023;25:2587–606.

113. Sain A, Khamrai D, Kandasamy T, Naskar D. Apigenin exerts anti-cancer effects in colon cancer by targeting HSP90AA1. J Biomol Struct Dyn. 2023. https://doi.org/10.1080/07391102.2023.2299305.

114. Szczuka I, Wierzbicki J, Serek P, Szczęśniak-Sięga BM, Krzystek-Korpacka M. Heat shock proteins HSPA1 and HSP90AA1 are upregulated in

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 30 of 31

colorectal polyps and can be targeted in cancer cells by anti-inflammatory oxicams with arylpiperazine pharmacophore and benzoyl moiety substitutions at thiazine ring. Biomolecules. 2021;11:1588.

115. Phelan JD, Staudt LM. Double-headed molecule activates cell-death pathways in cancer cells. Nature. 2023;620:285–6.

116. Debnath J, Gammoh N, Ryan KM. Autophagy and autophagy-related pathways in cancer. Nat Rev Mol Cell Biol. 2023;24:560–75.

117. Stanilov N, Velikova T, Stanilova S. Navigating the cytokine seas: targeting cytokine signaling pathways in cancer therapy. Int J Mol Sci. 2024;25:1009.

118. Zhang KM, Zhao DC, Li ZY, Wang Y, Liu JN, Du T, Zhou L, Chen YH, Yu QC, Chen QS, et al. Inactivated cGAS-STING signaling facilitates endocrine resistance by forming a positive feedback loop with AKT kinase in ER+HER2- breast cancer. Adv Sci. 2024;11: e2403592.

119. Khawar MB, Yu S, Afzal A, Sun H. Unraveling the complex landscape of endocrine resistance in breast cancer: emerging role of long noncoding RNA AGPG and beyond. Chin Med J (Engl). 2024;137:1985–7.

120. Liu B, Liu Y, Yang S, Ye J, Hu J, Chen S, Wu S, Liu Q, Tang F, Liu Y, et al. Enhanced desmosome assembly driven by acquired high-level desmoglein-2 promotes phenotypic plasticity and endocrine resistance in ER(+) breast cancer. Cancer Lett. 2024;600: 217179.

121. Khoshdel F, Mottaghi-Dastjerdi N, Yazdani F, Salehi S, Ghorbani A, Montazeri H, Soltany-Rezaee-Rad M, Goodarzy B. CTGF, FN1, IL-6, THBS1, and WISP1 genes and PI3K-Akt signaling pathway as prognostic and therapeutic targets in gastric cancer identified by gene network modeling. Discov Oncol. 2024;15:344.

122. Benichou E, Seffou B, Topçu S, Renoult O, Lenoir V, Planchais J, Bonner C, Postic C, Prip-Buus C, Pecqueur C, et al. The transcription factor ChREBP Orchestrates liver carcinogenesis by coordinating the PI3K/AKT signaling and cancer metabolism. Nat Commun. 1879;2024:15.

123. Tian Y, Zhao L, Gui Z, Liu S, Liu C, Yu T, Zhang L. PI3K/AKT signaling activates HIF1α to modulate the biological effects of invasive breast cancer with microcalcification. NPJ Breast Cancer. 2023;9:93.

124. Wang Y, Sun X, Yang Q, Guo C. Cucurbitacin IIb attenuates cancer cachexia induced skeletal muscle atrophy by regulating the IL-6/STAT3/FoxO signaling pathway. Phytother Res. 2023;37:3380–93.

125. Farhan M, Wang H, Gaur U, Little PJ, Xu J, Zheng W. FOXO signaling pathways as therapeutic targets in cancer. Int J Biol Sci. 2017;13:815–27.

126. Mei Q, Li X, Zhang K, Wu Z, Li X, Meng Y, Guo M, Luo G, Fu X, Han W. Genetic and methylation-induced loss of miR-181a2/181b2 within chr9q33.3 facilitates tumor growth of cervical cancer through the PIK3R3/Akt/FoxO signaling pathway. Clin Cancer Res. 2017;23:575–86.

127. Motokura T, Bloom T, Kim HG, Jüppner H, Ruderman JV, Kronenberg HM, Arnold A. A novel cyclin encoded by a bcl1-linked candidate oncogene. Nature. 1991;350:512–5.

128. Rimokh R, Berger F, Bastard C, Klein B, French M, Archimbaud E, Rouault JP, Santa Lucia B, Duret L, Vuillaume M, et al. Rearrangement of CCND1 (BCL1/PRAD1) 3' untranslated region in mantle-cell lymphomas and t(11q13)-associated leukemias. Blood. 1994;83:3689–96.

129. Chesi M, Bergsagel PL, Brents LA, Smith CM, Gerhard DS, Kuehl WM. Dysregulation of cyclin D1 by translocation into an IgH gamma switch region in two multiple myeloma cell lines. Blood. 1996;88:674–81.

130. Lew DJ, Dulić V, Reed SI. Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast. Cell. 1991;66:1197–206.

131. Santra MK, Wajapeyee N, Green MR. F-box protein FBXO31 mediates cyclin D1 degradation to induce G1 arrest after DNA damage. Nature. 2009;459:722–5.

132. Dougherty MK, Müller J, Ritt DA, Zhou M, Zhou XZ, Copeland TD, Conrads TP, Veenstra TD, Lu KP, Morrison DK. Regulation of Raf-1 by direct feedback phosphorylation. Mol Cell. 2005;17:215–24.

133. Brondello JM, Pouysségur J, McKenzie FR. Reduced MAP kinase phosphatase-1 degradation after p42/p44MAPK-dependent phosphorylation. Science. 1999;286:2514–7.

134. Walsh JG, Cullen SP, Sheridan C, Lüthi AU, Gerner C, Martin SJ. Executioner caspase-3 and caspase-7 are functionally distinct proteases. Proc Natl Acad Sci U S A. 2008;105:12815–9.

135. Peng T, Tao X, Xia Z, Hu S, Xue J, Zhu Q, Pan X, Zhang Q, Li S. Pathogen hijacks programmed cell death signaling by arginine ADPR-deacylization of caspases. Mol Cell. 2022;82:1806-1820.e1808.

136. Nicholson DW, Ali A, Thornberry NA, Vaillancourt JP, Ding CK, Gallant M, Gareau Y, Griffin PR, Labelle M, Lazebnik YA, et al. Identification and inhibition of the ICE/CED-3 protease necessary for mammalian apoptosis. Nature. 1995;376:37–43.

137. Shi X, Sun Q, Hou Y, Zeng H, Cao Y, Dong M, Ding J, Shao F. Recognition and maturation of IL-18 by caspase-4 noncanonical inflammasome. Nature. 2023;624:442–50.

138. Goldberg YP, Nicholson DW, Rasper DM, Kalchman MA, Koide HB, Graham RK, Bromm M, Kazemi-Esfarjani P, Thornberry NA, Vaillancourt JP, Hayden MR. Cleavage of huntingtin by apopain, a proapoptotic cysteine protease, is modulated by the polyglutamine tract. Nat Genet. 1996;13:442–9.

139. Matsuura I, Denissova NG, Wang G, He D, Long J, Liu F. Cyclin-dependent kinases regulate the antiproliferative function of Smads. Nature. 2004;430:226–31.

140. Huang L, Guo Z, Wang F, Fu L. KRAS mutation: from undruggable to druggable in cancer. Signal Transduct Target Ther. 2021;6:386.

141. Schubbert S, Zenker M, Rowe SL, Böll S, Klein C, Bollag G, van der Burgt I, Musante L, Kalscheuer V, Wehner LE, et al. Germline KRAS mutations cause noonan syndrome. Nat Genet. 2006;38:331–6.

142. Niihori T, Aoki Y, Narumi Y, Neri G, Cavé H, Verloes A, Okamoto N, Hennekam RC, Gillessen-Kaesbach G, Wieczorek D, et al. Germline KRAS and BRAF mutations in cardio-facio-cutaneous syndrome. Nat Genet. 2006;38:294–6.

143. Wang J, Beauchemin M, Bertrand R. Bcl-xL phosphorylation at Ser49 by polo kinase 3 during cell cycle progression and checkpoints. Cell Signal. 2011;23:2030–8.

144. Bruey JM, Bruey-Sedano N, Luciano F, Zhai D, Balpai R, Xu C, Kress CL, Bailly-Maitre B, Li X, Osterman A, et al. Bcl-2 and Bcl-XL regulate proinflammatory caspase-1 activation by interaction with NALP1. Cell. 2007;129:45–56.

145. Di Donato M, Di Zazzo E, Salvati A, Sorrentino C, Giurato G, Fiore D, Proto MC, Rienzo M, Casamassimi A, Gazzerro P, et al. RIZ2 at the crossroad of the EGF/EGFR signaling in colorectal cancer. J Transl Med. 2023;21:736.

146. Ye F, Yuan Z, Tang Y, Li J, Liu X, Sun X, Chen S, Ye X, Zeng Z, Zhang XK, Zhou H. Endocytic activation and exosomal secretion of matriptase stimulate the second wave of EGF signaling to promote skin and breast cancer invasion. Cell Rep. 2024;43: 114002.

147. Fernandez MK, Sinha M, Kühnemuth R, Renz M. Repeated FRAP of the actin-binding protein CapG in the cell nucleus-a functional assay for EGF signaling in the single live breast cancer cell. Sci Rep. 2024;14:23159.

148. Liang XB, Dai ZC, Zou R, Tang JX, Yao CW. The therapeutic potential of CDK4/6 inhibitors, novel cancer drugs, in kidney diseases. Int J Mol Sci. 2023;24:13558.

149. Rocha AS, Paternot S, Coulonval K, Dumont JE, Soares P, Roger PP. Cyclic AMP inhibits the proliferation of thyroid carcinoma cell lines through regulation of CDK4 phosphorylation. Mol Biol Cell. 2008;19:4814–25.

150. Li W, Guo F, Zeng R, Liang H, Wang Y, Xiong W, Wu H, Yang C, Jin X. CDK4/6 Alters TBK1 Phosphorylation to Inhibit the STING Signaling Pathway in Prostate Cancer. Cancer Res. 2024. https://doi.org/10.1158/0008-5472.CAN-23-3704.

151. Liang H, Yang C, Zeng R, Song Y, Wang J, Xiong W, Yan B, Jin X. Targeting CBX3 with a Dual BET/PLK1 inhibitor enhances the antitumor efficacy of CDK4/6 inhibitors in prostate cancer. Adv Sci. 2023;10: e2302368.

152. Huang SP, Bao BY, Hour TC, Huang CY, Yu CC, Liu CC, Lee YC, Huang CN, Pao JB, Huang CH. Genetic variants in CASP3, BMP5, and IRS2 genes may influence survival in prostate cancer patients receiving androgen-deprivation therapy. PLoS ONE. 2012;7: e41219.

153. Köksal Karayildirim Ç, Nalbantsoy A, Karabay Yavaşoğlu N. Prunetin inhibits nitric oxide activity and induces apoptosis in urinary bladder cancer cells via CASP3 and TNF-α genes. Mol Biol Rep. 2021;48:7251–9.

154. Liang R, Chen W, Chen XY, Fan HN, Zhang J, Zhu JS. Dihydroartemisinin inhibits the tumorigenesis and invasion of gastric cancer by regulating STAT1/KDR/MMP9 and P53/BCL2L1/CASP3/7 pathways. Pathol Res Pract. 2021;218: 153318.

155. Guan X, Liu Z, Liu H, Yu H, Wang LE, Sturgis EM, Li G, Wei Q. A functional variant at the miR-885-5p binding site of CASP3 confers risk of both index and second primary malignancies in patients with head and neck cancer. Faseb j. 2013;27:1404–12.

156. Ohashi R, Angori S, Batavia AA, Rupp NJ, Ajioka Y, Schraml P, Moch H. Loss of CDKN1A mRNA and protein expression are independent

Xie *et al. Journal of Translational Medicine*   (2025) 23:32

Page 31 of 31

predictors of poor outcome in chromophobe renal cell carcinoma patients. Cancers. 2020;12:465.

157. Freeman MR, Washecka R, Chung LW. Aberrant expression of epidermal growth factor receptor and HER-2 (erbB-2) messenger RNAs in human renal cancers. Cancer Res. 1989;49:6221–5.

158. Stumm G, Eberwein S, Rostock-Wolf S, Stein H, Pomer S, Schlegel J, Waldherr R. Concomitant overexpression of the EGFR and erbB-2 genes in renal cell carcinoma (RCC) is correlated with dedifferentiation and metastasis. Int J Cancer. 1996;69:17–22.

159. Angori S, Banaei-Esfahani A, Mühlbauer K, Bolck HA, Kahraman A, Kara-kulak T, Poyet C, Feodoroff M, Potdar S, Kallioniemi O, et al. Ex vivo drug testing in patient-derived papillary renal cancer cells reveals EGFR and the BCL2 family as therapeutic targets. Eur Urol Focus. 2023;9:751–9.

160. Tang Y, Song T, Gao L, Mao F. Venetoclax synergizes sunitinib in renal cell carcincoma through inhibition of Bcl-2. Anticancer Agents Med Chem. 2023;23:2027–34.

161. Fang Z, Zhang N, Yuan X, Xing X, Li X, Qin X, Liu Z, Neo S, Liu C, Kong F, et al. GABPA-activated TGFBR2 transcription inhibits aggressiveness but is epigenetically erased by oncometabolites in renal cell carcinoma. J Exp Clin Cancer Res. 2022;41:173.

Publisher's Note