Yin et al. BMC Cancer (2025) 25:470
https://doi.org/10.1186/s12885-025-13903-9

BMC Cancer

RESEARCH Open Access

Exploring FAM13A-N-Myc interactions to uncover potential targets in MYCN-amplified neuroblastoma: a study of protein interactions and molecular dynamics simulations

Hongli Yin[1*†], Tianyi Liu[2†], Di Wu[1†], Xiaolu Li[1], Gen Li[1], Weiwei Song[3], Xiaodong Wang[6], Shan Xin[4*], Yisu Liu[5*] and Jian Pan[1*]

Abstract

Neuroblastoma (NB), a common infantile neuroendocrine tumor, presents a substantial therapeutic challenge when MYCN is amplified. Given that the protein structure of N-Myc is disordered, we utilized Alphafold for prediction and GROMACS for optimization of the N-Myc structure, thereby improving the reliability of the predicted structure. The publicly available datasets GSE49710 and GSE73517 were adopted, which contain the transcriptome data of clinical samples from 598 NB patients. Through various machine learning algorithms, FAM13A was identified as a characteristic gene of MYCN. Cell functional experiments, including those on cell proliferation, apoptosis, and cell cycle, also indicate that FAM13A is a potential risk factor. Additionally, Alphafold and GROMACS were employed to predict and optimize the structure of FAM13A. Protein-protein docking and molecular dynamic modeling techniques were then used to validate the enhanced protein stability resulting from the interaction between N-Myc and FAM13A. Consequently, targeting FAM13A holds the potential to reduce the stability of N-Myc, hinder the proliferation of NB cells, and increase the infiltration of immune cells. This multi-faceted approach effectively combats tumor cells, making FAM13A a prospective therapeutic target for MYCN-amplified NB.

Keywords Neuroblastoma, MYCN, Protein structure prediction, Protein–protein docking, Single cell transcriptomes, Immune cell infiltration

†Hongli Yina, Tianyi Liu and Di Wua contributed equally to this work.

*Correspondence:
Hongli Yin
yinhongli1992@outlook.com
Shan Xin
shan.xin@yale.edu
Yisu Liu
Sophie1liu@foxmail.com
Jian Pan
panjian2008@163.com
[1]Institute of Pediatric Research, Children's Hospital of Soochow University, Suzhou 215003, China
[2]Department of Pharmaceutics, Dalian Women and Children's Medical Group, Dalian, Liaoning 116012, China
[3]Department of Pharmacy, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing 102218, China
[4]Genetics and Cellular Engineering Group, Research Unit Signaling and Translation, Helmholtz Zentrum Munich, Ingolstaedter Landstr. 1, Neuherberg 85764, Germany
[5]China College of Life Sciences and Chemistry, Hunan University of Technology, Zhuzhou 412007, China
[6]Department of Orthopaedics, Children's Hospital of Soochow University, Suzhou 215003, China

Introduction

Neuroblastoma (NB) is a widespread solid tumor originating from neural crest cells during early infancy. It accounts for around 15% of pediatric cancer-related mortalities, with a median age of diagnosis at 18 months [1]. NB exhibits conspicuous heterogeneity, encompassing a range of clinical outcomes, including spontaneous remission, progressive disease, and the development of metastases [2]. About 25% of NB patients have MYCN oncogene amplification [3], and stage 4 patients with MYCN amplification had a 5-year overall survival rate of 34% [4]. The identification of amplification of MYCN, which encodes N-Myc, has been recognized as a major factor with prognostic significance. This is particularly relevant in the prediction of unfavorable clinical outcomes and poor survival rates among high-risk patients diagnosed with NB [3]. MYCN belongs to the MYC oncogene family and demonstrates prominent expression in neural tissue [5]. The protein being examined functions as a key transcriptional regulator in the context of neural development, exerting influence over a variety of biological processes such as cellular differentiation, proliferation, survival, self-renewal, and metabolic functions [3]. However, therapeutic efforts have proven challenging in combating MYCN-amplified NB, and as a result, N-Myc has been labeled as "undruggable" with limited treatment options [3]. Determining the indirect regulators of MYCN expression is also urgently required.

FAM13A (family with sequence similarity 13 member A) isoform 1 (v2) and isoform 2 (v1) have been identified as splice variants in humans. The Ras homologous GTPase-activating protein (RhoGAP) domain in FAM13A isoform 1 is critical for the control of cell proliferation and survival [6]. In genome-wide association studies (GWAS), the presence of genetic variations in the FAM13A gene has been observed to have a correlation with lung function in various prevalent chronic lung diseases, including asthma, chronic obstructive pulmonary disease (COPD), and idiopathic interstitial pneumonia (IIP) [7]. In addition, FAM13A improves insulin sensitivity in mice by modulating insulin signaling in adipocytes and preserving body homoeostasis [8]. Studies have been showed that knocking down FAM13A promotes adipocyte development and glucose absorption and improves lipogenic differentiation [9]. A recent investigation has provided further evidence on the role of FAM13A in the regulation of lipid metabolism. Notably, elimination of this gene has been found to confer protection to mice against the development of fatty liver induced by a high-fat diet [10]. However, the biological role of FAM13A in NB remains unclear, and there is limited knowledge regarding the underlying mechanism of this protein activity.

The objective of this study was to evaluate the interaction between N-Myc and FAM13A using bioinformatics, experimental methods, and single-cell analysis. Additionally, we aimed to clarify the biological role of FAM13A in NB. The optimization of the structure of N-Myc and FAM13A was performed using GROMACS. Subsequently, protein-protein docking and molecular dynamics modeling techniques were employed to validate the improved protein stability resulting from the interaction between N-Myc and FAM13A. Hence, the targeting of FAM13A has the potential to decrease the stability of N-Myc, impede the proliferation of NB cells, and enhance the infiltration of immune cells. This multifaceted approach effectively combats tumor cells, making FAM13A a prospective therapeutic target for MYCN-amplified NB.

Materials and methods

Machine learning algorithm analysis of MYCN-related characteristic genes

In this study, the publicly available datasets GSE49710 and GSE73517 were utilized. We analyzed them using data - mining techniques such as Support Vector Machine Recursive Feature Elimination (SVM - RFE) and Random Forest.

Support vector machine recursive feature elimination (SVM-RFE)

This research method is mainly based on support vector machine (SVM) and recursive feature elimination (RFE) technology, aiming to screen out important feature genes from MYCN-related data. Through the evaluation and visualization of different feature numbers, the optimal feature genes were determined and output for further analysis and research. The three R packages "e1071" [11],"kernlab" [12] and "caret" [13] were introduced using the library function to provide support for subsequent machine learning calculations. During the feature selection process using the rfe function, several key parameters play important roles. First, y = as.numeric(as.factor(group)) is used to extract the group variable from the row names of the data. It is then converted into a factor and further into a numeric type to serve as the response variable. This operation transforms the grouping information in the data into a numeric form suitable for model processing, enabling it to participate in subsequent feature selection and model evaluation processes. Secondly, sizes = c(2, 4, 6, 8, seq(10, 40, by = 3)) defines the range of the number of features to be evaluated during the feature elimination process. Specifically, feature subsets with feature numbers of 2, 4, 6, 8, as well as those starting from 10 and incrementing by 3 up to 40 will be considered in sequence. This setting aims to systematically screen feature combinations of

different scales to find the optimal combination, thereby optimizing the performance of the model. Finally, rfeControl = rfeControl(functions = caretFuncs, method = "cv") uses caretFuncs in the caret package as the default set of evaluation functions and adopts cross - validation (method = "cv") to evaluate the performance of different feature subsets in the support vector machine model. During the cross - validation process, the dataset is divided into multiple subsets. In each iteration, one subset is used as the validation set, and the remaining subsets form the training set. Through multiple iterations, the performance of the model under different feature combinations can be evaluated more comprehensively and reliably.

In terms of validation techniques, the support vector machine uses the Radial Basis Function kernel (methods = "svmRadial"). The Radial Basis Function kernel is commonly used to handle non - linearly separable data because it can map the original data into a high - dimensional space, making the data that is non - linearly separable in the low - dimensional space become linearly separable in the high - dimensional space, thus laying the foundation for constructing a more accurate classification or regression model. Regarding performance evaluation metrics, the Root Mean Square Error (RMSE) is chosen as the main measure. By using the rfe function to build support vector machine models under different feature subsets and perform cross - validation, the corresponding RMSE is calculated. The graph of the relationship between the number of variables (representing the size of different feature subsets) and RMSE is stored in the SVM - RFE.pdf file, where the x - axis represents the number of features and the y - axis represents the RMSE. The point with the minimum RMSE is marked on the graph, and the feature subset corresponding to this point is the optimal feature subset selected through cross - validation.

Random forest

This research method mainly utilizes the random forest algorithm for model construction and screening of feature genes. Firstly, the randomForest package is imported, and a random seed is set to ensure the reproducibility of the results. The input file and working directory are specified, and after reading the input file, data processing is carried out. Next, the randomForest function is used to build a random forest model with the number of trees set to 500. A chart of the random forest model is plotted to visually observe the model performance. Then, a five - fold cross - validation is adopted to evaluate and optimize the model to find the point with the minimum error. Based on this point, the random forest model is reconstructed. At the same time, the importance of genes is calculated, and feature genes are screened according to the scores. A chart of gene importance is drawn to intuitively understand the importance degree of each gene. Disease - related feature genes are selected according to the importance scores and written into a file. Finally, the expression levels of important genes are output and the results are written into a file. Through the above steps, the aim is to improve the accuracy and reliability of the model and provide strong support for the screening and research of disease - related feature genes.

Cell culture

The SK-N-BE(2), SK-N-SH human NB cell lines and embryonic kidney cells HEK293T were acquired from the National Collection of Authenticated Cell Cultures (Shanghai, China). The cells were incubated at a temperature of 37 °C in a controlled environment with 5% CO_2. For SK-N-BE(2), the culture medium employed was DMEM/F12 (Thermo Fisher Scientific, USA), while for SK-N-SH, alpha-MEM (Thermo Fisher Scientific, USA) was used. HEK293T were maintained in DMEM (Thermo Fisher Scientific, USA). Both media were supplemented with 10% foetal bovine serum (FBS) (Biological Industries, USA), and 1% penicillin/streptomycin (Beyotime, China).

RNA isolation and qRT-PCR

The RNA was extracted via the RNeasy Mini Kit (Qiagen, Germany). The Reverse Transcription System (Promega, USA) was utilized for reverse transcription. On a LightCycler 480 Real-Time System (Roche, Switzerland), LightCycler 480 SYBR Green I Master mix (Roche, Switzerland) was utilized for PCR. The $2^{-\Delta\Delta Ct}$ method was employed to determine the expression, with GAPDH expression utilized as an internal reference [14]. Tsingke Biological Technology (Beijing, China) synthesized the qPCR primers that were used. Primer sequences of the *FAM13A*, *MYCN*, *HLA-A*, *HLA-B*, *HLA-C* and *GAPDH* are as follows (forward and reverse, respectively): *FAM13A*, 5′-ACCCTGTTTGAAGTAGAGTATACAG-3′ and 5′-AGACCTCTTTTACTATGATAAGCCT-3′; *MYCN*, 5′-ACCCGGACGAAGATGACTTCT-3′ and 5′-CAGCTCGTTCTCAAGCAGCAT-3′; *HLA-A*, 5′-AAAAGGAGGGAGTTACACTCAGG-3′ and 5′-GCTGTGAGGGACACATCAGAG-3′; *HLA-B*, 5′-CTACCCTGCGGAGATCA-3′ and 5′-ACAGCCAGGCCAGCAACA-3′; *HLA-C*, 5′-CACACCTCTCCTTTGTGACTTCAA-3′ and 5′-CCACCTCCTCACATTATGCTAACA-3′; *GAPDH*, 5′-GAAGGTGAAGGTCGGAGTC-3′ and 5′-GAAGATGGTGATGGGATTTC-3′.

Cell viability assay

The cell counting kit-8 (CCK8) test (Dojindo Molecular Technologies, Japan) was employed to assess cell viability. SK-N-BE(2) cells were prepared as a single cell

suspension using Trypsin (Beyotime Biotech, China). Then, the cells were seeded into 96-well plates at a density of 2×10^3 cells per well. The plates were incubated for 0, 1, 3, 5, and 7 days respectively. After each incubation period, CCK8 solution was added to the culture medium in the 96-well plates and further incubated for 4 h. Subsequently, the absorbance of the cells was detected using spectrophotometry (Thermo Fisher Scientific, USA) at a specific wavelength of 450 nm. The cell viability was calculated by comparing the absorbance values between the sh-NC and sh-FAM13A groups. Statistical analysis was performed on data obtained from three independent replicate experiments.

Cell apoptosis and cell cycle analysis

For apoptosis analysis, cells were extracted using a trypsin solution devoid of EDTA. Then, the cells were washed and stained with an Annexin V-FITC kit (Miltenyi Biotec, Germany), specifically with Annexin V and propidium iodide solution. Subsequently, cytometry analysis was performed. For cell cycle analysis, the cells were removed, subsequently washed, and then preserved in ice-cold 70% ethanol at 4 °C overnight. The following day, cells were stained with propidium iodide/RNase A staining solution (Sigma-Aldrich, USA) for 20 min at room temperature. Flow cytometry was used to examine the samples (Beckman Gallios, Germany), and cell apoptosis was analyzed by FlowJo v10 (Tree Star, USA), the Q2 and Q3 areas were labeled as apoptotic. The cell cycle distribution was analyzed with ModFit LT (Verity Software House, USA). All data are expressed as the mean ± standard error of the mean (SEM) of three independent experiments.

Colony formation assay

NB cells capable of stable gene silencing were inoculated at a density of 2,000 per well in six-well plates and cultured for approximately two weeks. The media was discarded after the culture, rinsed three times with cool, sterile 1xPBS, preserved with 4% formaldehyde, and stained with 0.1% crystal violet. The number of colonies was then counted macroscopically.

Immunoprecipitation and Immunoblotting

For immunoblotting, cell lysates were isolated using RIPA buffer containing 1 mM PMSF and 1% phosphatase inhibitor cocktail. Solubilized proteins were collected by centrifugation and quantified in 3×10^6 cells per sample. Samples with equal amounts of protein were incubated on ice for 30 min and centrifuged at 13,000xg for 10 min prior to immunoblotting. The samples were resolved by electrophoresis on a 10% gel and transferred to PVDF membranes for western blot analysis. The membrane was blocked with a 5% skim milk solution in TBST buffer (10 mM Tris-HCl, 150 mM NaCl, and 1% Tween 20). Then, the primary antibody was added and incubated overnight at 4 °C. The next day, after three washes with TBST buffer, the PVDF membrane was incubated with a secondary antibody in 5% skim milk at room temperature for 1 h. The PVDF membrane was washed three times with TBST buffer, and proteins were detected by standard enhanced chemiluminescence (ECL) (Millipore, USA). The antibodies used were specific for FAM13A (55401-1-AP, Proteintech, USA), N-Myc (9405 S, Cell Signaling Technology, USA), Cleaved caspase-3 (9661 S, Cell Signaling Technology, USA), PARP (9542 S, Cell Signaling Technology, USA), CDK2 (2546T, Cell Signaling Technology, USA), CDK4 (12790T, Cell Signaling Technology, USA), CDK6 (13331T, Cell Signaling Technology, USA), MCL-1 (4572 S, Cell Signaling Technology, USA), BCL-2 (15071T, Cell Signaling Technology, USA), HA (3724 S, Cell Signaling Technology, USA), Flag (14793T, Cell Signaling Technology, USA) and GAPDH (MAB373, Millipore, USA).

The Co-IP assay was performed using the Beaver-Beads™ Protein A/G Immunoprecipitation Kit (22202-100, Beaver, China). Briefly, cells in 6-well plates were collected (1×10^7 SK-N-BE(2) for endogenous IP). IP binding buffer (500 ml PBS, 0.3% Tween 20, and 75 mM NaCl) was added for lysis on ice for 40 min. Supernatant, IgG, and IP were used as inputs. The target antibody was added and incubated overnight at 4 °C. The next day, after washing with IP binding solution, 20 µL of magnetic beads were incubated at 4 °C for 2 h. Subsequently, the supernatant was removed and the magnetic beads were washed again. For denaturation, 100 µL loading buffer was added, mixed with the magnetic beads, and heated at 100 °C for 6 min. Finally, a Western blot assay was used to measure the protein. The antibodies used were specific for HA (3724 S, Cell Signaling Technology, USA).

ShRNA infection and plasmid DNA transfection

For shRNA-mediated knockdown, the sequences of shRNA were synthesized by IGE Biotechnology, Ltd. (Guangzhou, China), sequences of shFAM13A are as follows: sh-FAM13A #1, 5′-TAATAACTCTGGAGGTCAAAG-3′; sh-FAM13A #2, 5′-GGAGAACTCTTAGAAAGAACC-3′; sh-FAM13A #3, 5′-GCCGGTAACAAAGAACGAACG-3′. Lentiviral packaging was performed as previously described [15]. If necessary, lentiviral medium was concentrated by PEG-8000 (Beyotime, China) precipitation. Lentiviral particles were applied to infect cells in the presence of polybrene (Sigma-Aldrich, USA). After puromycin (Sigma-Aldrich, USA) screening, stable depletion cells were established. For overexpression experiments, the HEK293T cell line was transfected with 4 µg pcDNA3.1-MYCN-Flag or/and pcDNA3.1-FAM13A-HA. Cells were transfected after 48 h. Transfections were

performed using Lipofectamine 3000 (Thermo Fisher Scientific, USA) following the manufacturer´s protocol.

Cleavage under target and release using nuclease (CUT&RUN)

CUT&RUN is a sophisticated technique used to study protein-DNA interactions. It involves the use of Protein G-fused MNase nucleases, which are guided by antibodies to specifically target proteins and induce DNA fragmentation near the desired location. The experiment adhered to the procedures specified in the Hyperactive pG-MNase CUT&RUN Assay Kit for qPCR (HD101-01; Vazyme, China). DNA was purified using a MiniElute PCR Purification Kit (Qiagen, Germany). Real-time PCR was performed using the LightCycler® 480 instrument with purified DNA, Power SYBR Green PCR Master Mix (Roche, Switzerland). The N-Myc antibody (Catalog # 61185, Active Motif, USA) and control IgG antibodies (EPR25A, Abcam, UK) were used. Real-time PCR was performed with primers designed to cover the regions of the FAM13A gene containing MYCN-responsive motifs or remote negative control regions. Fold enrichment of the FAM13A gene containing the binding regions by the anti-N-Myc antibody was calculated by dividing the PCR product from this region by the PCR product from the negative control region, relative to the input. FAM13A forward primer: 5´-TCTCTTTCCGCTGAACCCAC-3´, FAM13A reverse primer: 5´-CTCCCTCTACTTGCCA GCAC-3´; Negative Control forward primer: 5´-CTAG CTTTTGGATTAGTTT-3´, Negative Control reverse primer: 5´-AGCACTAAATGCCCAC-3´.

Single cell transcriptomes analysis

We utilized a previously published sc-RNA-seq dataset, available in H5AD format from the Neuroblastoma Cell Atlas, which can be accessed from https://neuroblastoma-cell-atlas.cog.sanger.ac.uk/nb_GOSH_cellxgene.h5ad [16]. This dataset was processed and analyzed using the Scanpy software package to explore cellular landscapes and gene dynamics [17]. Following standard preprocessing practices recommended by Scanpy, the data were normalized using the 'normalize_per_cell' function and then log-transformed using the 'log1p' function. Highly variable genes were selected using the 'highly_variable_genes' function. Subsequently, cluster identification and visualization were performed using the UMAP technique, with distinct clusters detected using the Louvain and Leiden algorithms. Marker gene analysis was employed to characterize the identified clusters, and pseudo-time analysis was utilized to infer developmental trajectories by Scanpy's 'scanpy.tl.paga' function.

Bioinformatic analysis

Data mining

Expression matrix of tumor cell lines: Obtained from the Cancer Cell Line Encyclopedia (CCLE) data set at the website https://portals.broadinstitute.org/ccle. RNA sequence data and clinical information of neuroblastoma: Obtained from the Therapeutically Applicable Research to Generate Effective Treatments (Target) project database at the website https://ocg.cancer.gov/programs/target, which contains tissues from 249 tumor patients. At the same time, the mRNA profile data and clinical characteristics of neuroblastoma are publicly available on an open - access platform. Microarray data: Retrieved through the Gene Expression Omnibus (GEO) at the website https://www.ncbi.nlm.nih.gov/geo/, and the orig inal data is obtained by downloading MINiML files.

Data analysis

Enrichment analysis (GO and KEGG): Conducted using R software and related bioinformatics analysis packages. After obtaining data from specific databases, the gene set is pre - processed, and then the analysis is carried out to determine the enrichment of genes in biological functions and metabolic pathways. Overall survival analysis: The R v4.0.3 software package ggplot2 is used for the preliminary study, and the Kaplan - Meier curve is constructed using the online platform (https://r2platform.com). According to the gene expression level grouping, the relationship with the overall survival of patients is evaluated. Difference analysis: Performed using the limma package combined with R software. After obtaining data from the database, the mRNA is analyzed, and the differential expression threshold is determined for judgment. Gene correlation analysis: The R software package ggstatsplot is used to visualize gene correlation. After processing the data, the association between genes is displayed through this package.

Regarding the differential expression threshold for mRNAs, it is determined as "Adjusted $P < 0.05$ and Log (Fold Change) > 1 or Log (Fold Change) < -1".

Immune infiltration analysis

CIBERSORT [18] is an algorithm widely used to characterize the cellular composition of the different complex tissues by gene expression values in the solid tumors. LM22 (a special genetic marker) signature algorithm was employed as a special genetic marker, which contains 547 distinct genes. It can distinguish 22 immune cell subtypes downloaded from the CIBERSORT portal (http://cibersort.stanford.edu/). In this study, CIBERSORT package and LM22 algorithm were used to calculate the infiltration abundance of 22 immune cell subtypes in the NB patient samples of Target database between the high and low FAM13A expression groups containing the different

T cells, B cells, plasma cells, natural killer cells, and different myeloid subsets. We also analyzed the correlation of expression distribution of 22 immune cell subtypes between FAM13A high and low expression groups. And we also investigated the correlations between FAM13A expression and immunodulators, MHC molecules, chemokines, and chemokine receptors in NB patients' samples from the Target database.

Gene enrichment analysis

The cluster Profiler package [19] (FDR < 0.1) was employed for KEGG pathway analysis and GO analysis of the DEGs. The expression matrix of the differential genes grouped by high and low expressions of FAM13A was used for GSEA analysis [20], with msigdb.v7.0.entrez.gmt as the selected reference gene set. Then ggPlot2 package was utilized to visualize GO, KEGG, and GSEA pathways by creating the related bar charts, bubble charts, and enrichment maps.

Protein structure prediction

The National Center for Biotechnology Information (NCBI) was used to acquire the amino acid sequences of N-Myc (https://www.ncbi.nlm.nih.gov/protein/NP_005369.2, NP_005369.2) and FAM13A (https://www.ncbi.nlm.nih.gov/protein/NP_055698.2, NP_055698.2). Using the monomer_casp14 model from Alphafold v2.3.2 [21], we predicted the high-resolution structures of N-Myc and FAM13A. The following database versions were used: uniport 2023-03-01, uniref90 2023-03-01, pdb_mmcif 2023-03-03, and pdb_seqres 2023-03-03. Additionally, pTM (Predicted Targeting Metric) is employed to assess the quality of the protein structures predicted by Alphafold.

Protein structure assessment

The SWISS-MODEL online instrument (https://swissmodel.expasy.org/assesssment) was utilized to evaluate the predicted protein structure. This tool utilizes MolProbity version 4.4 to generate Ramachandran diagrams and calculate Ramachandran Favoured values. The Ramachandran plot is a graph that depicts the energetically favorable locations for backbone dihedral angles of amino acid residues within the structure of a protein. Typically, if the Ramachandran Favoured value of 90% or more, we can conclude that the predicted protein model conforms to stereochemical principles.

Protein-Protein Docking

This study used Schrödinger software for protein-to-protein docking. First, preprocess the protein. Then, bond-level assignment, hydrogenation, zero-order bond assignment to metal atoms, and disulfide bond creation were performed on the two proteins. Then, the hydrogen bond network was optimized and protein energy minimization was performed using the OPLS_4 force field. Then, the Protein-Protein docking (Piper) module in Schrödinger was used for molecular docking. The standard mode is used for docking, and the number of rotatable probes for the ligand is defined as 70,000, which enables sufficient conformational sampling of the ligand protein, and the number of generated conformations is defined as 30. For proteins with multiple chains, we only use one chain for docking. Finally, among all the conformations generated, Piper clusters the top 1000 rotational conformations based on the Root-mean-square deviation (RMSD) between each atom, and the representative conformation in each class is selected from the conformation with the most neighbors in this class. Piper ranks the generated conformations based on the number of clusters in each category. The top-ranked conformation is the conformation with the largest number of clusters, which is the optimal binding mode predicted for the FAM13A and N-Myc protein interaction.

Molecular dynamics (MD) simulation

On an Ubuntu 20.04.01 platform with an Intel Core i9-12900k CPU, GeForce RTX 4070 GPU, and 64 GB of RAM, GROMACS (2023.1 single precision) was used to conduct the MD simulation. The software SPDBV 4.10 was initially used to modify the heavy ions and small molecules of the proteins. Using the CHARMM36 force field, the topology structure of the protein was subsequently calculated. The TIP3P water model was used to solve the complex, which was encircled by a cube-shaped cage reaching at least 1.2 nm on all sides. The system was neutralized by adding Na^+ and Cl^- ions, followed by 0.15 M NaCl, which restored it back to a near-physiological state. Following that, for 5000 iterations, the steepest descent technique was utilized to accomplish energy minimization and a maximum force of less than 1000 kJ/mol/nm. To achieve a well-equilibrated and stable system, the system was subjected to 100 ps of restricted NVT (number of particles, volume, temperature) and NPT (number of particles, pressure, temperature) equilibration at 310 K and 1 bar. Finally, the MD simulation of the complex was conducted for 100 ns and 50,000,000 steps. During the simulation, the Verlet cut-off scheme and Leap-frog integrator with a 2 fs step size were utilized, and trajectory data was saved every 10 ps. Unless explicitly stated otherwise, the RMSD calculations pertain solely to the protein backbone. we utilized the "gmx rms" command to compute RMSD, which initially employs the least-squares method to fit the structure to the reference structure ($t_2 = 0$), and subsequently quantifies RMSD based on the following formula:

$$RMSD\left(t_1, t_2\right) = \left[\frac{1}{M}\sum {}_{i=1}^{N} m_i \left\| r_i\left(t_1\right) - r_i\left(t_2\right)\right\|^2\right]^{1/2} \quad (1)$$

Where, $M = \sum {}_{i=1}^{N} m_i$, $\mathrm{r_i}$ (t) is the the position of the atom *i* at frame *t.* Unless explicitly stated otherwise, the Root Mean Square Fluctuation (RMSF) calculations pertain solely to the protein backbone. The RMSF value is determined using the following formula:

$$RMSF_i = = \sqrt{\frac{1}{T}\sum {}_{t=1}^{T}\left(r_i\ (t) - \ \bar{r}\right)^2} \quad (2)$$

Where, ri (t) is the position vector of the atom i at frame t and is the average atom position over all T frames. Radius of Gyration (RoG) plays a pivotal role in the evaluation of protein structure and its conformational alterations. In this research, the RoG value is determined using the following formula:

$$RoG = \left(\frac{\sum {}_i \left\| r_i\right\|^2 m_i}{\sum {}_i m_i}\right)^{1/2} \quad (3)$$

Where, mi is the molecular weight of atom i, ri is the position of atom i relative to the center of the molecule. The Solvent Accessible Surface Area (SASA) was estimated using GROMACS commands. In this research, the SASA value is determined using the following formula:

$$SASA = \sum \frac{R}{\sqrt{R^2 - Z_i^2}}\ x\ D\ x\ L_i \quad (4)$$

Where, Li is the length of the arc drawn on a given section i; Zi is the perpendicular distance of section i from the centre of the sphere. Free energy landscape (FEL) of receptor-ligand binding represents a vivid and informative two- or three-dimensional depiction of the changes in free energy that occur during the binding process between receptor and ligand. The translational and rotational motion of the MD simulation trajectory was corrected using the gmx trjconv command. Upon obtaining the computed results of the RMSD and RoG, the FEL of protein-ligand complex stability was calculated using the gmx sham command. After analyzing the energy well, the gmx trjconv command was used to extract the conformation of the energy minimum point.

Statistical analysis

The statistical analysis was conducted using GraphPad Prism software (version 8.0, GraphPad Prism Inc.). The experimental data is displayed as mean ± standard error of the mean (SEM). The Pearson correlation coefficients were computed to ascertain the bivariate correlation between the variables being studied. The student's t-test was used to assess differences between two distinct groups, whereas ANOVA was used for more than two groups. We used $P<0.05$ to evaluate statistical significance. The symbols $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, and $^{****}P<0.0001$ indicate increasing significance.

Results

Multiple machine learning algorithms identified MYCN related feature genes

The presence of multiple copies of the MYCN gene is a significant prognostic factor associated with an unfavorable outcome in individuals diagnosed with neuroblastoma (NB). Through data mining, we acquired data from 623 samples (475 MYCN non-amplified samples and 148 MYCN amplified samples) from two data sets (GSE49710, GSE73517) in the GEO database. By examining the differences between the MYCN amplified and non-amplified groups, 169 MYCN-related differential genes were identified, of which 74 were up-regulated and 95 were down-regulated. Heat maps and volcano plots were depicted in Fig. 1A and B, respectively. As illustrated in Fig. 1C-E, support vector machine-recursive feature elimination (SVM-RFE), random forest, and other machine learning approaches were employed to screen the distinctive genes associated with MYCN. As shown in Fig. 1F, there are five MYCN feature genes that can be screened by both machine learning algorithms: PHGDH, FAM13A, KLRG2, RAB3B, and NPW. To further screen out the genes that can indirectly affect the expression of MYCN in NB for this study, we used the R2 database to analyze the GSE62564 data set. The results showed that except for the RAB3B gene, when the other four genes (PHGDH, FAM13A, KLRG2, and NPW) were highly expressed, the prognosis of the patient was poor (Fig. S1). Among them, the research on PHGDH has demonstrated its role in NB [22]. The research on KLRG2, a receptor expressed on natural killer cells and T cells, mainly focuses on its role in immune responses [23]. NPW is a secreted protein that is typically secreted by cells into the extracellular space or into the blood circulation [24]. Since the N-Myc protein is predominantly located in the nucleus, the interaction between NPW and the N-Myc protein may require specific signal transduction pathways or molecular mechanisms. FAM13A is a protein that may play a significant role in tumor initiation, progression, and metastasis, but its role in NB remains unclear [25]. Based on the above considerations, we ultimately decided to select FAM13A as the research object of this paper.

Positive association between FAM13A and MYCN

An in-depth investigation was conducted on the Target and GEO databases to explore the potential correlation between FAM13A and MYCN amplification. The findings of our analysis demonstrate a statistically significant

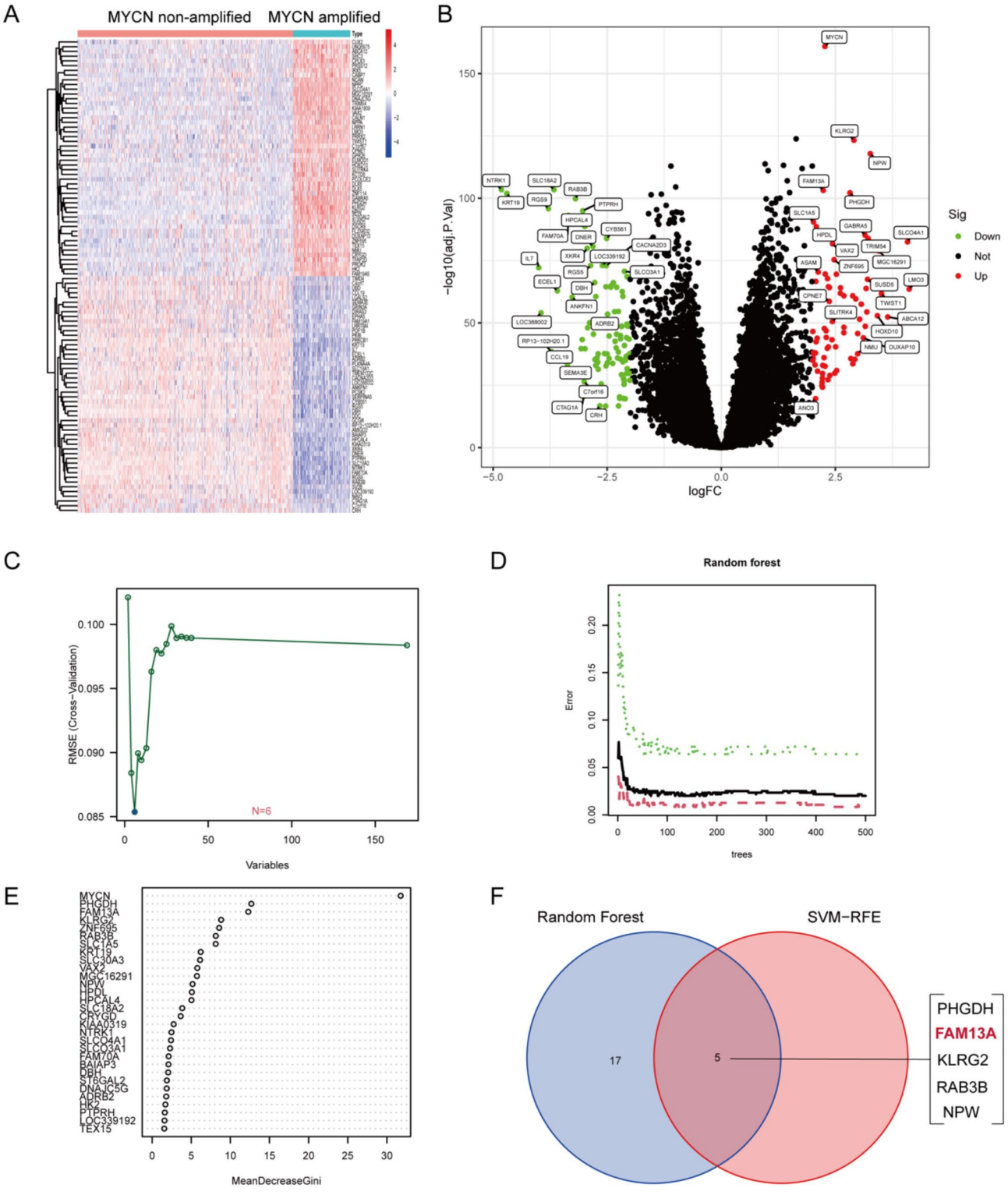


Fig. 1 Identification of MYCN-related feature genes through screening with multiple machine learning algorithms. **(A-B)** Heat map and volcano plot respectively depict differential genes related to MYCN. **(C-E)** Screen MYCN-related feature genes using SVM-feature recursive elimination, random forest, and other machine learning methods. **(F)** The Venn graphic indicates that both machine learning methods can choose 5 MYCN-related feature genes. *$p < 0.05$; asterisks (*) represent levels of statistical significance. Two groups' statistical differences were contrasted using the student´s t test

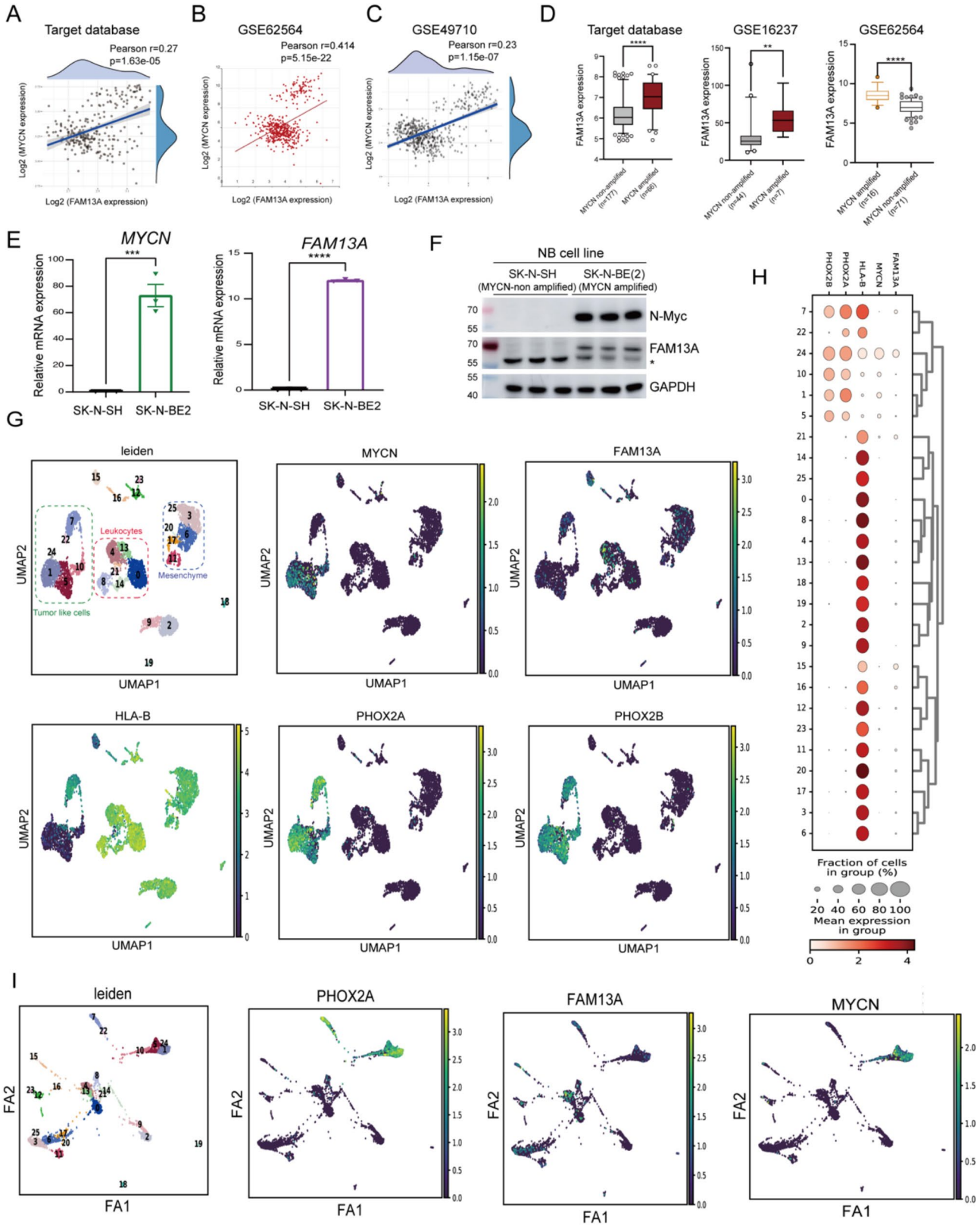


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Co-expression and integrated analysis of FAM13A and MYCN in neuroblastoma. **(A-C)** Expression of FAM13A and MYCN are correlated in NB patients (r = Pearson´s correlation coefficient). **(D)** FAM13A is overexpressed in NB with MYCN amplification compared to NB without MYCN amplification. Cases were divided based on the status of MYCN amplification, and FAM13A expression is measured. **(E-F)** qRT-PCR and Western blot revealing FAM13A and MYCN expression in MYCN amplified and nonamplified NB cell lines. **(G)** Dimensionality reduction (UMAP) of sc-RNA-seq data from untreated NB patients (6,442 cells). The five panels show the expression of NB associated gene. **(H)** Hierarchical clustering analysis was performed on the sc-RNA-seq data, focusing on the expression levels of FAM13A, MYCN, PHOX2A, PHOX2B, and HLA-B. The dengrogram illustrates the relationships between cells based on their gene expression profiles. **(I)** Pseudo-time analysis and visualization of the sc-RNA-seq unveil crucial branching points in cell deterioration of the NB tumor cells. $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$; asterisks (*) represent levels of statistical significance. Two groups' statistical differences were contrasted using the student´s t test

positive association between the FAM13A gene and the MYCN gene (Fig. 2A-C). Patients who displayed MYCN amplification revealed a significantly elevated expression of FAM13A as compared to patients without MYCN amplification (Fig. 2D). Subsequently, we procured the expression matrix of the MYCN gene and the FAM13A gene from NB cell lines by utilizing the Cancer Cell Line Encyclopedia (CCLE) dataset. Our analysis revealed a noteworthy co-expression of the MYCN gene and FAM13A gene in cell lines with MYCN amplification (Fig. S2). Furthermore, the quantitative reverse transcription polymerase chain reaction (qRT-PCR) study demonstrated a considerable elevation in FAM13A levels inside the MYCN-amplified cell line (SK-N-BE(2)) compared to the MYCN-nonamplified cell line (SK-N-SH) (Fig. 2E). The protein expression levels of N-Myc and FAM13A were evaluated by Western Blot. The obtained results indicated a significant overexpression of both N-Myc and FAM13A in the NB cell line with MYCN amplification (Fig. 2F).

In order to establish the existence of a positive association between FAM13A and MYCN expression, we performed a thorough study using single cell transcriptomes derived from a dataset of 6,442 cells (10X) collected from untreated individuals with NB [16]. After applying cluster analysis to the sc-RNA-seq data and utilizing UMAP dimension reduction (Fig. 2G), we proceeded to examine the expression levels of two prominent NB susceptibility genes, namely PHOX2A and PHOX2B, along with the oncogene MYCN, and other genes exhibiting differential expression (Fig. S3), within these distinct cellular subtypes. In addition, an annotation analysis was conducted on the discovered cell clusters, as described in the annotation analysis conducted by Sam Behjati et al. The sc-RNA-seq data analysis unveils that tumor-like cells prominently express the NB susceptibility genes PHOX2A and PHOX2B. Intriguingly, this expression pattern demarcates distinct cell subtypes characterized by divergent MYCN expression levels. Hierarchical clustering analysis (Fig. 2H) discerns a segregation of tumor-like cells, with one branch exemplifying heightened MYCN expression (Cell Cluster 1, 5, 10, and 24), while the other represents cells exhibiting subdued MYCN expression (Cell Cluster 7 and 22). Notably, within the subset displaying elevated MYCN expression, HLA-B experiences a notable decline in expression. Concurrently, the expression profiles of FAM13A correspond closely to those of MYCN, and within Cell Cluster 22, both FAM13A and MYCN manifest significantly diminished expression levels. Subsequently, we conducted further pseudo-time analysis (Fig. 2I), revealing a bifurcation trajectory within the tumor-like cells. In this trajectory, the initial cell cluster of the MYCN-low expressing branch is characterized by the absence of both MYCN and FAM13A, denoted as cluster 22. This cluster then progresses to cluster 7, exhibiting worsening malignant features. While a subset of cells in this branch does show lower expression of MYCN and FAM13A, an interesting observation is the elevated expression of HLA-B, indicative of a certain level of activated immune response. This phenomenon suggests a potentially more favorable prognosis for this branch. Conversely, the other branch, characterized by relatively lower proportions of HLA-B-expressing cells, demonstrates a gradual enhancement in the expression of MYCN and FAM13A with increasing malignancy. This comprehensive analysis underscores the potential pivotal roles of MYCN, FAM13A, and MHC molecules in dictating the trajectory of malignant progression in tumor cells.

Correlation between the expression of FAM13A and the clinical features of NB patients

Furthermore, based on the analysis of the Target database and R2 Platform, it was shown that the expression of FAM13A exhibited a significant correlation with the prognostic outcomes of patients with NB. Specifically, a greater level of FAM13A expression was found to be linked to a decrease in overall survival (OS) among NB patients (Fig. 3A-B, Fig. S1B). The expression level of FAM13A in patients with a high Children´s Oncology Group (COG) risk was significantly greater compared to patients with a low COG risk, as observed in both the Target and GEO databases (Fig. 3C). Furthermore, it was shown that patients with unfavorable NB exhibited considerably elevated levels of FAM13A expression compared to those with favorable NB (Fig. 3D). Moreover, FAM13A was predominantly expressed in NB patients with a high mitosis-karyorrhexis index (MKI) (Fig. 3E). And patients with diploid chromosomes expressed higher levels of FAM13A than patients with

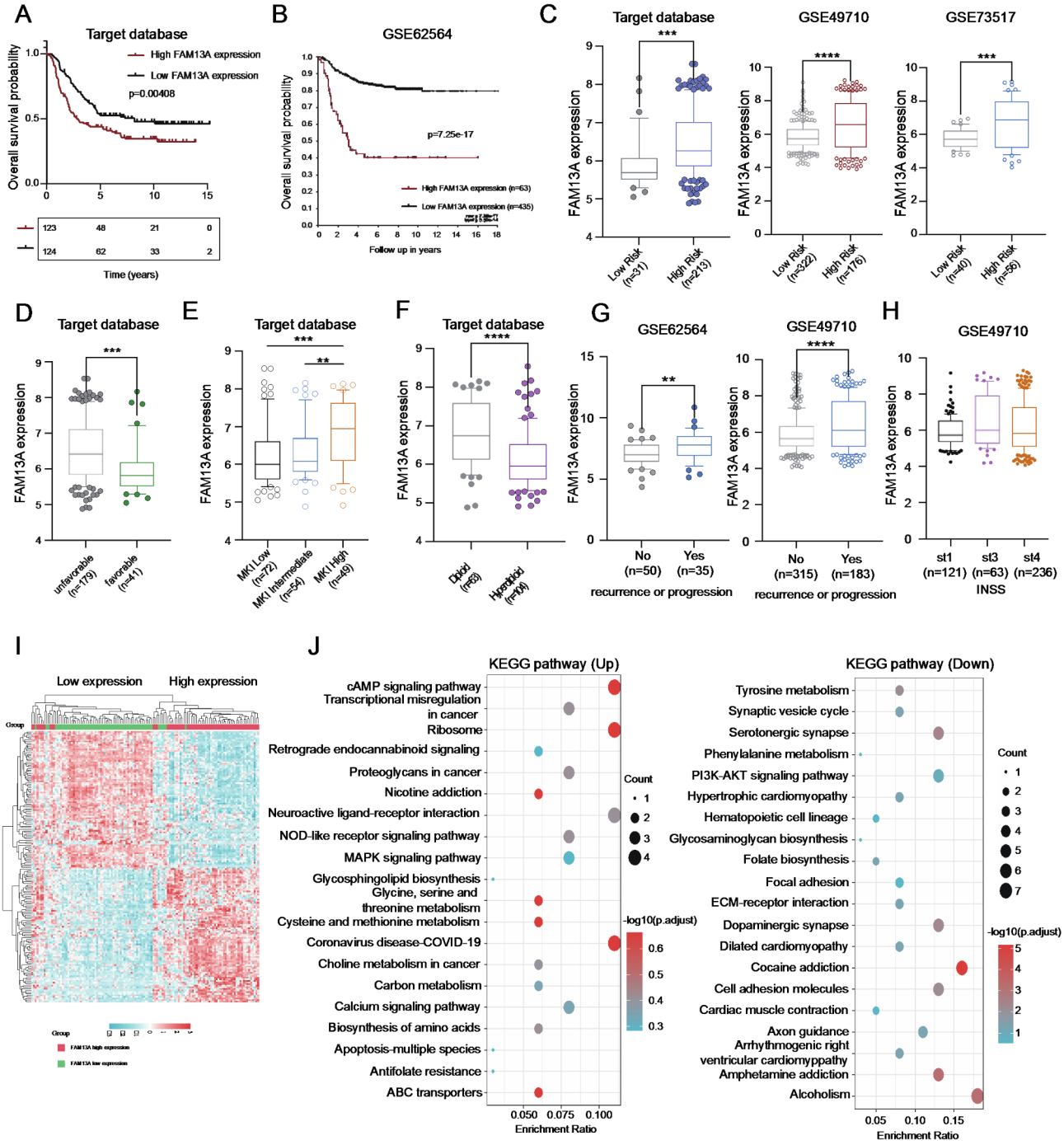


Fig. 3 FAM13A as a prognostic biomarker in neuroblastoma: correlation with clinical features, downstream pathways, and functional enrichment analysis. **(A-B)** Kaplan-Meier curves show the probability of overall survival based on FAM13A expression. **(C-H)** The correlation of FAM13A expression in patients with NB and various clinical stages. **(I)** Distinct colors indicate gene expression trends in distinct groups in the differential gene expression heatmap. This figure shows the top 25% up and down-regulated genes. **(J)** The enriched KEGG signaling pathways were chosen to demonstrate the primary biological actions of potential main mRNA. The abscissa represents gene ratio, while the ordinate displays enriched pathways. $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$; asterisks (*) represent levels of statistical significance. Two groups' statistical differences were contrasted using the student´s t test

hyper diploid chromosomes (Fig. 3F). Additionally, the expression of FAM13A was significantly elevated in those exhibiting progression of NB compared to those without progression (Fig. 3G). Further, it is worth noting that the expression of FAM13A is comparatively higher in stage 4 NB tumors as compared to stage 1 tumors. However, it is important to mention that this difference does not reach statistical significance ($p > 0.05$) (Fig. 3H). The above data

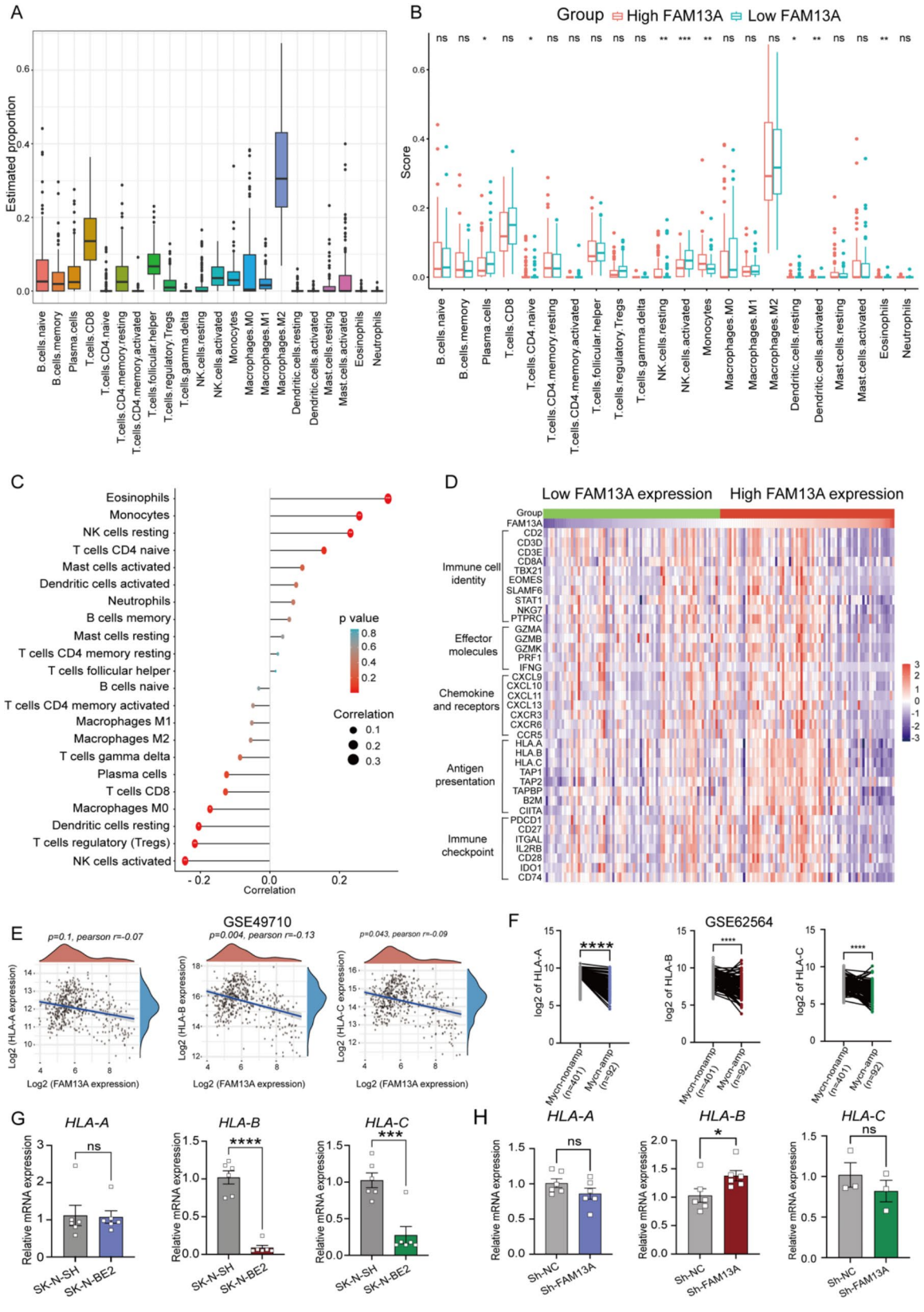


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 FAM13A Regulates Immune Infiltration and HLA Class I Expression in Neuroblastoma: Implications for Immunotherapy. **(A)** The proportions of 22 immune cells in NB patients based on Target database. **(B)** Graph showing the difference in CIBERSORT TIICs scores between high FAM13A and low FAM13A samples. **(C)** Relationship between FAM13A expression and immune cells. **(D)** Biomarkers for immunotherapy response between high FAM13A and low FAM13A samples. **(E)** FAM13A is inversely proportional to HLA-A, HLA-B and HLA-C. **(F)** The expression of FAM13A according to the MYCN status in NB patients. **(G)** qRT-PCR revealing HLA-A, HLA-B and HLA-C mRNA expression in MYCN-amplified and MYCN-non amplified cell lines. **(H)** qRT-PCR revealing HLA-A, HLA-B and HLA-C mRNA expression in NB stably transfected with sh-NC or sh-FAM13A. ns, not significant; *$p < 0.05$, ***$p < 0.001$, ****$p < 0.0001$; asterisks (*) represent levels of statistical significance. Two groups' statistical differences were contrasted using the student´s t test. Data were representative of at least three independent experiments

suggest that FAM13A has the potential to function as a novel prospective tumor marker for NB.

To examine the potential downstream genes of FAM13A and gain deeper insights into its mechanism in neuroblastoma (NB), we conducted an RNA-seq analysis comparing two groups of patients: the top 25% with high FAM13A expression and the bottom 25% with low FAM13A expression. This analysis was performed on a cohort of 249 NB patients from the Target database. Our goal was to identify differentially expressed genes (DEGs) between these two groups, which could shed light on the molecular pathways regulated by FAM13A in NB progression (Fig. 3I). To investigate the potential involvement of aberrant FAM13A expression in various biological pathways or processes, we conducted Gene Ontology (GO) terms and KEGG pathway analyses. These analyses were performed to evaluate the correlation of genes ($\log_2$ Fold Change > = 0.5, $p < 0.05$) based on Target database. The results of the enrichment analysis indicate that the most significantly enriched pathway and GO terms are the MAPK signaling pathway and cAMP-mediated signaling (Fig. 3J). It was revealed that genes associated with FAM13A exhibited enrichment in pathways pertaining to several biosynthetic, metabolic, and proliferative components of cellular physiology. These pathways include ABC transport, ribosome biogenesis, and biosynthesis of amino acids (Fig. 3J, Fig. S4A-B).

FAM13A expression negatively correlates with the level of immune infiltration in NB patients

Increased immune cell infiltration into malignancies is correlated with increased patient survival and predicts the efficacy of immune therapies. To determine whether FAM13A expression is associated with the degree of immune infiltration in NB, we examined the correlation between FAM13A expression and the degree of immune infiltration in NB samples obtained from the Target database. Figure 4A illustrates the estimated proportions of various immune cell types in NB samples, as determined by CIBERSORT using the LM22 algorithm. The presented bar graph depicted the relative proportions of immune cells for each patient, highlighting variations in the distribution of immune cells between individuals exhibiting high and low FAM13A expression levels (Fig. 4B, Fig. S5A). It is evident that the low FAM13A expression group exhibits greater proportions of $CD8^+$ T cells, resting $CD4^+$ memory T cells, activated NK cells, regulatory T cells, gamma delta T cells, M0 macrophages, M1 macrophages and M2 macrophages (Fig. 4C, Fig. S5B). Further research revealed that immune cell identity markers, effector molecules, chemokine and receptors, antigen presentation markers, and immune checkpoint markers were identified in immune infiltrated NB samples distinguished by FAM13A expression level (Fig. 4D). All immune cell markers (CD2, CD3D, CD3E, CD8A, TBX21, and NKG7) demonstrated a negative correlation with FAM13A expression (Fig. S6A). In addition, we examined HLA molecules, also referred to as MHC molecules, which include HLA class I, HLA class II, and HLA class III. Since NB cells lack surface HLA class I and II molecules, the T cell compartment of the host is likely to ignore them. As shown in Fig. 4E, the expression of HLA class I (HLA-A, HLA-B, and HLA-C) was significantly higher in the FAM13A-low expression group compared to the FAM13A-high expression group. And FAM13A and HLA class II (HLA-DQ) expressions were negatively correlated (Fig. S6B-D). Subsequently, we employed the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm to forecast the potential immune checkpoint blockade (ICB) response. This was accomplished by acquiring RNA-sequencing expression profiles (level 3) and the accompanying clinical data for NB from the Target database [26]. The TIDE methodology utilizes a set of gene expression signatures to assess two specific processes involved in the evasion of the immune system by tumors. These mechanisms include the dysfunction of cytotoxic T lymphocytes (CTLs) that infiltrate the tumor and the resistance of CTLs against immunosuppressive substances. A high TIDE score is indicative of suboptimal efficacy of immune checkpoint blockade (ICB) therapy and reduced overall survival following ICB treatment. As shown in Fig. S6E, the high FAM13A expression group had a high TIDE score, which is consistent with the fact that NB samples with low FAM13A expression exhibited robust immune infiltration. The preceding information suggests that patients with limited FAM13A expression may be more likely to respond to immunotherapy. Although the low-FAM13A expression group exhibited higher levels of HLA class I and II expressions, there was no statistically significant disparity observed in the expression of immunological checkpoint-related genes, including LAG3, TIGIT, and SIGLEC15 (Fig. S6F).

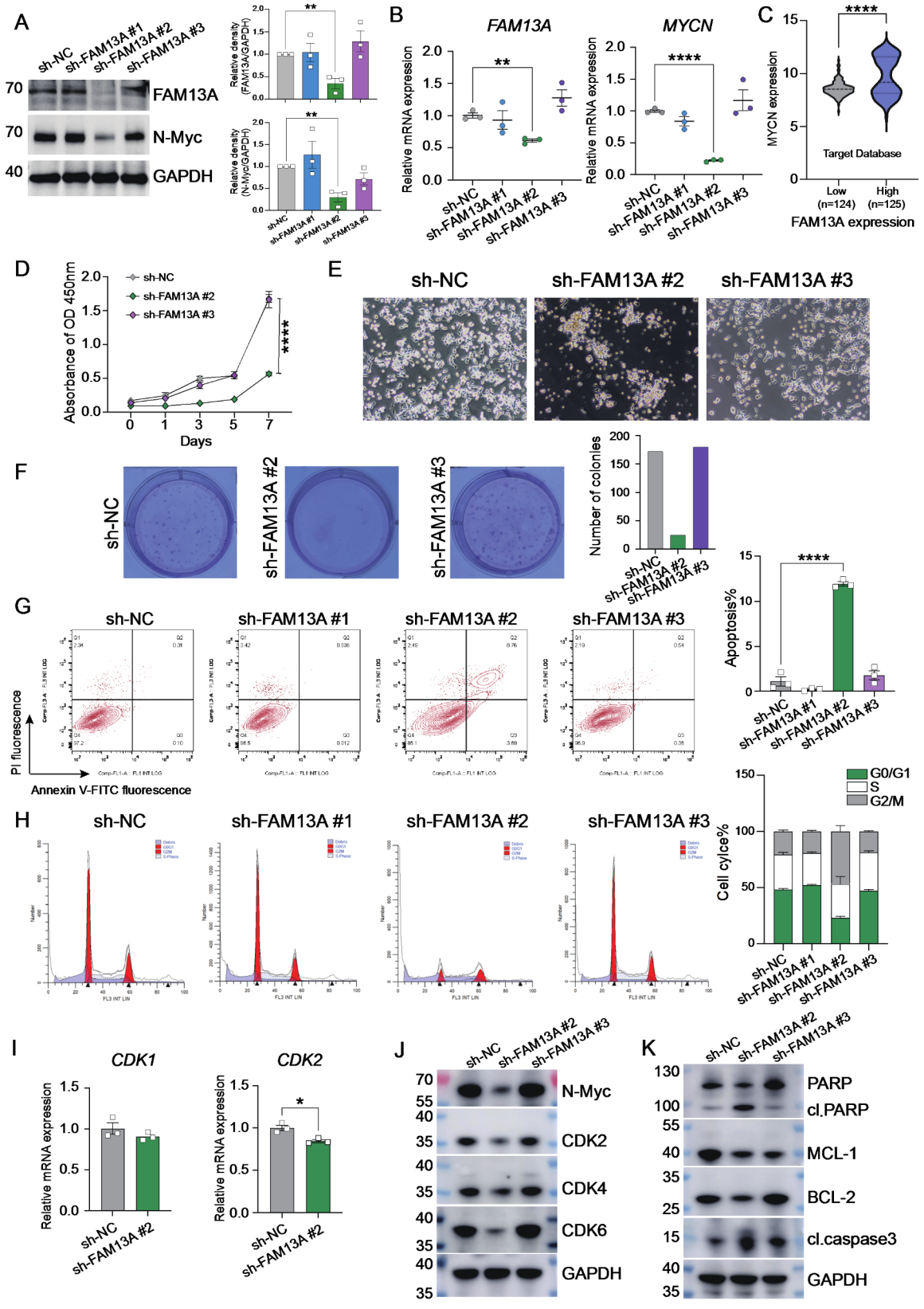


Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 FAM13A knockdown suppresses proliferation, induces apoptosis, and disrupts cell cycle in MYCN-amplified neuroblastoma cells. **(A-B)** FAM13A and N-Myc expression was detected by Western blot and qRT-PCR in NB stably transfected with sh-NC or sh-FAM13A. **(C)** Public datasets demonstrate the differential expression of MYCN in NB patients with differing FAM13A expression. **(D)** The CCK-8 assay reveals the difference in cell viability between NB cells transfected with sh-NC or sh-FAM13A. **(E)** Colony formation assay images and quantification depicting the growth of NB cells transfected with sh-NC or sh-FAM13A. **(F)** Representative images of NB cells infected with sh-NC or sh-FAM13A. **(G)** Flow cytometry analysis demonstrating the apoptosis of NB cells transfected with sh-NC or sh-FAM13A. Apoptotic index is defined as the ratio between Q2 and Q3 over total cells. **(H)** Flow cytometry depicting cell cycle distribution of SK-N-BE(2) cells transfected with either sh-NC or sh-FAM13A. The experiment is carried out in three replicates. **(I)** RT-qPCR showed CDK1 and CDK2 mRNA expression in NB cells transfected with sh-NC or sh-FAM13A. **(J)** Western blot revealing cell cycle related markers in NB cells transfected with sh-NC or sh-FAM13A. **(K)** Western blot revealing apoptosis related markers in NB cells transfected with sh-NC or sh-FAM13A. ns, not significant; *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$; asterisks (*) represent levels of statistical significance. Two groups' statistical differences were contrasted using the student´s t test. Data were representative of at least three independent experiments

Therefore, the high immune cell infiltration in the group with low FAM13A expression may be primarily attributable to the high MHC expression on the surface of NB cells. The downregulation of HLA class I is a commonly seen immune evasion mechanism in cancer cells. Notably, NB has a high frequency and severity of HLA loss, making it a prominent illustration of this phenomena. Following an examination of the GEO database, it was observed that patients diagnosed with NB and exhibiting MYCN amplification displayed notably reduced levels of surface HLA class I expression (Fig. 4F). Checking HLA-A, B, and C mRNA expression levels in both MYCN-amplified and MYCN-nonamplified cell lines, we discovered that HLA-B and HLA-C were extremely lowly expressed in the MYCN-amplified cell line (Fig. 4G). Knocking down FAM13A in the MYCN-amplified cell line could restore HLA class I (HLA-B) expression, which may be an especially promising strategy for enhancing T cell-mediated antitumor immunity in patients with MYCN-amplified NB (Fig. 4H). The results may have indicated an immune-inflamed microenvironment in the group with low FAM13A expression, suggesting immunotherapy susceptibility.

FAM13A sustains MYCN-amplified NB cells growth and survival

To investigate how FAM13A knockdown affects the proliferation of NB cell lines with MYCN amplification. First, we transfected FAM13A shRNA into SK-N-BE(2) cells and showed efficient knockdown of FAM13A expression in both protein and mRNA level (Fig. 5A-B). Considering the positive correlation between FAM13A and MYCN expression, our study aimed to examine the phenotypic impact of FAM13A on MYCN expression in NB cells with MYCN amplification. Our findings revealed a similar reduction in both MYCN mRNA and protein expression in SK-N-BE(2) cells (Fig. 5A-B). The results obtained from the Target database are consistent with the experimental results, and MYCN also has a higher expression level in the FAM13A high expression group (Fig. 5C). Analysis of the FAM13A gene promoter sequence and published chromatin immunoprecipitation (ChIP)-sequencing data revealed MYCN binding sites near the FAM13A transcription start site (TSS) [27]. Using the motif database and investigating the FAM13A gene promoter region with JASPAR (http://jaspardev.genereg.net) predicted a potential binding site for MYCN. We performed CUT&RUN assays with an anti-N-Myc antibody or control IgG antibody followed by real-time PCR with primers targeting a negative control and the MYCN peak summit near the FAM13A TSS. CUT&RUN assays showed significant MYCN binding at its putative binding site. (Fig. S7). The knockdown of FAM13A resulted in a decrease in cell growth when compared to cells transfected with scrambled shRNA (Fig. 5D). Upon microscopic examination, it was seen that the suppression of FAM13A led to a notable suppression of cellular growth in the NB cell line (Fig. 5E). Moreover, when compared to the transfection of scrambled negative control (sh-NC), the suppression of FAM13A exhibited a significant reduction in NB proliferation as observed in the colony formation experiment (Fig. 5F). Collectively, these data suggest that NB cell proliferation is functionally dependent on FAM13A signaling. Flow cytometry for apoptosis studies demonstrated that FAM13A-knockdown could hamper tumor cell proliferation and induce apoptosis (Fig. 5G). The FAM13A-knockdown group consistently exhibited heightened caspase-3 activation and PARP cleavage, both of which serve as indicators of cellular death. Furthermore, the expression levels of BCL2 and MCL1, which are well-known pro-survival proteins involved in inhibiting cell death (apoptosis) in cancer, were seen to be significantly low in cells where FAM13A had been knocked down (Fig. 5K). The results of cell cycle research demonstrated that the downregulation of FAM13A in SK-N-BE(2) cells resulted in a significant reduction in the number of cells in the G1 phase, while concurrently exhibiting an elevation in the number of cells in the G2 phase (Fig. 5H). Consistently, using qRT-PCR to detect cell cycle-related genes in each group of cells, following the knockdown of FAM13A, there was a notable reduction in the mRNA levels of CDK1 and CDK2 (Fig. 5I). Additionally, the WB analysis revealed a decrease in the expression levels of CDK2, CDK4, and CDK6 cell cycle proteins in the FAM13A knockdown group (Fig. 5J). Collectively, the findings of our investigation collectively

demonstrate that the suppression of FAM13A has a detrimental impact on the proliferation and viability of NB cells.

Protein structure optimization of N-Myc and FAM13A

MYC family proteins are intrinsically disordered proteins, and the structure of N-Myc is unstable, thus its experimental structure is presently unknown. As shown in Fig. 6A, we obtained the amino acid sequence of N-Myc from NCBI and used DeepMind's artificial intelligence system Alphafold v2.3.2 to predict the protein structure of N-Myc (pTM = 0.27). The Ramachandran plots, a common instrument for analyzing protein secondary structure abnormalities, were then utilized to

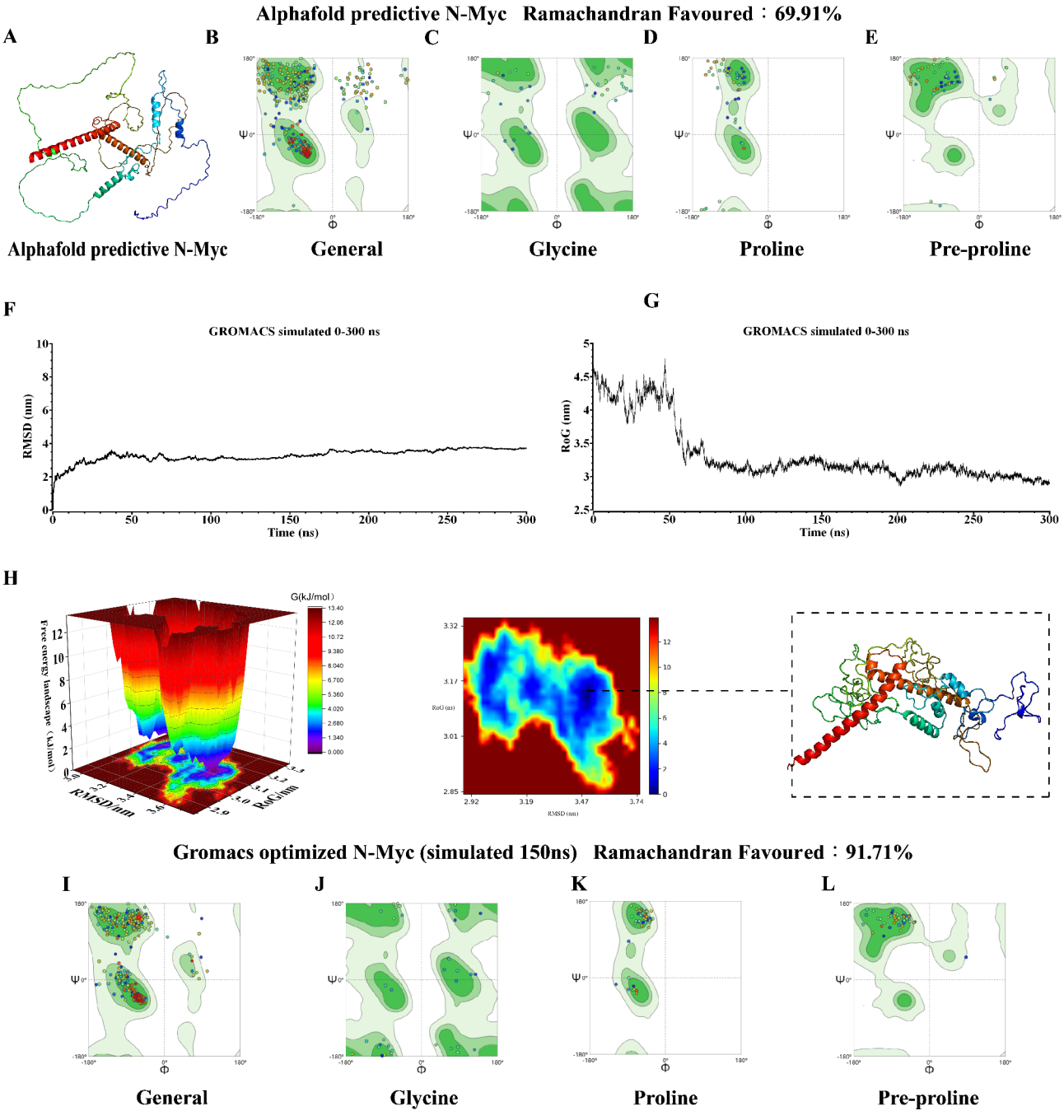


Fig. 6 Molecular dynamics simulation optimizes the predicted structure of N-Myc for accurate protein-protein docking studies. **(A)** Alphafold v2.3.2 to predict the protein structure of N-Myc. **(B-E)** The Ramachandran plots were then utilized to assess the accuracy of the N-Myc predicted structures. **(F-G)** The results of RMSD and RoG show that the protein structure of N-Myc tends to be stable. **(H)** Free energy topography of N-Myc after stabilization. **(I-L)** The Ramachandran plots were then utilized to assess the accuracy of the N-Myc optimized structures

assess the accuracy of the N-Myc predicted structures. Each point in the Ramachandran plots represents the ϕ and ψ angles of an amino acid residue, and various amino acid residues occupy different locations on the graph. In most cases, reasonable amino acid dihedral angles are distributed in a particular region of the graph known as the tolerable region. Points that descend in the wild represent deviations in the dihedral angle of the amino acid residue. As shown in Fig. 6B-E, only 69.91% of all amino acid residues fell within the permitted region, and the dihedral angles of a significant number of amino acid residues were aberrant. The abnormal dihedral angle of amino acid residues will lead to the aberrant spatial conformation of the protein, indicating that the N-Myc protein structure predicted by Alphafold is insufficiently accurate and cannot be used directly in experiments. Therefore, in this study, the GROMACS molecular dynamics simulation method was utilized to place the Alphafold structure of the N-Myc protein into the water box of the simulated 1.2 nm^3 CHARMM36 viewpoint. Adding 0.15 M Na^+ and Cl^- to the simulation space makes the entire simulation system electrically balanced, which is as similar as possible to the actual environment of N-Myc in the cell. Through the thermal movement of N-Myc in the simulated system, the predicted structure of N-Myc progressively folds into a more stable and real-space conformation under the influence of water, ions, and the force field. As shown in Fig. 6F-G, after about 50 ns of simulation, the protein structure of N-Myc tends to be stable. Based on the results of Root-mean-square deviation (RMSD) and Radius of Gyration (RoG), we plotted the free energy topography. As shown in Fig. 6H, the free energy topography of N-Myc after stabilization has a single energy well. In the free energy topography diagram, a single energy well means the lowest energy point and the most stable conformation of N-Myc during the simulation. Because when a protein is in a single energy well, it has the lowest energy at that position and is therefore most likely to stay there. We intercept the N-Myc conformation at the lowest point of the energy well at the simulation time of 150 ns as our optimized N-Myc conformation. As depicted in Fig. 6I-L, the simulation-optimized N-Myc has 91.71% of the amino acid residues within the permissible region and the aberrant dihedral angles of the majority of the amino acid residues have been optimized. The optimized N-Myc is utilized for subsequent protein-protein docking experiments as it matches the experimental requirements.

The experimental structure of FAM13A is currently unknown. In this study, the amino acid sequence of FAM13A was obtained from NCBI and the artificial intelligence system Alphafold v2.3.2 developed by DeepMind was used to predict the protein structure of FAM13A (pTM = 0.33), as depicted in Fig. 7A. The precision of the FAM13A predicted structure was evaluated using Ramachandran plots. As depicted in Fig. 7B-E, only 70.52% of all amino acid residues fell within the permissible region, indicating that the FAM13A protein structure predicted by Alphafold was insufficiently accurate and could not be utilized directly in experiments. In this investigation, the Alphafold-predicted FAM13A protein structure was optimized using the GROMACS molecular dynamics simulation technique. As depicted in Fig. 7F-G, the protein structure of FAM13A tends to be stable after approximately 20 ns of simulation. Based on the results of RMSD and RoG, we plotted the free energy topography. As shown in Fig. 7H, the free energy topography of FAM13A after stabilization has a single energy well. We intercept the FAM13A conformation at the lowest point of the energy well at the simulated time of 230 ns as our optimized FAM13A conformation. Consequently, the spatial conformation of FAM13A at 230 ns interception simulation is depicted in Fig. 7H, and the stability and authenticity of the FAM13A structure are once more confirmed using Ramachandran plots. As shown in Fig. 7I-L, after simulation optimization, 89.68% of FAM13A's amino acid residues fell within the permissible region, and the aberrant dihedral angles of most amino acid residues were optimized. The optimized and simulated FAM13A was used for subsequent protein-protein interaction experiments as it essentially met the experimental requirements.

Simulation of N-Myc and FAM13A interaction by protein-protein Docking and GROMACS

Schrödinger 2023-1 was used in this study to calculate the protein-protein docking of N-Myc and FAM13A, as shown in Fig. 8A N-Myc and FAM13A are stably combined in the form of hydrogen bonds and salt bridges. The molecular dynamics simulation method was used to simulate the 100 ns motion trajectory of the binding conformation of N-Myc and FAM13A. As shown in Fig. 8B-C, the structures of RMSD and RoG show that N-Myc and FAM13A can form a stable complex. As shown in the free energy topography map drawn based on the RMSD and RoG results (Fig. 8D), the free energy topography map of the N-Myc and FAM13A complex has a single energy well. The conformation of the complex intercepting the lowest point of the energy well is shown in Fig. 8D. The conformation and binding mode of the complex did not change significantly during the simulation process. This shows that the combination of N-Myc and FAM13A is stable. The results of Root Mean Square Fluctuation (RMSF) (Fig. 8E) and Solvent Accessible Surface Area (SASA) (Fig. 8F) also support the above conclusion. As we have shown via protein docking that FAM13A and N-Myc can bind stably, we aimed to explore the possibility of a physical interaction between

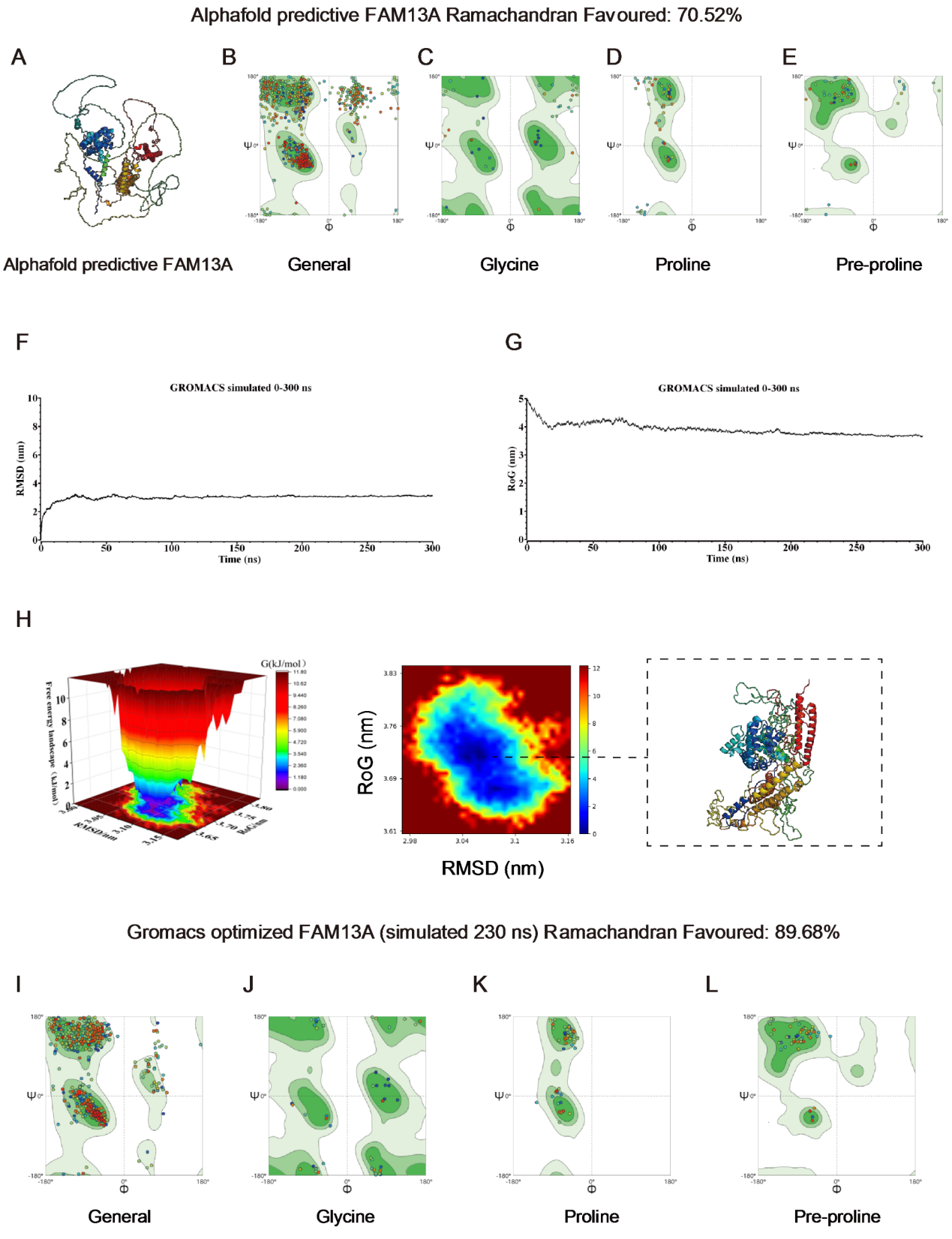
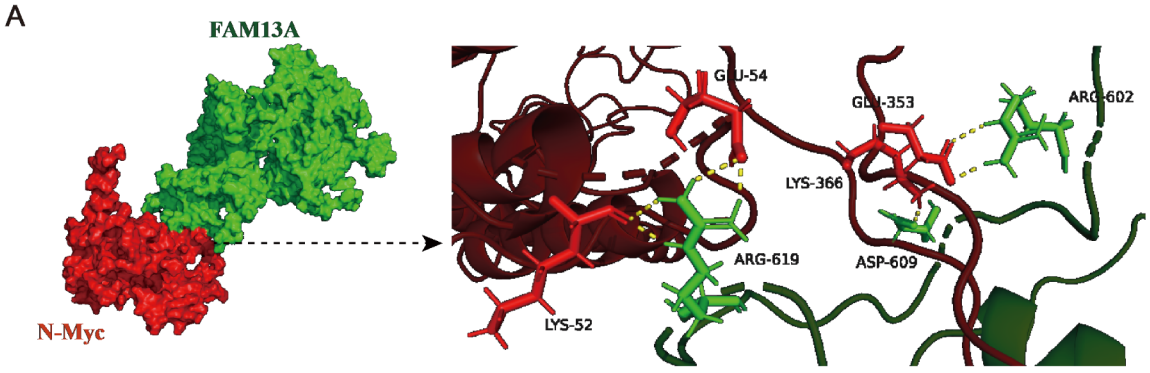


Fig. 7 Molecular dynamics simulation optimizes the predicted structure of FAM13A for accurate protein-protein docking studies. **(A)** Alphafold v2.3.2 to predict the protein structure of FAM13A. **(B-E)** The Ramachandran plots were then utilized to assess the accuracy of the FAM13A predicted structures. **(F-G)** The results of RMSD and RoG show that the protein structure of FAM13A tends to be stable. **(H)** Free energy topography of FAM13A after stabilization. **(I-L)** The Ramachandran plots were then utilized to assess the accuracy of the FAM13A optimized structures



Hydrogen Bonding

FAM13A (Receptor)	MYCN(ligand)	Distance(Å)	Number
A:Arg 619	B:Lys 52	2	2
A:Arg 619	B:Glu 54	1.9	2
A:Asp 609	B:Lys 366	1.9	1
A:Arg 602	B:Glu 353	2.1	1

Salt Bridges

FAM13A (Receptor)	MYCN(ligand)	Distance(Å)	Number
A:Arg 619	B:Glu 54	1.9	1
A:Asp 609	B:Lys 366	1.9	1
A:Arg 602	B:Glu 353	2.1	1

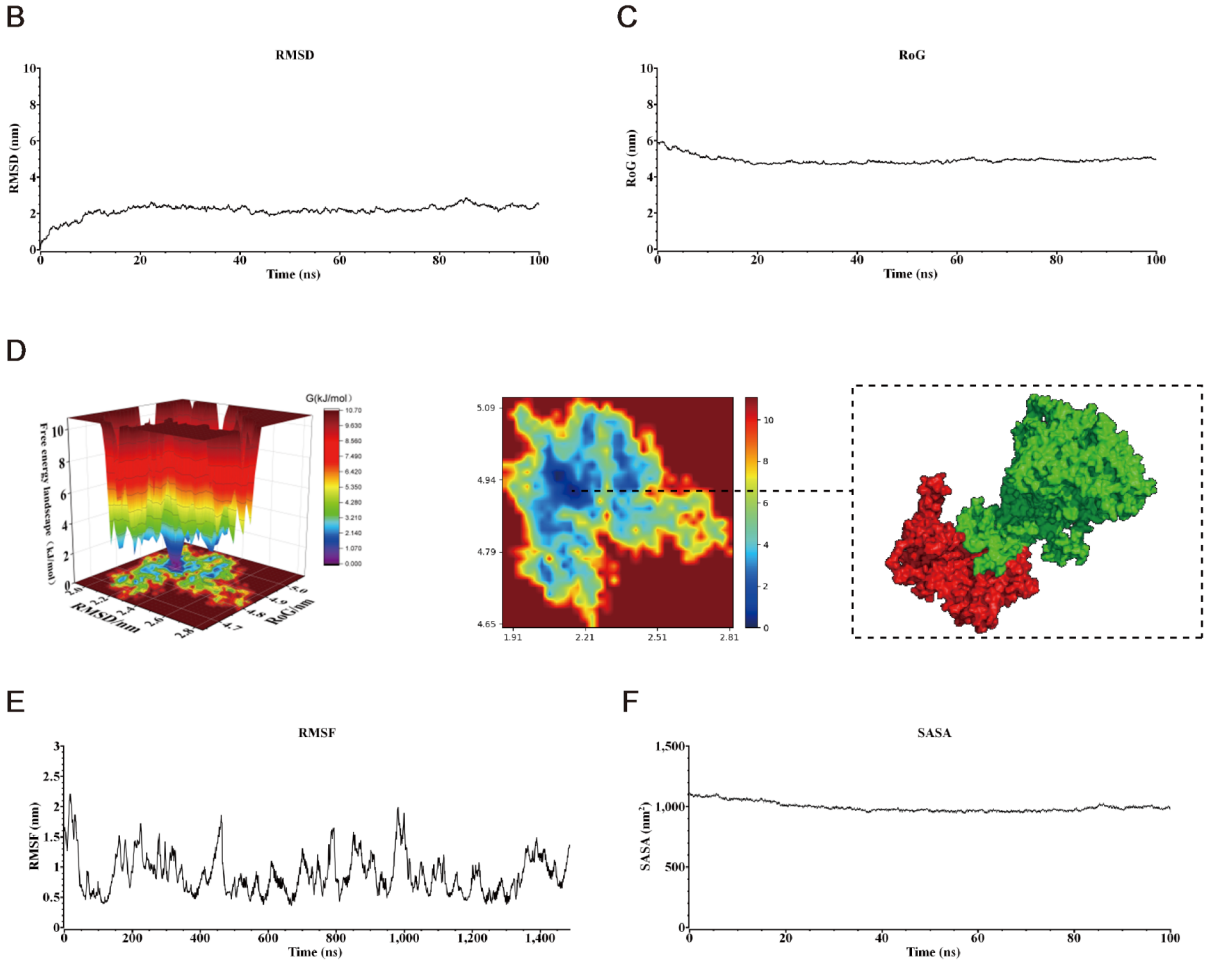


Fig. 8 Simulation of N-Myc and FAM13A interaction by protein-protein docking and GROMACS. **(A)** The binding model of the complex FAM13A with N-Myc. The detail binding model of FAM13A with N-Myc (right). **(B-C)** The RMSD and RoG values of 100 ns-long trajectories of molecular dynamics simulations were calculated. **(D)** Free energy topography plot based on RMSD and RoG results. **(E-F)** RMSF and SASA results show that the combination of N-Myc and FAM13A is stable

them. We transfected HEK293T human embryonic kidney cells with a construct expressing Flag-tagged MYCN either alone, or one expressing HA-tagged FAM13A. A co-immunoprecipitation (Co-IP) of total cellular protein with an HA-tag antibody for FAM13A demonstrated that FAM13A-bound N-Myc (Fig. S8).

Discussion

MYCN amplification diminishes the efficacy of neuroblastoma (NB) and is the main cause of treatment failures. Silencing MYCN triggers NB differentiation and apoptosis [28, 29]. However, RNA interference therapies are limited by the lack of suitable drug delivery vehicles [30]. Current research focuses on synthesizing and screening chemicals interacting with MYC protein [31, 32]. Small-molecule antagonists can block MYC-induced transcription, but direct pharmacological suppression of MYCN is challenging due to lack of specificity and in vivo efficiency [33]. Only modest progress has been made in controlling MYCN-dependent expression and activities. Indirect targeting strategies for MYCN-interacting molecules have advanced and modifying indirectly targeted MYCN genes may block malignant traits [34, 35]. Exploring regulatory elements and developing new techniques centered on MYCN amplification´s activation mechanisms, especially "selective inhibition", is highly relevant.

Researchers analyzed MYCN-amplified cells via co-immunoprecipitation and identified MYC(N) interactors. EZH2 is a direct and reliable binding partner of MYC(N). Depleting EZH2 induces MYC(N) degradation and suppresses tumor cell growth in MYC(N)-driven NB and small cell lung cancer. EZH2 may be a new target for NB treatment by indirectly regulating MYCN stability [36]. After analyzing NB patients´ whole transcriptome in Target database, ALYREF was differentially expressed in those with 17q21-ter gain or MYCN amplification. Down-regulating USP3 by inhibiting ALREF-MYCN complex disrupts N-Myc protein homeostasis, changes K48-linked polyubiquitin chain abundance on MYCN and causes proteasome degradation. The ALYREF-USP3-MYCN complex is a potential therapeutic target for NB by exploiting obligatory ALYREF dependence in MYCN-amplified cells [37]. In addition, studies have identified ALDH18A1 as a potential risk factor with prognostic value for NB patients, especially those with MYCN amplification. ALDH18A1 is a regulator of NB cell proliferation, self-renewal, and tumorigenicity. It regulates MYCN expression and forms a positive feedback loop with MYCN [38].

In this study, we initially utilized machine learning algorithms to screen for MYCN signature genes to discover new targets that can indirectly target MYCN. In screening disease characteristic genes from differentially expressed genes, machine learning methods exhibit notable advantages over traditional approaches. Random forest, integrating multiple decision trees, can utilize data well to capture complex gene-disease association patterns and has much higher accuracy than traditional simple statistical index-based screening methods. The bootstrap resampling technique reduces single model error and improves overall accuracy [39]. Support Vector Machine Recursive Feature Elimination (SVM-RFE) shows excellence in handling high-dimensional gene data. It finds the optimal hyperplane and uses recursive elimination to screen key genes, solving the problems of dimension disaster and overfitting to noisy data that traditional methods encounter [40]. By combining random forest and SVM-RFE, this study achieves unique advantages. In terms of accuracy, random forest initially screens key genes, and SVM-RFE further removes redundant features, enhancing accuracy and reducing misjudgment. In handing high-dimensional data, SVM-RFE selects features to reduce dimensions, while random forest adapts to many gene features and mines complex relationships. Together, they form a multi-level processing system for representative and biologically significant characteristic genes. In terms of adaptability, the combination can respond flexibly to different disease data and find the most representative disease characteristic genes. In our study, there are five MYCN feature genes that can be screened by both machine learning algorithms: PHGDH, FAM13A, KLRG2, RAB3B, and NPW. Through subsequent analysis, we targeted FAM13A for further experimental verification. FAM13A has a RhoGAP domain that activates small GTPases [41]. Rho GTPase signaling is important for cell proliferation and survival [6]. FAM13A is involved in cancers and tumor cell proliferation [6, 25]. However, the function of FAM13A in MYCN-amplified NB remains unclear. Our study identified FAM13A as a risk factor and key regulator of NB cell growth in MYCN-amplified NB. FAM13A can influence MYCN expression. Exploring FAM13A as a potential cell-based technique for restoring HLA expression on NB cells is a novel therapeutic approach for patients with MYCN-amplified NB, aiming to enhance T cell-mediated anti-tumor immune responses. The innovation lies in using GROMACS molecular dynamic simulation technique to optimize the protein structures of MYCN and FAM13A, which were first predicted by the Alphafold algorithm. Then, Schrödinger software was used to implement protein-protein docking to confirm the stability of the interaction between MYCN and FAM13A. Experiments prove they can bind. The results show that FAM13A binding to MYCN enhances its stability and reduces proteasomal degradation. Targeting the FAM13A-MYCN complex is a potential therapeutic avenue for MYCN-amplified NB patients.

Numerous studies have shown that destabilizing MYCN could be a promising therapeutic approach. However, challenges in specificity and clinical translation persist. We have uncovered that FAM13A, functioning as a cofactor, can stabilize MYCN and promote the growth of NB cells, suggesting that it may potentially serve as a specific therapeutic target for MYCN. As of now, a definite FAM13A inhibitor has not been developed. We believe that the development of small molecule compounds aimed at disrupting the interaction between FAM13A and MYCN holds significant research value, and these compounds are expected to act as potential inhibitors. The potential of FAM13A as a therapeutic target for MYCN-amplified NB is exhilarating. Nonetheless, the absence of animal experiments constitutes a major limitation of this study. Animal models can offer crucial insights into the efficacy and safety of targeting FAM13A.

In summary, although our study indicates that FAM13A may be a MYCN-specific therapeutic target for MYCN-amplified NB, further research employing animal models is required to validate these findings and pave the way for clinical translation. This represents an important area for further research and also a shortcoming of the current study. In addition, due to its own characteristics, MD simulation has many influencing factors, such as floating - point precision and calculation order. These factors can cause differences in results between different runs, thus affecting the reproducibility of molecular dynamics simulations. Different from real-world experiments, the simulation experiment process is not continuous. Instead, the simulation process is truncated into specific time points for analysis. Each truncation will introduce truncation errors and rounding errors. The current floating-point representation precision of computers is limited, resulting in the non - commutativity of addition operations. In computers, is not always equal to. Therefore, when performing cumulative operations, if the cumulative order cannot be precisely specified, different results are likely to be obtained. Especially to improve computing power, molecular dynamics simulations are usually performed in parallel on a Graphics Processing Unit (GPU). Since GPUs have many cores that can run simultaneously, it is even more difficult to ensure the cumulative order. Limited by the above-mentioned precision and reproducibility issues of molecular dynamics simulations, the molecular dynamics simulation experiments used in this study are only a preliminary verification of the binding situation between FAM13A and N-Myc. Although it is only a preliminary verification, it is of certain significance for promoting the progress of the experiment. At the same time, this study also carried out a Co-IP experiment to supplement and orthogonally verify the results of virtual modeling verification. The results of the Co-IP experiment showed that there is a binding between FAM13A and N-Myc.

Abbreviations

NB	Neuroblastoma
FAM13A	Family with sequence similarity 13 member A
SVM	Support vector machine
RFE	Recursive feature elimination
RMSE	Root mean square error
CCK8	Cell counting kit-8
CCLE	Cancer Cell Line Encyclopedia
TARGET	Therapeutically Applicable Research to Generate Effective Treatments
GEO	Gene Expression Omnibus
RMSD	Root mean square deviation
MD	Molecular Dynamics
RMSF	Root mean square fluctuation
RoG	Radius of Gyration
SASA	Solvent Accessible Surface Area
FEL	Free energy landscape
SEM	Mean ± standard error of the mean
COG	Children´s Oncology Group
OS	Overall survival
MKI	Mitosis-karyorrhexis index
GO	Gene Ontology
TIDE	Tumor Immune Dysfunction and Exclusion
ICB	Immune checkpoint blockade
CTLs	Cytotoxic T lymphocytes
ChIP	Chromatin immunoprecipitation
CUT&RUN	Cleavage Under Target and Release Using Nuclease

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-025-13903-9.

Supplementary Material 1

Supplementary Material 2

Acknowledgements
Not applicable.

Author contributions
Hongli Yin: Conceptualization, Methodology, Software, Writing. Tianyi Liu: Methodology, Investigation, Writing. Di Wu: Experiments, Writing. Xiaolu Li: Software, Methodology. Gen Li: Methodology. Weiwei Song: Methodology. Xiaodong Wang: Methodology. Shan Xin: Software, Supervision. Yisu Liu: Single cell analysis, Project administration. Jian Pan: Conceptualization, Supervision, Project administration, Writing.

Funding
This work was supported by grants from the National Natural Science Foundation (82404067); Science and Technology Development Project of Suzhou City (SKJY2021111, SKJY2021112, SKY2023193); Jiangsu Provincial Health Commission Scientific Research Project (Z2022031); Jiangsu Province Key R&D Plan (Social Development) Project (BE2022732); The Natural Science Foundation of Jiangsu Province (BK20230217); Suzhou Applied Basic Research (Medical and Health) Science and Technology Innovation Project (SYW2024105); Dalian Science and Technology Innovation Fund (2024JJ13PT056).

Data availability
The protein structures predicted by AlphaFold and those optimized by MD are publicly accessible on Github at https://github.com/Hongli-Yin/Hongli-Yin. However, due to the large size of the MD simulation trajectory files, they could not be uploaded. Interested parties may contact the corresponding author（Hongli Yin） to request access to these trajectory files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 14 November 2024 / Accepted: 10 March 2025
Published online: 14 March 2025

References

1. London WB, Castleberry RP, Matthay KK, Look AT, Seeger RC, Shimada H, Thorner P, Brodeur G, Maris JM, Reynolds CP, Cohn SL. Evidence for an age cutoff greater than 365 days for neuroblastoma risk group stratification in the children's oncology group. JCO. 2005;23:6459–65. https://doi.org/10.1200/JCO.2005.05.571.
2. Brodeur GM. Neuroblastoma: biological insights into a clinical enigma. Nat Rev Cancer. 2003;3:203–16. https://doi.org/10.1038/nrc1014.
3. Huang M, Weiss WA, Neuroblastoma. MYCN Cold Spring Harb Perspect Med. 2013;3:a014415. https://doi.org/10.1101/cshperspect.a014415.
4. Bartolucci D, Montemurro L, Raieli S, Lampis S, Pession A, Hrelia P, Tonelli R. MYCN impact on High-Risk neuroblastoma: from diagnosis and prognosis to targeted treatment. Cancers (Basel). 2022;14:4421. https://doi.org/10.3390/cancers14184421.
5. Ruiz-Pérez MV, Henley AB, Arsenian-Henriksson M. The MYCN protein in health and disease. Genes (Basel). 2017;8:113. https://doi.org/10.3390/genes8040113.
6. Ziółkowska-Suchanek I, Podralska M, Żurawek M, Łaczmańska J, Iżykowska K, Dzikiewicz-Krawczyk A, Rozwadowska N. Hypoxia-Induced FAM13A regulates the proliferation and metastasis of Non-Small cell lung cancer cells. Int J Mol Sci. 2021;22:4302. https://doi.org/10.3390/ijms22094302.
7. Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Hersh CP, DeMeo DL, Hunninghake GM, Litonjua AA, Sparrow D, Lange C, Won S, Murphy JR, Beaty TH, Regan EA, Make BJ, Hokanson JE, Crapo JD, Kong X, Anderson WH, Tal-Singer R, Lomas DA, Bakke P, Gulsvik A, Pillai SG. Silverman, variants in FAM13A are associated with chronic obstructive pulmonary disease. Nat Genet. 2010;42:200–2. https://doi.org/10.1038/ng.535.
8. Tang J, Zhou H, Sahay K, Xu W, Yang J, Zhang W, Chen W. Obesity-associated family with sequence similarity 13, member A (FAM13A) is dispensable for adipose development and insulin sensitivity. Int J Obes. 2019;43:1269–80. https://doi.org/10.1038/s41366-018-0222-y.
9. Fathzadeh M, Li J, Rao A, Cook N, Chennamsetty I, Seldin M, Zhou X, Sangwung P, Gloudemans MJ, Keller M, Attie A, Yang J, Wabitsch M, Carcamo-Orive I, Tada Y, Lusis AJ, Shin MK, Molony CM, McLaughlin T, Reaven G, Montgomery SB, Reilly D, Quertermous T, Ingelsson E, Knowles JW. FAM13A affects body fat distribution and adipocyte function. Nat Commun. 2020;11:1465. https://doi.org/10.1038/s41467-020-15291-z.
10. Lin X, Liou Y-H, Li Y, Gong L, Li Y, Hao Y, Pham B, Xu S, Jiang Z, Li L, Peng Y, Qiao D, Lin H, Liu P, Wei W, Zhang G, Lee C-H, Zhou X. FAM13A Represses AMPK Activity and Regulates Hepatic Glucose and Lipid Metabolism, iScience (2020). https://www.semanticscholar.org/paper/FAM13A-Represses-AMPK-Activity-and-Regulates-and-Lin-Liou/b216f132c056d0b38b2e9594a287149a06d3f218 (accessed June 11, 2023).
11. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F, e, Wien TU. (1999) 1.7–14. https://doi.org/10.32614/CRAN.package.e1071
12. Karatzoglou A, Smola A, Hornik K, Zeileis A. kernlab - An S4 package for kernel methods in R. J Stat Softw. 2004;11:1–20. https://doi.org/10.18637/jss.v011.i09.
13. Kuhn M. Building predictive models in R using the caret package. J Stat Softw. 2008;28:1–26. https://doi.org/10.18637/jss.v028.i05.
14. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta delta C(T)) method. Methods. 2001;25:402–8. https://doi.org/10.1006/meth.2001.1262.
15. Monteverde T, Sahoo S, La Montagna M, Magee P, Shi L, Lee D, Sellers R, Baker AR, Leong HS, Fassan M, Garofalo M. CKAP2L promotes Non–Small cell lung cancer progression through regulation of transcription elongation. Cancer Res. 2021;81:1719–31. https://doi.org/10.1158/0008-5472.CAN-20-1968.
16. Kildisiute G, Kholosy WM, Young MD, Roberts K, Elmentaite R, van Hooff SR, Pacyna CN, Khabirova E, Piapi A, Thevanesan C, Bugallo-Blanco E, Burke C, Mamanova L, Keller KM, Langenberg-Ververgaert KPS, Lijnzaad P, Margaritis T, Holstege FCP, Tas ML, Wijnen MHWA, van Noesel MM, del Valle I, Barone G, van der Linden R, Duncan C, Anderson J, Achermann JC, Haniffa M, Teichmann SA, Rampling D, Sebire NJ, He X, de Krijger RR, Barker RA, Meyer KB, Bayraktar O, Straathof K, Molenaar JJ, Behjati S. Tumor to normal single-cell mRNA comparisons reveal a pan-neuroblastoma cancer cell. Sci Adv. 2021;7:eabd3311. https://doi.org/10.1126/sciadv.abd3311.
17. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19:15. https://doi.org/10.1186/s13059-017-1382-0.
18. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. Methods Mol Biol. 2018;1711:243–59. https://doi.org/10.1007/978-1-4939-7493-1_12.
19. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters, OMICS 16 (2012) 284–287. https://doi.org/10.1089/omi.2011.0118
20. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for gene set enrichment analysis. Bioinformatics. 2007;23:3251–3. https://doi.org/10.1093/bioinformatics/btm369.
21. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, Petersen S, Reiman D, Clancy E, Zielinski M, Steinegger M, Pacholska M, Berghammer T, Bodenstein S, Silver D, Vinyals O, Senior AW, Kavukcuoglu K, Kohli P, Hassabis D. Highly accurate protein structure prediction with alphafold. Nature. 2021;596:583–9. https://doi.org/10.1038/s41586-021-03819-2.
22. Watanabe K, Kimura S, Seki M, Isobe T, Kubota Y, Sekiguchi M, Sato-Otsubo A, Hiwatari M, Kato M, Oka A, Koh K, Sato Y, Tanaka H, Miyano S, Kawai T, Hata K, Ueno H, Nannya Y, Suzuki H, Yoshida K, Fujii Y, Nagae G, Aburatani H, Ogawa S, Takita J. Identification of the ultrahigh-risk subgroup in neuroblastoma cases through DNA methylation analysis and its treatment exploiting cancer metabolism. Oncogene. 2022;41:4994–5007. https://doi.org/10.1038/s41388-022-02489-2.
23. Miranda A, Hamilton PT, Zhang AW, Pattnaik S, Becht E, Mezheyeuski A, Bruun J, Micke P, De Reynies A, Nelson BH. Cancer stemness, intratumoral heterogeneity, and immune response across cancers, Proc. Natl. Acad. Sci. U.S.A. 116 (2019) 9020–9029. https://doi.org/10.1073/pnas.1818210116
24. Mondal MS, Yamaguchi H, Date Y, Tsuruta T, Shimbara T, Toshinai K, Shimomura Y, Mori M, Nakazato M. Ontogeny of a new enteric peptide, neuropeptide W (NPW), in the developing rat stomach. Regul Pept. 2008;145:141–6. https://doi.org/10.1016/j.regpep.2007.09.029.
25. Eisenhut F, Heim L, Trump S, Mittler S, Sopel N, Andreev K, Ferrazzi F, Ekici AB, Rieker R, Springel R, Assmann VL, Lechmann M, Koch S, Engelhardt M, Warnecke C, Trufa DI, Sirbu H, Hartmann A, Finotto S. FAM13A is associated with non-small cell lung cancer (NSCLC) progression and controls tumor cell proliferation and survival. OncoImmunology. 2017;6:e1256526. https://doi.org/10.1080/2162402X.2016.1256526.
26. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, Wucherpfennig KW, Liu XS. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24:1550–8. https://doi.org/10.1038/s41591-018-0136-1.
27. Zeid R, Lawlor MA, Poon E, Reyes JM, Fulciniti M, Lopez MA, Scott TG, Nabet B, Erb MA, Winter GE, Jacobson Z, Polaski DR, Karlin KL, Hirsch RA, Munshi NP, Westbrook TF, Chesler L, Lin CY, Bradner JE. Enhancer invasion shapes MYCN dependent transcriptional amplification in neuroblastoma. Nat Genet. 2018;50:515–23. https://doi.org/10.1038/s41588-018-0044-9.
28. Kang J-H, Rychahou PG, Ishola TA, Qiao J, Evers BM, Chung DH. MYCN Silencing induces differentiation and apoptosis in human neuroblastoma cells. Biochem Biophys Res Commun. 2006;351:192–7. https://doi.org/10.1016/j.bbrc.2006.10.020.
29. Burkhart CA, Cheng AJ, Madafiglio J, Kavallaris M, Mili M, Marshall GM, Weiss WA, Khachigian LM, Norris MD, Haber M. Effects of MYCN antisense oligonucleotide administration on tumorigenesis in a murine model of neuroblastoma. JNCI J Natl Cancer Inst. 2003;95:1394–403. https://doi.org/10.1093/jnci/djg045.
30. Whitehead KA, Langer R, Anderson DG. Knocking down barriers: advances in SiRNA delivery. Nat Rev Drug Discov. 2009;8:129–38. https://doi.org/10.1038/nrd2742.

31. Brooks TA, Hurley LH. The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. Nat Rev Cancer. 2009;9:849–61. https://doi.org/10.1038/nrc2733.
32. Evan G. Taking a back door to target Myc. Science. 2012;335:293–4. https://doi.org/10.1126/science.1217819.
33. Lu X, Vogt PK, Boger DL, Lunec J. Disruption of the MYC transcriptional function by a small-molecule antagonist of MYC/MAX dimerization. Oncol Rep. 2008;19:825–30. https://doi.org/10.3892/or.19.3.825.
34. Otto T, Horn S, Brockmann M, Eilers U, Schüttrumpf L, Popov N, Kenney AM, Schulte JH, Beijersbergen R, Christiansen H, Berwanger B, Eilers M. Stabilization of N-Myc is a critical function of Aurora A in human neuroblastoma. Cancer Cell. 2009;15:67–78. https://doi.org/10.1016/j.ccr.2008.12.005.
35. Durbin AD, Zimmerman MW, Dharia NV, Abraham BJ, Iniguez AB, Weichert-Leahey N, He S, Krill-Burger JM, Root DE, Vazquez F, Tsherniak A, Hahn WC, Golub TR, Young RA, Look AT, Stegmaier K. Selective gene dependencies in MYCN-amplified neuroblastoma include the core transcriptional regulatory circuitry. Nat Genet. 2018;50:1240–6. https://doi.org/10.1038/s41588-018-0191-z.
36. Wang L, Chen C, Song Z, Wang H, Ye M, Wang D, Kang W, Liu H, Qing G. EZH2 depletion potentiates MYC degradation inhibiting neuroblastoma and small cell carcinoma tumor formation. Nat Commun. 2022;13:12. https://doi.org/10.1038/s41467-021-27609-6.
37. Nagy Z, Seneviratne JA, Kanikevich M, Chang W, Mayoh C, Venkat P, Du Y, Jiang C, Salib A, Koach J, Carter DR, Mittra R, Liu T, Parker MW, Cheung BB, Marshall GM. An ALYREF-MYCN coactivator complex drives neuroblastoma tumorigenesis through effects on USP3 and MYCN stability. Nat Commun. 2021;12:1881. https://doi.org/10.1038/s41467-021-22143-x.
38. Guo Y-F, Duan J-J, Wang J, Li L, Wang D, Liu X-Z, Yang J, Zhang H-R, Lv J, Yang Y-J, Yang Z-Y, Cai J, Liao X-M, Tang T, Huang T-T, Wu F, Yang X-Y, Wen Q, Bian X-W, Yu S-C. Inhibition of the ALDH18A1-MYCN positive feedback loop attenuates MYCN-amplified neuroblastoma growth. Sci Transl Med. 2020;12:eaax8694. https://doi.org/10.1126/scitranslmed.aax8694.
39. Yang Y, Yi X, Cai Y, Zhang Y, Xu Z. Immune-Associated gene signatures and subtypes to predict the progression of atherosclerotic plaques based on machine learning. Front Pharmacol. 2022;13. https://doi.org/10.3389/fphar.2022.865624.
40. Zhao S, Zhang L, Ji W, Shi Y, Lai G, Chi H, Huang W, Cheng C. Machine learning-based characterization of cuprotosis-related biomarkers and immune infiltration in Parkinson's disease. Front Genet. 2022;13:1010361. https://doi.org/10.3389/fgene.2022.1010361.
41. Corvol H, Hodges CA, Drumm ML, Guillot L. Moving beyond genetics: is FAM13A a major biological contributor in lung physiology and chronic lung diseases? J Med Genet. 2014;51:646–9. https://doi.org/10.1136/jmedgenet-2014-102525.

Publisher's note