

RESEARCH

Open Access



# Characterization of G2/M checkpoint classifier for personalized treatment in uterine corpus endometrial carcinoma

Yiming Liu<sup>1</sup>, Yusi Wang<sup>2</sup>, Shu Tan<sup>1</sup>, Xiaochen Shi<sup>1</sup>, Jinglin Wen<sup>1</sup>, Dejia Chen<sup>1</sup>, Yue Zhao<sup>1</sup>, Wenjing Pan<sup>3</sup>, Zhaoyang Jia<sup>3</sup>, Chunru Lu<sup>4\*</sup> and Ge Lou<sup>1\*</sup>

## Abstract

**Background** Uterine Corpus Endometrial Carcinoma (UCEC) is a highly heterogeneous tumor, and limitations in current diagnostic methods, along with treatment resistance in some patients, pose significant challenges for managing UCEC. The excessive activation of G2/M checkpoint genes is a crucial factor affecting malignancy prognosis and promoting treatment resistance.

**Methods** Gene expression profiles and clinical feature data mainly came from the TCGA-UCEC cohort. Unsupervised clustering was performed to construct G2/M checkpoint (G2MC) subtypes. The differences in biological and clinical features of different subtypes were compared through survival analysis, clinical characteristics, immune infiltration, tumor mutation burden, and drug sensitivity analysis. Ultimately, an artificial neural network (ANN) and machine learning were employed to develop the G2MC subtypes classifier.

**Results** We constructed a classifier based on the overall activity of the G2/M checkpoint signaling pathway to identify patients with different risks and treatment responses, and attempted to explore potential therapeutic targets. The results showed that two G2MC subtypes have completely different G2/M checkpoint-related gene expression profiles. Compared with the subtype C2, the subtype C1 exhibited higher G2MC scores and was associated with faster disease progression, higher clinical staging, poorer pathological types, and lower therapy responsiveness of cisplatin, radiotherapy and immunotherapy. Experiments targeting the feature gene KIF23 revealed its crucial role in reducing HEC-1A sensitivity to cisplatin and radiotherapy.

**Conclusion** In summary, our study developed a classifier for identifying G2MC subtypes, and this finding holds promise for advancing precision treatment strategies for UCEC.

**Keywords** Uterine Corpus Endometrial Carcinoma, G2/M Checkpoint, Precision Treatment, Drug Resistance, G2MCS

\*Correspondence:

Chunru Lu

64579401@qq.com

Ge Lou

louge@ems.hrbmu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Introduction

Uterine corpus endometrial carcinoma (UCEC) is a group of epithelial malignancies originating from the endometrium, commonly occurring in perimenopausal and postmenopausal women. UCEC stands as one of the most prevalent malignancies of the female reproductive system and has the third highest mortality rate among common gynecological tumors [1, 2]. Despite two-thirds of patients presenting with early stage disease at diagnosis and the overall five-year survival rate reaches 81%, patients in stages IVA and IVB have significantly lower five-year survival rates of 17% and 15%, respectively [3].

The poor prognosis of UCEC stems from the heterogeneity of tumors and limitations of the current clinical diagnostic system. The existing diagnostic model based on clinical stage, pathological type, and genomic characterization still has considerable prognostic heterogeneity among patients when dealing with some high-level UCEC clinical decisions, which present challenges in clinical diagnosis and treatment [4, 5]. In addition, the resistance of UCEC to radiotherapy and chemotherapy is another great therapeutic challenge [6], and existing therapies often fail to produce a complete and durable tumor response, ultimately leading to treatment resistance and tumor recurrence. Therefore, there is an urgent need for new molecular subtypes to more accurately distinguish patients with different clinical characteristics and provide personalized treatment. Since the publication of the Cancer Genome Atlas data (TCGA, <https://portal.gdc.cancer.gov/>) [7, 8], the molecular landscape of tumor has been continuously studied, providing a clearer and broader perspective on the biological heterogeneity of UCEC and its impact on prognosis and response to anti-tumor therapy. In recent years, it has been found that the heterogeneity of tumor microenvironment (TME) plays an important role in poor prognosis and treatment resistance. Exploring new molecular pathways from TME may bring new possibilities for the precision diagnosis and treatment of UCEC.

Cancer is a group of diseases characterized by uncontrolled cell division, which is tightly regulated by several conserved cell cycle regulation mechanisms to ensure accurate replication and division of genetic material [9]. Cell cycle checkpoints function as regulatory procedures for DNA replication, preventing the accumulation and spread of faulty genetic information during cell division. At least three cell cycle checkpoints, G1/S, S, and G2/M, are known to strictly control cell cycle progression. Due to the prevalence of P53 gene mutations resulting in G1 checkpoint defects, tumor cells rely primarily on S and G2/M checkpoints to repair DNA damage [10, 11]. As the last checkpoint before a cell enters mitosis, the main role of the G2 checkpoint is to monitor the state

of the cell as it prepares to enter the M phase, ensuring that all DNA has been properly repaired and replicated before the cell divides. Chk1 is activated by ATR in the G2 phase results in CDC25A, -B, and -C phosphorylation, which prevents cyclin B/CDK1 activation, resulting in G2 phase arrest [12]. CDC25B/C inactivation caused by stress-induced p38MAPK/MMK activation is another important pathway leading to G2 phase arrest [13]. In addition, G2/M phase is also the most sensitive cell phase to a variety of anti-tumor drugs and ionizing radiation. When tumor cells are exposed to Cytotoxic drugs and ionizing radiation, G2/M cell checkpoints are inhibited, and the shutdown of this “relief” mechanism forces normal repair and leads to programmed cell death in the presence of DNA damage, which has become an effective strategy for sensitizing cancer anti-tumor therapy. Several drugs have been developed to target specific checkpoint components, with Wee1 tyrosine kinase playing a key role in G2/M checkpoint regulation of DNA damage repair. High levels of Wee1 have been observed in multiple cancer species such as breast cancer, leukemia, glioma, and melanoma [14–17]. Preclinical studies have shown that inhibiting Wee1 to disrupt G2/M checkpoint damage repair function can enhance the anti-tumor activity of radiotherapy and certain cytotoxic drugs. Several clinical studies have been conducted to investigate the clinical value of Wee1 inhibitors in combination with chemotherapy and radiation therapy [18–21]. Checkpoint kinase-1 (Chk1), an active transduction kinase at S and G2 checkpoints, is a potential target for anti-tumor drug development. Inactivation of Chk1 disrupts a key signaling pathway for G2-phase damage repair, reducing stress in tumor cells’ resistance to DNA damage caused by radiotherapy and chemotherapy [22]. Studies have revealed that Chk1 inhibitors effectively enhance the efficacy of chemotherapy and radiotherapy without increasing cytotoxicity, demonstrating their potential to improve responsiveness in drug-resistant cell lines [23, 24]. It is important to explore the distribution of G2/M checkpoint-related pathway activity in UCEC patients and identify potential prognostic markers and therapeutic targets for UCEC sensitization.

In this study, we have identified five differentially expressed genes (DEGs) and constructed an artificial neural network (ANN) classifier to identify subtypes with high G2/M activity scores in UCEC, which are associated with poor prognosis and lower responsiveness to most anti-tumor therapies. Further clinical cohort analysis and cell experiments revealed that groups with high expression of DEGs, including KIF23, were associated with worse prognosis, and inhibiting the expression level of KIF23 could effectively increase the sensitivity of UCEC cells to radiotherapy and cisplatin treatment. These findings may

provide a new strategy for subtype diagnosis and treatment sensitization in UCEC.

## Materials and methods

### Data collection and processing

354 G2/M checkpoint-related genes (G2MCRGs) were downloaded from three G2/M checkpoint-related pathways of the Molecular Signatures Database (MsigDB, <https://www.gsea-msigdb.org/gsea/msigdb>): “HALL-MARK\_G2M\_CHECKPOINT”, “REACTOME\_G2\_M\_CHECKPOINTS” and “BIOCARTA\_G2\_PATHWAY”. The “TCGAAbiologics” R package was used to download data of the TCGA-UCEC cohort from the GDC database (<https://portal.gdc.cancer.gov/>). The study includes 580 cases of gene expression profiling data, comprising 545 cancer cases and 35 cancer-adjacent cases, out of 548 cancer patients, three patients lacked expression profiling data. ‘Cancer-adjacent’ refers to tissue located near the tumor that is not cancerous but may exhibit early changes relevant to tumor development. Additionally, the dataset features 548 cases of clinical characteristics and 507 cases of TCGA molecular subtype data. Gene expression profiling data were converted to log2 (TPM+1) format for subsequent analysis. The clinical characteristics distribution of the patients from TCGA is shown in Supplementary Table 1. Complete prognostic data of these patients were obtained from the study by Liu et al. [25]. The simple nucleotide variation (SNV) data of the TCGA-UCEC cohort were obtained directly from the GDC database, stored in “maf” format, and used to calculate the tumor mutation burden (TMB) for each sample, which was calculated using the following formula: TMB (mut/mb)=total mutation amount (including synonymous, non-synonymous, substitution, insertion, and deletion mutations)/size of target coding area. Copy number variation (CNV) data of the TCGA-UCEC cohort were downloaded from the UCSC Xena database (<https://xena.ucsc.edu/>).

Supplementary Tab S1 Baseline Data Sheet about the clinical characteristics of the TCGA-UCEC cohort.

### Quantification of G2/M checkpoint pathway activity

The ssGSEA algorithm [26] based on the “GSEABase” R package and the “GSVA” R package calculated the G2/M checkpoint score (G2MCS) for each sample using 354 G2MCRGs as the input gene set. G2MCS was used to reflect G2/M checkpoint pathway activity.

### Identification and evaluation of G2/M checkpoint-related NMF clustering

Based on the expression profiles of G2MCRGs, the Non-negative matrix factorization (NMF) algorithm [27] was used to divide the TCGA-UCEC cohort into different

clusters (molecular subtypes). The NMF algorithm used the “Brunet” method with 10 iterations. The top point with the fastest cophenetic decline was used to determine the optimal number of clusters. t-Distributed Stochastic Neighbor Embedding (t-SNE) was used to downscale and visualize the distinguishability of the G2MCRGs expression profiles between different clusters. Difference analysis was used to compare the G2MCS between different subtypes. Kaplan–Meier (K-M) survival analysis was used to compare the differences in overall survival (OS), disease-specific survival (DSS), progression-free interval (PFI), and disease-free interval (DFI) between different G2MC subtypes. Gene set variation analysis (GSVA) was performed to compare the variation of metabolic pathways between different G2MC subtypes using 70 metabolism-related pathways screened from “c2.cp.kegg.v7.5.1.symbols.gmt” [28].

### Assessment of cell infiltration abundance in TME

A marker gene set of 23 immune cells was downloaded from the TISIDB database (<http://cis.hku.hk/TISIDB/data/download/CellReports.txt>) [29] and assessed relative infiltration abundance of various immune cells for the TCGA-UCEC cohort by the ssGSEA algorithm. CIBERSORT is an expression profile-based deconvolution algorithm for assessing the infiltration abundance of 22 immune cells in the TME based on the “e1701” R package [30]. MCPcounter algorithm was used to assess the population abundance of the 10 immune and stromal cells infiltrating in the TME [31]. The “ESTIMATE” R package was used to calculate the StromalScore, ImmuneScore and ESTIMATEScore in the TME, where StromalScore and ImmuneScore characterize the stromal and immune components, respectively. ESTIMATEScore is the sum of them [32]. A list of 47 immune checkpoint genes [33] and 46 cytokines [34] was collected from previous study to compare the expression differences of them between different G2MC subtypes.

### Subgroup analysis of clinical characteristics

TCGA typing is an important molecular typing for UCEC [35] and patients are categorized into four subtypes: copy-number high (CN\_HIGH), copy-number low (CN\_LOW), microsatellite instability hypermutated (MSI-H) and POLE ultramutated (POLE). CN\_LOW type represents the majority of patients with histologic grades G1 and G2, POLE type has the best prognosis, and CN\_HIGH type has the worst prognosis. POLE type has poorer sensitivity to conventional radiotherapy and chemotherapy. POLE type and MSI-H type are more responsive to immunotherapy. The subtype distribution graph demonstrated the differences in the distribution of the four TCGA subtypes between different G2MC

subtypes. The distributions of various important clinical characteristics in G2MC subtypes were compared, and visualized by pie charts, including "Grade", "Stage", "BMI", "Age" and "Histological type". In addition, differences of G2MCS between subgroups with different clinical characteristics were compared.

### Treatment responsiveness analysis

Chemotherapy, radiotherapy, immunotherapy, targeted therapy and endocrine therapy are important non-surgical treatment modalities for patients with UCEC. The "oncoPredict" R package constructed ridge regression models based on the "GDSC2" dataset to predict the AUC values of patients for several common UCEC drugs, including cisplatin, docetaxel, paclitaxel, tamoxifen and temsirolimus [36]. The smaller the predicted AUC value, the more sensitive the patient is likely to be to the drug. Radiotherapy-related efficacy assessment based on the TCGA-UCEC cohort compared the response to radiotherapy in patients with different G2MC subtypes, and subtype-related K-M survival analysis were used to determine the impact of radiotherapy on OS. The TIDE algorithm from the Tumor Immune Dysfunction and Exclusion (TIDE) database (<http://tide.dfc.harvard.edu/>) was utilized to predict the response to immunotherapy [37]. The TIDE score and Exclusion score reflected the level of tumor immune escape, with higher scores representing less responsiveness to immunotherapy. The Cancer Immunome Atlas (TCIA) database (<https://tcia.at/home>) provided four Immunophenoscores (IPS) reflecting the responsiveness of patients in the TCGA cohort to different immunotherapies [38]. Higher IPS represents higher immunotherapy responsiveness.

### Screening of subtype differential characteristic genes (SDCGs) by machine learning algorithms

Firstly, gene expression profiles difference analysis was performed to obtain subtype differentially expressed genes with  $|logFC| > 1$  and  $p < 0.05$  as the screening threshold, where the p-value was corrected by False Discovery Rate. Subsequently, the "glmnet" R package was used to perform the Least absolute shrinkage and selection operator (LASSO) regression algorithm [39]. LASSO achieved the purpose of downscaling and feature screening by constructing a penalty function and compressing the zero regression coefficients. The "randomForest" R package was used to perform the Random Forest (RF) algorithm for feature screening. The default number of iterations for RF was 100. The RF model was considered robust enough when 500 decision trees were constructed. The genes were scored for importance based on "mean decrease in accuracy", and those with importance scores greater than 2 were filtered as SDCGs. The support

vector machine-recursive feature elimination (SVM-RFE) algorithm based on the "e1071" R package and the "caret" R package is a posterior term selection algorithm for sequences based on the maximum interval principle of SVM, suitable for the screening of low-dimensional features [40]. Lastly, the SDCGs obtained from LASSO and RF are taken as intersection and entered into SVM-RFE model for final screening. After tenfold cross-validation, the genes corresponding to the model with the highest accuracy and the lowest error were selected as the final SDCGs.

### Construction, evaluation and validation of ANN subtype classifier

Firstly, we compared the expression levels of the SDCGs (classifier genes) in a single sample with the median of all samples. For down-regulated genes, the value was 0 if the expression level is higher than the median, otherwise it was 1. The opposite was true for up-regulated genes. The gene expression profiles of the patients were transformed into [0,1] normalized "gene signatures". Then, we constructed an ANN subtype classifier based on the "gene signatures" using the "neuralnet" R package and visualized it using the "NeuralNetTools" R package [41]. The number of neurons in the hidden layer was set to two-thirds of the sum of the neurons numbers in the input layer and the output layer. The "pROC" R package was used to construct Receiver operating characteristic (ROC) curves and compute Area under curve (AUC) to evaluate the prediction accuracy of the ANN subtype classifier. The GSE120490 cohort was used to validate the clinical application significance of subtype classifier for treatment responsiveness prediction.

### Genetic variation analysis

The "maftools" R package was used to analyze SNV data in "maf" format from the TCGA-UCEC cohort and to present the somatic mutation landscape by waterfall plots. RNA stemness scores (RNAss) data were downloaded from Pan-Cancer Atlas Hub (<https://pancanatlas.xenahubs.net>) to characterize the tumor stemness levels of samples [42]. The expression levels of four mismatch repair genes (MSH2, MSH6, MLH1, and PMS2) were used to characterize the microsatellite instability (MSI) of the samples. In addition, the CNV "gain" and "loss" frequencies of the classifier genes were calculated and displayed on the chromosomes.

### Expression and prognosis exploration of classifier genes

Differences in mRNA expression of classifier genes between UCEC and normal endometrial tissues were compared based on the GEPIA2 database (<http://GEPIA2>)

([gepia2.cancer-pku.cn/#index](http://gepia2.cancer-pku.cn/#index)). The effect of classifier gene expression on OS and recurrence-free progression (RFS) in UCEC patients was explored based on mRNA sequencing data from the Kaplan–Meier Plotter database (<https://kmplot.com/analysis/>). Prognostic data from the TCGA cohort were used to complement the association of classifier gene expression with DSS, PFI, and DFI.

### Clinical sample collection

The study included 23 UCEC tissue samples collected from January 2019 to August 2024. All samples were obtained from the operating room of the Department of Gynecology, Harbin Medical University Cancer Hospital. The pathological diagnosis was confirmed independently by two pathologists using the latest FIGO staging criteria (2023 edition). All samples were stored in liquid nitrogen and neutral tissue fixative within 30 min after separation for subsequent studies. This study was approved by the Ethics Committee of Harbin Medical University. Informed consent was obtained from all patients. The clinical information of the patients is listed in Table 1.

### Immunohistochemistry

Tumor tissues collected from 23 UCEC patients were fixed with neutral fixative at room temperature and embedded in paraffin. Paraffin-embedded tissue sections were dewaxed and rehydrated, and then blocked with 3% hydrogen peroxide at 25 °C for 25 min. Subsequently, the sections were first incubated with primary antibodies CDC7 (1:200, CUSABIO, China), ASPM (1:200, CUSA-BIO, China), CENPE (1:500, Sanying, China), KIF23 (1:200, CUSABIO, China), and DEPDC1B (1:200, CUSA-BIO, China) at 4 °C overnight, and then incubated with multimerized anti-rabbit IgG-HRP secondary antibodies at room temperature for 90 min. Finally, all samples were imaged at the same magnification under an optical microscope.

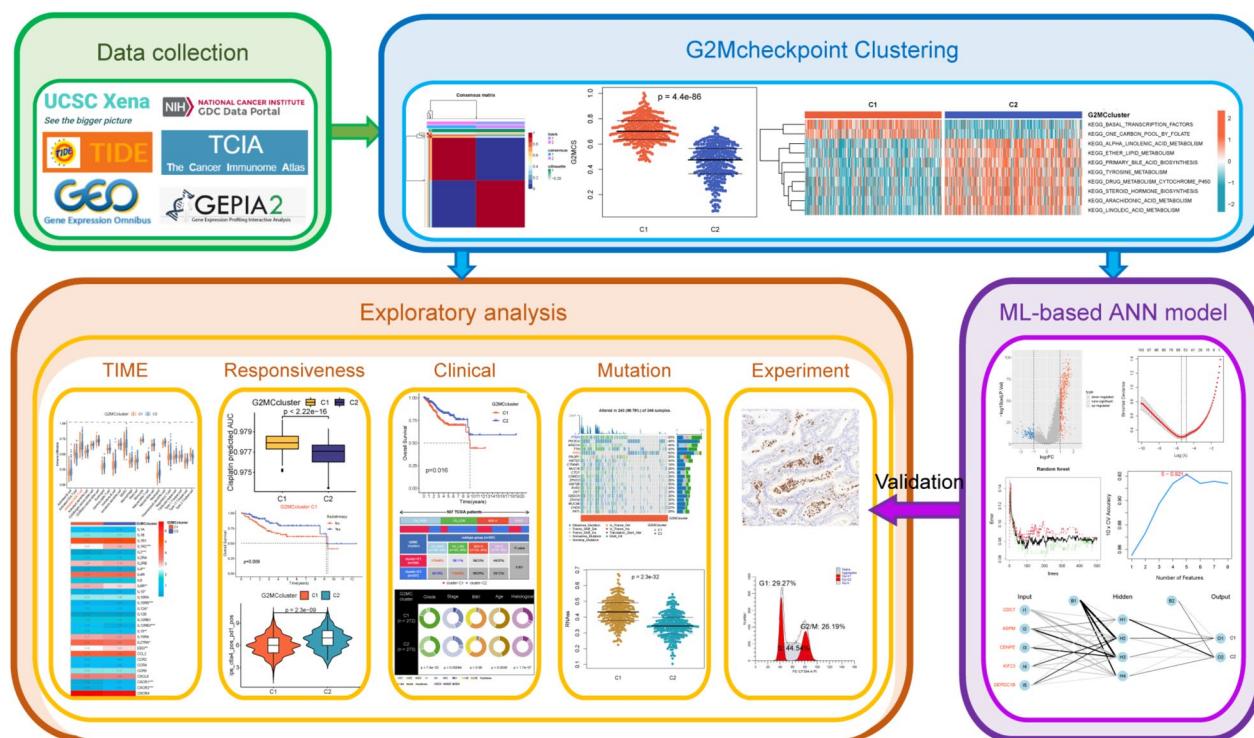
### Cell culture and transfection

Human UCEC cell line (HEC-1A) was purchased from Wuhan Procell Life Technology (Wuhan, China). The cell line was cultured in McCoy's 5A medium, and the complete medium was supplemented with 10% fetal bovine serum and 1% penicillin–streptomycin. The cells were cultured in a constant temperature incubator at 37 °C and 5% CO<sub>2</sub>. SiRNA was transfected to knock down the expression of KIF23 in UCEC cell line, and the targeted

**Table 1** Baseline data sheet for the real-world clinical cohort

| Characteristic         | Levels                                  | Numbers (%) |            |
|------------------------|---|-------------|------------|
|                        |   | Subtype C1  | Subtype C2 |
| Age                    | <64 years old                           | 10 (77%)    | 8 (80%)    |
|                        | ≥64 years old                           | 3 (23%)     | 2 (20%)    |
| Grade                  | G1                                      | 1 (8%)      | 3 (30%)    |
|                        | G2                                      | 4 (31%)     | 7 (70%)    |
| Clinical_Stage         | G3                                      | 8 (61%)     | 0 (0)      |
|                        | I                                       | 4 (31%)     | 8 (80%)    |
|                        | II                                      | 2 (15%)     | 0 (0)      |
|                        | III                                     | 7 (54%)     | 2 (20%)    |
| Bokhman_Classification | IV                                      | 0 (0)       | 0 (0)      |
|                        | I                                       | 5 (39%)     | 9 (90%)    |
| Histological_Type      | II                                      | 8 (61%)     | 1 (10%)    |
|                        | Endometrioid endometrial adenocarcinoma | 5 (39%)     | 9 (90%)    |
|                        | Clear cell carcinoma                    | 2 (15%)     | 0 (0)      |
|                        | Mixed serous and endometrioid           | 3 (23%)     | 1 (10%)    |
|                        | Serous endometrial adenocarcinoma       | 2 (15%)     | 0 (0)      |
|                        | Carcinosarcoma                          | 1 (8%)      | 0 (0)      |

G1: well differentiated carcinoma with tumor solid growth area ≤ 5%; G2: moderately differentiated carcinoma with solid growth area accounting for 6%–50%; G3: poorly differentiated carcinoma with solid growth area > 50%. Stage:—Stage I: Cancer is limited to the uterus.—Stage II: Cancer has spread from the uterus to the cervix.—Stage III: Cancer has extended beyond the uterus but is still within the pelvic area, possibly involving nearby lymph nodes.—Stage IV: Cancer has spread to the bladder or rectum, or to distant organs such as the liver or lungs. Bokhman classification: Type I: Estrogen-dependent, also known as endometrioid adenocarcinoma, which accounts for approximately 90% of endometrial cancer cases with a better prognosis. Type II: Non-estrogen-dependent, also known as non-endometrioid adenocarcinoma or special histologic subtypes of endometrial cancer, mainly including serous carcinoma, clear cell carcinoma, mucinous carcinoma, etc., accounting for about 10% of all endometrial cancers, with a poor prognosis



**Fig. 1** The workflow diagram of this study

siRNA construct and negative control si-NC were purchased from Tsingke Biotechnology Co., Ltd. (Beijing, China). The siRNA was transfected into the cells using jetPRIME® siRNA transfection reagent (Polyplus), and the follow-up study was conducted 24 h later. The three sequences of si-KIF23 are: 5'-UGGAUUUGUACCAUUCUUCUG-3', 5'-ACUCAUUGGUCCUUUAAGGG-3' and 5'-AAGUUUCGUUGAUACCUGUC-3'. The sequence of si-NC is: 5'-UGGUAUUGUACCAUUCUUUCUG-3'.

#### RNA extraction and quantitative real-time polymerase chain reaction (qRT-PCR)

Total RNA was extracted using the Accurate Steady-Pure Fast RNA Extraction Kit (AG21023, Hunan, China), and the concentration and purity of the extracted RNA were measured using NanoDrop (Thermofisher, USA).

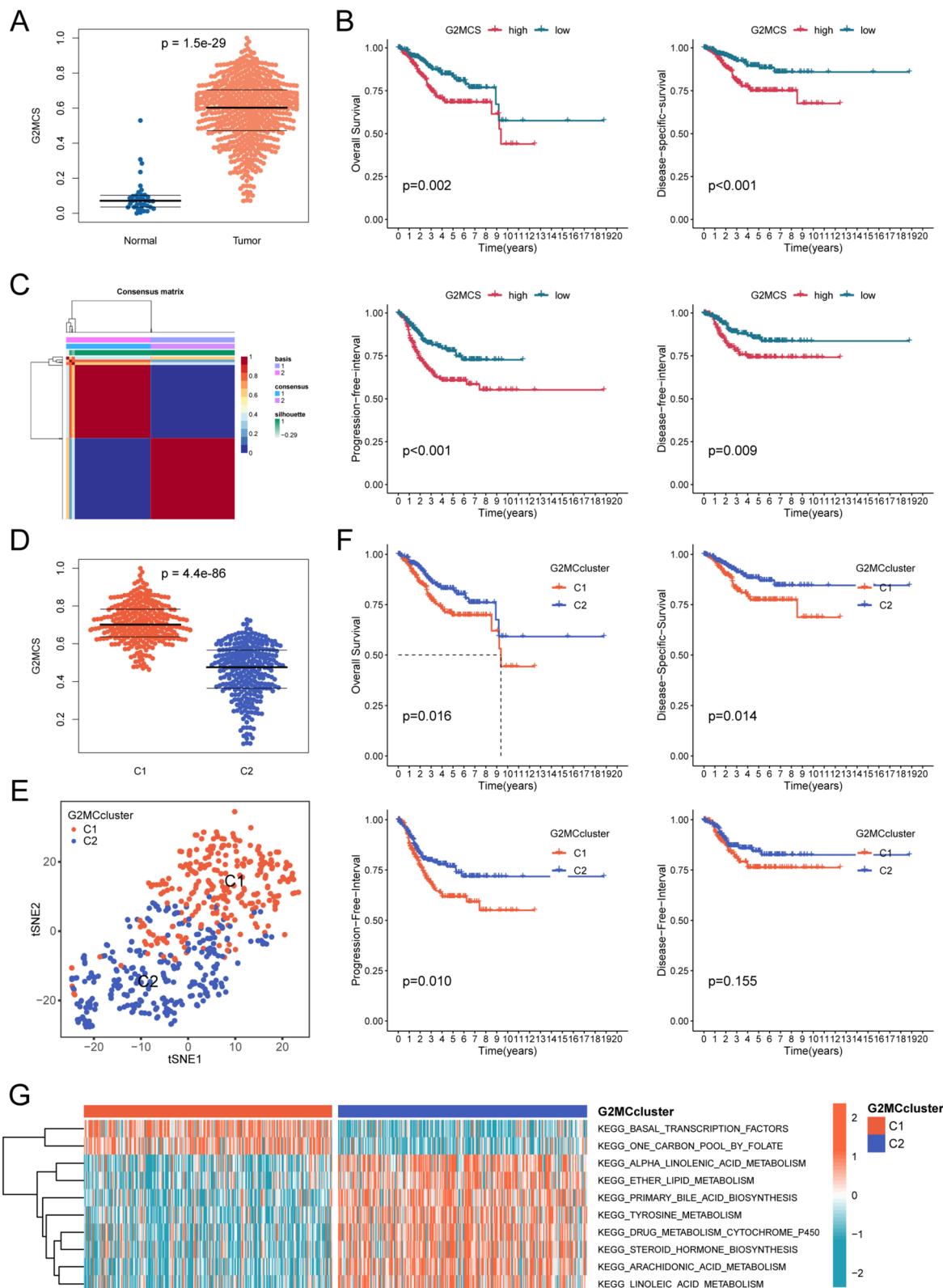
The HiScript III RT SuperMix for qPCR (#R323-01, Nanjing, China) was used for cDNA reverse transcription, and ChamQ Universal SYBR qPCR Master Mix (Vazyme#Q711) was used for real-time quantification. GAPDH was selected as the internal reference, and the  $2^{-\Delta\Delta Ct}$  method was applied to standardize the comparative expression levels of target genes. The primers for qRT-PCR are listed in Table S3.

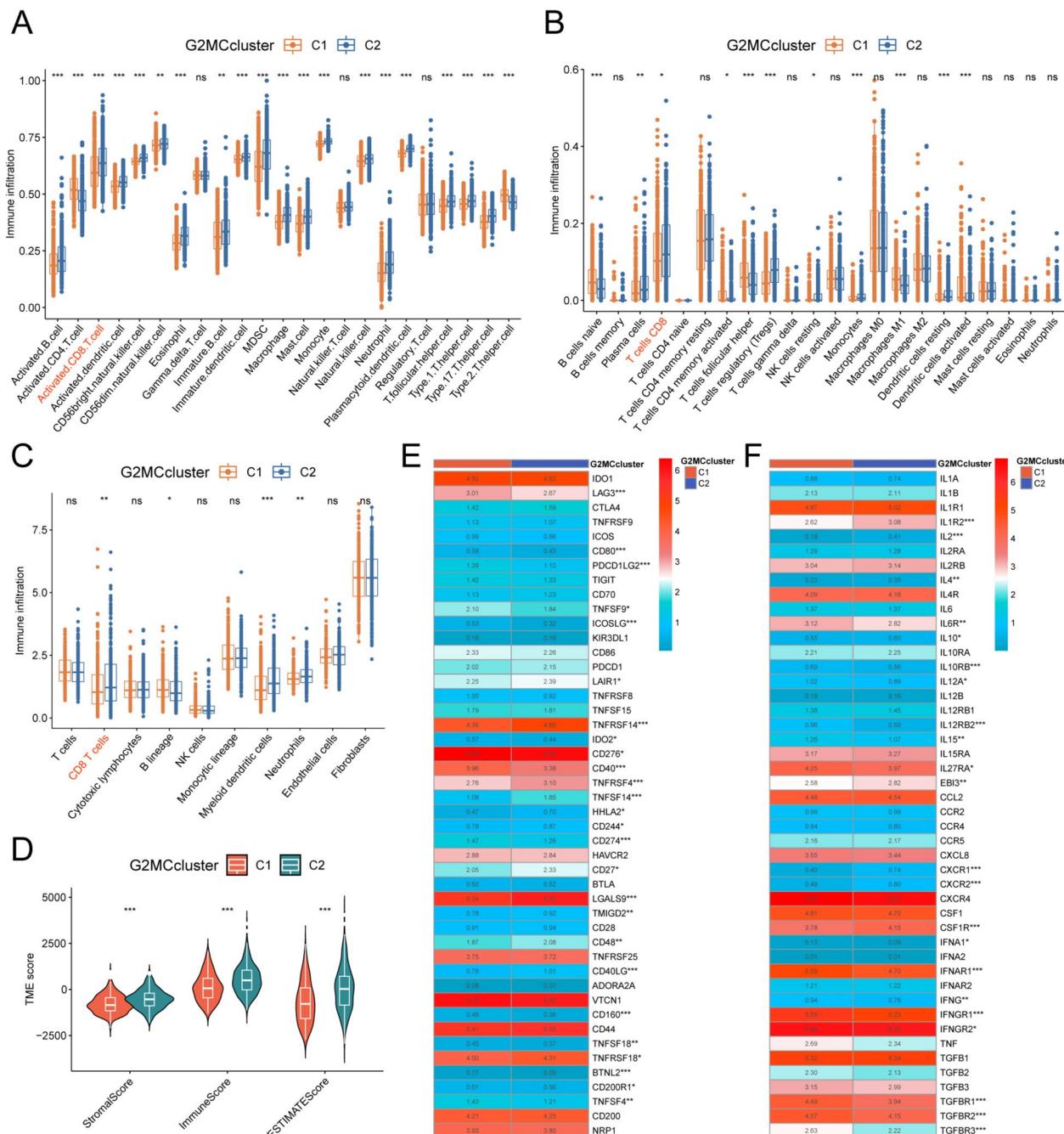
#### Cell cycle analysis

HEC-1A cells transfected with siRNA were harvested and fixed in 75% ethanol at 4 °C for 24 h. After centrifugation, the cells were stained with PI for 20 min, and the fluorescence generated was measured using flow cytometry. The acquired data was further analyzed using ModFit LT software (V4.1.7) to assess cell cycle distribution.

(See figure on next page.)

**Fig. 2** Identification and evaluation of G2MC Subtypes based on 353 G2MCRGs in the TCGA-UCEC cohort. **A** Differential analysis of activity of G2/M checkpoint pathway between UCEC and normal endometrial tissues based on G2MCS. **B** K-M survival analysis of G2MCS based on optimal cutoff value grouping. The ending events are OS, DSS, PFI, and DFI. **C** NMF clustering divides UCEC samples into two subtypes ( $k=2$ ) based on 353 G2MCRGs. **D** Differential analysis of activity of G2/M checkpoint pathway between the two G2MC subtypes based on G2MCS. **E** tSNE descending dimension analysis of the two G2MC subtypes. **F** K-M survival analysis reveals differences in OS, DSS, PFI, and DFI between the two G2MC subtypes. **G** GSVA between the two G2MC subtypes of the 70 metabolic pathway gene sets from KEGG. The color of the bar represents the GSVA score

**Fig. 2** (See legend on previous page.)



**Fig. 3** TME analysis of G2MC subtypes based on the TCGA-UCEC cohort. **A–C** Box plots based on ssGSEA (**A**), CIBERSORT (**B**) and MCPcounter (**C**) algorithms illustrated immune cells infiltration landscapes in the TME of patients with two G2MC subtypes. Red represents CD8(+) T cells. **D** Comparison of the stromal score, immune score and ESTIMATE score between patients with subtypes C1 and C2. **E** Expression differences of 47 immune checkpoint genes between the two G2MC subtypes. The scale bar in the legend is the expression value of the gene. ns, no significant difference; \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001

## Anti-tumor treatment

After 24 h of cell adhesion, the treatment group was treated with cisplatin monotherapy and radiotherapy, respectively. For the chemotherapy group, cisplatin (MedChemexpress, USA) was added to the culture medium to a final concentration of 50  $\mu$ mol/L for cell incubation. Cells in the radiotherapy group were irradiated using the Faxitron MultiRad 225 X-ray irradiation system (USA) at a total dose of 8 Gy, with a dose rate of 6 Gy/min. During irradiation, only the cells in the radiotherapy group were exposed, and other groups were shielded using three lead-equivalent lead plates. All cells had their culture medium replaced after six hours of treatment.

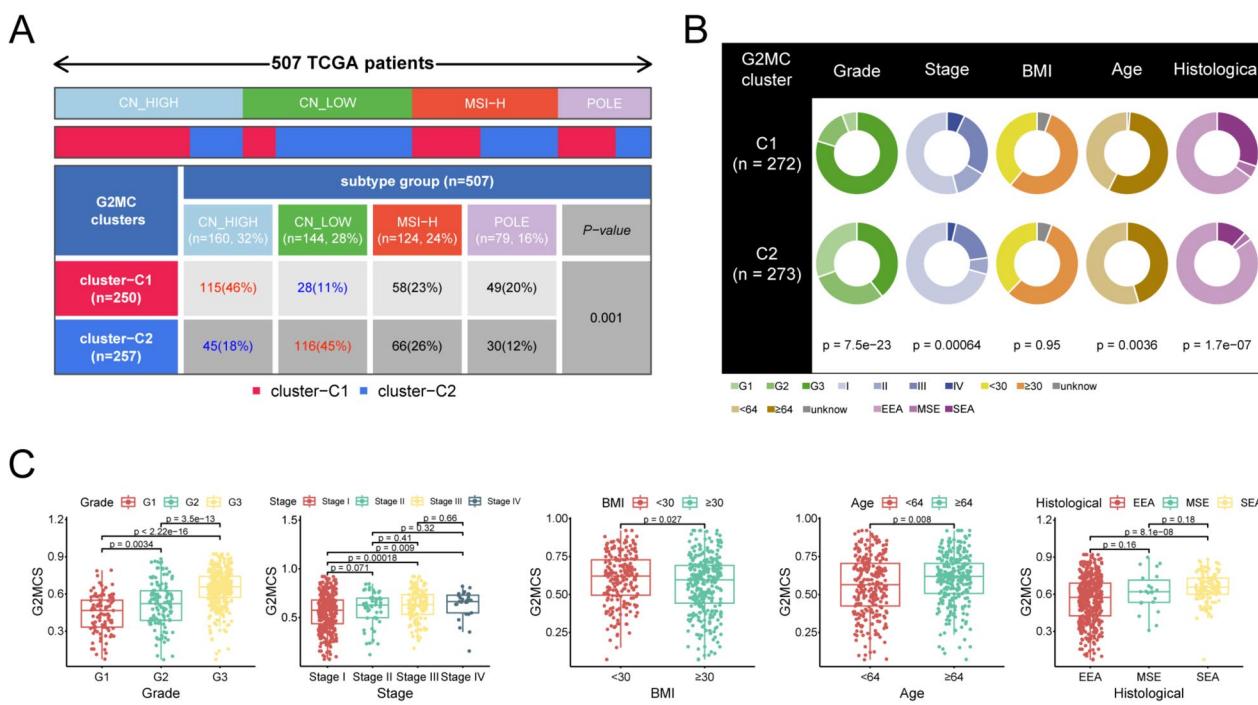
## CCK-8 assay

The cells were inoculated in 96-well plates at a density of 2000 cells per well and cultured in 100  $\mu$ L of complete medium. After the cells in different groups were subjected to their respective treatments for a specified period, 10  $\mu$ L of CCK-8 reagent (Beyotime, Shanghai, China) was added to each well and incubated for two hours. Absorbance values at 450 nm were measured using a microplate reader (Thermofisher, USA)

to determine the proliferative capacity of cells in each group.

## Statistical analysis

All bioinformatics analysis in this study were performed using R (version 4.2.1), and the Perl language was used for batch processing and cleaning of data. Unless otherwise stated, difference analysis in this study was performed with the “limma” R package. The Wilcoxon test was used to compare differences between two groups, while the Kruskal-Wallis test was for three or more groups. Correlation analysis defaulted to the “pearson” method. The chi-square test was used to compare differences in rates and component ratios between groups. K-M survival analysis and log-rank tests based on the “survival” R package and “survminer” R package were used to compare survival differences between groups. A two-tailed p value of  $<0.05$  was considered statistically significant. Statistical analysis of all cell experiments was performed using GraphpadPrism(version9.0), and independent Student’s t-test was used to analyze the differences between groups. All statistical tests were bilateral tests, and p values  $<0.05$  were considered statistically significant.



**Fig. 4** Clinical characteristics analysis of G2MC subtypes based on the TCGA-UCEC cohort. **A** Distribution proportion of four TCGA molecular subtypes in two G2MC subtypes. Red represent significant high proportions, blue represents low significant proportions. **B** Pie charts combined with chi-square tests show differences in the distribution of various clinical characteristics in the two G2MC subtypes. The median age of the cohort patients is 64 (years). **C** Differences comparison of G2MCS in patients with different clinical characteristics. EEA, Endometrioid endometrial adenocarcinoma; MSE, Mixed serous and endometrioid; SEA, Serous endometrial adenocarcinoma

## Results

### Identification and evaluation of UCEC subtypes based on 354 G2MCRGs

The workflow of this study is shown in Fig. 1. We calculated G2MCS to reflect the activity of the G2/M checkpoint pathway for all samples of the TCGA-UCEC cohort and found that G2MCS was significantly higher in UCEC samples than in paracancer samples (Fig. 2A). Survival analysis based on optimal cutoff value grouping showed that patients in the high G2MCS group had significantly lower OS, DSS, PFI, and DFI than those in the low G2MCS group (Fig. 2B). NMF clustering based on the expression profiles of 354 G2MCRGs classified UCEC patients into two G2MC subtypes, C1 and C2 (Fig. 2C). G2MCS was significantly higher in patients with subtype C1 than in patients with subtype C2 (Fig. 2D). tSNE results showed that the G2MCRGs expression profiles of the two subtypes were distinguishable (Fig. 2E). Patients with subtype C1 had significantly lower OS, DSS, PFI, and DFI than patients with subtype C2 (Fig. 2F). GSVA results showed significant differences in the metabolic pathways between the two G2MC subtypes. The basal transcriptional activity was higher in subtype C1, and the metabolic pathways of tyrosine,  $\alpha$ -linolenic acid, linoleic acid, and arachidonic acid were more active in subtype C2 (Fig. 2G).

### TME characterization based on G2MC subtypes

The heterogeneity of TME is not only closely related to tumor immunity, but also an important factor affecting the prognosis and treatment responsiveness of patients. We evaluated TME using three algorithms, ssGSEA (Fig. 3A), CIBERSORT (Fig. 3B) and MCP-counter (Fig. 3C), which showed significant differences in immune cell infiltration abundance between the two G2MC subtypes. ssGSEA showed that the infiltration abundance of most of the immune cells in subtype C2 was significantly higher than that in subtype C1. The results of CD8<sup>+</sup> T cells, the main effector cell of tumor immunity, showed a consistent trend across the three algorithms, with higher infiltration abundance in the

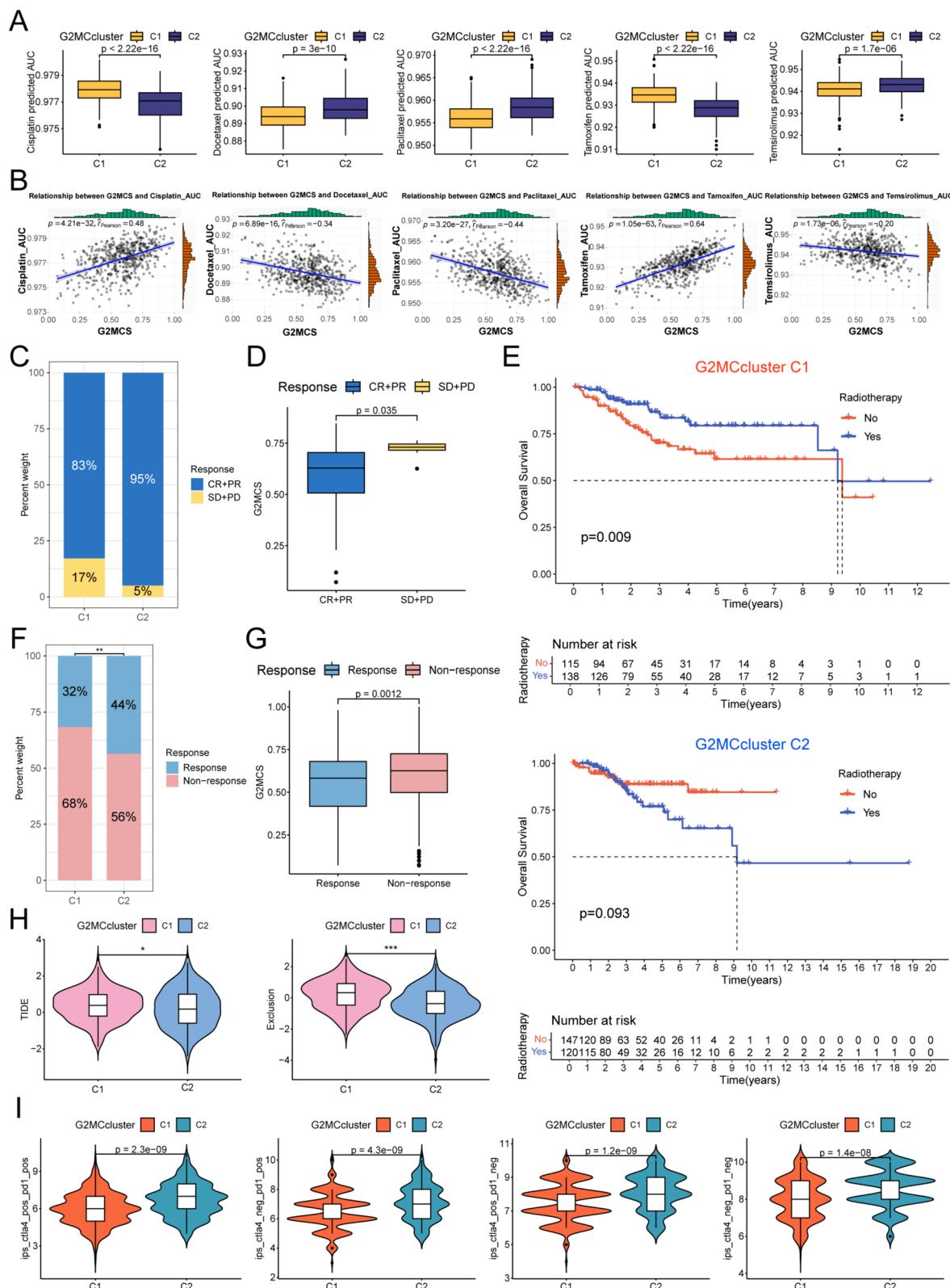
TME of subtype C2. We comprehensively assessed the stromal and immune components of TME by the ESTIMATE algorithm, and the StomatalScore, ImmuneScore, and ESTIMATEScore were significantly higher in subtype C2 than in subtype C1 (Fig. 3D). Further, we compared the differences in the expression of 47 immune checkpoint genes and 46 cytokines in TME between the two G2MC subtypes. The results showed that the expression of immune checkpoint genes such as LAG3, CD80, PDCD1LG2, CD274, and CD40 was significantly higher in the TME of subtype C1 than that of subtype C2, whereas the opposite results were observed for TNFSF14, TNFRSF4, TNFRSF14, and TNFRSF18 (Fig. 3E). The expression of multiple inhibitory cytokines was significantly higher in the TME of subtype C1 than that of subtype C2, including IL10RB, IL12A, IL12RB2, IL27RA, TGFBR1, TGFBR2, and TGFBR3 (Fig. 3F).

### Subgroup analysis of clinical characteristics associated with G2MC subtypes

TCGA subtype is an important molecular typing pattern for UCEC, guiding the precise treatment of patients. We found that the proportion of CN\_HIGH subtype was the highest in patients with subtype C1 and significantly higher than subtype C2. And the proportion of CN\_LOW subtype was the highest in patients with subtype C2 and significantly higher than subtype C1 (Fig. 4A). The results of the clinical characteristics combined analysis showed a predominance of G3 in subtype C1, with a significantly lower proportion of G1 and G2 than in subtype C2. Patients with subtype C1 had a lower proportion of clinical stage I and a higher proportion of the remaining clinical stages than patients with subtype C2. Patients with subtype C1 had a lower proportion of endometrioid endometrial adenocarcinoma (EEA) and a higher proportion of serous endometrial adenocarcinoma (SEA). In addition, older patients were more often clustered in subtype C1. There was no significant difference in the proportion of obesity ( $BMI \geq 30$ ) [43] between the two G2MC subtypes (Fig. 4B). Further, we

(See figure on next page.)

**Fig. 5** Therapeutic responsiveness analysis based on G2MC subtypes. **A** The oncoPredict algorithm (based on GDSC2 datasource) were used to predict and compare the AUC value between two G2MC subtypes to five therapeutic drugs in the TCGA-UCEC cohort. **B** Pearson correlation analysis of G2MCS with AUC values of 5 therapeutic drugs. **C** Proportional distribution of radiotherapy response status in the two G2MC subtypes based on TCGA-UCEC-radiotherapy cohort. **D** Differential analysis of G2MCS in patients with different radiotherapy response status. **E** K-M survival analysis reveals effect of radiotherapy or not on OS in different G2MC subtypes. **F** Proportional distribution of immunotherapy response status in the two G2MC subtypes based on the TCGA-UCEC-TIDE cohort. **G** Differential analysis of G2MCS in patients with different immunotherapy response status. **H** Differential analysis of TIDE and Exclusion scores closely associated with immune escape between the two G2MC subtypes based on TIDE algorithm. **I** Differential analysis of four IPS based on the TCIA database was used to compare responsiveness to PD1 inhibitors and CTLA4 inhibitors between different G2MC subtypes. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ . CR, Complete response; PR, Partial response; SD, Stable disease; PD, Progressive disease

**Fig. 5** (See legend on previous page.)

compared the differences in G2MCS between subgroups with different clinical characteristics and found that the higher the Grade, the higher the G2MCS. Patients with clinical stages III and IV had significantly higher G2MCS compared with stage I. Patients with SEA had significantly higher G2MCS compared with patients with EEA. In addition, older patients and non-obese patients had higher G2MCS (Fig. 4C).

### Prediction of treatment responsiveness based on G2MC subtypes

Patients with UCEC are treated in a very diverse way, and systematic assessment of antitumor therapy responsiveness can help to develop precise treatment strategies based on G2MC subtypes. The results of drug sensitivity analysis showed that patients with subtype C2 had lower AUC values for cisplatin and tamoxifen and higher AUC values for docetaxel, paclitaxel, and teicoplanin compared with patients with subtype C1 (Fig. 5A). G2MCS was positively correlated with the AUC values of cisplatin and tamoxifen and negatively correlated with the AUC values of docetaxel, paclitaxel, and temsirolimus (Fig. 5B). In the TCGA-UCEC-radiotherapy cohort, patients with subtype C2 had a higher rate of objective remission (CR + PR) than subtype C1 (Fig. 5C), and patients without objective remission (SD + PD) had a significantly higher G2MCS compared with patients with objective remission (Fig. 5D). To assess the long-term benefit of radiotherapy, we used a subgroup K-M survival analysis to explore the effect of radiotherapy on OS. In patients with subtype C1, radiotherapy significantly improved patients' OS in patients with type C2, radiotherapy instead may decrease patients' OS (Fig. 5E). We assessed immune escape and predicted immunotherapy responsiveness of the TCGA-UCEC cohort based on the TIDE algorithm. Patients with subtype C1 had lower immunotherapy response rates (Fig. 5F) and higher TIDE scores and Exclusion scores compared with subtype C2 (Fig. 5G). Immunotherapy-responsive patients had significantly lower G2MCS compared with non-responders (Fig. 5H). Moreover, we validated immunotherapy

responsiveness through the TCIA database. The results showed that patients with subtype C1 had significantly lower all IPS than patients with subtype C2 (Fig. 5I).

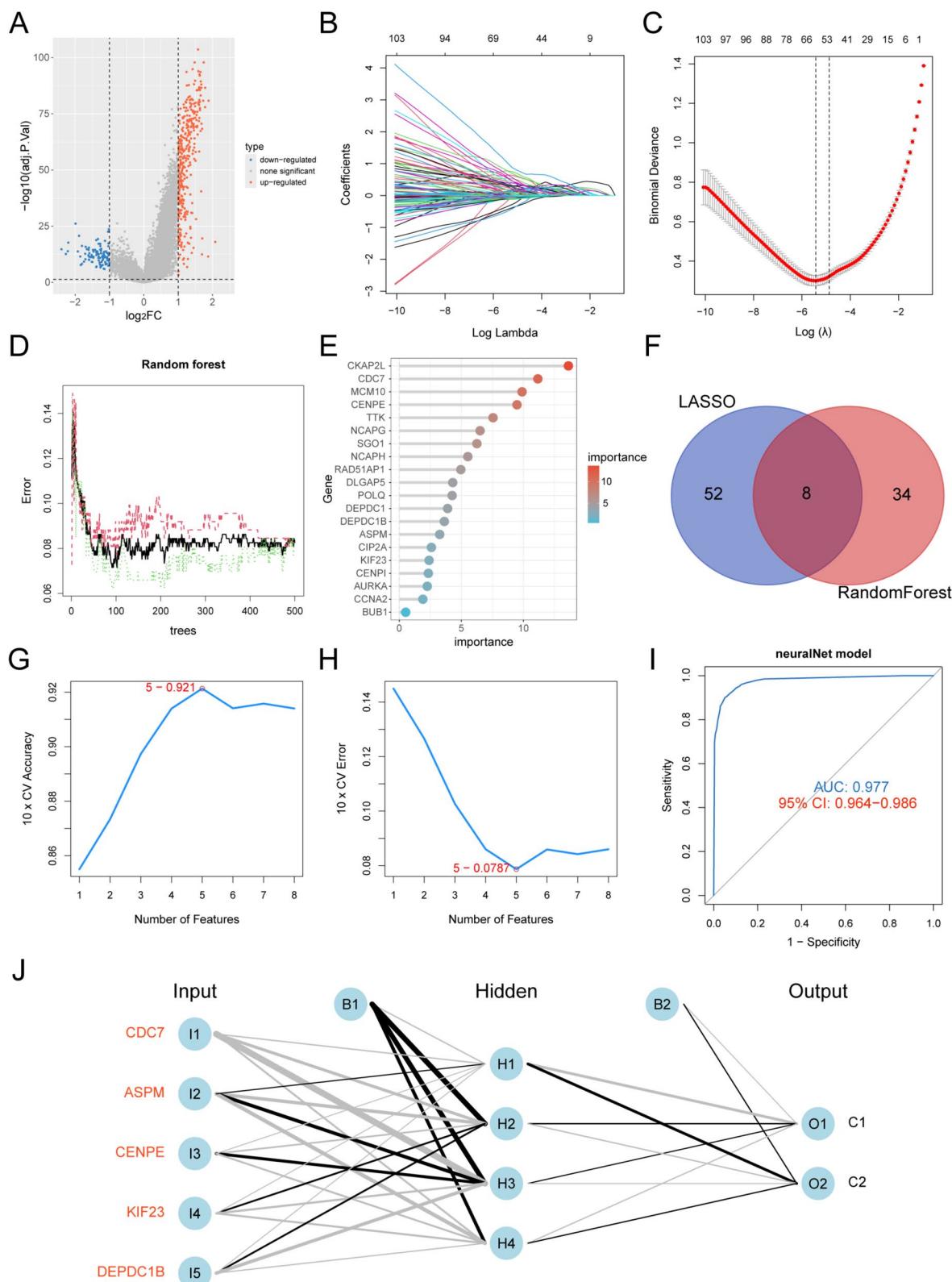
### Construction, evaluation and validation of G2MC subtype classifier

In order to map the results of large-sample NMF clustering to small-sample or single-sample clinical applications, we constructed a G2MC subtype classifier based on machine learning and ANN. When the expression profiles of the classifier genes are input, the classifier can accurately classify UCEC patients into the corresponding G2MC subtypes. Firstly, we obtained 450 subtypes differentially expressed genes by difference analysis. Using subtype C2 as a control, subtype C1 had 106 down-regulated genes and 344 up-regulated genes (Fig. 6A). Subsequently, we screened and obtained 60 SDCGs by LASSO regression algorithm (Fig. 6B, C) and 42 SDCGs by RF algorithm (Fig. 6D, E), with eight overlapping genes between the two algorithms (Fig. 6F). Further screening was performed by the SVM-RFE algorithm, and the model had the highest accuracy (Fig. 6G) and the lowest error (Fig. 6H) when the number of characteristic genes were five. Five genes (CDC7, ASPM, CENPE, KIF23 and DEPDC1B) were finally screened as classifier genes. Finally, we constructed the ANN classifier with an AUC value of 0.977 for distinguishing G2MC subtypes (Fig. 6I). The specific workflow of the ANN classifier is shown in Fig. 6J.

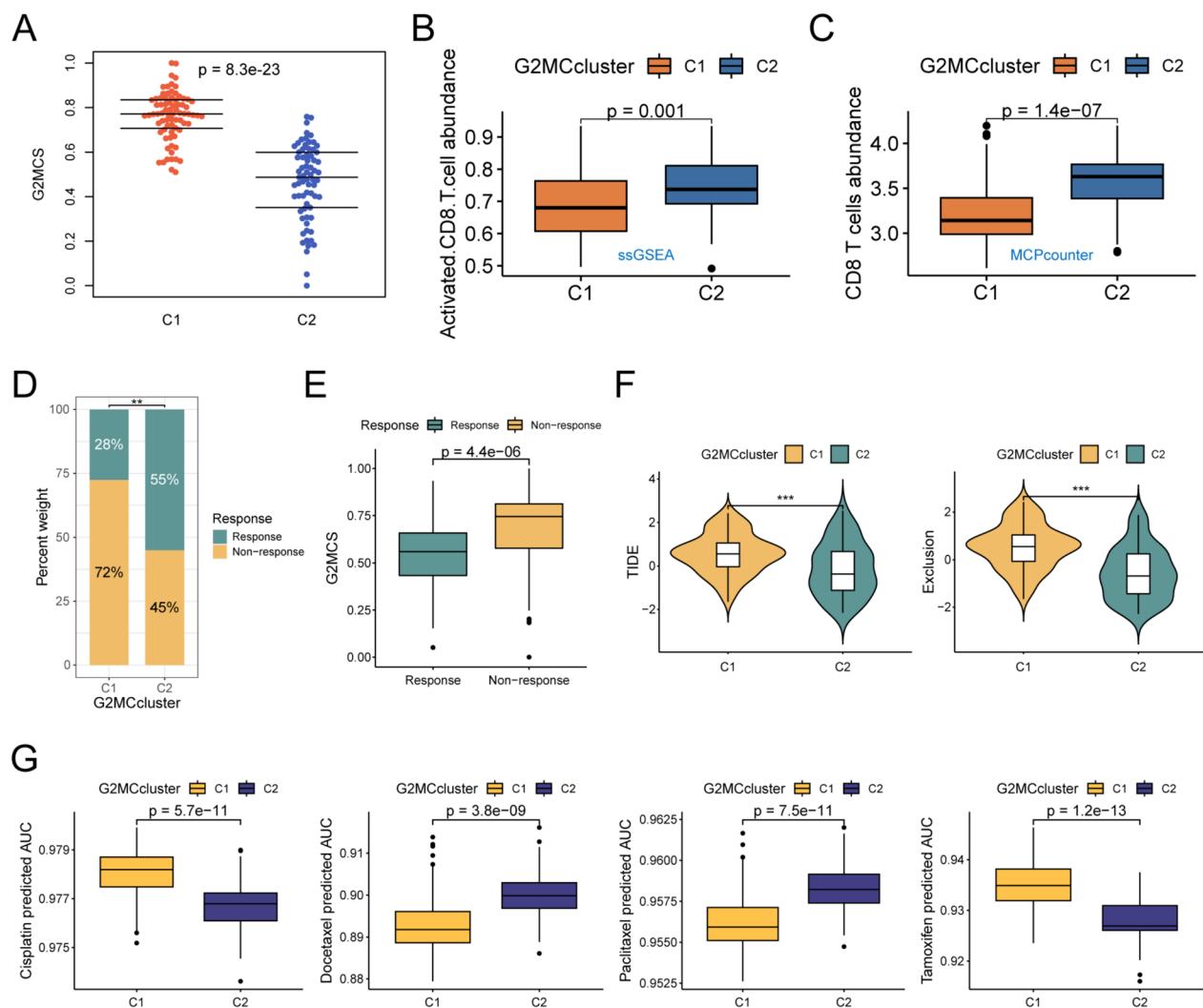
We used the GSE120490 cohort to validate the predictive role of the G2MC subtype classifier on treatment responsiveness. Firstly, we used the subtype classifier to categorize patients of the GSE120490 cohort into subtype C1 and subtype C2, with subtype C1 having a higher G2MCS (Fig. 7A). Evaluation results based on the ssGSEA algorithm (Fig. 7B) and the MCPcounter algorithm (Fig. 7C) showed that subtype C2 had a higher infiltration level of CD8<sup>+</sup> T cells. Validation of the TIDE algorithm showed that patients with subtype C1 had a lower response rate to immunotherapy compared with

(See figure on next page.)

**Fig. 6** Construction of ANN classifiers by multiple machine learning for identification of G2MC subtypes. **A** Differential analysis identified 450 DEGs of subtype C1 compared with subtype C2. **B** Ten-fold cross-validation of the coefficients of 450 DEGs in the LASSO regression model. **C** Robustness test of the LASSO regression model with varying number of DEGs revealed the highest stability when the number was 60. **D** The relationship between the number of trees and model error in the RF model. The model has the smallest error when the number of trees is 91. Green for subtype C1, red for subtype C2, black for all samples. **E** The top 20 genes were ranked by gene importance score based on mean decrease in accuracy. **F** Venn diagram illustrating the intersection of the characteristic genes of LASSO model with those of RF model. **G, H** Correspondence between the number of feature genes and the tenfold cross-validated accuracy (**G**) and error (**H**) in the SVM-RFE model. The model has the highest accuracy and the smallest error when the number of genes is five. **I** The ROC of ANN classifier was used to verify the predictive efficacy. **J** Schematic diagram of ANN classifier with four hidden layers. Red represents higher expression of the model gene in subtype C1 compared with subtype C2



**Fig. 6** (See legend on previous page.)

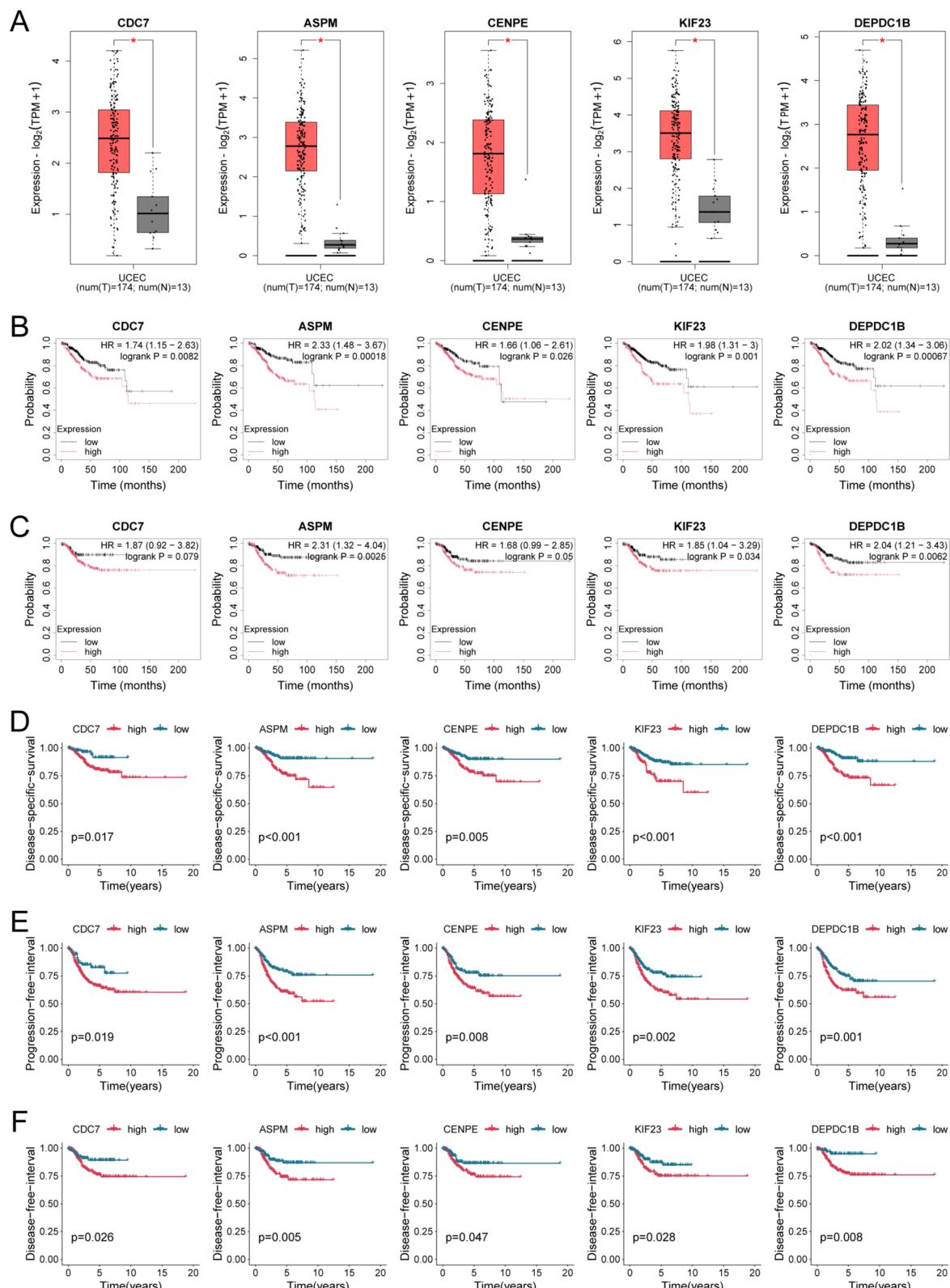


**Fig. 7** Validation of ANN classifiers to identify G2MC subtypes based on GSE120490 cohort. **A** Differential analysis of G2MCS between the two G2MC subtypes. **B, C** Differences comparison of CD8(+) T cells' infiltration abundance between the two G2MC subtypes by ssGSEA (B) and MCPcounter (C) algorithm. **D** Proportional distribution of immunotherapy response status in the two G2MC subtypes based on the GSE120490-TIDE cohort. **E** Differential analysis of G2MCS in patients with different immunotherapy response status. **F** Differential analysis of TIDE and Exclusion scores between the two G2MC subtypes. **G** Validation of the drug sensitivity analysis showed that the AUC value of four therapeutic drugs had a trend of difference consistent with the TCGA-UCEC cohort between the two G2MC subtypes. \*\* $p < 0.01$ ; \*\*\* $p < 0.001$

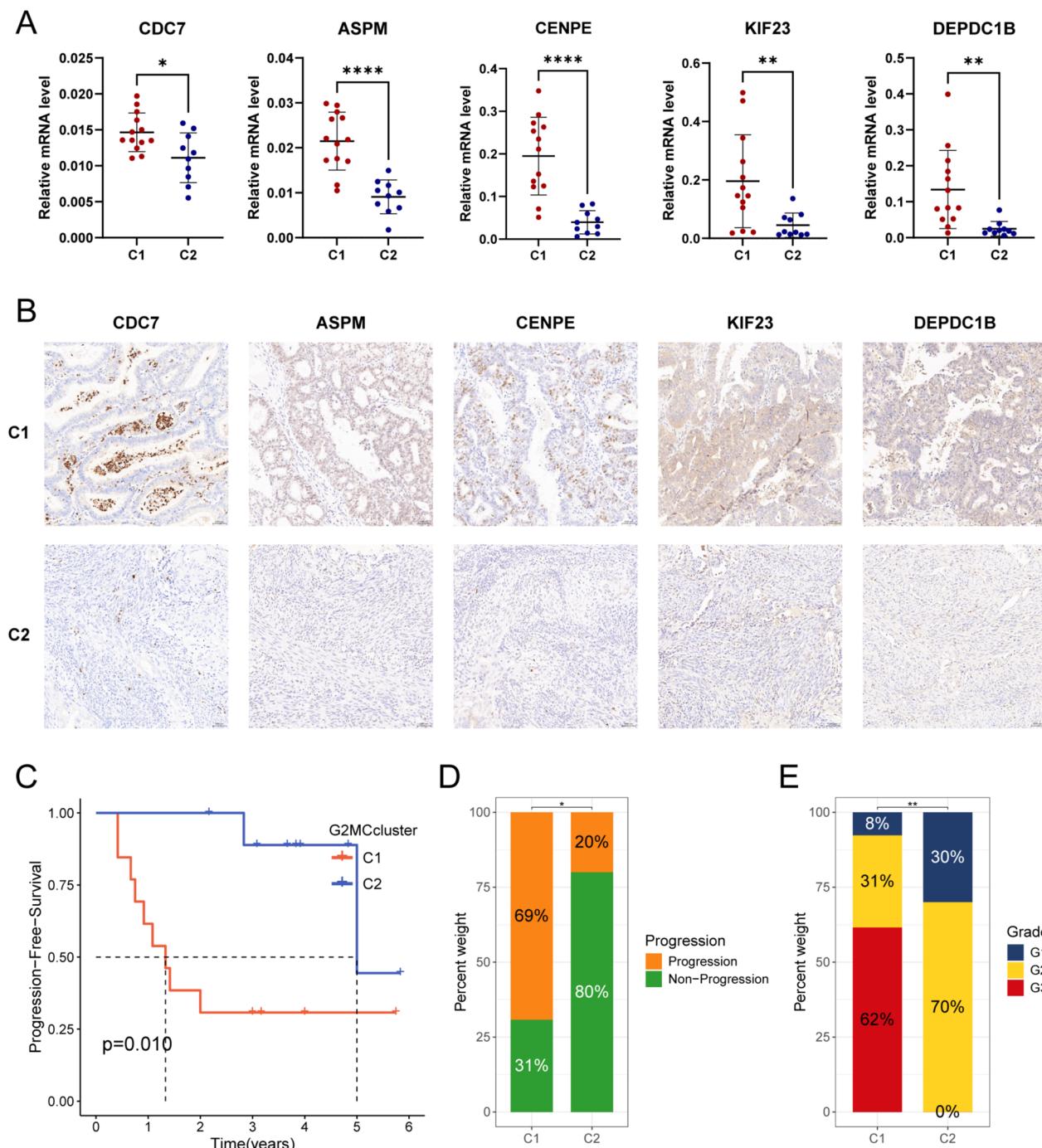
subtype C2 (Fig. 7D), higher TIDE scores and Exclusion scores (Fig. 7F). Immunotherapy-responsive patients had significantly lower G2MCS compared with non-responders (Fig. 7E). Validation of the drug sensitivity analysis showed that patients with subtype C2 had lower AUC values for cisplatin and tamoxifen and higher AUC values for docetaxel and paclitaxel compared with subtype C1 (Fig. 7G). This is consistent with the results of the TCGA cohort.

### Genetic variation analysis associated with G2MC subtypes

The waterfall plots demonstrated the somatic mutational landscape of the two G2MC subtypes, including the 20 genes with the highest mutation frequencies in UCEC. Compared with subtype C2 (Figure S1B), patients with subtype C1 (Figure S1A) had a higher frequency of TP53 mutation (52% vs. 19%) and a lower frequency of PTEN mutation (53% vs. 74%). The results of tumor stemness



**Fig. 8** Differential expression and prognostic analysis of five model genes. **A** Comparison of mRNA expression differences of five model genes between UCEC and normal endometrial tissues based on the GEPIA2 database. **B, C** K-M survival analysis of five model genes by optimal cutoff value grouping based on the Kaplan–Meier Plotter database. The ending events are OS (**B**) and RFS (**C**). **D, F** K-M survival analysis of five model genes by optimal cutoff value grouping based on the TCGA-UCEC cohort. The ending events are DSS (**D**), PFI (**E**), and DFI (**F**). \* $p < 0.05$



**Fig. 9** Validation of real-world cohort based on the classifier. **A** The expression levels of CDC7, ASPM, CENPE, KIF23 and DEPDC1B on qRT-PCR results. **B** By inputting the mRNA levels of the 5 model genes into the classifier, the real-world UCEC cohort was identified as two G2MC subtypes. Representative immunohistochemistry images of CDC7, ASPM, CENPE, KIF23, and DEPDC1B in the two G2MC subtypes, magnification 40x. **C** K-M survival analysis of two G2MC subtypes based on optimal cutoff value grouping. The ending events are PFS. **D** The proportion of progression after treatment in the two groups of G2MC subtypes. **E** Comparing the proportion of patients with pathological grades G1 to G3 in the two G2MC subtypes in the real-world cohort. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$

analysis showed that patients with subtype C1 had significantly higher RNAss compared with subtype C2 (Figure S1C). The expression levels of all four mismatch repair genes were higher in patients with subtype C1 than subtype C2 (Figure S1D), and were positively correlated with G2MCS (Figure S1E). Five classifier genes were mutated in the UCEC samples, with ASPM having the highest mutation frequency of 15% (Figure S1F). The CNV "gain" frequency of ASPM was significantly higher than the CNV "loss" frequency, and the CNV "loss" frequency of DEPDC1B was significantly higher than the CNV "gain" frequency (Figure S1G). We then constructed the CNV landscape of all classifier genes at the chromosome (Figure S1H).

### Expression and prognosis exploration of classifier genes

The results of difference analysis based on the GEPIA database showed that the mRNA expression levels of all five classifier genes were significantly higher in UCEC tissues than in normal endometrial tissues (Fig. 8A). The results of survival analysis based on the Kaplan–Meier Plotter database showed that for all five classifier genes, high expression was a prognostic risk factor for OS (Fig. 8B) and RFS (Fig. 8C, with non-significant results for CDC7 and CENPE) of patients with UCEC. We used the TCGA cohort to explore the prognostic impact of classifier gene expression from different clinical outcomes and found that for all five classifier genes, patients in the high-expression group had significantly lower DSS (Fig. 8D), PFI (Fig. 8E), and DFI (Fig. 8F) compared with the low-expression group.

### Validation of the G2MC subtype classifier in the UCEC clinical cohort

Tumor tissues from a total of 23 UCEC patients were collected from the Department of Gynecology, Harbin Medical University Cancer Hospital between January 2019 and August 2024. qRT-PCR and immunohistochemistry were conducted to assess the expression levels of the characteristic genes CDC7, ASPM, CENPE, DEPDC1B, and KIF23, in order to evaluate the efficacy of the classifier. Patients were classified based on mRNA expression levels, with 13 patients classified as subtype

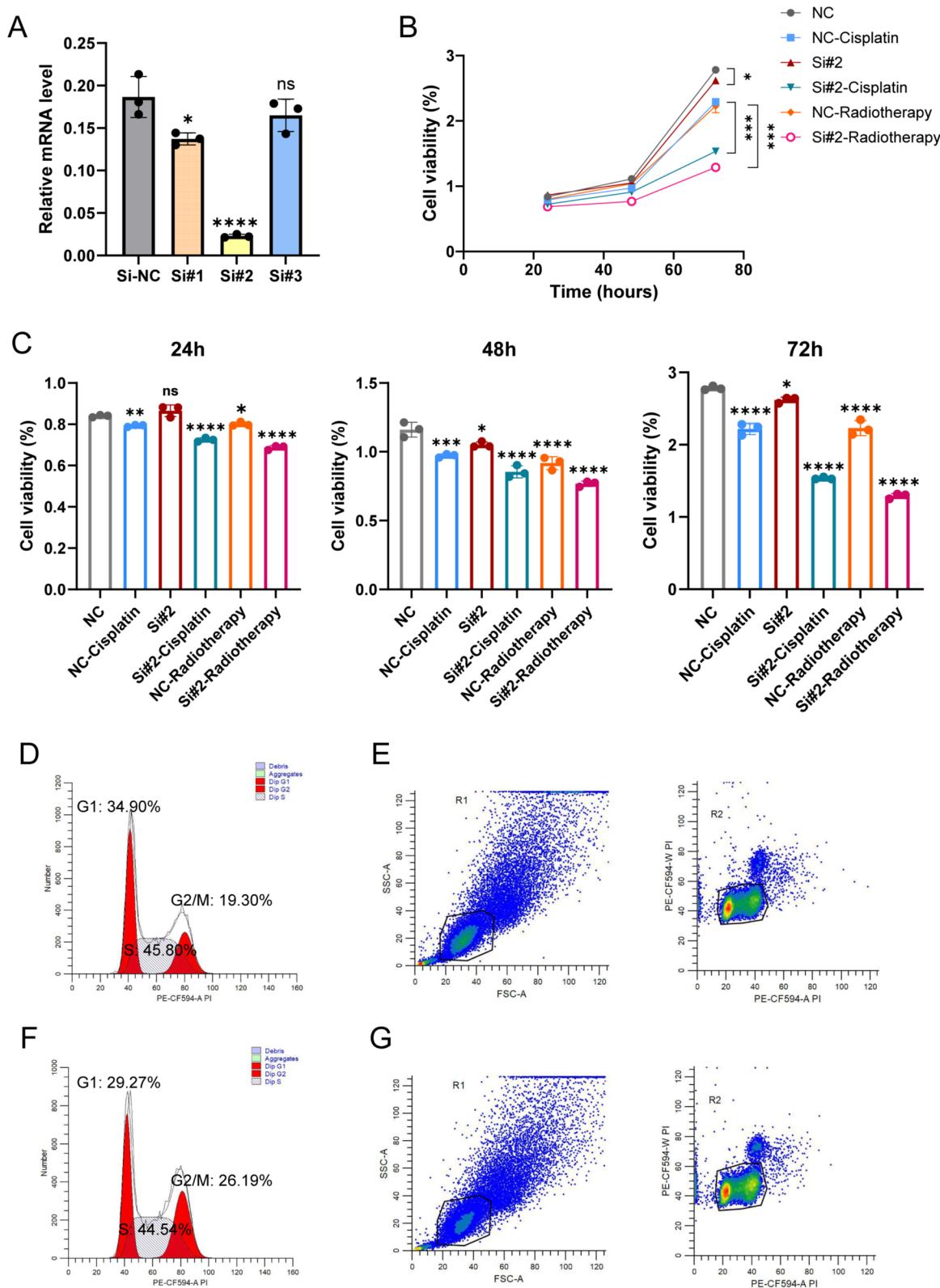
C1 and 10 patients classified as subtype C2. In our real-world cohort, we found that the expression levels of the five characteristic molecules were significantly higher at both the nucleic acid and protein levels in the subtype C1 compared to the subtype C2 (Fig. 9A, B). Selecting disease progression as the endpoint event for survival analysis, we observed that patients in the C1 subgroup had significantly shorter PFS than those in the C2 subgroup ( $p=0.01$ ), with a higher proportion of patients experiencing disease progression (Fig. 9C, D). Additionally, the proportion of patients with pathological grade G3 was significantly higher in the subtype C1 compared to the subtype C2 (Fig. 9E).

### Knockdown of KIF23 alters the cell cycle and treatment sensitivity of HEC-1A

As one of the feature gene members selected by our classifier, KIF23 has been confirmed as a cell cycle regulator in multiple tumors. However, research on KIF23 in the context of UCEC is still blank. In order to explore the potential influence of KIF23 on UCEC-related phenotypes, we successfully knocked down the expression level of KIF23 using siRNA (Fig. 10A). The results of CCK-8 assays showed that decreased expression of KIF23 significantly decreased the proliferation capacity of HEC-1A cells after cisplatin and radiotherapy (Fig. 10B, C), suggesting that reduced expression of KIF23 can increase the sensitivity of UCEC tumors to cisplatin and radiation. Furthermore, to explore the impact of regulating KIF23 expression on the G2M phase, we conducted cell cycle analysis using flow cytometry (Fig. 10D–G). The results revealed an increase in the proportion of the G2/M phase in the si-KIF23 group. Although there was an extension of the G2/M phase, the CCK-8 assay results suggested that the knockdown of KIF23 suppressed the activity of signaling molecules, leading to a significant reduction in the efficiency of DNA damage repair. These findings indicate that KIF23 participates in the G2/M-related signaling pathway in UCEC cell division, thereby enhancing the cells' resistance to cisplatin and radiation therapy.

(See figure on next page.)

**Fig. 10** Knockdown of KIF23 by specific siRNA in the HEC-1A cell line. **A** The knockdown efficiency of three siRNAs targeting KIF23 based on PCR validation. **B** Cell proliferation curves assessing by CCK-8 assay for si-KIF23 and si-NC at 24, 48, and 72 h after receiving cisplatin and radiation, respectively. **C** Column graphs of cell proliferation levels detected by CCK-8 assay for si-KIF23 and si-NC at 24, 48, and 72 h after receiving cisplatin and radiation, respectively. **D, F** Proportion of cell cycle for si-NC (**D**) and si-KIF23 (**F**). **E, G** Cell distribution and gating situation in si-NC (**E**) and si-KIF23 (**G**) in flow cytometry. Among them, ns: no significant statistical difference, \* $p<0.05$ , \*\* $p<0.01$ , \*\*\* $p<0.001$ , \*\*\*\* $p<0.0001$ . Data are means  $\pm$  SD, with  $n=3$

**Fig. 10** (See legend on previous page.)

## Discussion

UCEC is the most common gynecological malignancy in both high- and middle-income countries [5]. Despite an overall favorable prognosis, advanced-stage UCEC exhibits strong treatment resistance and a tendency for recurrence [44]. Recent exploration of the potential biological properties of UCEC has made some progress, revealing that most UCEC cases are caused by a series of somatic DNA mutations, with the most common mutations occurring in genes such as PTEN, mismatch repair genes, and TP53 [4, 5]. UCEC has been classified into four molecular subgroups based on mutation burden copy number variation, leading to the development of different treatment strategies and significantly improving the precision of treatment for this complex malignancy [45, 46]. While advancements in multi-omics technologies have greatly enhanced our understanding of UCEC heterogeneity, current molecular subtyping systems still result in different clinical outcomes for some patients with the same subtypes [47]. Therefore, the development of new molecular subtypes and the formulation of more precise anti-tumor treatment regimens are pressing clinical needs.

Tumor drug resistance results from the complex interplay of various factors, and the aberration of cell cycle regulatory mechanisms is a key driver in the development of tumor drug resistance. Overactivation of cell cycle checkpoints signaling molecules enables damaged cells to prolong their cycle arrest during replication, aiding in DNA damage repair and replication fork stability, ultimately reducing drug and radiation efficacy in killing tumor cells [9, 10]. G2/M checkpoint surveillance monitors the integrity of genetic information and inhibits damaged cells from entering mitosis, thereby maintaining efficient tumor cell division and proliferation [11]. Thus, investigating the activity of the G2/M checkpoint pathway and exploring potential signaling molecules may serve as biomarkers for predicting treatment responsiveness and prognosis in UCEC patients. At the same time, inhibiting the activity of key cell cycle molecules may prevent timely DNA damage repair in tumors, and increasing genomic stress via radiation and chemotherapy may enhance UCEC sensitivity to treatment. In this study, we utilized various machine learning methods to identify potential hub genes associated with the G2/M checkpoint pathway, with the aim of facilitating the identification of potential therapeutic targets and guiding the development of precision treatment drugs.

Based on the expression profile of 354 G2MCRGs, we scored all samples of TCGA-UCEC (G2MCS). The G2MCS of subtype C1 was significantly higher in subtype C1 than in subtype C2, indicating greater G2/M

checkpoint activity. Survival and clinical feature analyses revealed that patients in subtype C1 had worse outcomes and more advanced disease than those in subtype C2. The immune infiltration of TME not only reflects the immune therapy effect of UCEC but is also closely associated with patient prognosis [33, 48, 49]. The results of tumor microenvironment analysis showed that subtype C1 had lower CD8+ T cell infiltration and higher expression of immune checkpoint genes and inhibitory cytokines, correlating with lower responsiveness to immunotherapy. In radiotherapy assessments, subtype C1 showed a lower objective response rate, but this subtype benefited from improved overall survival over time. Drug sensitivity analysis indicated that subtype C1 patients were more responsive to docetaxel, paclitaxel, and temsirolimus, while subtype C2 derived greater benefit from cisplatin and tamoxifen. These findings suggest new avenues for personalized UCEC treatment. Cisplatin, as the most classical cytotoxic drug, primarily binds to and breaks cancer cell DNA, interrupting its normal gene replication and repair, thereby inducing apoptosis of tumor cells. Radiotherapy directly or indirectly damages the genetic information of tumor cells through high-energy rays or by generating oxygen free radicals. Our results show that G2MCS is negatively correlated with sensitivity to cisplatin and radiation. Subsequent experiments confirmed that downregulating the classifier gene KIF23 increased the effectiveness of cisplatin and radiotherapy in HEC-1A cells, indicating KIF23 as a potential therapeutic target.

To enhance the applicability of G2MCS in clinical settings, we used machine learning algorithms to identify five subtype-specific feature genes and constructed a G2M classifier based on artificial neural networks (ANN), which was validated for accuracy. The predictive value of the G2M subtypes for treatment response and prognosis was confirmed in both the GSE120490 dataset and our clinical cohort, consistent with TCGA results.

Finally, our study has several limitations. Further experiments are required to investigate the regulatory mechanisms of hub genes in the G2/M checkpoint pathway. Additionally, more sample data are necessary to validate the G2/M subtype classifier, as the available clinical variables for differential feature selection are limited. Due to data constraints, other clinical risk factors, such as estrogen levels and lifestyle habits, were not included; these may interact with G2/M checkpoint activity. We conducted a differential analysis and included age as a variable, but the improved classifier's ROC curve did not show significant enhancement (Figure S2). Future efforts should focus on collecting a broader range of clinical variables to enhance the classifier's predictive performance.

## Conclusion

In this study, we constructed an ANN classifier based on the activity of G2/M checkpoint pathway to identify different subtypes in UCEC patients. We found significant differences in prognosis and anti-tumor treatment responsiveness between patients with different G2MC subtypes. These research findings provide new personalized treatment strategies for UCEC patients.

## Abbreviations

|         |  |
|---------|--|
| ANN     | Artificial neural network                            |
| AUC     | Area under curve                                     |
| Chk1    | Checkpoint kinase-1                                  |
| CN_HIGH | Copy-number high                                     |
| CN_LOW  | Copy-number low                                      |
| CNV     | Copy number variation                                |
| DEGs    | Differentially expressed genes                       |
| DFI     | Disease-free interval                                |
| DSS     | Disease-specific survival                            |
| EEA     | Endometrioid endometrial adenocarcinoma              |
| G2MC    | G2/M checkpoint                                      |
| G2MCRGs | G2/M checkpoint-related genes                        |
| G2MCS   | G2/M checkpoint score                                |
| GSVA    | Gene set variation analysis                          |
| IPS     | Immunophenoscores                                    |
| K-M     | Kaplan-Meier   |
| LASSO   | Least absolute shrinkage and selection operator      |
| MSI     | Microsatellite instability                           |
| MsigDB  | Molecular Signatures Database                        |
| MSI-H   | Microsatellite instability hypermutated              |
| NMF     | Non-negative matrix factorization                    |
| OS      | Overall survival                                     |
| PFI     | Progression-free interval                            |
| qRT-PCR | Quantitative real-time polymerase chain reaction     |
| RF      | Random Forest  |
| RFS     | Recurrence-free progression                          |
| RNAss   | RNA stemness scores                                  |
| ROC     | Receiver operating characteristic                    |
| SEA     | Serous endometrial adenocarcinoma                    |
| SNV     | Simple nucleotide variation                          |
| SVM-RFE | Support vector machine-recursive feature elimination |
| TCGA    | Cancer Genome Atlas data                             |
| TCIA    | The Cancer Immunome Atlas                            |
| TIDE    | Tumor Immune Dysfunction and Exclusion               |
| TMB     | Tumor mutation burden                                |
| TME     | Tumor microenvironment                               |
| t-SNE   | T-Distributed Stochastic Neighbor Embedding          |
| UCEC    | Uterine corpus endometrial carcinoma                 |

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12935-025-03667-4>.

Supplementary materials 1.Fig. S1. Genetic variation analysis of the G2MC subtypes and five model genes.The waterfall chart showing the landscape of gene somatic cell mutations in patients with subtype C1 and subtype C2. Red represents a significant percentage increase of important mutations, blue represents a significant percentage decrease of important mutations.Differences in RNAss that symbolizes stemness of tumor cells between patients with two G2MC subtypes.Differences in expression level of four mismatch repair genes between patients with two G2MC subtypes.Correlation matrix of G2MCS and expression levels of four mismatch repair genes.The Waterfall chart showing the types and frequencies of genetic variation in five model genes.Frequencies about gain and loss of CNV in five model genes.The Chromosome localization and the CNV landscape of five model genes. \*p<0.05; \*\*\*p<0.001.

Additional file 2: Fig. S2. Classifier construction and validation of comprehensive clinical variables.Differences in BMI between patients with two G2MC subtypes.Differences in age between patients with two G2MC subtypes.Diagram of the improved classifier with five hidden layers, where age is included as a reference variable.The ROC of improved ANN classifier was used to verify the predictive efficacy.

Additional file 3: Tab S1. Baseline Data Sheet about the clinical characteristics of the TCGA-UCEC cohort.

Additional file 4: Tab S2. Weight parameters between nodes in ANN model to identify different G2MC subtypes.

Additional file 5: Tab S3. The primer sequences of genes.

## Acknowledgements

Not applicable.

## Author contributions

GL and CRL designed, supervised the study, and revision and final approval of the manuscript. YML conducted the data analysis, and wrote the manuscript. YSW, ST, XCS, JLW and DJC participated in and contributed to the experiments of this study. YZ participated in manuscript revision. All authors read and approved the final manuscript.

## Funding

This work was supported by the National Natural Science Foundation of China (No. 82173238), Key Program of Natural Science Foundation of Heilongjiang Province(ZD2020H007).

## Availability of data and materials

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article. The complete original data and code can be downloaded from <https://www.jianguoyun.com/p/DYDA1aYQyIXODBjRg9QFIAA..> No datasets were generated or analysed during the current study.

## Declarations

### Ethics approval and consent to participate

We confirm that the research has been carried out in accordance with the world Medical Association Declaration of Helsinki. All experimental protocols were approved by the Ethics Committee of Harbin Medical University Cancer Hospital (Harbin, China) and informed consent was obtained from all patients. The Ethical permission in this study is also included in the supplementary materials.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Author details

<sup>1</sup>Department of Gynecology, Harbin Medical University Cancer Hospital, Harbin, China. <sup>2</sup>Laboratory of Medical Genetics, Harbin Medical University, Harbin, China. <sup>3</sup>Second Affiliated Hospital of Harbin Medical University, Harbin, China.

<sup>4</sup>Department of Gynecology, Suihua Maternity and Health Care Hospital, Suihua, China.

Received: 18 September 2024 Accepted: 29 January 2025

Published online: 07 February 2025

## References

1. Makker V, MacKay H, Ray-Coquard I, Levine DA, Westin SN, Aoki D, et al. Endometrial cancer. *Nat Rev Dis Prim.* 2021;7(1):88.

2. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *Cancer J Clin.* 2023;73(1):17–48.
3. Miller KD, Nogueira L, Devasia T, Mariotto AB, Yabroff KR, Jemal A, et al. Cancer treatment and survivorship statistics, 2022. *Cancer J Clin.* 2022;72(5):409–36.
4. Cai Y, Wang B, Xu W, Liu K, Gao Y, Guo C, et al. Endometrial cancer: genetic, metabolic characteristics, therapeutic strategies and nanomedicine. *Curr Med Chem.* 2021;28(42):8755–81.
5. Crosbie EJ, Kitson SJ, McAlpine JN, Mukhopadhyay A, Powell ME, Singh N. Endometrial cancer. *Lancet.* 2022;399(10333):1412–28.
6. Mestraller G, Brown M, Bozkus CC, Bhardwaj N. Immune escape and resistance to immunotherapy in mismatch repair deficient tumors. *Front Immunol.* 2023;14:121016.
7. Talhouk A, McConechy MK, Leung S, Li-Chang HH, Kwon JS, Melnyk N, et al. A clinically applicable molecular-based classification for endometrial cancers. *Br J Cancer.* 2015;113(2):299–310.
8. Berger AC, Korkut A, Kanchi RS, Hegde AM, Lenoir W, Liu W, et al. A comprehensive pan-cancer molecular study of gynecologic and breast cancers. *Cancer Cell.* 2018;33(4):690.
9. Matthews HK, Bertoli C, de Bruin RAM. Cell cycle control in cancer. *Nat Rev Mol Cell Biol.* 2022;23(1):74–88.
10. Pedroza-Garcia JA, Xiang Y, De Veylder L. Cell cycle checkpoint control in response to DNA damage by environmental stresses. *Plant J.* 2022;109(3):490–507.
11. Engeland K. Cell cycle arrest through indirect transcriptional repression by p53: i have a DREAM. *Cell Death Differ.* 2018;25(1):114–32.
12. Smith HL, Southgate H, Tweddle DA, Curtin NJ. DNA damage checkpoint kinases in cancer. *Expert Rev Mol Med.* 2020;22: e2.
13. Reinhardt HC, Hasskamp P, Schmedding I, Morandell S, van Vugt MA, Wang X, et al. DNA damage activates a spatially distinct late cytoplasmic cell-cycle checkpoint network controlled by MK2-mediated RNA stabilization. *Mol Cell.* 2010;40(1):34–49.
14. Matheson CJ, Backos DS, Reigan P. Targeting WEE1 kinase in cancer. *Trends Pharmacol Sci.* 2016;37(10):872–81.
15. Barnaba N, LaRocque JR. Targeting cell cycle regulation via the G2-M checkpoint for synthetic lethality in melanoma. *Cell Cycle.* 2021;20(11):1041–51.
16. de Nonneville A, Finetti P, Birnbaum D, Mameissier E, Bertucci F. WEE1 dependency and pejorative prognostic value in triple-negative breast cancer. *Adv Sci.* 2021;8(17):2101030.
17. Zhang W, Li Q, Yin R. Targeting WEE1 kinase in gynecological malignancies. *Drug Des Dev Ther.* 2024;18:2449–60.
18. Yap TA, Ngoi N, Dumbrava EE, Karp DD, Ahnert JR, Fu S, et al. NCI10329: Phase Ib sequential trial of agents against DNA Repair (STAR) Study to investigate the sequential combination of the Poly (ADP-Ribose) Polymerase inhibitor (PARPi) olaparib (ola) and WEE1 inhibitor (WEE1i) adavosertib (ada) in patients (pts) with DNA Damage Response (DDR)-aberrant advanced tumors, enriched for BRCA1/2 mutated and CCNE1 amplified cancers. *Eur J Cancer.* 2022;174:7.
19. Cai SX, Ma N, Wang X, Jiang Y, Zhang H, Guo M, et al. Discovery and development of a potent and highly selective WEE1 inhibitor IMP7068. *Cancer Res.* 2023;83(7):ND07.
20. Schutte T, Embaby A, Steeghs N, van der Mierden S, van Driel W, Rijlaarsdam M, et al. Clinical development of WEE1 inhibitors in gynecological cancers: a systematic review. *Cancer Treatment Rev.* 2023;115:102531.
21. Yap TA, Miller C, Stenehjem D, Brown EJ, Carleton M, Mirza NQ. First-in-human phase 1 study of WEE1 inhibitor APR-1051 in patients with advanced solid tumors harboring cancer-associated gene alterations. *Cancer Res.* 2024;84(7):CT195.
22. Merry C, Fu K, Wang J, Yeh IJ, Zhang Y. Targeting the checkpoint kinase Chk1 in cancer therapy. *Cell Cycle.* 2010;9(2):279–83.
23. Bucher N, Britten CD. G2 checkpoint abrogation and checkpoint kinase-1 targeting in the treatment of cancer. *Br J Cancer.* 2008;98(3):523–8.
24. Ahmed S, Alam W, Aschner M, Alsharif KF, Albrakati A, Saso L, et al. Natural products targeting the ATR-CHK1 signaling pathway in cancer therapy. *Biomed Pharmacot.* 2022;155:113797.
25. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell.* 2018;173(2):400–16.e11.
26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–50.
27. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A.* 2004;101(12):4164–9.
28. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14:7.
29. Jia Q, Wu W, Wang Y, Alexander PB, Sun C, Gong Z, et al. Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. *Nat Commun.* 2018;9(1):5361.
30. Bense RD, Sotiriou C, Piccart-Gebhart MJ, Haanen J, van Vugt M, de Vries EGE, et al. Relevance of tumor-infiltrating immune cell composition and functionality for disease outcome in breast cancer. *J Natl Cancer Inst.* 2017;109(1):jw192.
31. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitpretz F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 2016;17(1):218.
32. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-García W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* 2013;4:2612.
33. Kalbasi A, Ribas A. Tumour-intrinsic resistance to immune checkpoint blockade. *Nat Rev Immunol.* 2020;20(1):25–39.
34. Propper DJ, Balkwill FR. Harnessing cytokines and chemokines for cancer therapy. *Nat Rev Clin Oncol.* 2022;19(4):237–53.
35. Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, Shen H, et al. Integrated genomic characterization of endometrial carcinoma. *Nature.* 2013;497(7447):67–73.
36. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Brief Bioinform.* 2021;22(6):bbab260.
37. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med.* 2018;24(10):1550–8.
38. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 2017;18(1):248–62.
39. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin. Epigenetics.* 2019;11(1):123.
40. Sanz H, Valim C, Vegas E, Oller JM, Reverter F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. *BMC Bioinform.* 2018;19(1):432.
41. Beck MW. NeuralNetTools: visualization and analysis tools for neural networks. *J Stat Softw.* 2018;85(11):1–20.
42. Li N, Zhan X. Integrated genomic analysis of proteasome alterations across 11,057 patients with 33 cancer types: clinically relevant outcomes in framework of 3P medicine. *Epma J.* 2021;12(4):605–27.
43. World Health Organization. Obesity and overweight. 2020. <https://www.who.int/about/accountability/results/who-results-report-2020-2021>. (Accessed 10 Dec 2024)
44. Braun MM, Overbeek-Wager EA, Grumbo RJ. Diagnosis and management of endometrial cancer. *Am Fam Physician.* 2016;93(6):468–74.
45. Yen TT, Wang TL, Fader AN, Shih IM, Gaillard S. Molecular classification and emerging targeted therapy in endometrial cancer. *Int J Gynecol Pathol.* 2020;39(1):26–35.
46. Urick ME, Bell DW. Clinical actionability of molecular targets in endometrial cancer. *Nat Rev Cancer.* 2019;19(9):510–21.
47. Yang Y, Wu SF, Bao W. Molecular subtypes of endometrial cancer: implications for adjuvant treatment strategies. *Int J Gynaecol Obstet.* 2024;164(2):436–59.

48. Zhang G, Yin Z, Fang J, Wu A, Chen G, Cao K. Construction of the novel immune risk scoring system related to CD8+ T cells in uterine corpus endometrial carcinoma. *Cancer Cell Int.* 2023;23(1):124.
49. Zhan L, Liu X, Zhang J, Cao Y, Wei B. Immune disorder in endometrial cancer: Immunosuppressive microenvironment, mechanisms of immune evasion and immunotherapy. *Oncol Lett.* 2020;20(3):2075–90.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.