

# High Expression of TTK As Colon Stem Cell Marker Correlated With M1 Macrophages Infiltration

**Hua Chen**

Southern Medical University

**Zhicong Wu**

Nanfang Hospital

**Nisha Wu**

Fifth Affiliated Hospital of southern medical University, China

**Zunya Ma**

Southern Medical University

**Yanling Liang**

Nanfang Hospital

**Fangyin Zeng**

Fifth Affiliated Hospital of southern medical University, China

**Fan Deng** (✉ [fandeng@smu.edu.cn](mailto:fandeng@smu.edu.cn))

Southern Medical University

---

## Research Article

**Keywords:** colon adenocarcinoma, stemness, WGCNA, mRNAsi scores, immune cell

**Posted Date:** July 7th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-580545/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Tumor stemness-related genes promote tumor progression and resistance to immunotherapy, but the relation between tumor-infiltrating immune cells and colon adenocarcinoma (COAD) stemness-related genes remains unclear.

**Methods:** The mRNAsi scores, a stemness index derived from transcriptomic data of the TCGA COAD cohort were used for integrated bioinformatics analysis. The Weighted Gene Co-expression Network Analysis (WGCNA) was applied to classify the modules correlated with mRNAsi scores. The Gene Expression Omnibus (GEO) cohorts GSE76402 was used to verify the DEGs between colon cancer and normal tissues. Real-time PCR and Flow-cytometry were used to validate the expression of genes and cancer stem cells marker in the spheres of colon cancer cells. The Single-sample Gene Set Enrichment Analysis (ssGSEA) and CIBERSORT algorithm were used to explore the correlation between significant genes and the immune cells in tumor microenvironment.

**Result:** The mRNAsi score was associated with tumor development and progression. *TTK* were identified as COAD stemness-related genes based on WGCNA analysis. *TTK* were negatively correlated with immunity. Interestingly, M1 Macrophages had a higher infiltration in COAD with high *TTK* expression. Higher infiltration of M1 Macrophages would result in the worse survival rate of COAD patients.

**Conclusion:** *TTK* may have the potential as a prognostic biomarker and adjuvant target to promote immunotherapy sensitivity and improve outcomes of COAD patients.

## Introduction

Colon adenocarcinoma (COAD) ranks the third and fourth of cancer-related incidence and mortality in the world, accounting for more than 700,000 death each year[1, 2]. Immunotherapy has become a powerful clinical strategy for cancer treatment. However, in contrast to Melanoma and non-small-cell lung cancer, immunotherapy has not yet become a relevant part of the treatment landscape of unselected colon cancer and it remains an unmet treatment needs for COAD patients[2]. Immunotherapy in previously-treated CRC patients resulted in no improvement post-treatment, also its use in other CRC subtypes has been either unsuccessful, or not extensively explored [3, 4]. Hence, it is vital to further unravel the in-depth mechanism of the resistance of immunotherapy for COAD.

Effective immunotherapy is limited in most patients by the immunosuppressive tumor environment. Cancer stem cells (CSCs), a dynamically proliferating population in the COAD microenvironment, play an immunosuppressive role in regulating Tumor infiltrating immune cells (TIICs)[5, 6]. Failure of immunotherapy in COAD patients are mainly attributed to CSCs self-renewal and immune surveillance[7, 8]. These features might be attributed to genes related to CSCs[9, 10]. The genes related to CSCs stemness maintenance not only promote proliferation of CSCs, but also might regulate TIICs[11]. CD44, as CSCs biomarker in colon cancer, would regulate T helper type 1 (Th1) cell survival and promote optimal survival of effector CD4 + T cells engaged in immune response[12]. The genes related to CSCs

and tumor stemness are critical for developing individualized treatment strategies. But the comprehensive landscape of immune cells infiltrating of COAD with high stemness has not yet been elucidated, and the association between TIICs and CSCs related genes is unclear.

At present, with the rapid development of high-throughput genome sequencing technology, the genome and transcriptome characteristics of tumor cells are considered closely related to tumorigenesis and maintenance of cancer stemness. Malta et al. developed a machine learning algorithm to compare genome maps and epigenetic feature sets of tumor cells and embryonic stem cells[13]. This algorithm calculated the stemness index (mRNAsi scores) by analyzing the expression of characteristic genes related to the tumor stemness maintenance. The mRNAsi scores have been used to evaluate the correlations with tumor metastasis and survival rate[14]. However, the correlation of mRNAsi scores with the infiltration of TIICs remains unknown.

Weighted Gene Expression Network Analysis (WGCNA) is a systematic biological method, which could cluster the modules of highly correlated genes and relate modules to external sample traits[15]. WGCNA was widely used to explore the associations between gene sets and clinical features, and to identify candidate biomarkers[16].

In our study, we identified key genes related to the stemness of COAD with WGCNA analysis and analyze the correlation between stemness of COAD and the infiltration of TIICs. Also, we investigated that the infiltration of activated memory CD4 + T cells was higher, while Tregs was lower in COAD patients with higher mRNAsi scores. Taken together, our result has identified genes that play a critical role in the biological functions of COAD stemness and reveal the correlation between stemness of COAD and infiltrating immune cells. The flow chart of this study is shown in supplement Fig. 1.

## Materials And Methods

### 2.1 Cell culture and Real-time (RT) PCR

The human colon cancer cell lines including HT29, RKO cells and human normal colonic epithelial cells FHC were purchased from American Type Culture Collection (ATCC, Manassas, VA). All cell lines had been identified after purchasing and lack of contamination before the experiment. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco, Lot:8120287) with 10% fetal bovine serum (FBS, ExCell Bio, Lot:111032). Total RNA was extracted from cells by the RNAiso Plus (TaKaRaBio Technology, Lot:9109) and real-time RT-PCR were performed by the Reverse Transcription Kit and SYBR Green Master Mix (Vazyme, Lot: Q341-02). The Ct values were normalized to the GAPDH expression level and  $2^{-\Delta\Delta Ct}$  method was used to measure the relative expression. RT-PCR was repeated at least 3 times to ensure reproducibility. The sequence of the primers was shown in Table 1.

**Table 1:** The primers of colon stem cell markers and COAD stemness related genes.

Gene		Primer (5'-3')
TTK	Forward	CCGAGATTTGGTTGTGCCTGGA
	Reverse	CATCTGACACCAGAGGTTCTTG
PLK4	Forward	GACACCTCAGACTGAAACCGTAC
	Reverse	GTCCTTCTGCAAATCTGGATGGC
CHEK1	Forward	GTGTCAGAGTCTCCCAGTGGAT
	Reverse	GTTCTGGCTGAGAACTGGAGTAC
KIF18A	Forward	CTTGGACCAGTTCAGCCTATTCC
	Reverse	GCACACTTTGAGATGGTGGAGAC
BUB1	Forward	GTGGAAGAGACTGCACAACAGC
	Reverse	TCAGACGCTTGCTGATGGCTCT
GAPDH	Forward	GTCTCCTCTGACTTCAACAGCG
	Reverse	ACCACCCTGTTGCTGTAGCCAA
CD44	Forward	CCAGAAGGAACAGTGGTTTGGC
	Reverse	ACTGTCCTCTGGGCTTGGTGT
CD133	Forward	CACTACCAAGGACAAGGCGTTC
	Reverse	CAACGCCTCTTTGGTCTCCTTG
CD90	Forward	GAAGGTCCTCTACTTATCCGCC
	Reverse	TGATGCCCTCACACTTGACCAG
Wnt3a	Forward	ATGAACCGCCACAACAACGAGG
	Reverse	GTCCTTGAGGAAGTCACCGATG

## 2.2 Sphere-forming assay and Flow-cytometry

HCT116 were cultured in serum-free DMEM: F12 medium with composed of 5ug/ml Insulin, 2% B27, 20ng/ml EGF, 20ng/ml FGF and 0.4% BSA. The HCT116 were plated in low-adherent flasks and grow for 14 days, and the half of medium was replaced every three days before spheroid formation was analyzed. The spheres of HCT116 grown for 7-days were evaluated and defined as primary spheres. The primary spheres were trypsinized and separated into individual cells and regrown for 7 days, defined as secondary sphere.

For detecting colon cancer cells surface maker, the spheres of HCT116 were stained with fluorochrome-conjugated monoclonal antibodies (mAbs) CD133 (Elabscience, Lot: E-AB-F1268D) and CD44

(Elabscience, Lot: E-AB-F1038E) for 30 min at room temperature and detected by flow cytometry.

## **2.3 Collection of mRNAsi scores and COAD expression data**

The clinical data of 452 samples and the RNA sequencing (RNA-Seq) expression data of 513 samples, including 473 colon adenocarcinoma samples and 41 normal colon samples, were downloaded from the TCGA database (<https://cancergenome.nih.gov/>). The mRNAsi scores and epigenetically regulated mRNAsi(EREG-mRNAsi) index were referenced from Malta et al' work[13]. The Gene Expression Omnibus (GEO) cohorts GSE76402 was download from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). GSE76402 expression profile was based on GPL10558 and contained 30 normal colon tissues samples and 30 colon cancer samples. The edgeR package was applied to screen the differentially expressed genes (DEGs) between normal samples and COAD samples with false discovery rate (FDR) < 0.05 and log forge change (FC) > 1 as the cutoff. The pheatmap package was used to visualize the DEGs.

## **2.4 Correlation between mRNAsi scores and clinical characteristic**

The comprehensive clinical features and clinical subtypes of samples were comprehensively analyzed and correlated with mRNAsi scores. The high scores group and low scores group were distinguished with median mRNAsi scores as the cut-off value, then the survival rate of patients was displayed with the Kaplan-Meier curve. We detected the protein expression of DEGs in the normal colon tissues and colon cancer tissues from The Human Protein Atlas (HPA) database (<https://www.proteinatlas.org/>).

## **2.5 Weighted Correlation Network Analysis (WGCNA) analysis and identification of significant modules and genes**

The WGCNA algorithm eliminates multiple hypothesis testing and correction issues by clustering thousands of genes with similar expression patterns and then screen gene sets associated with phenotypes[15]. The co-expression network was constructed by the WGCNA package with the DEGs from the above data and mRNAsi scores. Pearson's correlation between genes and the power values were used to construct the weighted adjacency matrix and scale-free co-expression network Topological overlap matrix (TOM). Based on TOM analysis and clustering of genes with similar functions, the module with a minimum size of 50 genes was built for further construction of module.

To ensure the accuracy of the module, the module eigengenes (MEs) as an expression pattern with a distinct expression feature was used to calculate the correlation between different modules. Similar modules were aggregated with cutoff (ME scores < 0.25) into one module. The gene significance (GS) was defined as the correlation between DEGs expression and sample traits. The module membership (MM) was calculated to identify the correlation between genes in a certain module and gene expression profiles. The key genes and significant modules were confirmed by cutoff (cor.gene MM > 0.8 and cor.gene GS > 0.5).

## **2.6 Functional Annotation of modules genes**

R packages cluster “Profiler”, “enrichplot”, and “ggplot2” were used to analyze gene ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG). The content of GO analysis consisted of cellular component (CC), molecular function (MF), biological process (BP).

## **2.7 Protein-Protein Interactions (PPI) network and Co-expression analysis of genes**

To visualize the interactions among proteins in significant modules, we used the STRING website(<https://string-db.org/>) and the Cytoscape software to construct the PPI network. The STRING database was used to predict protein interaction and construct an interaction network. The “corplot” package was used to visualize the correlation between genes.

## **2.8 Immune scores evaluation, immunity stage classification and immune cell infiltration analysis in COAD**

ESTIMATE package was applied to infer the immune score with RNA-Seq expression data. The “ggExtra” and “ggpubr” packages were applied to calculate the correlation between mRNAsi scores and immune scores of each sample by spearman correlation analysis. The correlation plot between mRNAsi scores and immune scores was drawn with “ggplot2” package. COAD patients with low immune, medium immune and high immune stage were distinguished by Single sample Gene Set Enrichment Analysis (ssGSEA) based on the genetic characteristics expressed by 29 immune cells populations. was used to deconvolve the proportion of 22 types of immune cells infiltration was deconvolved by the CIBERSORT algorithm with detecting 547 immune cell-related gene expression. 22 immune cells include naive B cells, memory B cells, plasma cells, CD8 + T cells, naive CD4 + T cells, resting CD4 + memory T cells, activated CD4 + memory T cells, follicular helper T cells, regulatory T cells (Tregs), gamma delta T cells, resting Natural Killer(NK) cells, activated NK cells, monocytes, M0 macrophages, M1 macrophages, M2 macrophages, resting dendritic cells, activated dendritic cells, resting mast cells, activated mast cells, eosinophils, neutrophils. The 22 immune cells infiltration was visualized by the pheatmap package.

## **2.9 Statistical analysis**

All statistical analyses were performed on R (3.6.2) and corresponding packages. The difference of mRNAsi scores between cancer samples and normal samples and the infiltration of the immune cells in COAD with different mRNAsi scores were analyzed by the Wilcoxon rank sum test the Pearson correlation analysis was used for co-expression analysis of genes. The Kaplan Meier analysis was conducted to the overall survival of each group using the “survimer” package. For DEGs analysis, the difference expressed genes between groups were analyzed by the Wilcoxon rank-sum test. The correlation between miRNA scores and immune scores of samples was calculated by Spearman correlation analysis. The p-value < 0.05 was considered statistically significant.

## **Results**

## 3.1 Association of mRNAsi scores with COAD development and progression

To explore the correlation of COAD stemness with patient's clinical characteristics, we first analyzed the difference of mRNAsi score in COAD and colon tissues. As shown in Fig. 1A, the mRNAsi scores were much higher in COAD than in normal tissues.

Moreover, the Kaplan Meier survival curve showed that there was no difference in the overall survival rate between COAD patients with high mRNAsi scores and patients with low mRNAsi scores (Fig. 1B). However, the mRNAsi scores in COAD without distant metastasis was significantly higher than that in distant metastatic COAD (Fig. 1C). Meanwhile, with the progression of COAD lymph node metastasis, the mRNAsi scores in COAD decreased first and then increased (Fig. 1D). The above result suggested that the mRNAsi scores were significantly higher than that in normal colon tissue and increased with COAD metastasis.

## 3.2 Identification of the key modules and genes related to mRNAsi score via WGCNA analysis

Tumor stemness related genes are significantly higher in the tumor, which are important in maintaining tumor self-renewal properties[14]. To screen the genes related to mRNAsi score, we first analyzed and filtered the genes between COAD and normal colon tissues to identify differential expression genes (DEGs). We identified 6477 DEGs between normal tissues and COAD, consisted of 1916 down-regulated DEGs and 4561 up-regulated DEGs. we further used WGCNA analysis to identify biologically significant modules correlated with mRNAsi scores. For accurate analysis, 4 outliers were removed and the DEGs were divided into different modules with the clusters. Then the appropriate power value ( $\beta = 4$ , scale-free  $R^2 = 0.950$ ) was selected as a soft threshold for the scale-free network and MEDissThres as 0.25 was set to merge similar modules into one module (Fig. 2A). Furthermore, 11 modules were screened out and the heat map of the modules was drawn to reveal the correlation scores between COAD characteristics and the patient's mRNAsi scores and EREG-mRNAs scores (Fig. 2B).

Given that mRNAsi showed better consistency for most tumors compared to the EREG-mRNAsi, based on both RNA expression and epigenetics, elucidates the discrepancy between mRNAsi and shows a positive correlation with mRNAsi[13]. The green module was most positively correlated with the mRNAsi scores of COAD patients (Fig. 2C), while the brown module and the red module exhibited a significantly negative correlation with the mRNAsi scores of COAD patients (Fig. 2D-E). In addition, the brown and green modules exhibited a high correlation with mRNAsi scores. Thus, we chose the genes ( $GS > 0.5$  and  $MM > 0.8$ ) in the brown and green modules for subsequent analyses. The genes in the brown modules were defined as genes related mRNAsi score that was positively correlated with stemness of COAD.

## 3.3 Functional enrichment and co-expression analysis of COAD stemness related genes

In order to explore whether the genes in the module have similar functions and internal connections, we calculated the correlation between genes and performed GO and KEGG pathway enrichment analysis on these genes in the brown and green modules. As shown in Fig. 3A, GO analysis revealed that the genes in the brown module were associated with contractile fiber and actin binding, while the genes in the green module were associated with mitotic nuclear division, nuclear division, organelle fission, regulation of cell cycle phase transition, chromosomal region and protein serine/threonine kinase activity (Fig. 3B). In terms of KEGG analysis, the genes in the brown module mainly focused on cGMP – PKG signaling and the genes in the green module were associated with cell cycle (Fig. 3C-D). As shown in Fig. 3E-F, there was a strongly positive correlation between the genes in the green and the brown module.

### **3.4 Validation of COAD stemness related genes and construction of PPI network**

In order to demonstrate that the genes related mRNasi scores are higher in COAD with high stemness, we performed real-time RT-PCR to detect the expression of PLK4, TTK, CHEK1, KIF18A and BUB1 in HT29 and RKO colon cancer cell lines, which was derived from carcinoma and commonly [17, 18]. The RT-PCR result showed that the expression level of these COAD stemness related genes was significantly increased in HT29 and RKO cells than in normal colon epithelial cells FHC (Fig. 4A). Also, the expression of PLK4, TTK, CHEK1, KIF18A and BUB1, as well as cancer stem cell marker genes, including CD44, CD90 CD133 and WNT-3A, was higher in the secondary spheres of HCT116 (Fig. 4B). In the GSE76402 cohort, PLK4, TTK, CHEK1 and BUB1 were higher expression in colon cancer while KIF18A was higher expression in colon tissues (Fig. 4C). Furthermore, the spheres of HCT116 were capable of the sphere-forming ability and expressed high CD133 and CD44 (Fig. 4D-E). According to the protein expression in colon cancer and normal tissues from HPA database, colon cancer has higher protein expression of TTK, PLK4 and KIF18A (Fig. 4F). BUB1 and CHEK1 were not reported in HPA database. These results were indicating that these COAD stemness related genes might be biomarker for colon cancer stemness. Interplays among the genes in the same module were analyzed by the online tool STRING and constructed with the PPI network. As shown in Supplementary Fig. 2A, 124 nodes and 182 edges in the brown module were formed in the network and the number of remarkable nodes was shown in bar plot. Moreover, 5 nodes were formed in the green module and each node had a connection with another (Supplementary Fig. 2B).

### **3.5 Correlation of COAD stemness-related genes with immune cells infiltration.**

Mounting evidence showed that high stemness of tumor is negatively correlated with antitumor immunity, leading to poor immunotherapeutic efficacy [19, 20]. We first calculated the immune cell purity in COAD using an estimation algorithm based on RNA-seq expression data and analyzed the correlation between COAD stem related genes and immune scores. As shown in Supplementary Fig. 3, immune scores were also inversely correlated with the expression of the COAD stemness related genes and mRNasi scores. Moreover, we also interrogated the different level of COAD stemness related genes and mRNasi scores in different immune stages (Fig. 5A). Interestingly, the expression of TTK and mRNasi

scores decreased with the immune stages (Fig. 5B-C), there was no difference of expression of PLK4, CHEK1, KIF18A and BUB1 with the immune stage (Fig. 5D-G).

To further explore the association between the TIICs and COAD stemness, we analyzed the difference of infiltrating immune cells in different TTK level patients. CIBERSORT algorithm analysis showed that M1 Macrophages were significantly increased in patients with high TTK expression than in patients with low TTK expression, while memory resting CD4 + T cells were reduced in high TTK expression of patients than in low TTK expression of patients (Fig. 6A-B).

Finally, the Kaplan Meier curve was established to analyze the overall survival rate of COAD patients in different infiltration of immune cells and TTK expression. In Fig. 6C, the expression of TTK not affect overall survival of COAD patients. There was no statistical difference in survival rate between COAD patients in a different proportion of memory resting CD4 + T cells (Fig. 6D), but patients with low M1 Macrophages had a better prognosis (Fig. 6E). As shown in Supplementary Fig. 4, M1 Macrophages were significantly higher in patients with high COAD stemness related genes. This result confirmed that COAD stemness and related genes would affect the infiltration of immune cells.

## Discussion

COAD is the fourth most deadly cancer in the world, with a mortality rate of approximately 50% every year worldwide[2]. Surgical treatment and chemotherapy of early COAD can significantly improve the survival time of patients, but advanced COAD with chemotherapy and radiotherapy often have drug resistance and liver metastasis. With the in-depth research on the tumor microenvironment, CSCs, a type of cell subset, in COAD have been found to play an important role in the drug resistance of patients and the regulating immune cells in the tumor microenvironment[7, 8]. Tumor with high stemness has the characteristics of CSCs, which inhibit T cells and NK cell activation and finally leads to a poor outcome of immunotherapy[21–24]. But the in-depth mechanism of COAD stemness related genes resisting immunotherapy is not clear. In this study, we focus on identifying the genes related the COAD stemness and evaluating the correlation between COAD stemness-related genes and the immune cells in tumor environment.

We identified that as normal colon tissue developed into tumor tissue and deteriorates into lymph node metastasis status, mRNAsi scores become higher. Tumors with high stemness are more likely to migrate to other organs and lead to disease progression in patients[14], which is consistent with the increased mRNAsi scores in COAD. The result indicated that the stemness of COAD might be related to COAD progression and metastasis, but COAD patients with different mRNAsi scores had no significant difference in survival rate.

WGCNA analysis was applied to hierarchical cluster. the modules related the COAD stemness, and the brown and green modules were defined as key modules. Functional annotations of genes from the brown module mainly focused on contractile fiber, myofibril, actin binding and cGMP-PKG signaling pathway, which is down-expressed in COAD and closely related to COAD metastasis and recurrence[25, 26].

Besides, functional analysis of key genes from the green module was only enriched in cell cycle. The genes from the green module were highly positively correlated with mRNAsi scores, which might remain the stemness of COAD. The PPI network and co-expression diagrams indicated that genes in the same module had a high co-expression and interaction correlation with each other. The genes in green module had a highly positive relation with COAD stemness, and they were up-expressed in the spheres of HCT116. Hence, PLK4, TTK, CHEK1, KIF18A and BUB1 were identified as COAD stemness-related genes from the green module.

The functions of COAD stemness-related genes were mainly enriched in the cell cycle, which explained the phenomenon that CSCs promote COAD in a highly self-renewing state[27]. The Polo-like kinases (PLKs) play a critical role in centrosome replication and cell division and checkpoint regulation of mitosis, and Plk4 over-expression or activity can cause chromosomal instability and mitotic errors which would promote colon cancer development [28, 29]. TTK is an essential kinase to regulate chromosome alignment and error correction, which is positively correlated with the worse prognosis in several cancer[30, 31]. High expression level of TTK in colon cancer might regulate the functions of spindle assembly checkpoint and mitochondria to enhance the survival of aneuploid cancer[32]. BUB1 mitotic checkpoint serine/threonine kinase (BUB1) could regulate the chromosome segregation and was reported over-expressing in many cancers and was also positively correlated with poor prognosis[33, 34]. Bub1 insufficiency accelerates the rate at which whole chromosomes are lost or gained in ApcMin/+ mice, thereby perhaps accelerating Apc LOH and driving colonic tumorigenesis[35]. KIF18A, a member of the kinesin superfamily of molecular motor proteins, is a microtubule depolymerase and plays an important role in regulating the stability of microtubule plus ends in the mitotic spindle, and colon cancer cells with high expression of KIF18A show stronger migration[36].

Previous studies have found that colon CSCs inhibit immune cells activity in the tumor microenvironment and led to the poor outcome of immunotherapy[37–39]. we used the ESTIMATE algorithm to infer immune score and explored the correlation between immune score and mRNAsi score and COAD stemness-related genes[40]. The immune scores were found to be negatively correlated with the mRNAsi scores and COAD stemness-related genes. This would suggest that high expression of COAD stemness-related genes prevents the immunity of patients, which might cause the failure of the clinical effectiveness of immunotherapy.

Subsequently, we analyzed the association between TIICs and COAD stemness-related genes, and high TTK in COAD might lead to less infiltration of memory resting CD4 + T cells and more infiltration of M1 Macrophages. CD4 + T-cell activation is required for effective CD8 T-cell responses, and the higher infiltration of resting CD4 + T cells might lead to low CD8 T cells responses and poor immunotherapy[41]. This suggested that COAD with high stemness promoted poor outcome of immunotherapy by affecting the inactivation of memory CD4 + T cells. Interestingly, in COAD with high expression of TTK, the infiltration of M1 Macrophages was higher. Macrophages as a key driver of tumor inflammation, contribute to nurturing cancer stem cells and modulate CSC formation and maintenance by transferring into the M1 phenotype [42, 43]. This evidence suggests that inactivation of CD4 + T cells and low

infiltration of M1 Macrophages might be detrimental to cancer stemness related gene expression and immune escape. Overall survival analysis had elucidated significant roles for high M1 Macrophages in predicting prognosis for COAD patients, the COAD patients with low M1 Macrophages infiltration were considered to have a better prognosis. The result above demonstrated that the stemness of COAD was related to infiltrating immune cell in the tumor microenvironment.

In conclusion, PLK4, TTK, CHEK1, KIF18A and BUB1 were positively correlated with stemness of COAD, and TTK might be associated with CD4 T cells activation and M1 Macrophages infiltration in the tumor microenvironment. Also, the increased infiltration of memory activated CD4 + T cells and M1 Macrophages might further contribute to formation and immune escape of COAD with high stemness. TTK may have the potential as a prognostic biomarker and adjuvant target to promote immunotherapy sensitivity to improve outcomes in COAD. But conclusion derived from our study was mainly based on bioinformatic analysis, but it was not enough to reflect the gene expression in COAD with high stemness.

## Abbreviations

COAD, colon adenocarcinoma; TCGA, The Cancer Genome Atlas; WGCNA, The Weighted Gene Co-expression Network Analysis; RNA-Seq, RNA sequencing; GEO, Gene Expression Omnibus; HPA, Human Protein Atlas database; RT-PCR, Real-time PCR; CSCs, Cancer stem cells; EREG-mRNAsi, Epigenetically regulated mRNAsi; TOM, Topological overlap matrix; CC, Cellular Component; MF, Molecular Function; BP, Biological Process; ssGSEA, Single Sample Gene Set Enrichment Analysis; TIICs, Tumor infiltrating immune cells; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, Differential Expression Genes; DCLK1, Doublecortin-like kinase 1; BUB1, BUB1 mitotic checkpoint serine/threonine kinase; PLKs, Polo-like kinases;

## Declarations

### Ethics approval and consent to participate

Not applicable

### Consent for publication

Not applicable

### Data Availability Statement

Publicly available datasets were analyzed in this study, data can be download from TCGA(<https://portal.gdc.cancer.gov/>) and GEO (<https://www.ncbi.nlm.nih.gov/geo/>) database.

### Competing interests

There is not any other conflict of interest by the all of the authors.

## Funding

This work was supported financially by the National Natural Science Foundation of China (Grant No. 81772761, 81672540, 81472407); Science and Technology Foundation of Guangzhou in China (Grant No. 201607010351, 201707010303). President Foundation of The Fifth Affiliated Hospital, Southern Medical University (YZ2017ZD002). Funding bodies did not have any influence in the design of the study and data collection, analysis and interpretation of data or in writing the manuscript.

## Author contributions

Hua Chen and Zhicong Wu designed this research, performed data analysis, and wrote the manuscript. Nisha Wu, Zunya Ma and Yanling Liang performed experiments and data analysis. Fangyin Zeng and Fan Deng designed the study, interpreted the data and approved the final version of the manuscript for publication. All authors read and approved the final manuscript.

## Acknowledgment

We gratefully thank Ming Liu (Nanfang Hospital, Southern Medical University) to provide the information about colon cancer cells culture.

## References

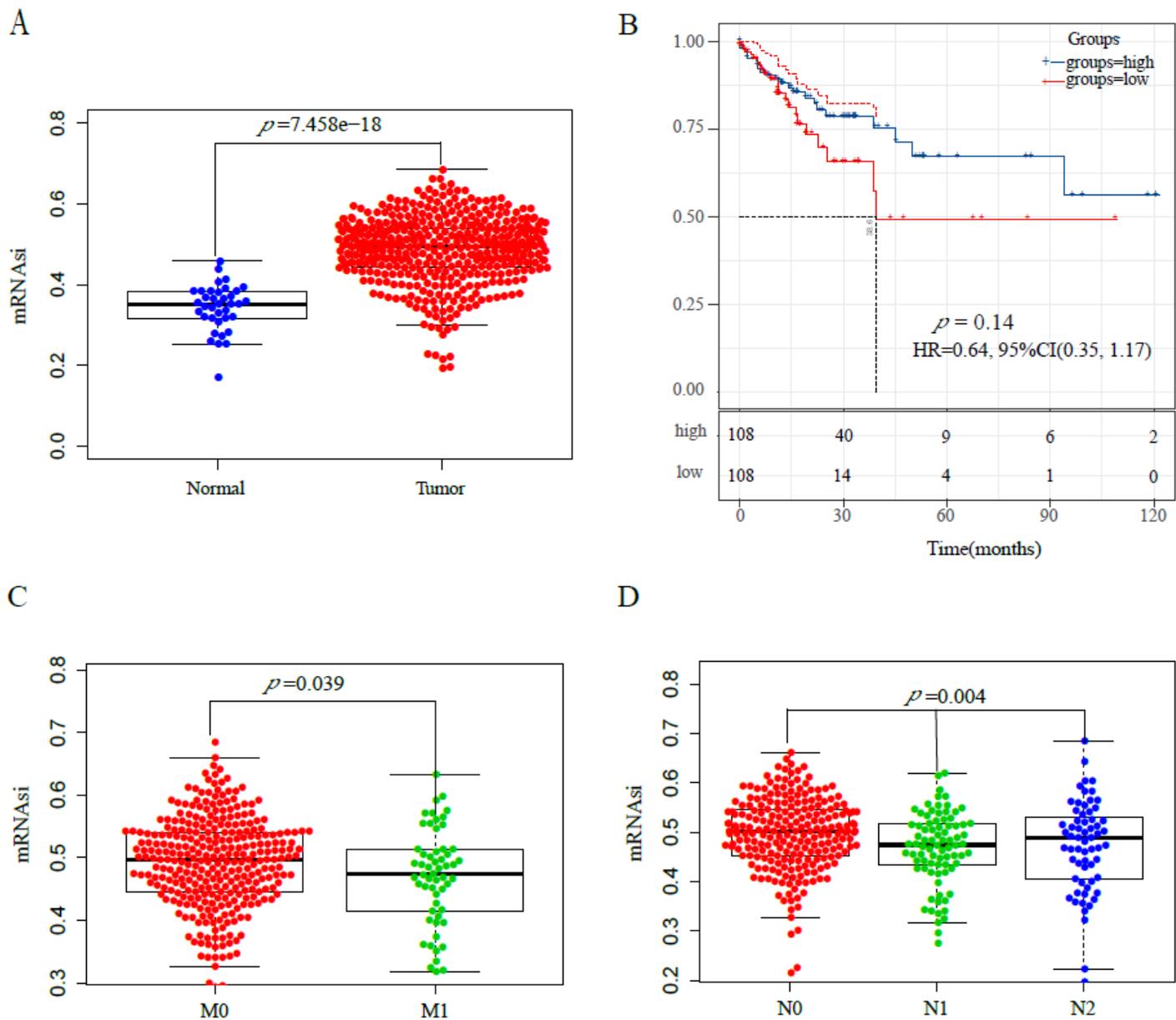
1. Haggard FA, Boushey RP: **Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors.** *Clin Colon Rectal Surg* 2009, **22**(4):191–197.
2. Siegel RL, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, Jemal A: **Colorectal cancer statistics, 2017.** *CA Cancer J Clin* 2017, **67**(3):177–193.
3. Chung KY, Gore I, Fong L, Venook A, Beck SB, Dorazio P, Criscitiello PJ, Healey DI, Huang B, Gomez-Navarro J *et al*: **Phase II study of the anti-cytotoxic T-lymphocyte-associated antigen 4 monoclonal antibody, tremelimumab, in patients with refractory metastatic colorectal cancer.** *J Clin Oncol* 2010, **28**(21):3485–3490.
4. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, Powderly JD, Carvajal RD, Sosman JA, Atkins MB *et al*: **Safety, activity, and immune correlates of anti-PD-1 antibody in cancer.** *N Engl J Med* 2012, **366**(26):2443–2454.
5. Nassar D, Blanpain C: **Cancer Stem Cells: Basic Concepts and Therapeutic Implications.** *Annu Rev Pathol* 2016, **11**:47–76.
6. Shimizu M, Tanaka N: **IL-8-induced O-GlcNAc modification via GLUT3 and GFAT regulates cancer stem cell-like properties in colon and lung cancer cells.** *Oncogene* 2019, **38**(9):1520–1533.
7. Lazarus J, Maj T, Smith JJ, Perusina Lanfranca M, Rao A, D'Angelica MI, Delrosario L, Girgis A, Schukow C, Shia J *et al*: **Spatial and phenotypic immune profiling of metastatic colon cancer.** *JCI Insight* 2018, **3**(22).

8. Zhang Y, Kang M, Zhang B, Meng F, Song J, Kaneko H, Shimamoto F, Tang B: **m(6)A modification-mediated CBX8 induction regulates stemness and chemosensitivity of colon cancer via upregulation of LGR5.** *Mol Cancer* 2019, **18**(1):185.
9. Paczulla AM, Rothfelder K, Raffel S, Konantz M, Steinbacher J, Wang H, Tandler C, Mbarga M, Schaefer T, Falcone M *et al*: **Absence of NKG2D ligands defines leukaemia stem cells and mediates their immune evasion.** *Nature* 2019, **572**(7768):254+.
10. Was H, Czarnecka J, Kominek A, Barszcz K, Bernas T, Piwocka K, Kaminska B: **Some chemotherapeutics-treated colon cancer cells display a specific phenotype being a combination of stem-like and senescent cell features.** *Cancer Biol Ther* 2018, **19**(1):63–75.
11. Clara JA, Monge C, Yang Y, Takebe N: **Targeting signalling pathways and the immune microenvironment of cancer stem cells - a clinical update.** *Nat Rev Clin Oncol* 2020, **17**(4):204–232.
12. Baaten BJ, Li CR, Deiro MF, Lin MM, Linton PJ, Bradley LM: **CD44 regulates survival and memory development in Th1 cells.** *Immunity* 2010, **32**(1):104–115.
13. Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kaminska B, Huelsken J, Omberg L, Gevaert O *et al*: **Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation.** *Cell* 2018, **173**(2):338–354 e315.
14. Liao Y, Xiao H, Cheng M, Fan X: **Bioinformatics Analysis Reveals Biomarkers With Cancer Stem Cell Characteristics in Lung Squamous Cell Carcinoma.** *Front Genet* 2020, **11**:427.
15. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis.** *BMC Bioinformatics* 2008, **9**:559.
16. Giulietti M, Occhipinti G, Righetti A, Bracci M, Conti A, Ruzzo A, Cerigioni E, Cacciamani T, Principato G, Piva F: **Emerging Biomarkers in Bladder Cancer Identified by Network Analysis of Transcriptomic Data.** *Front Oncol* 2018, **8**:450.
17. Saltanatpour Z, Johari B, Alizadeh A, Lotfinia M, Majidzadeh AK, Nikbin B, Kadivar M: **Enrichment of cancer stem-like cells by the induction of epithelial-mesenchymal transition using lentiviral vector carrying E-cadherin shRNA in HT29 cell line.** *J Cell Physiol* 2019, **234**(12):22935–22946.
18. Liu C, Banister CE, Weige CC, Altomare D, Richardson JH, Contreras CM, Buckhaults PJ: **PRDM1 silences stem cell-related genes and inhibits proliferation of human colon tumor organoids.** *Proc Natl Acad Sci U S A* 2018, **115**(22):E5066-E5075.
19. 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* 2019, **116**(18):9020–9029.
20. Pan S, Zhan Y, Chen X, Wu B, Liu B: **Identification of Biomarkers for Controlling Cancer Stem Cell Characteristics in Bladder Cancer by Network Analysis of Transcriptome Data Stemness Indices.** *Front Oncol* 2019, **9**:613.
21. Lenos KJ, Miedema DM, Lodestijn SC, Nijman LE, van den Bosch T, Romero Ros X, Lourenco FC, Lecca MC, van der Heijden M, van Neerven SM *et al*: **Stem cell functionality is microenvironmentally**

- defined during tumour expansion and therapy response in colon cancer. *Nat Cell Biol* 2018, **20**(10):1193–1202.
22. Manic G, Sistigu A, Corradi F, Musella M, De Maria R, Vitale I: **Replication stress response in cancer stem cells as a target for chemotherapy.** *Semin Cancer Biol* 2018, **53**:31–41.
  23. Miao Y, Yang H, Levorse J, Yuan S, Polak L, Sribour M, Singh B, Rosenblum MD, Fuchs E: **Adaptive Immune Resistance Emerges from Tumor-Initiating Stem Cells.** *Cell* 2019, **177**(5):1172–1186 e1114.
  24. Pei J, Wang Y, Li Y: **Identification of key genes controlling breast cancer stem cell characteristics via stemness indices analysis.** *J Transl Med* 2020, **18**(1):74.
  25. Lv Y, Wang X, Li X, Xu G, Bai Y, Wu J, Piao Y, Shi Y, Xiang R, Wang L: **Nucleotide de novo synthesis increases breast cancer stemness and metastasis via cGMP-PKG-MAPK signaling pathway.** *PLoS Biol* 2020, **18**(11):e3000872.
  26. Mierke CT: **The matrix environmental and cell mechanical properties regulate cell migration and contribute to the invasive phenotype of cancer cells.** *Rep Prog Phys* 2019, **82**(6):064602.
  27. Gulaia V, Kumeiko V, Shved N, Cicinskas E, Rybtsov S, Ruzov A, Kagansky A: **Molecular Mechanisms Governing the Stem Cell's Fate in Brain Cancer: Factors of Stemness and Quiescence.** *Front Cell Neurosci* 2018, **12**:388.
  28. Fan G, Sun L, Shan P, Zhang X, Huan J, Zhang X, Li D, Wang T, Wei T, Zhang X *et al.*: **Loss of KLF14 triggers centrosome amplification and tumorigenesis.** *Nat Commun* 2015, **6**:8450.
  29. Zhao Y, Wang X: **PLK4: a promising target for cancer therapy.** *J Cancer Res Clin Oncol* 2019, **145**(10):2413–2422.
  30. Daniel J, Coulter J, Woo JH, Wilsbach K, Gabrielson E: **High levels of the Mps1 checkpoint protein are protective of aneuploidy in breast cancer cells.** *Proc Natl Acad Sci U S A* 2011, **108**(13):5384–5389.
  31. Tighe A, Staples O, Taylor S: **Mps1 kinase activity restrains anaphase during an unperturbed mitosis and targets Mad2 to kinetochores.** *J Cell Biol* 2008, **181**(6):893–901.
  32. Zhang X, Ling Y, Guo Y, Bai Y, Shi X, Gong F, Tan P, Zhang Y, Wei C, He X *et al.*: **Mps1 kinase regulates tumor cell viability via its novel role in mitochondria.** *Cell Death Dis* 2016, **7**(7):e2292.
  33. Chou CK, Wu CY, Chen JYF, Ng MC, Wang HMD, Chen JH, Yuan SSF, Tsai EM, Chang JG, Chiu CC: **BubR1 Acts as a Promoter in Cellular Motility of Human Oral Squamous Cancer Cells through Regulating MMP-2 and MMP-9.** *Int J Mol Sci* 2015, **16**(7):15104–15117.
  34. Han JY, Han YK, Park GY, Kim SD, Lee CG: **Bub1 is required for maintaining cancer stem cells in breast cancer cell lines (vol 5, 15993, 2016).** *Sci Rep-Uk* 2016, **6**.
  35. Baker DJ, Jin F, Jeganathan KB, van Deursen JM: **Whole Chromosome Instability Caused by Bub1 Insufficiency Drives Tumorigenesis through Tumor Suppressor Gene Loss of Heterozygosity.** *Cancer Cell* 2009, **16**(6):475–486.
  36. Nagahara M, Nishida N, Iwatsuki M, Ishimaru S, Mimori K, Tanaka F, Nakagawa T, Sato T, Sugihara K, Hoon DS *et al.*: **Kinesin 18A expression: clinical relevance to colorectal cancer progression.** *Int J Cancer* 2011, **129**(11):2543–2552.

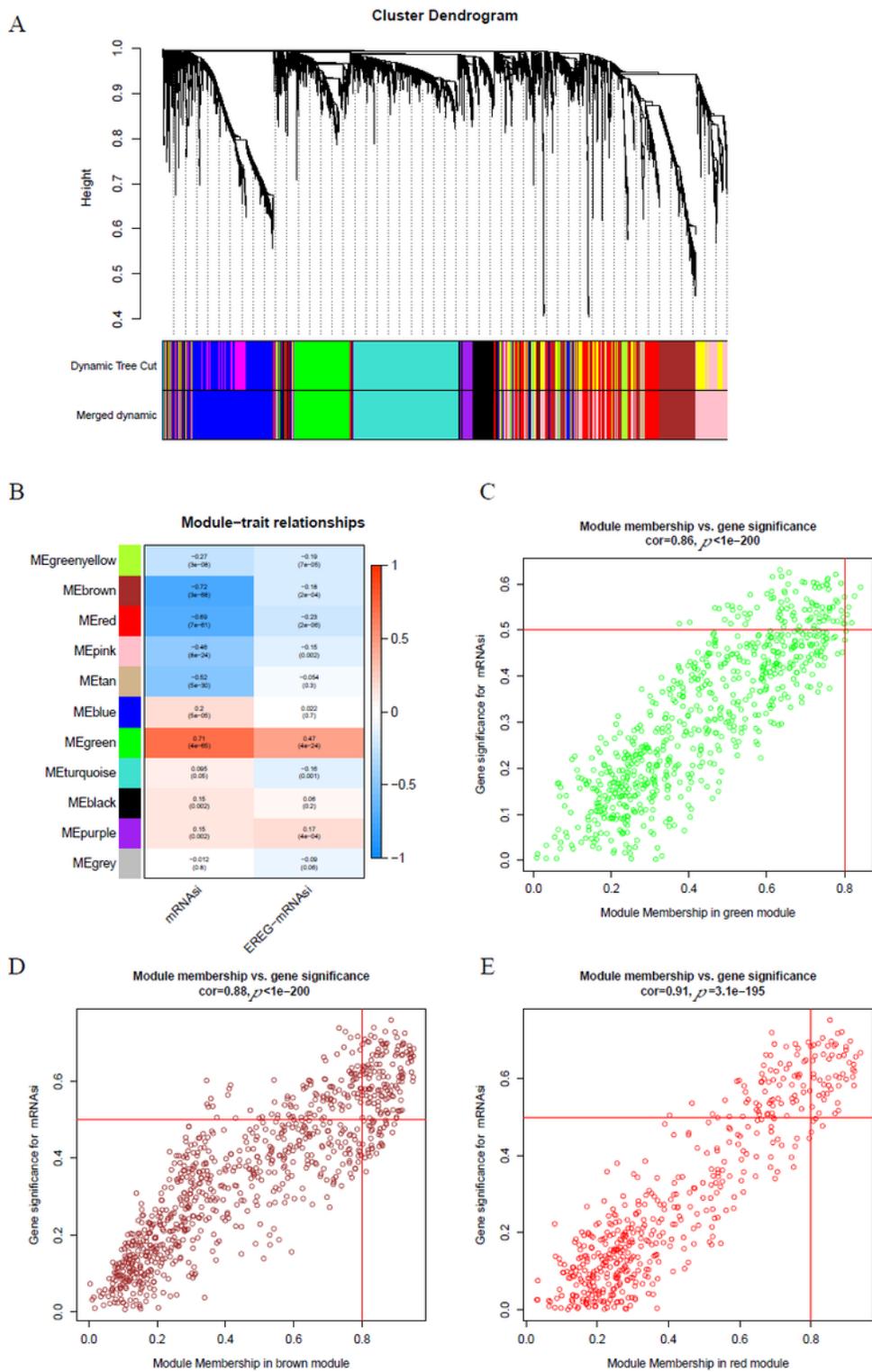
37. Mathonnet M, Perraud A, Christou N, Akil H, Melin C, Battu S, Jauberteau MO, Denizot Y: **Hallmarks in colorectal cancer: angiogenesis and cancer stem-like cells.** *World J Gastroenterol* 2014, **20**(15):4189–4196.
38. Tauriello DVF, Palomo-Ponce S, Stork D, Berenguer-Llergo A, Badia-Ramentol J, Iglesias M, Sevillano M, Ibiza S, Canellas A, Hernando-Momblona X *et al.*: **TGFbeta drives immune evasion in genetically reconstituted colon cancer metastasis.** *Nature* 2018, **554**(7693):538–543.
39. Wu X, Qu D, Weygant N, Peng J, Houchen CW: **Cancer Stem Cell Marker DCLK1 Correlates with Tumorigenic Immune Infiltrates in the Colon and Gastric Adenocarcinoma Microenvironments.** *Cancers (Basel)* 2020, **12**(2).
40. Chen H, Luo J, Guo J: **Development and validation of a five-immune gene prognostic risk model in colon cancer.** *BMC Cancer* 2020, **20**(1):395.
41. MacLeod MK, Kappler JW, Marrack P: **Memory CD4 T cells: generation, reactivation and re-assignment.** *Immunology* 2010, **130**(1):10–15.
42. Mantovani A, Marchesi F, Malesci A, Laghi L, Allavena P: **Tumour-associated macrophages as treatment targets in oncology.** *Nat Rev Clin Oncol* 2017, **14**(7):399–416.
43. Guo L, Cheng X, Chen HY, Chen CG, Xie S, Zhao M, Liu D, Deng Q, Liu YJ, Wang XM *et al.*: **Induction of breast cancer stem cells by M1 macrophages through Lin-28B-let-7-HMGA2 axis.** *Cancer Lett* 2019, **452**:213–225.

## Figures



**Figure 1**

mRNAasi scores is related to the survival and progression of COAD patients. (A) The comparison of mRNAasi scores in 473 tumor cases and in 41 normal cases ( $p < 0.001$ ). (B) The Kaplan-Meier curve of high mRNAasi score group ( $n=216$ ) and low mRNAasi score group ( $n=216$ ). Comparison between mRNAasi scores and clinical characteristics in COAD, including M stage(C) and N stage(D).

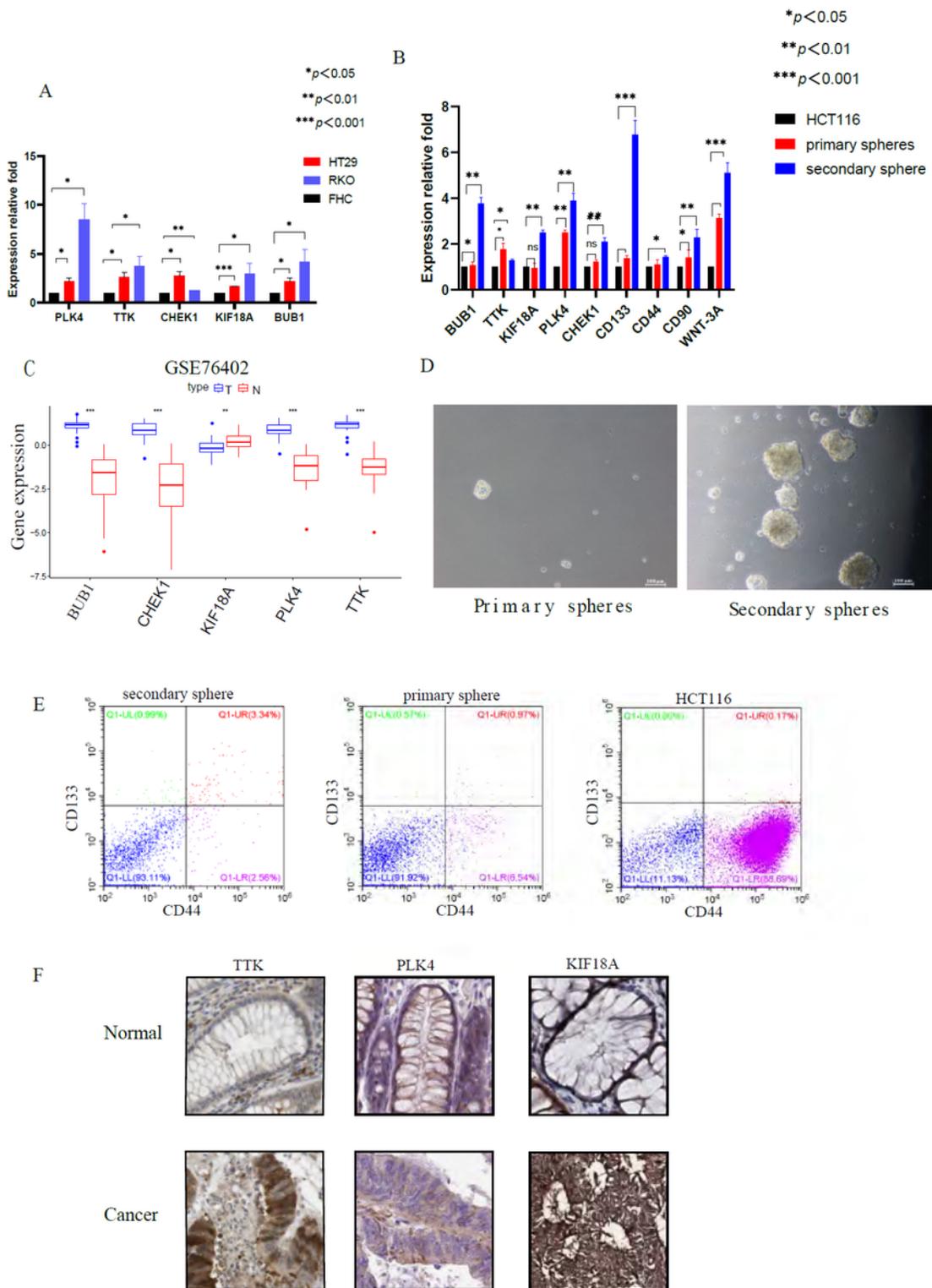


**Figure 2**

Identification of the key genes and modules via WGCNA network. (A) Each piece of the leaves on the cluster dendrogram corresponded to a gene, and those genes with similar expression patterns compose a branch. (B) Heat map of correlation between gene module and mRNAsi scores or EREG-mRNAsi. The number in the upper row of each cell is the correlation coefficient between the specific gene module and mRNAsi or EREG-mRNAsi, while the number in the second row indicates the p-value. The correlation

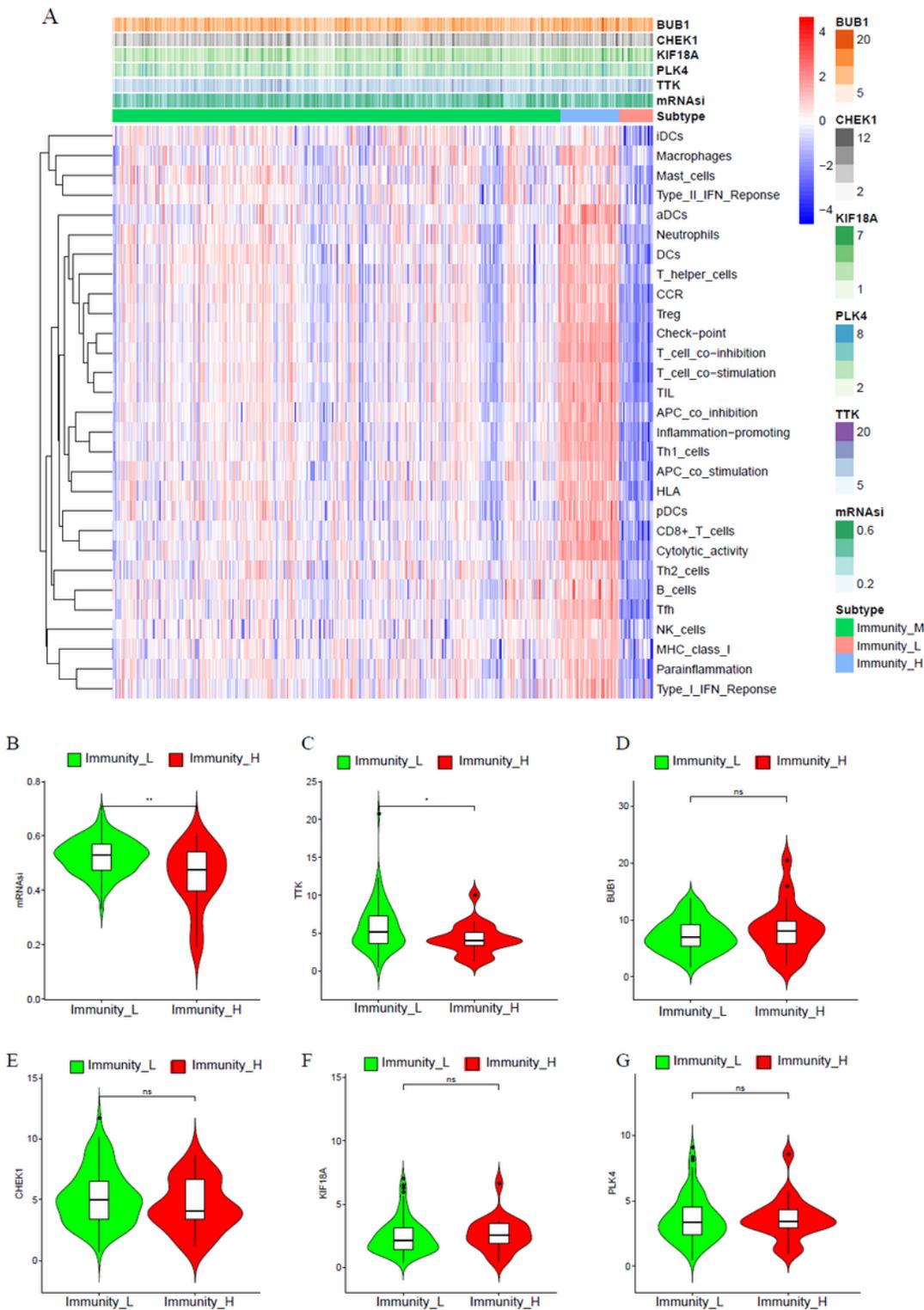


significant genes from the green(E) and brown modules(F).



**Figure 4**

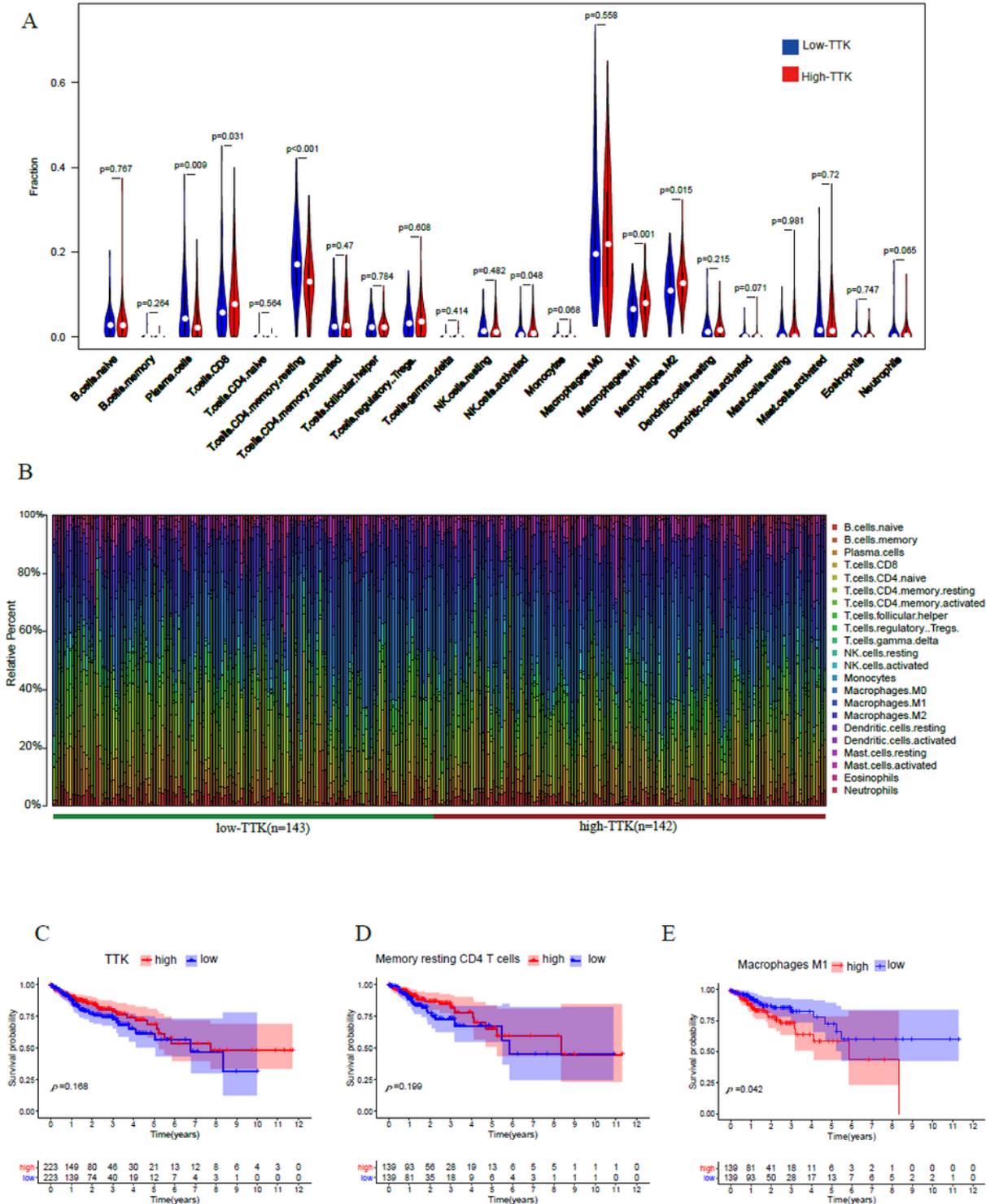
Validation of COAD stemness related genes. The expression of PLK4, TTK, CHEK1, KIF18A and BUB1 in colon cancer cell lines HT29, RKO and normal colon cell lines FHC (A). The expression of PLK4, TTK, CHEK1, KIF18A and BUB1 in colon cancer and colon normal tissues from GSE76402 (B). The spheres of HCT116 and expression of COAD stemness related genes and cancer stem cell marker genes (C, D).



**Figure 5**

TTK and mRNAsi scores are negatively correlated to immunity stage. (A) 29 immune-related gene sets were used to cluster and classify the COAD patients into high immunity stage (n =22), middle immunity stage (n=372) and low immunity stage (n=88), respectively; (B-E) The difference of the expression level of BUB1, CHEK1, KIF18A and PLK4 in low immunity stage and high immunity stage ( $p > 0.05$ ); (F) The difference of the expression level of BUB1, CHEK1, KIF18A and PLK4 in low immunity stage and high

immunity stage( $p < 0.05$ ); (G) The difference of the mRNAsi scores in in low immunity stage and high immunity stage( $p < 0.001$ ).



**Figure 6**

TTK expression is associated with the immune cell infiltration in COAD. (A) Differences of the proportion of 22 types of immune cells infiltration are shown by the violin plot in low and high group of TTK expression. (B) The proportions of immune cells in each COAD sample are presented in bar plot and each

bar represents the proportion of the immune cells. The Kaplan–Meier survival curve of the patients with different TTK expression (C), different infiltration of resting memory CD4+ T cells(D) and M1 macrophages (E).

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigure1.pdf](#)
- [SupplementaryFigure2.pdf](#)
- [SupplementaryFigure3.pdf](#)
- [SupplementaryFigure4.pdf](#)
- [SupplementaryFigurelegend.docx](#)