



Characteristic Analysis of Featured Genes Associated With Stemness Indices in Colorectal Cancer

Yongqu Lu, Xin Zhou*, Zhenzhen Liu, Wendong Wang, Fei Li and Wei Fu*

Department of General Surgery, Peking University Third Hospital, Beijing, China

OPEN ACCESS

Edited by:

Kamran Ghaedi,
University of Isfahan, Iran

Reviewed by:

Nikolay Mikhaylovich Borisov,
Moscow Institute of Physics
and Technology, Russia
Mahendra K. Singh,
University of Miami, United States

*Correspondence:

Xin Zhou
zhouxinasd@sina.cn
Wei Fu
fuwei@bjmu.edu.cn

Specialty section:

This article was submitted to
Molecular Diagnostics
and Therapeutics,
a section of the journal
Frontiers in Molecular Biosciences

Received: 20 May 2020

Accepted: 14 September 2020

Published: 06 October 2020

Citation:

Lu Y, Zhou X, Liu Z, Wang W, Li F
and Fu W (2020) Characteristic
Analysis of Featured Genes
Associated With Stemness Indices
in Colorectal Cancer.
Front. Mol. Biosci. 7:563922.
doi: 10.3389/fmolb.2020.563922

Cancer stem cells (CSCs) with self-renewal play an important role in tumor initiation and progression and are associated with drug resistance in cancer therapy. Here, we investigated the characteristics of stem cell-related genes in colorectal cancer (CRC) based on datasets from The Cancer Genome Atlas (TCGA) and Oncomine. We found that the stemness indices were significantly overexpressed in CRC tissues and were associated with patient survival. Weighted gene co-expression network analysis (WGCNA) was performed to determine the modules of stemness and featured genes. Significant modules and 8 genes (*BUB1*, *BUB1B*, *CHEK1*, *DNA2*, *KIF23*, *MCM10*, *PLK4*, and *TTK*) were selected according to the inclusion criteria. Expression analyses of transcription and protein levels confirmed internal correlation and their relevance with the tumor. Functional analysis of these genes demonstrated their enrichment in pathways, including checkpoint, chromosomal region and protein serine/threonine kinase activity. These results suggested that the characteristics of the featured genes fit well with CRC pathology and could provide new strategies for individual prevention and treatment.

Keywords: biomarker, colorectal cancer, cancer stem cell, stemness index, WGCNA

INTRODUCTION

Cancer stem cells (CSCs) that fuel tumor hierarchies through minor self-renewal are believed to be dominant in many tumor types, with mechanisms similar to the protection of normal stem cells and its resistance to standardized chemotherapy and radiation treatment partly accounts for disease relapse after cancer treatment (Clevers, 2011). Studies have confirmed that CSCs can generate tumors efficiently upon xenotransplantation without modified stimuli in the environment (Batlle and Clevers, 2017). Programs that focus on eliminating tumor cells have been launched, including proposed therapeutic strategies to interfere with CSCs and some conceptions have been progressively verified in different cancers from hematological malignancies to solid tumors (Harris et al., 2012; Boumahdi et al., 2014). The idea of anti-CSC therapy, aimed at key signaling pathways, epigenetic alterations or other CSC-targeting approaches, benefits from the exploration of biological markers in great depth (Nervi et al., 2015; Takebe et al., 2015). CSC-associated signals and epigenetic mechanisms, as well as regulation

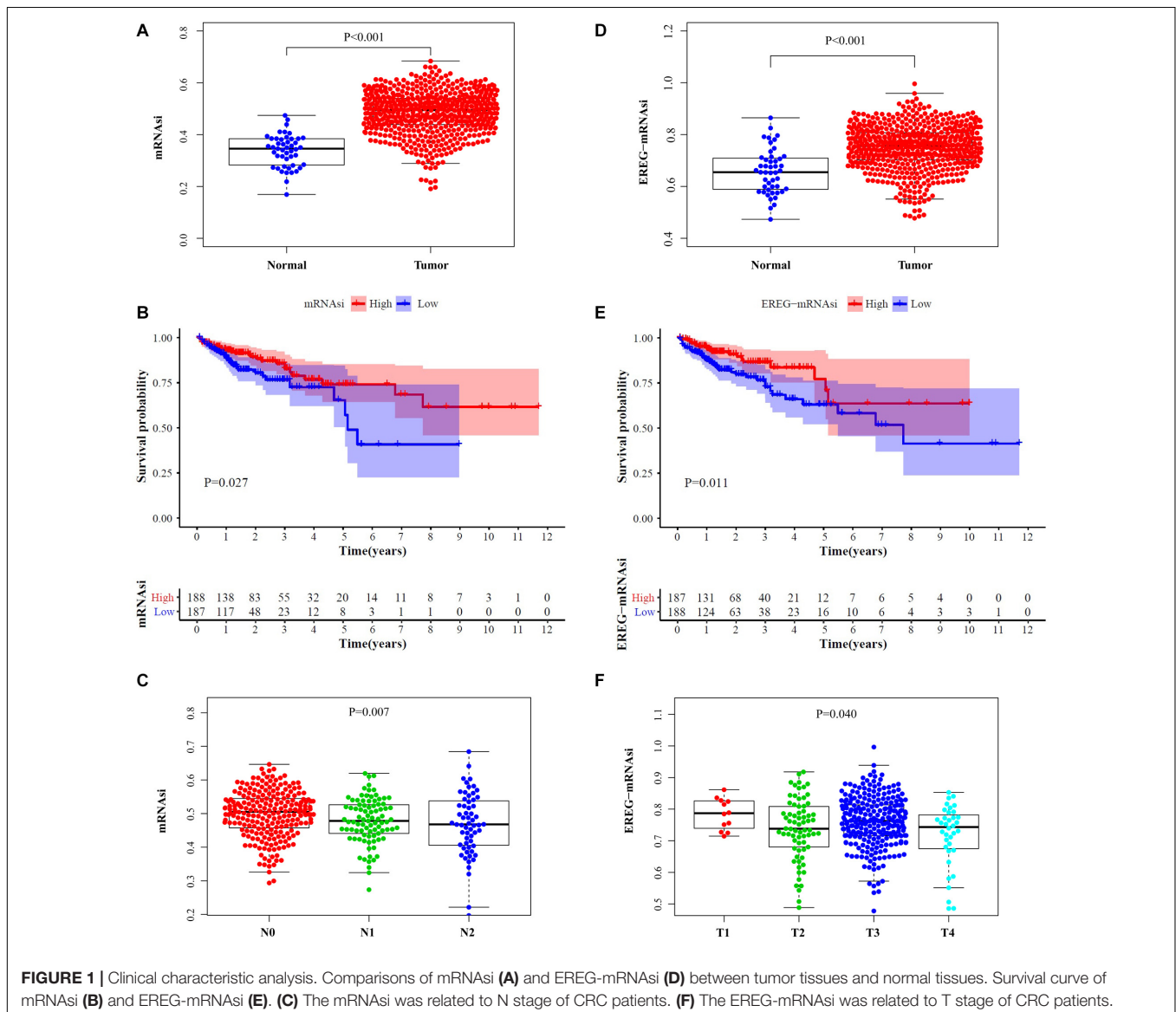
from specific microenvironments, contribute to depict stemness potential and classify the functional characterization of CSCs that are isolated from the genetic background of tumor cells (Maccalli and De Maria, 2015; Nassar and Blanpain, 2016).

As a common malignancy of the digestive tract, colorectal cancer (CRC) is ranked second in terms of mortality worldwide and individual treatments are notably various in different

patients as a result of personal heterogeneities (Bray et al., 2018; Miller et al., 2019). Serological tests, like enzymes and factors, can easily be completed during the initial screening and monitoring of tumor, but have a considerable false positive and false negative rate. For better clinical benefit, genomics research, as a new direction distinguished from traditional physiology, is being applied in CRC patients to uncover its underlying mechanism for more precise prevention, diagnosis and therapy. Characteristic analysis of CRC patients is an urgent process required to find molecular markers with a high specificity and sensibility.

CSCs in CRC can be isolated from tumor tissues with specific identifications, such as *LGR5*, and have self-renewal and differentiation capacity as stemness (Medema, 2017). Malta et al. (2018) assessed the mRNA expression-based signatures based on the expression values of available stem cell samples with different degrees of stemness from the Progenitor Cell Biology

Abbreviations: CRC, colorectal cancer; CSC, cancer stem cell; DAB, 3,3N-Diaminobenzidine Tetrahydrochloride; DEG, differentially expressed gene; EREG, epigenetically regulated; FC, fold change; FDR, false discovery rate; GO, gene ontology; GS, gene significance; HRP, horseradish peroxidase; ME, module eigengene; MM, module membership; mRNAsi, mRNA expression-based stemness index; MS, module significance; OCLR, one-class logistic regression; OS, overall survival; PCA, principal components analysis; PBS, phosphate buffer saline; PCBC, Progenitor Cell Biology Consortium; PPI, protein-protein interaction; STRING, Search Tool for the Retrieval of Interacting Genes; TCGA, The Cancer Genome Atlas; TNM, Tumor Node Metastasis; TOM, topological overlap matrix; WGCNA, weighted gene co-expression network analysis.



Consortium (PCBC) dataset that compatible with The Cancer Genome Atlas (TCGA) and used one-class logistic regression (OCLR) to derive a predictive model of stemness indices, mRNA expression-based stemness index (mRNAsi), that identified stem cell signatures (Hutter and Zenklusen, 2018). A transcriptomic set of stemness-related epigenetically regulated genes was used to generate the epigenetically regulated-mRNAsi (EREG-mRNAsi) with OCLR that exhibited the methylation/expression relationships and the index showed a positive correlation with both RNA expression and epigenetics. The mRNAsi and EREG-mRNAsi were confirmed to be related with biological processes in CSCs and tumor state for CRC and other tumor samples. These efforts provided data support as a measurement of CSCs for our study.

In this study we aimed to identify which featured genes correlated with CSC stemness based on CRC mRNAsi and EREG-mRNAsi from the results of Malta et al. via bioinformatic analysis. Differentially expressed genes (DEGs) in separate modules were determined through weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath, 2008). Genes in the significant module were selected for further confirmation and functional analysis. The integrated assessment of key genes associated with stemness in CRC could inform the initiation and progression of tumor and enlighten novel therapeutic target strategies for patients.

MATERIALS AND METHODS

Data Extraction

Expression profiles of RNA-sequencing data from TCGA¹ and corresponding clinical data were downloaded on January 15, 2020. The stemness indices of CRC patients, including mRNAsi and EREG-mRNAsi, were obtained from the study of Malta et al. (2018). After removal of indices-absent cases, only patients with a definitive diagnosis of CRC and a no less than 30 days overall survival (OS) were brought into the analysis. Furthermore, cases without full clinical parameters on age, sex and Tumor Node Metastasis (TNM) stage were removed from the correlation testing of clinical factors. Finally, 396 CRC patients were included for survival analysis, among these 376 patients had full clinical parameters.

Survival and Clinical Parameter Correlations Analysis

Survival analyses for the mRNAsi and EREG-mRNAsi were performed on R to identify the prognostic values. The correlations between mRNAsi or and EREG-mRNAsi and the clinical parameters were investigated on the same platform.

Screening of DEGs

The limma package in R package was used to compare expression levels between normal samples and tumor samples. The criteria for DEG selection were set as follows: fold change (FC) > 2 and false discovery rate (FDR) < 0.05.

¹<https://portal.gdc.cancer.gov>

WGCNA

Genes with the highest 25% of DEG variance were selected to ensure the heterogeneity and accuracy of the bioinformatics statistics for co-expression network analysis. Outliers in RNA-sequencing data were filtered and a Pearson correlation matrix was constructed for the co-expression analysis of paired genes. A weighted adjacency matrix was constructed using the function $a_{mn} = |c_{mn}|^b$ (a_{mn} as adjacency between gene m and gene n ; c_{mn} as Pearson's correlation between gene m and gene n ; b as power showing a strong correlation between genes and penalized the weak correlation). An appropriate b -value was selected to build a co-expression network and the adjacency matrix was converted into a topological overlap matrix (TOM) to measure the connectivity of the genes in the network. Average linkage hierarchical clustering was carried out according to a TOM-based dissimilarity, and the minimum size of the gene dendrogram was 30 for the construction of module dendrograms.

Featured Gene Identification

The main components were composed of module eigengenes (MEs) and the single expression signature in each given module was summarized by the expression patterns of all genes in the principal components analysis (PCA). Gene significance (GS) was calculated as the log10 conversion of the P -value in linear regression between gene expression and mRNAsi or EREG-mRNAsi. Module significance (MS) was calculated as the average GS in a specific module to stand for the correlation between the module and sample traits. A cutoff of less than 0.25 was used to merge similar modules. The correlation between the genes in corresponding modules and gene expression profiles was defined as module membership (MM). After the modules of interest were selected, the GS and MM of each gene were calculated and thresholds for the selection of key genes were set as $MM > 0.8$ and $GS > 0.5$ in modules.

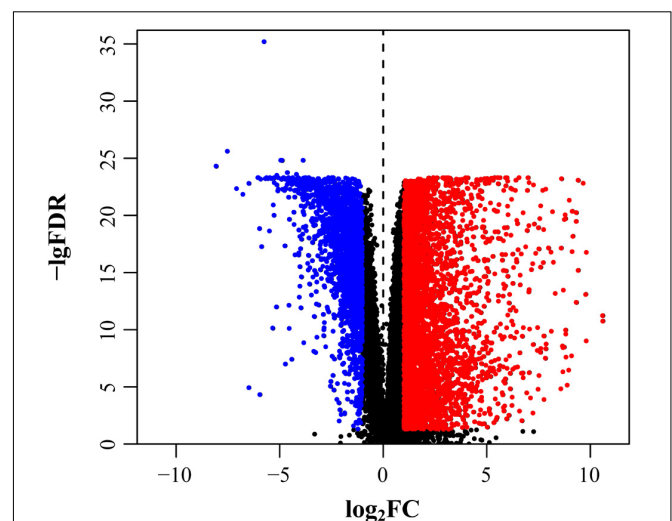
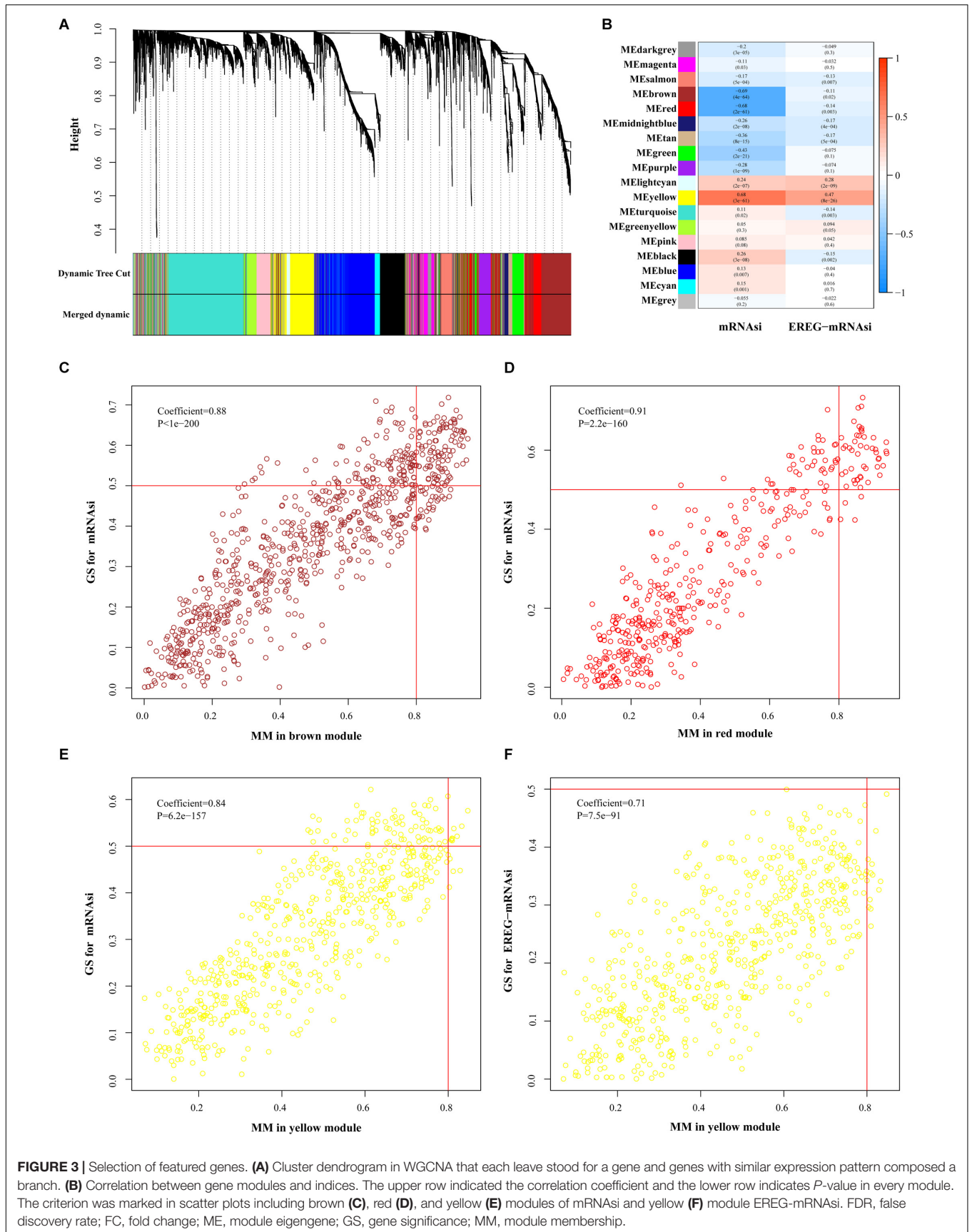


FIGURE 2 | Selection of DEGs. DEGs after screening including upregulation (red) and downregulation (blue).



Differential Expression and Co-expression Analysis

Expression values of the selected genes were compared between the normal samples and tumor samples based on data from the TCGA. The validation utilized mRNA profile in another public database, Oncomine², and the threshold was set as follows: *P*-value, 1E-4; FC, 2; gene level, top 10%.

R package was used to analyze co-expression relationships between selected genes.

Human Samples

All CRC patients provided written informed consent for research study that was approved by the ethical committees. Tissues including normal mucosae, adenomas, carcinomas and liver metastases were sampled immediately after surgical resection of the specimens from Peking University Third Hospital. The samples were assessed by independent pathologists.

Immunohistochemical Staining

Tissues (5 μm thick) were deparaffinized and treated with 3% H₂O₂-CH₃OH for 15 min to block endogenous peroxidase. Samples were submerged in pH 6.0 or 9.0 buffer in a pressure cooker for antigen retrieval and then incubated at 37°C for 2 h with antibody *CHEK1* (Abcam, ab40866, 1:400), *BUB1B* (Abcam, ab183496, 1:100), *BUB1* (Abcam, ab195268, 1:100), *TTK* (Proteintech, 10381-1-AP, 1:400), *PLK4* (Proteintech, 12952-1-AP, 1:400), *DNA2* (Proteintech, 18727-1-AP, 1:400), *MCM10* (SAB, 43512, 1:200) and *KIF23* (CUSABIO, CSB-PA23569A0Rb, 1:400). After washing with phosphate buffer saline (PBS), slices were incubated with horseradish peroxidase (HRP)-conjugated IgG (ZSGB-Bio, PV-6000) at room temperature for 30 min and then stained with 3,3'-Diaminobenzidine Tetrahydrochloride (DAB) detection system kit (ZSGB-Bio, ZLI-9018). Protein expression and localization were detected under light microscopy and analyzed by NDP view (version 2.6.8).

Protein-Protein Interaction (PPI) Network and Functional Enrichment Analysis

A PPI network was constructed on the Search Tool for the Retrieval of Interacting Genes (STRING)³.

R package was used to perform gene ontology (GO) analysis for investigation of the biological functions behind selected genes.

Statistical Analysis

All statistical analyses were performed in R (version 3.6.0)⁴. Wilcoxon testing and Kruskal-Wallis testing were used in the differential analyses. Survival curves were generated through the Kaplan-Meier method and the survival difference of OS was evaluated with log-rank testing. *P* < 0.05 was considered significant in all statistical analyses.

²<https://www.oncomine.org>

³<https://string-db.org>

⁴<https://www.r-project.org>

RESULTS

Stemness Indices and Clinical Significance in CRC

After comparing the scores of mRNasi and EREG-mRNasi between the normal and tumor tissues of CRC patients, we found that mRNasi and EREG-mRNasi were both significantly higher in the tumor samples (*P* < 0.001, **Figures 1A,D**). For investigation of prognostic value, we simultaneously stratified the CRC cohorts into high- and low-score groups on the basis of the medians of mRNasi and EREG-mRNasi and constructed survival curves. The high-scores of the mRNasi and EREG-mRNasi were obviously associated with a better OS in CRC patients with 5-year survival rates of 0.741 and 0.772, respectively (*P* < 0.05, **Figures 1B,E**). Age, sex and TNM stage were included for analysis and the results showed that mRNasi had an overall decreasing trend in N stages, and there was a significant difference in EREG-mRNasi for different T stages (*P* < 0.05, **Figures 1C,F**). The two indices were not correlated with age or sex.

Identification of DEGs and Featured Genes

As mRNasi and EREG-mRNasi in tumor samples were significantly different from the normal samples, we screened out stemness-related DEGs based on the CRC indices. After differential analysis, 6,476 DEGs were screened, of which, 4,414 were upregulated and 2,062 were downregulated (**Figure 2**).

For further filtering, we used WGCNA to select the DEGs, which assigned genes with similar expression patterns into one module and 36 modules were obtained (**Figure 3A**). To find correlations between the gene module and indices, we applied MS as the overall gene expression level of the corresponding module. The results showed that the correlations in the brown, red and yellow modules of mRNasi and the yellow module of EREG-mRNasi were relatively high, indicating more close associations between genes and indices (*P* < 0.001, **Figure 3B**). Key genes associated with indices were screened out using the threshold as defined. In the mRNasi group, 106 genes were in the brown module, 46 genes in the red module and 8 genes in the yellow

TABLE 1 | Featured genes after screening.

Gene	Name
<i>BUB1</i>	BUB1 mitotic checkpoint serine/threonine kinase
<i>BUB1B</i>	BUB1 mitotic checkpoint serine/threonine kinase B
<i>CHEK1</i>	Checkpoint kinase 1
<i>DNA2</i>	DNA replication helicase/nuclease 2
<i>KIF23</i>	Kinesin family member 23
<i>MCM10</i>	Minichromosome maintenance 10 replication initiation factor
<i>PLK4</i>	Polo like kinase 4
<i>TTK</i>	TTK protein kinase

module, while none of the genes were included in the EREG-mRNasi module (Figures 3C–F). We selected the yellow module in the mRNasi group for the subsequent analysis of featured genes (Table 1).

Analysis and Validation of Featured Genes

Expressions of the featured genes were analyzed with TCGA and Oncomine datasets. Results from the TCGA dataset showed that these genes were indeed obviously overexpressed in CRC tissues ($P < 0.001$, Figure 4A). Additionally, we found that the featured genes were overexpressed not only

in CRC but also in many other cancers using the Oncomine dataset (Figure 4B).

We performed a co-expression network analysis for further exploration and validated the evident relationship among the featured genes with a coefficient of at least 0.63. The correlation between *BUB1B* and *KIF23* had the highest coefficient of 0.87 (Figure 5).

For further confirmation at protein levels for these genes, we developed immunohistochemical staining in human samples. Comparing of corresponding adenomas, the expressions of *BUB1*, *BUB1B*, *PLK4*, and *DNA2* in carcinomas were higher. The protein levels of *CHEK1* and *MCM10* were more significant in carcinomas than normal tissues and the differential expressions

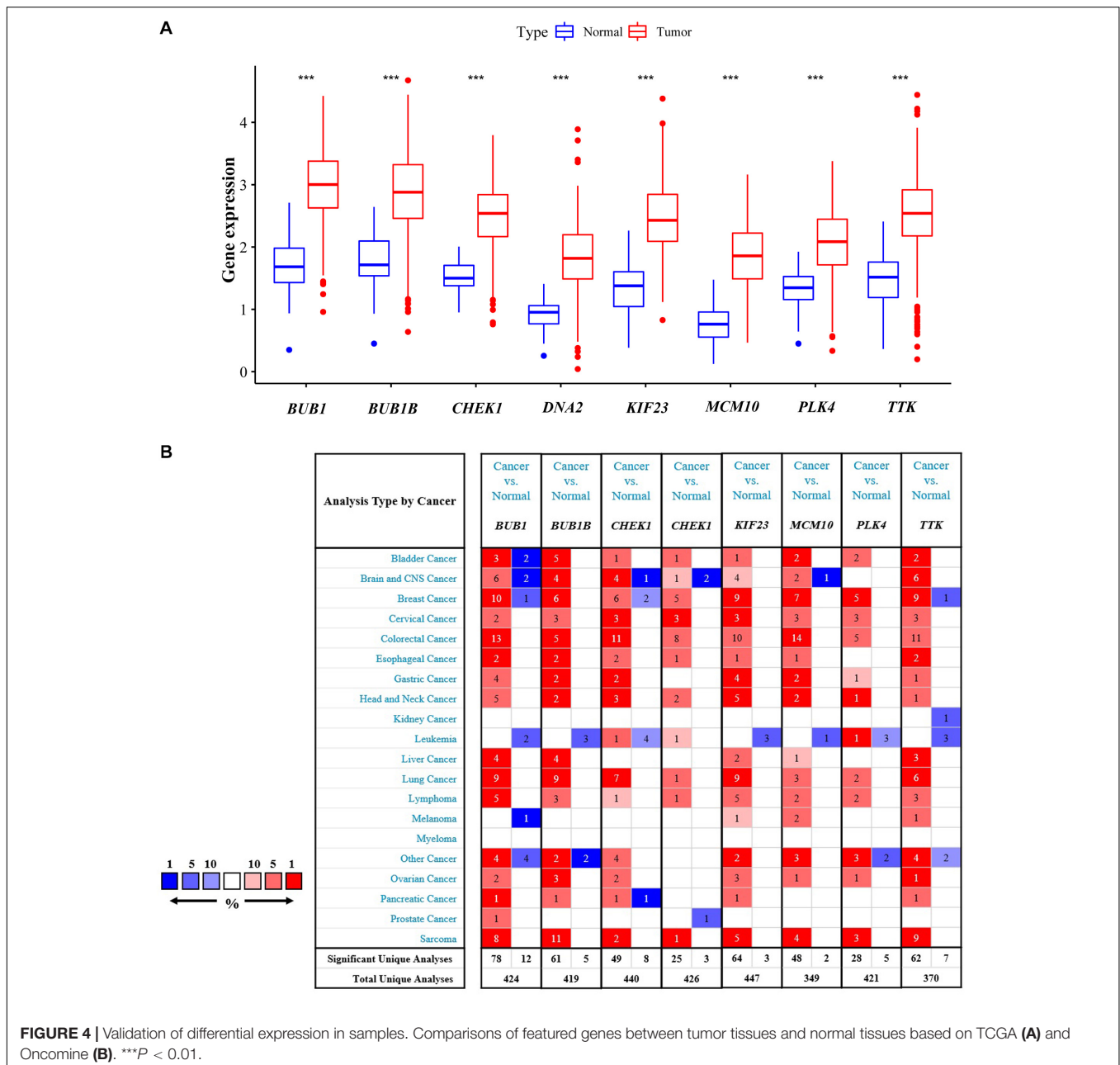
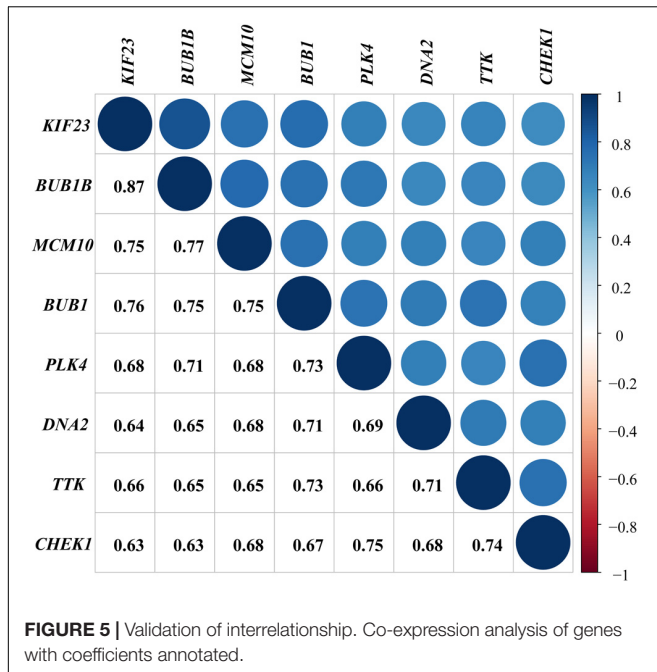


FIGURE 4 | Validation of differential expression in samples. Comparisons of featured genes between tumor tissues and normal tissues based on TCGA (A) and Oncomine (B). *** $P < 0.01$.



of *KIF23* and *TTK* could be observed between primary lesions and liver metastases (Figure 6).

Analysis at Protein Level and Functional Enrichment

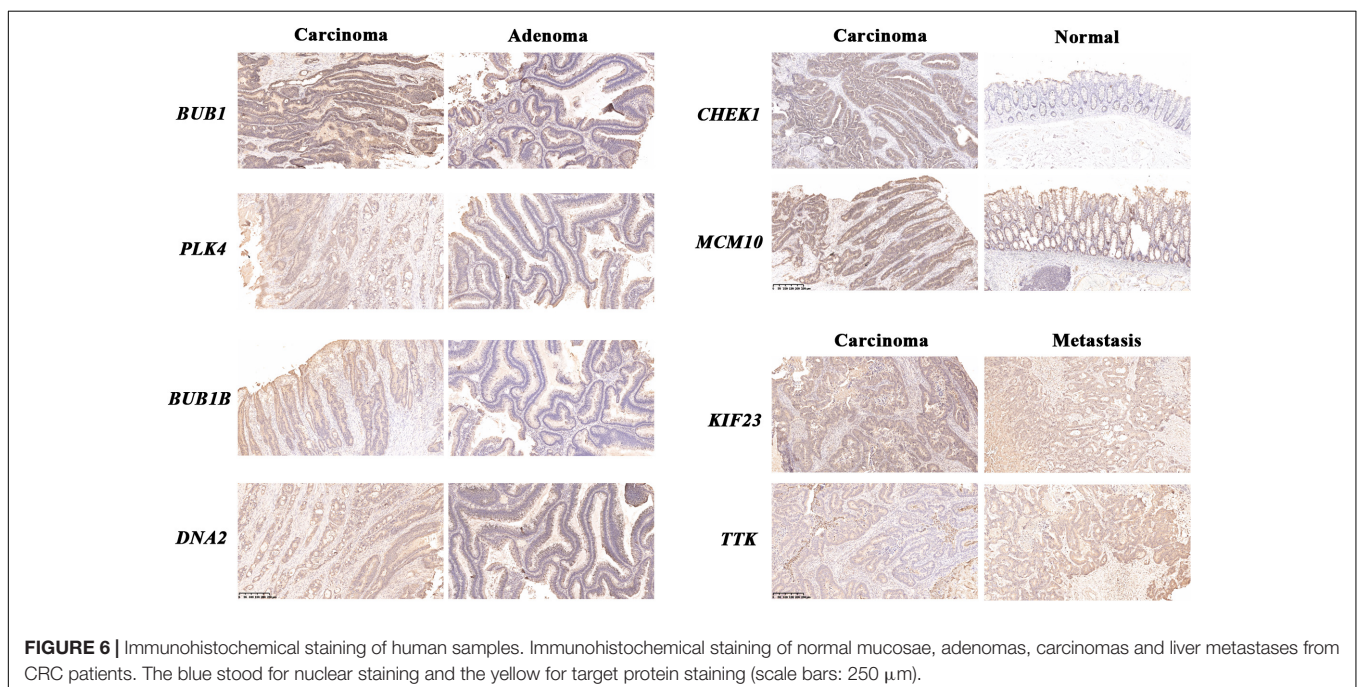
We obtained the mutual relationship of the 8 genes at protein level through the STRING and mapped the PPI network, which showed a strong correlation among the featured genes

(Figure 7A). Four genes (*BUB1*, *CHEK1*, *MCM10*, and *TTK*) had the highest node number of 7, which indicates that these genes composed a dense interaction in the network (Figure 7B).

Gene ontology analysis was performed and the results revealed that the major biological processes where featured genes are enriched were cell cycle checkpoint, chromosomal region, protein serine/threonine kinase activity and other cell cycle functions closely associated with the proliferation of tumor cells (Figure 7C).

DISCUSSION

Colorectal cancer exhibits a highly complex tumor heterogeneity in patient populations that are affected by various inter-factors and the mechanisms of oncogenesis and tumor progression need to be explored from different aspects, including microenvironment, metabolism and stemness. The classical theory points out that the initiation of CRC is closely associated with the disorder of stem cells in the crypts, which change into adenoma generation and then canceration (Barker et al., 2009). Normal self-renewal gradually deflects from wild-type stem cells as a result of oncogenic mutations in a long process of accumulation, and genetic and environmental factors play an important role in tumor progression. CSCs remain an increasingly competitive condition and demonstrate advantages over time during tumor evolution, with a confused environment composed of multiple cells, which provide them with different signals (Zeuner et al., 2014). The WNT pathway, Notch pathway and several other signals have been identified as vital factors in the regulation of colorectal CSC (Sikandar et al., 2010; Vermeulen et al., 2010). Precise markers under different classifications used to assess the capacity of stemness are crucial



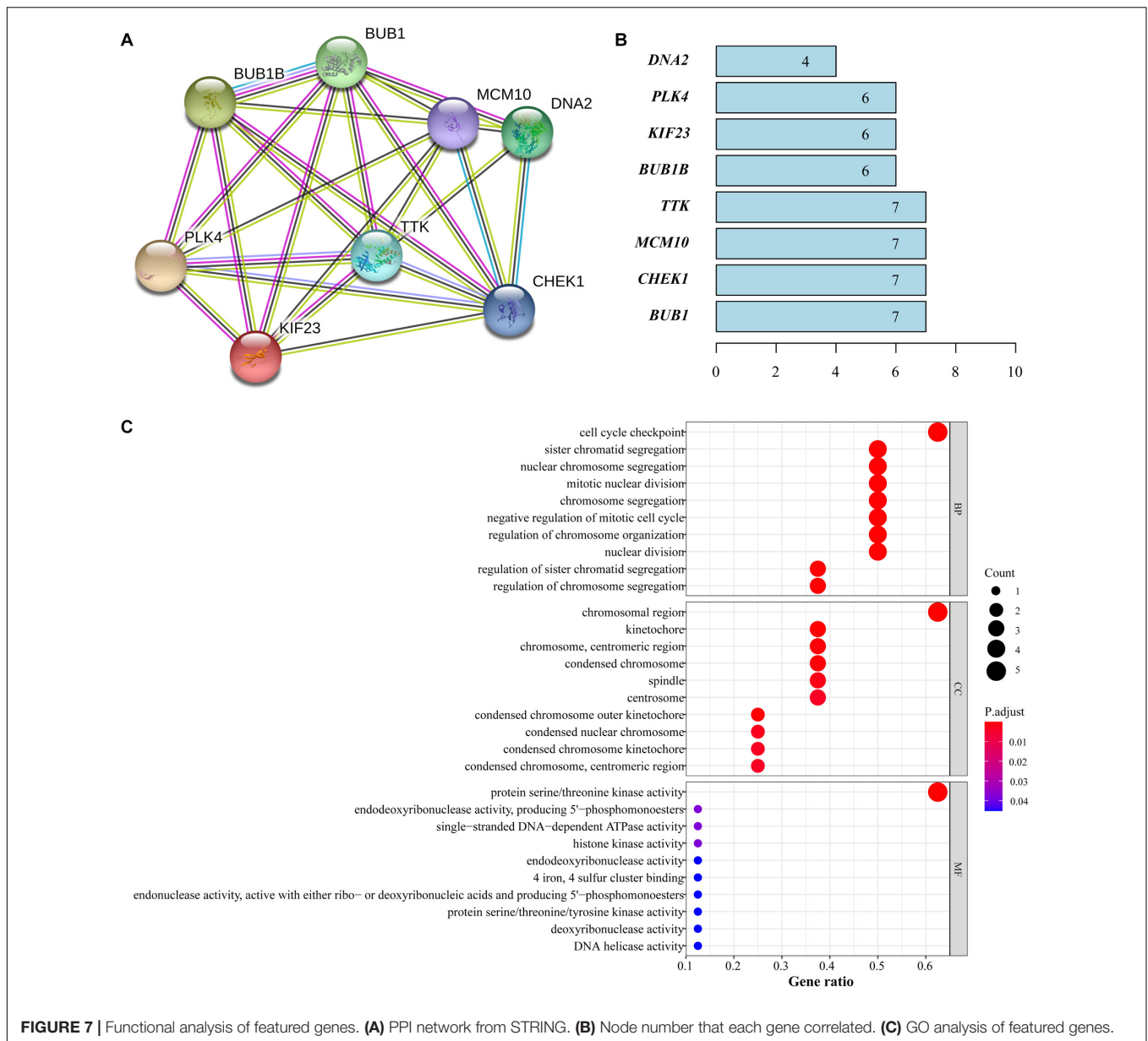


FIGURE 7 | Functional analysis of featured genes. **(A)** PPI network from STRING. **(B)** Node number that each gene correlated. **(C)** GO analysis of featured genes.

for the translation of CSC discovery into practical clinics and its development for therapy in CRC patients.

Our study focused on the excavation of featured genes that were characterized with stemness in CRC using the proven-effective indices system for in-depth exploration. Based on clinical parameters, mRNasi was related to lymphatic metastasis and EREG-mRNasi could reflect tumor infiltration. Both indices showed significant relationships with patient prognosis in accordance with the characteristics of stemness in CRC. WGCNA classified the DEGs with indices and conjugated genes for integral study in a small scale considering the interaction of multiple genes. For the mRNasi group we obtained optimum genes after filtering, while none of the genes met the threshold in the EREG-mRNasi group, as clustering with a high correlation was not completed.

Although epigenetic regulators were confirmed to be effective in curation targeting at stemness and the EREG-mRNasi was calculated from a set of stemness-related genes that controlled the epigenome in original study, featured genes about EREG-mRNasi still failed to summarize under strict criteria (Kreso et al., 2014).

The selected genes were validated to have a strong inner-connection and were tightly associated with the process of tumor progression. *BUB1*, *CHEK1*, *MCM10*, and *TTK* were confirmed to be in the central position of the network. Mutation of *BUB1*, which is known as the spindle assembly checkpoint gene, causes mosaic variegated aneuploidy and gives rise to an increased CRC risk (de Voer et al., 2013). *CHEK1* is a DNA damage sensor at a transcriptional level with the ability to arrest cell cycle for the repairment of damage and participates in defense against

evasion. Overexpression of *CHEK1* generally predicts a more serious malignancy and a worse prognosis for CRC patients (Gali-Muhtasib et al., 2008). *MCM10* plays an important role in assembly of complex during DNA replication and its unbalance with other factors broken by DNA damage might contribute to the initiation of cancer (Yoshida and Inoue, 2004). The frameshift mutations of mitotic checkpoint kinase *TTK* in CRC with microsatellite-unstable are frequent and tend to prompt tumor growth because of a triggered alteration in the cell cycle (Ahn et al., 2009). These identified functions have been investigated in other studies and are in accordance with our GO analysis results, illustrating that the set of our featured genes could impact CRC through regulation of the cell cycle and proliferation.

Current evaluations for CRC patients are frequently based on the clinical characteristics such as age, sex and tumor size considering the easy access to data (Colditz et al., 2000; Yeoh et al., 2011). Sequencing technology has drawn gene alteration into disease assessment as precision medicine. Stemness index describes the characteristics of related genes from an independent point of view and is comparatively associated with tumor progression. Our study mined CRC-related biomarkers from this aspect and that is of great significance to be used in clinic. The excavation of stemness index has been reported in other cancer and is proved to have high application value. Pan et al. (2019) identified 13 key genes which played important roles in the maintenance of bladder CSCs in their study. The 12 genes that Pei et al. (2020) found to be closely related in breast cancer stem cell characteristics showed prognosis-oriented effects in patients. Our study in CRC has also acquired 8 meaningful genes and verified by methods of laboratory experiments and bioinformatics.

The featured genes in our study might have several limitations that affect their application and need improvement. Firstly, all of our data came from public databases and further prospective

studies need to be implemented to obtain the stemness indices and parameters of CRC patients for the validation of accurate calculation and clinical value. Secondly, the intrinsic association among the featured genes and their relationship with CSC can be experimented on *in vitro* and *in vivo*.

In conclusion, the set of featured genes that we obtained was based on the stemness indices and validated with further analysis. Eight genes were found to be closely related to CSC characteristics in CRC patients and the significant function related to genes was center on the cell cycle, which might support new ideas in individualized preventions and treatments.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

WF designed the study. YL and XZ collected the data. ZL, FL, and WW analyzed and interpreted the data. All authors were involved in writing manuscript and approved of the submitted and published versions.

FUNDING

This work was supported by grants from the National Natural Science Foundation of China (Nos. 91959110 and 81972702) and Natural Science Foundation of Beijing (No. 7204324). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

- Ahn, C. H., Kim, Y. R., Kim, S. S., Yoo, N. J., and Lee, S. H. (2009). Mutational analysis of *TTK* gene in gastric and colorectal cancers with microsatellite instability. *Cancer Res. Treat.* 41, 224–228. doi: 10.4143/crt.2009.41.4.224
- Barker, N., Ridgway, R. A., van Es, J. H., van de Wetering, M., Begthel, H., van den Born, M., et al. (2009). Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* 457, 608–611. doi: 10.1038/nature07602
- Battle, E., and Clevers, H. (2017). Cancer stem cells revisited. *Nat. Med.* 23, 1124–1134. doi: 10.1038/nm.4409
- Boumahdi, S., Driessens, G., Lapouge, G., Rorive, S., Nassar, D., Le Mercier, M., et al. (2014). SOX2 controls tumour initiation and cancer stem-cell functions in squamous-cell carcinoma. *Nature* 511, 246–250. doi: 10.1038/nature13305
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Clevers, H. (2011). The cancer stem cell: premises, promises and challenges. *Nat. Med.* 17, 313–319. doi: 10.1038/nm.2304
- Colditz, G. A., Atwood, K. A., Emmons, K., Monson, R. R., Willett, W. C., Trichopoulos, D., et al. (2000). Harvard report on cancer prevention volume 4: harvard cancer risk index. Risk index working group, harvard center for cancer prevention. *Cancer Causes Control* 11, 477–488. doi: 10.1023/a:1008984432272
- de Voer, R. M., Geurts van Kessel, A., Weren, R. D., Ligtenberg, M. J., Smeets, D., Fu, L., et al. (2013). Germline mutations in the spindle assembly checkpoint genes *BUB1* and *BUB3* are risk factors for colorectal cancer. *Gastroenterology* 145, 544–547. doi: 10.1053/j.gastro.2013.06.001
- Gali-Muhtasib, H., Kuester, D., Mawrin, C., Bajbouj, K., Diestel, A., Ocker, M., et al. (2008). Thymoquinone triggers inactivation of the stress response pathway sensor *CHEK1* and contributes to apoptosis in colorectal cancer cells. *Cancer Res.* 68, 5609–5618. doi: 10.1158/0008-5472.CAN-08-0884
- Harris, W. J., Huang, X., Lynch, J. T., Spencer, G. J., Hitchin, J. R., Li, Y., et al. (2012). The histone demethylase *KDM1A* sustains the oncogenic potential of *MLL-AF9* leukemia stem cells. *Cancer Cell* 21, 473–487. doi: 10.1016/j.ccr.2012.03.014
- Hutter, C., and Zenklusen, J. C. (2018). The cancer genome atlas: creating lasting value beyond its data. *Cell* 173, 283–285. doi: 10.1016/j.cell.2018.03.042
- Kreso, A., van Galen, P., Pedley, N. M., Lima-Fernandes, E., Frelin, C., Davis, T., et al. (2014). Self-renewal as a therapeutic target in human colorectal cancer. *Nat. Med.* 20, 29–36. doi: 10.1038/nm.3418
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559
- Maccalli, C., and De Maria, R. (2015). Cancer stem cells: perspectives for therapeutic targeting. *Cancer Immunol. Immunother.* 64, 91–97. doi: 10.1007/s00262-014-1592-1

- Malta, T. M., Sokolov, A., Gentles, A. J., Burzykowski, T., Poisson, L., Weinstein, J. N., et al. (2018). Machine Learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* 173, e315. doi: 10.1016/j.cell.2018.03.034
- Medema, J. P. (2017). Targeting the colorectal cancer stem cell. *N. Engl. J. Med.* 377, 888–890. doi: 10.1056/NEJMcibr1706541
- Miller, K. D., Nogueira, L., Mariotto, A. B., Rowland, J. H., Yabroff, K. R., Alfano, C. M., et al. (2019). Cancer treatment and survivorship statistics, 2019. *CA Cancer J. Clin.* 69, 363–385. doi: 10.3322/caac.21565
- Nassar, D., and Blanpain, C. (2016). Cancer stem cells: basic concepts and therapeutic implications. *Annu. Rev. Pathol.* 11, 47–76. doi: 10.1146/annurev-pathol-012615-044438
- Nervi, C., De Marinis, E., and Codacci-Pisanelli, G. (2015). Epigenetic treatment of solid tumours: a review of clinical trials. *Clin. Epigenetics* 7:127. doi: 10.1186/s13148-015-0157-2
- Pan, S., Zhan, Y., Chen, X., Wu, B., and Liu, B. (2019). Identification of biomarkers for controlling cancer stem cell characteristics in bladder cancer by network analysis of transcriptome data stemness indices. *Front. Oncol.* 9:613. doi: 10.3389/fonc.2019.00613
- Pei, J., Wang, Y., and Li, Y. (2020). Identification of key genes controlling breast cancer stem cell characteristics via stemness indices analysis. *J. Transl. Med.* 18:74. doi: 10.1186/s12967-020-02260-9
- Sikandar, S. S., Pate, K. T., Anderson, S., Dizon, D., Edwards, R. A., Waterman, M. L., et al. (2010). NOTCH signaling is required for formation and self-renewal of tumor-initiating cells and for repression of secretory cell differentiation in colon cancer. *Cancer Res.* 70, 1469–1478. doi: 10.1158/0008-5472.CAN-09-2557
- Takebe, N., Miele, L., Harris, P. J., Jeong, W., Bando, H., Kahn, M., et al. (2015). Targeting notch, hedgehog, and Wnt pathways in cancer stem cells: clinical update. *Nat. Rev. Clin. Oncol.* 12, 445–464. doi: 10.1038/nrclinonc.2015.61
- Vermeulen, L., De Sousa, E. M. F., van der Heijden, M., Cameron, K., de Jong, J. H., Borovski, T., et al. (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nat. Cell Biol.* 12, 468–476. doi: 10.1038/ncb2048
- Yeoh, K. G., Ho, K. Y., Chiu, H. M., Zhu, F., Ching, J. Y., Wu, D. C., et al. (2011). The Asia-Pacific colorectal screening score: a validated tool that stratifies risk for colorectal advanced neoplasia in asymptomatic Asian subjects. *Gut* 60, 1236–1241. doi: 10.1136/gut.2010.221168
- Yoshida, K., and Inoue, I. (2004). Expression of MCM10 and TopBP1 is regulated by cell proliferation and UV irradiation via the E2F transcription factor. *Oncogene* 23, 6250–6260. doi: 10.1038/sj.onc.1207829
- Zeuner, A., Todaro, M., Stassi, G., and De Maria, R. (2014). Colorectal cancer stem cells: from the crypt to the clinic. *Cell Stem Cell* 15, 692–705. doi: 10.1016/j.stem.2014.11.012

Conflict of Interest: 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.

Copyright © 2020 Lu, Zhou, Liu, Wang, Li and Fu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.