



OPEN ACCESS

EDITED BY

Simin Li,
Southern Medical University, China

REVIEWED BY

Yupei Deng,
China Tibetology Research Center,
China
Zhixiang Yu,
Fourth Military Medical University, China

*CORRESPONDENCE

Dongfang Wu,
wudf@yucebio.com
Yu Peng,
pengyu20220718@163.com

[†]These authors have contributed equally to this work

SPECIALTY SECTION

This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 19 July 2022

ACCEPTED 17 October 2022

PUBLISHED 16 November 2022

CITATION

Liang H, Zhao Y, Liu K, Xiao Y, Chen K, Li D, Zhong S, Zhao Z, Wu D and Peng Y (2022), The mechanism of lncRNAs in the crosstalk between epithelial-mesenchymal transition and tumor microenvironment for early colon adenocarcinoma based on molecular subtyping. *Front. Genet.* 13:997739. doi: 10.3389/fgene.2022.997739

COPYRIGHT

© 2022 Liang, Zhao, Liu, Xiao, Chen, Li, Zhong, Zhao, Wu and Peng. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

The mechanism of lncRNAs in the crosstalk between epithelial-mesenchymal transition and tumor microenvironment for early colon adenocarcinoma based on molecular subtyping

Hanlin Liang^{1†}, Yi Zhao^{2†}, Kai Liu^{3†}, Yajie Xiao^{4†}, Kexu Chen¹, Delan Li¹, Shupeng Zhong¹, Zhikun Zhao⁴, Dongfang Wu^{4*} and Yu Peng^{5*}

¹Chemotherapy Department, Zhongshan City People's Hospital, Zhongshan, China, ²GI Medicine, The Third Hospital Affiliated to Naval Medical University, Shanghai, China, ³Department of Colorectal Oncology, National Clinical Research Center for Cancer, Key Laboratory of Cancer Prevention and Therapy of Tianjin, Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China, ⁴Department of Medicine, YuceBio Technology Co., Ltd., Shenzhen, China, ⁵Oncology Department, Jiangmen Central Hospital, Jiangmen, China

A large number of colon adenocarcinoma (COAD) patients are already advanced when diagnosed. In this study, we aimed to further understand the mechanism of tumor development in early COAD by focusing on epithelial-mesenchymal transition (EMT) and long non-coding RNAs (lncRNAs). Expression profiles of early COAD patients were obtained from public databases. EMT-related lncRNAs were used as a basis for constructing molecular subtypes through unsupervised consensus clustering. Genomic features, pathways and tumor microenvironment (TME) were compared between two subtypes. LncAtlas database was applied to analyze the relation between lncRNAs and transcription factors (TFs). First order partial correlation analysis was conducted to identify key EMT-related lncRNAs. C1 and C2 subtypes with distinct prognosis were constructed. Oncogenic pathways such as EMT, KRAS signaling, JAK-STAT signaling, and TGF- β signaling were significantly enriched in C2 subtype. Higher immune infiltration and expression of immune checkpoints were also observed in C2 subtype, suggesting the key EMT-related lncRNAs may play a critical role in the modulation of TME. In addition, JAK-STAT signaling pathway was obviously enriched in upregulated TFs in C2 subtype, which indicated a link between key lncRNAs and JAK-STAT signaling that may regulate TME. The study further expanded the research on the role of EMT-related lncRNAs in the early COAD. The six identified EMT-related lncRNAs could serve as biomarkers for early screening COAD.

KEYWORDS

colon adenocarcinoma, epithelial-mesenchymal transition, long non-coding RNAs, tumor microenvironment, transcription factors, biomarkers, bioinformatics analysis

Introduction

According to the global cancer statistics, colon adenocarcinoma (COAD) is the sixth most diagnosed cancer with 1,148,515 new cases in 2020, contributing 6.0% of all new diagnosed cancer cases (Sung et al., 2021). Simultaneously, the death cases contribute 5.8% (576,858) of all deaths by cancer, which is the fifth leading cause of cancer death in 2020. Although the 5-year overall survival (OS) is upon 75% of American Joint Committee on Cancer (AJCC) stage I and II, a dramatically decreased survival of distant metastasis is shown with only less than 20% of AJCC stage IV (Ulanja et al., 2019). Prognostic difference is shown between right-sided and left-sided colon cancer, and left-sided colon cancer has a lower death risk than right-sided colon cancer (Petrelli et al., 2017), which may result from their genetic and immunological differences (Lee et al., 2015). Screening techniques for colon cancer include invasive and non-invasive tests. Colonoscopic tests are recommended for high-risk individuals and non-colonoscopy tests are recommended in average-risk individuals according to European Society for Medical Oncology (ESMO) clinical practice guidelines (Argilés et al., 2020). However, they are not sensitive in the diagnosis of early colon cancer. Actually, a number of patients are already advanced when diagnosed as COAD. Therefore, early screening of COAD is of great value for improving prognosis and releasing cancer burden worldwide.

To reach accurate screening, comprehensive understanding of COAD tumorigenesis and development is a basis for identifying effective biomarkers for COAD screening. Of the hallmarks of cancers, epithelial-mesenchymal transition (EMT) is one of the most important features contributing for metastasis (Dongre and Weinberg, 2019). Tumor microenvironment (TME), another important component in cancer tissue, has been demonstrated to have a strong correlation with EMT through the linkages of proinflammatory factors such as TGF- β , TNF- α , and IL-6 (Jung et al., 2015). Among these interactions, long non-coding RNAs (lncRNAs) are considered as critical regulators for modulating TME and managing EMT (Sun et al., 2018; O'Brien et al., 2020). Oncogenic pathways involving in EMT such as WNT signaling, JAK-STAT3 signaling (Xue et al., 2018), mTOR signaling and MAPK/ERK signaling have been illustrated to be regulated by various lncRNAs (O'Brien et al., 2020). In colon cancer, lncRNA-HOTAIR associated with EMT was identified as a predictor of metastasis and prognosis (Wu et al., 2014).

As lncRNAs are of potential to serve as biomarkers for COAD prognosis, we consider that lncRNAs involving in EMT process may also be effective predictors for COAD. Therefore, in this study, we tried to construct a novel molecular subtyping system based on EMT-related lncRNAs. Compared to pathological subtyping or clinical features, molecular subtyping is more

accurate for classifying cancer patients into different classes with different prognosis. By exploring the mechanisms of EMT-related lncRNAs in tumorigenesis in COAD, we further identified six key EMT-related lncRNAs that could serve as biomarkers for early screening COAD.

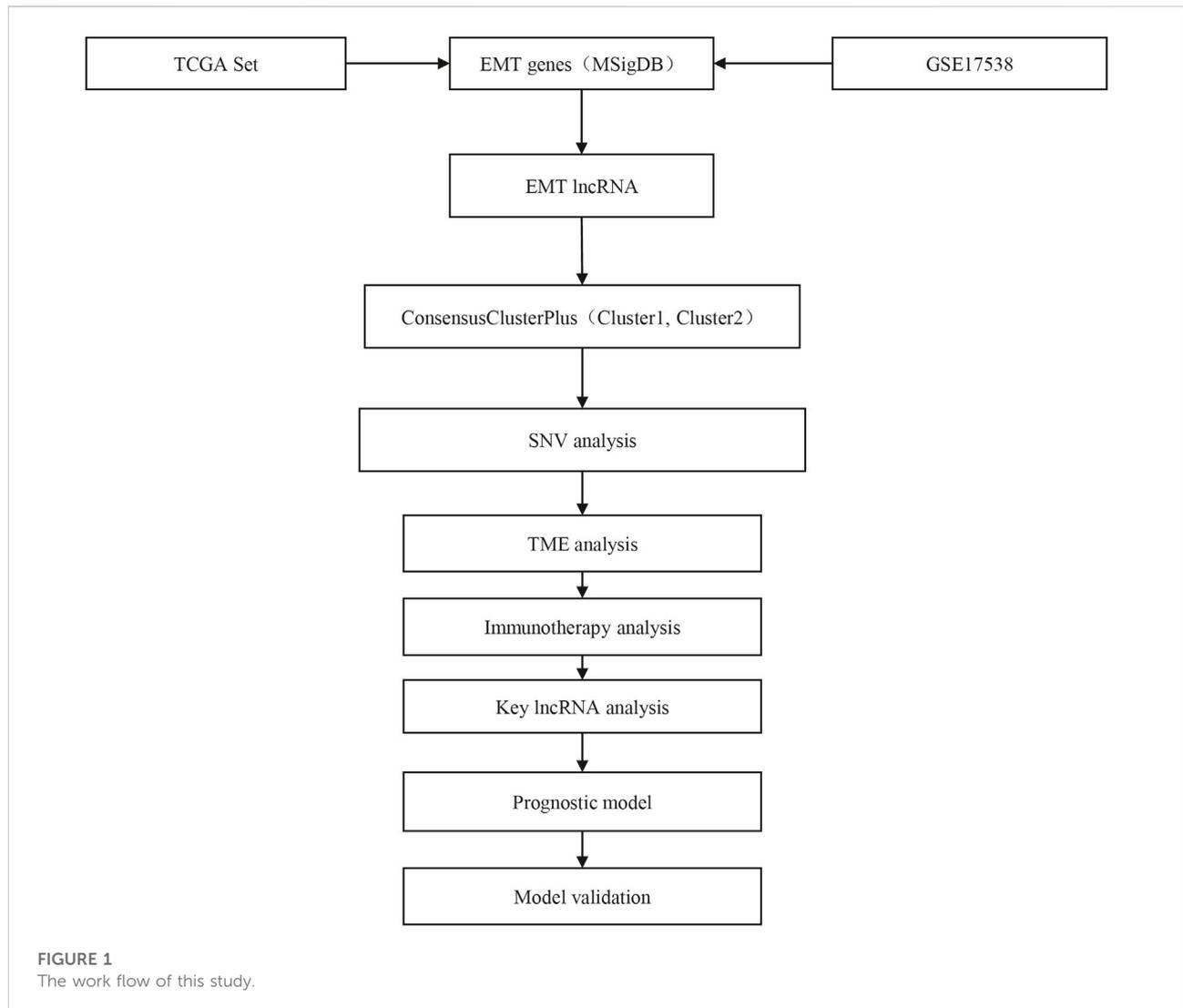
Materials and methods

Sample collection

COAD samples were obtained from public databases through Sangerbox platform (Shen et al., 2022). RNA-seq data with clinical information was downloaded from The Cancer Genome Atlas (TCGA) database. GSE17538 with gene expression profiles was downloaded from Gene Expression Omnibus (GEO) database. In TCGA-COAD and GSE17538 cohorts, only samples of stage I and II were retained, and samples without survival information were removed. The clinical information of COAD samples was shown in Table 1.

TABLE 1 The clinical information of COAD samples.

Clinical features	TCGA-COAD	GSE17538
OS		
0 (alive)	199	80
1 (dead)	26	20
T Stage		
T1	8	
T2	61	
T3	145	
T4	10	
TX	1	
N Stage		
N0	225	
M Stage		
M0	204	
MX	21	
Stage		
I	70	28
II	155	72
Gender		
Female	99	47
Male	126	53
Age		
<=70	117	53
>70	108	47



Identification of epithelial-mesenchymal transition-related lncRNAs

EMT-related genes in hallmark EMT pathway were obtained from Molecular Signatures Database (MSigDB, v7.4, <https://www.gsea-msigdb.org/gsea/msigdb/>) (Liberzon et al., 2015). lncRNAs and mRNAs in TCGA-COAD and GSE17538 cohorts were annotated by gene transfer format (GTF, v32) file which was downloaded from GENCODE (<https://www.genencodegenes.org/>).

EMT score of each sample was calculated by single sample gene set enrichment analysis (ssGSEA) in GSVA R package (Hänzelmann et al., 2013). Then Pearson correlation analysis was employed to calculate correlation coefficients between EMT score and expression of lncRNAs. EMT-related lncRNAs were determined by the conditions of $|\text{coefficient}| > 0.25$ and $p < 0.05$.

Identification of molecular subtypes based on epithelial-mesenchymal transition-related lncRNAs

Screened EMT-related lncRNAs that were overlapped in TCGA-COAD and GSE17538 cohorts were used as a basis to construct molecular subtypes. Unsupervised consensus clustering in ConsensusClusterPlus R package was implemented to construct consensus matrix (Wilkerson and Hayes, 2010). KM algorithm and Euclidean distance were used to conduct 500 bootstraps with each bootstrap containing 80% samples. Cluster number k from 2 to 10 was included to screen the optimal cluster according to consensus matrix and cumulative distribution function (CDF).

Gene set enrichment analysis

GSEA is a powerful analytical method for interpreting biological processes based on gene expression profiles (Subramanian et al., 2005), which was applied to assess hallmark pathways for two subtypes. Hallmark pathways with a series of gene sets were obtained from MSigDB (Liberzon et al., 2015). The proportion of 28 immune cells was estimated by GSEA based on gene signatures of different cell types (Şenbabaoğlu et al., 2016). Estimation of STromal and Immune cells in Malignant Tumours using Expression data (ESTIMATE) (Yoshihara et al., 2013), also based on GSEA was used to calculate stromal score and immune score of two subtypes. ClusterProfiler R package was used to annotate Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways on TFs and EMT-related lncRNAs (Yu et al., 2012).

Localization of lncRNAs and calculation of transcription factor activity

Relative concentration index (RCI) based on LncAtlas database was introduced to measure the localization of lncRNAs (Mas-Ponte et al., 2017). RCI >0 indicates lncRNAs localizing in the cytoplasm and RCI <0 indicates the nuclear. The TF activity was assessed according to the algorithm from Garcia-

Alonso et al. Garcia-Alonso et al. (2018). Pearson correlation analysis was conducted to analyze the association between EMT-related lncRNAs and TFs.

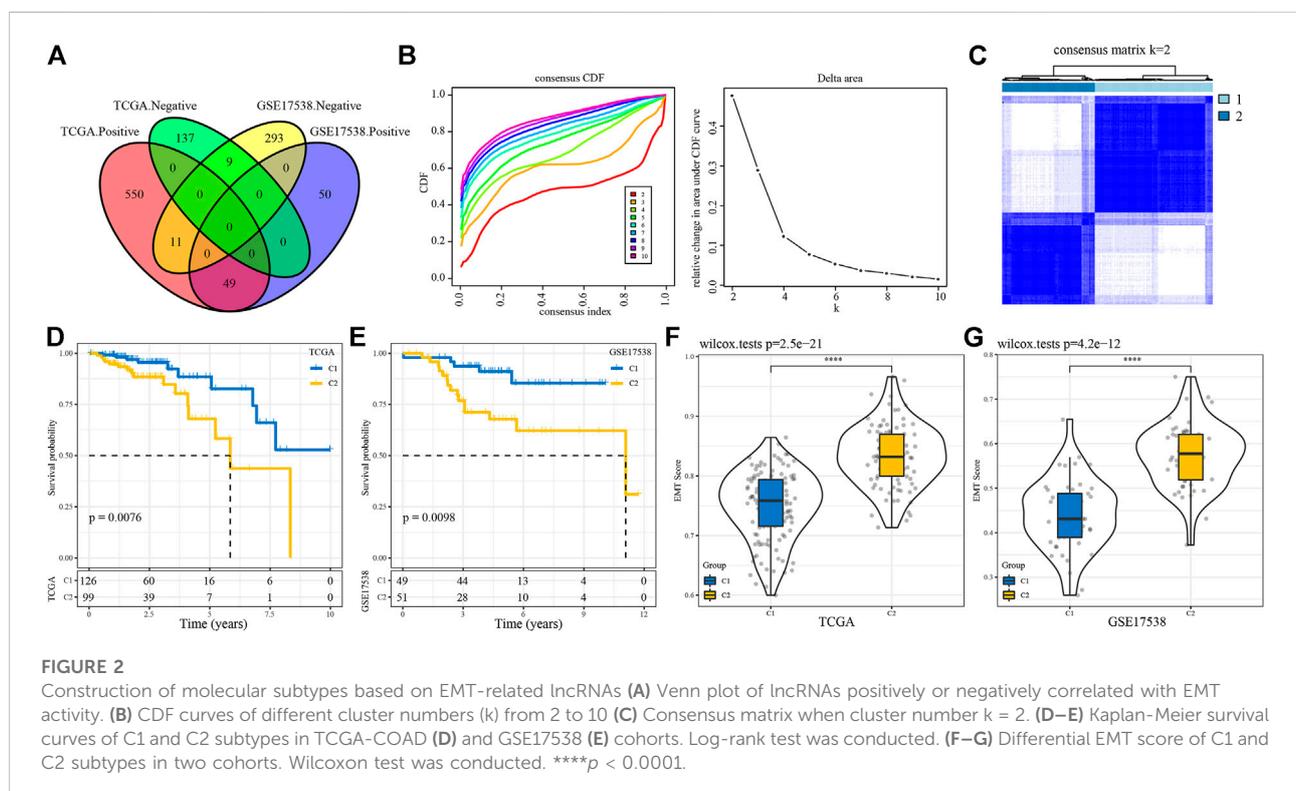
Identification of key epithelial-mesenchymal transition-related lncRNAs

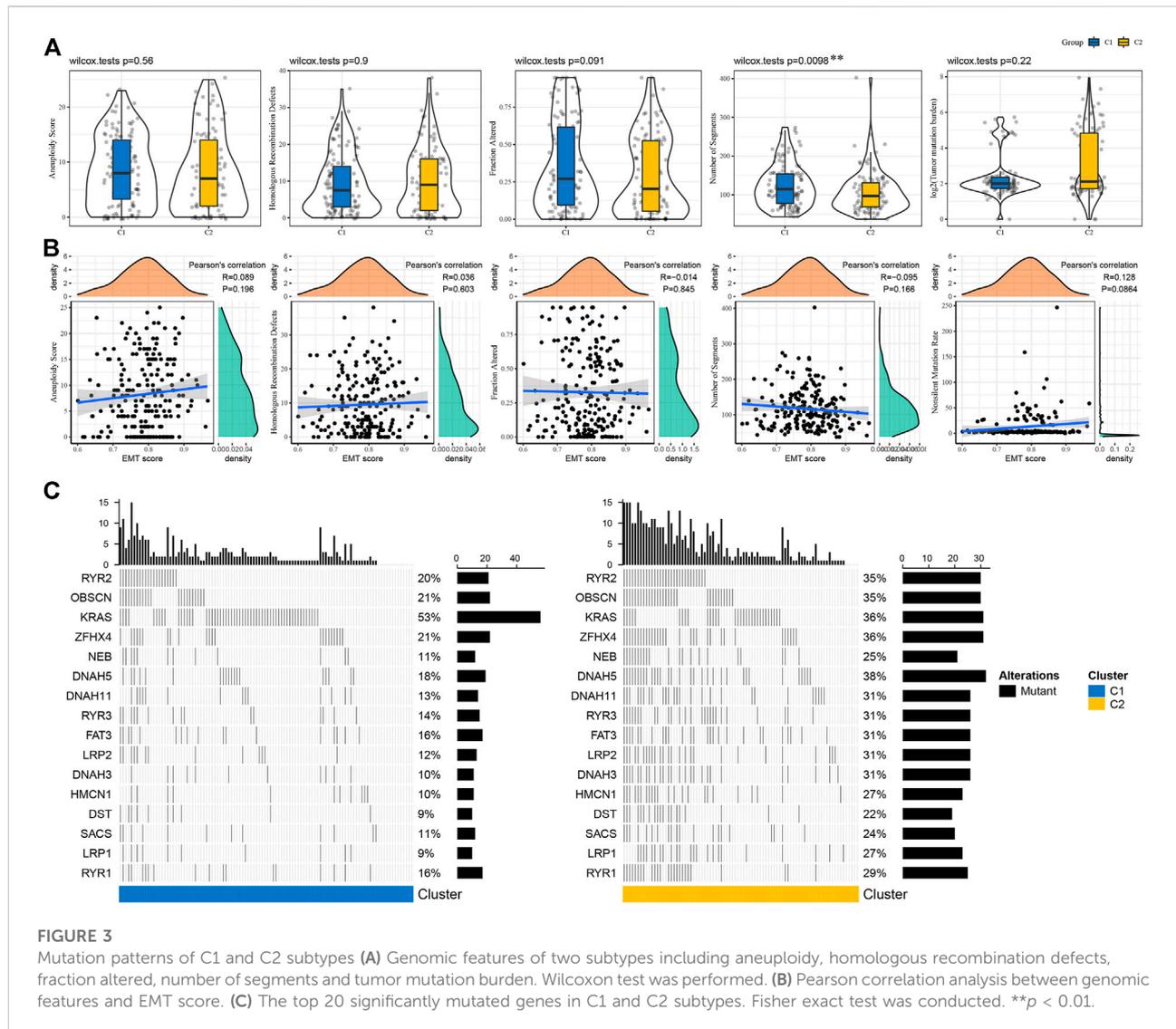
First order partial correlation analysis was used to evaluate the linkage among EMT-related lncRNAs, EMT score and EMT-related genes (Reverter and Chan, 2008). The association between two variables were largely decreased when eliminating the effect of another variable, and the variable was considered as key EMT-related lncRNA strongly associated with EMT score and EMT-related genes.

The identified key EMT-related lncRNAs were used to construct a prognostic model for predicting OS. Univariate Cox regression analysis was conducted on these lncRNAs and coefficients were generated for building the model, defining as: risk score = $\sum (\beta_i \times \text{exp } i)$. Beta represents coefficients and exp represents the expression of lncRNAs.

Statistical analysis

Statistical analysis was performed in R (4.1.1) software. Parameters of R packages and software were default if no





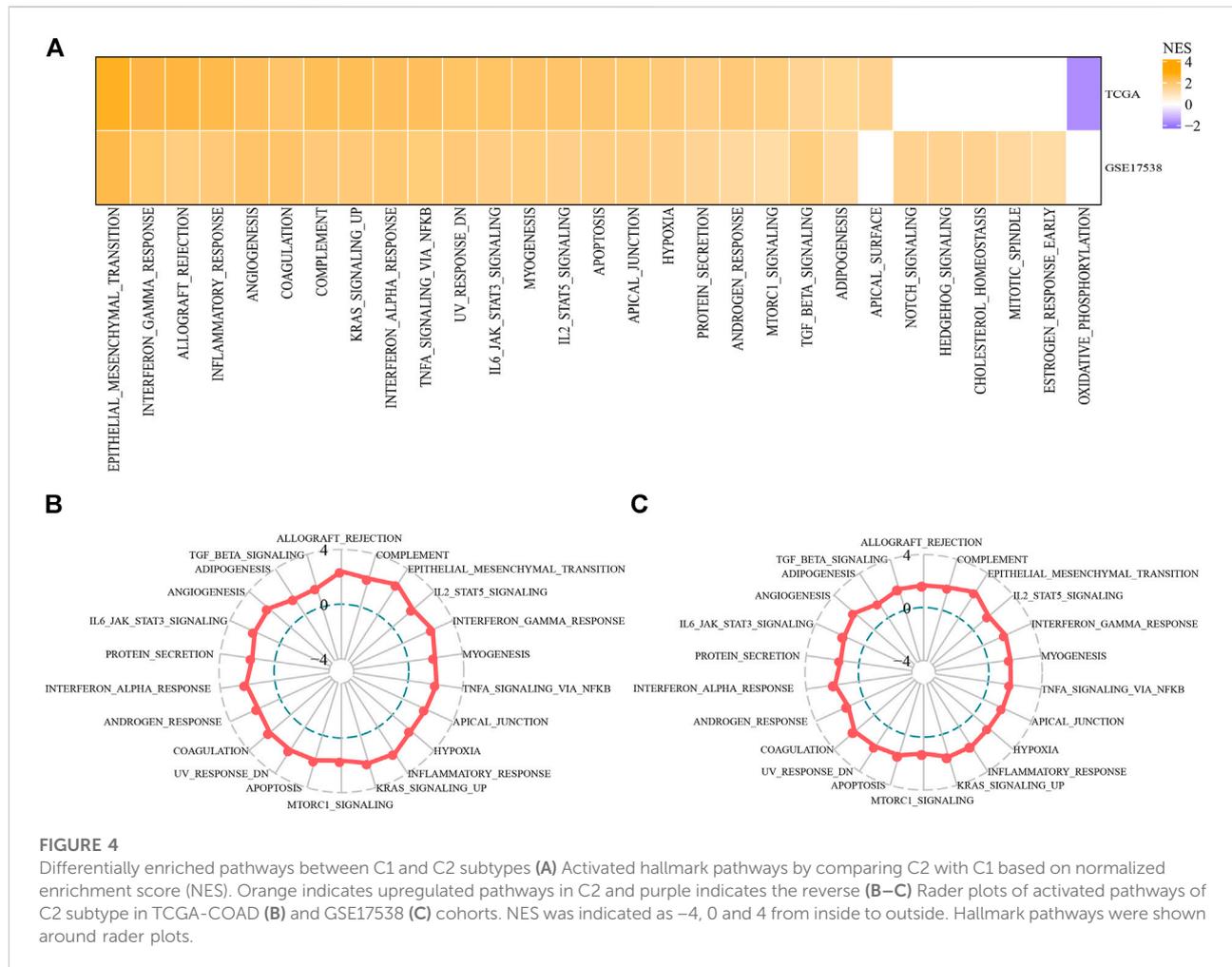
introduce. Statistical methods were indicated in the corresponding figure legends. $p < 0.05$ was considered as significant. ns, no significance. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Results

Constructing two molecular subtypes based on epithelial-mesenchymal transition-related lncRNAs

The work flow of this study was shown in Figure 1. To identify EMT-related lncRNAs, Pearson correlation analysis was conducted between EMT activity and lncRNA expression. A total of 756 and 412 EMT-related lncRNAs were identified

in TCGA-COAD and GSE17538 cohorts respectively (Figure 2A). Then the intersected part of 58 EMT-related lncRNAs were used as a basis for unsupervised consensus clustering. The optimal cluster number (k) was determined according to CDF curve and consensus matrix (Figures 2B,C). When $k = 2$, samples were obviously divided into two groups. Kaplan-Meier survival analysis of two groups showed a significance of OS in both two cohorts ($p = 0.0076$ and $p = 0.0098$ in TCGA-COAD and GSE17538 respectively, Figures 2D,E). Finally, COAD samples were classified into two molecular subtypes (C1 and C2), with C1 subtype had a superior OS than C2 subtype. EMT activity shown as EMT score also varied largely between two subtypes. Not surprisingly, C2 subtype exhibited a significantly higher EMT score than C1 subtype, indicating that EMT pathway was more activated in C2 subtype ($p < 0.0001$, Figures 2F,G).



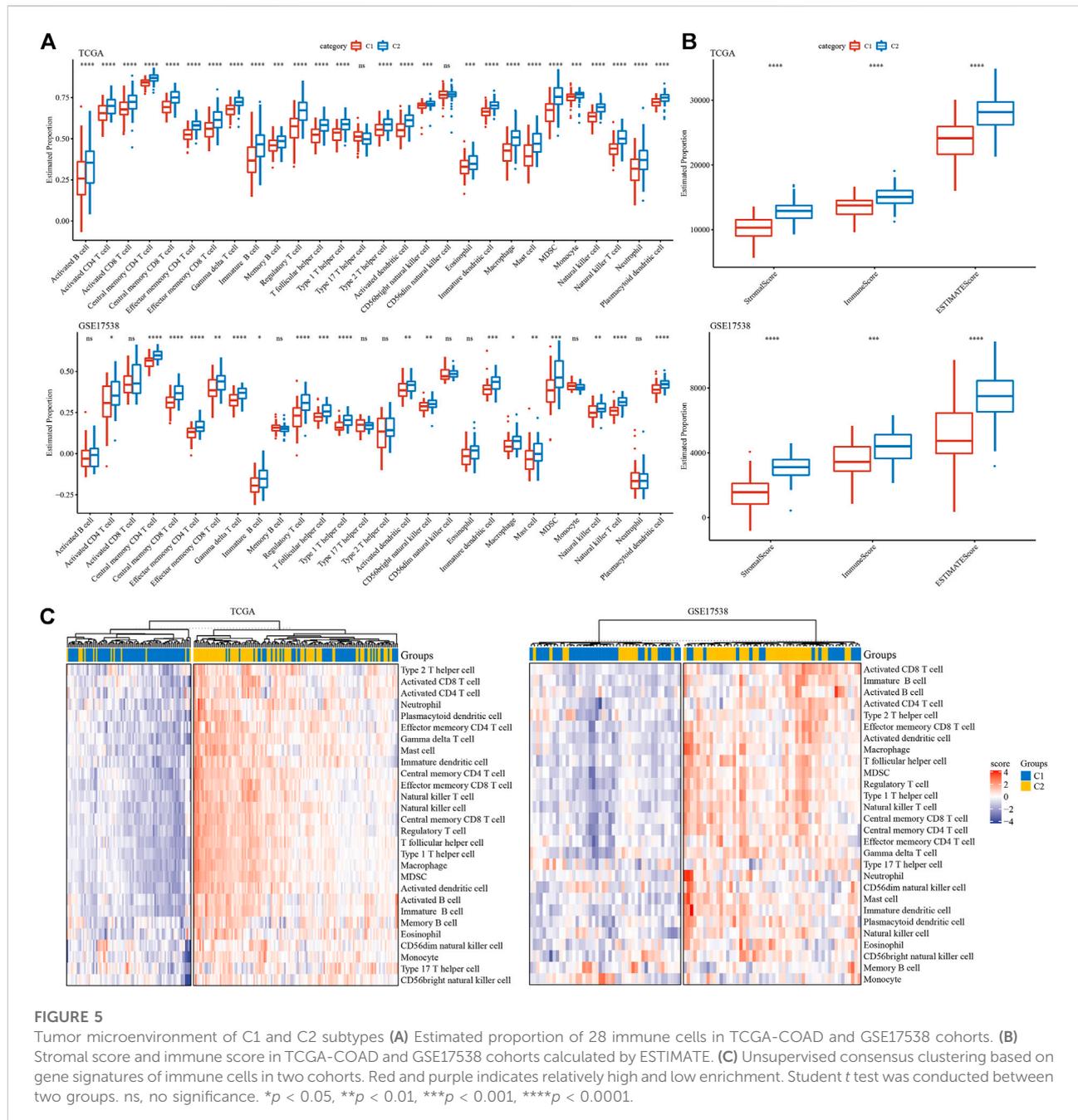
In addition, the result also demonstrated that these 58 EMT-related lncRNAs may play a key role in regulating EMT activity.

Characterizing gene mutations of two molecular subtypes

We compared the genomic features between C1 and C2 subtypes in TCGA-COAD cohort on five aspects including aneuploidy, fraction altered, tumor mutation burden, homologous recombination defects and number of segments. We observed that only a significant difference was shown in number of segments between two subtypes (Figure 3A). No obvious correlation was manifested between genomic features and EMT score (Figure 3B). Gene mutation analysis revealed that samples in C2 subtype had a higher mutated proportion than C1 subtype, except for *KRAS* mutations contributing for 53% samples in C1 subtype (Figure 3C).

Oncogenic pathways were more enriched in C2 subtype

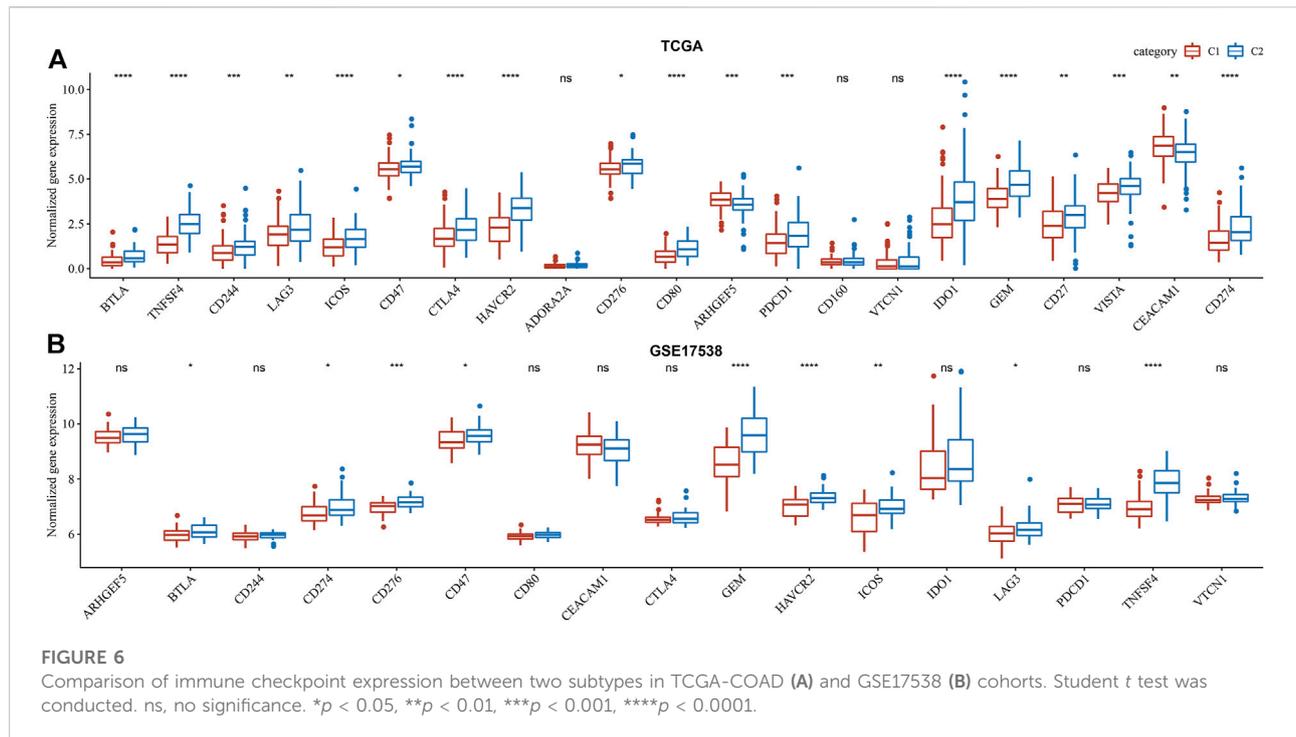
Next we tried to know if there was a difference of activated pathways between two subtypes. Hallmark pathways from MSigDB were included for GSEA and significantly enriched pathways were outputted with false discovery rate (FDR) < 0.05. By comparing C2 subtype with C1 subtype, we observed that C2 subtype had 23 activated and 11 suppressed pathways in TCGA-COAD cohort, and 27 activated and 9 suppressed pathways in GSE17538 cohort. Of these activated pathways, we found that oncogenic pathways and immune-related pathways were greatly enriched, such as EMT, angiogenesis, *KRAS* signaling, hypoxia, interferon response, TNF- α signaling, IL6-JAK-STAT3 signaling, IL2-STAT5 signaling, and TGF- β signaling pathways (Figure 4A). Overall, C2 subtype had significantly higher normalized enrichment scores (NES) of these pathways than C1 in both two cohorts (Figures 4B,C), suggesting that EMT-related lncRNAs may be extensively involved in the regulation of these pathways especially EMT.



C2 subtype had higher infiltration of immune cells

To evaluate the tumor microenvironment of C1 and C2 subtypes, we assessed the estimated proportions of a series immune cells based on gene signatures from Şenbabaoğlu et al. Şenbabaoğlu et al. (2016). To our surprise, C2 subtype had extremely higher proportions of most immune cells in both two cohorts (Figure 5A). Specifically, activated CD4 T cells, activated CD8 T cells,

regulatory T cells, dendritic cells, macrophages, myeloid-derived suppressor cells (MDSCs) and natural killer cells were all more enriched in C2 subtype. ESTIMATE evaluation also supported the result that C2 subtype had higher stromal score and immune score than C1 subtype in both two cohorts ($p < 0.001$, Figure 5B). Furthermore, unsupervised consensus clustering based on these immune cells clearly divided samples into two groups of high and low immune infiltration (Figure 5C). Obviously, samples in high immune infiltration group largely belonged to C2 subtype.



These results implicated a close link between EMT-related lncRNAs and TME modulation.

Commonly, high immune infiltration of cytotoxic immune cells has favorable prognosis. However, immunosuppressive immune cells such as regulatory T cells and MDSCs were simultaneously increased in C2 subtype. In addition, we analyzed the expression of immune checkpoints obtained from HisgAtlas database (Liu et al., 2017). Higher expression of many important immune checkpoints was observed in C2 subtype, such as LAG3, ICOS, CTLA4, CD276, PDCD1, IDO1 and CD274 (Figure 6).

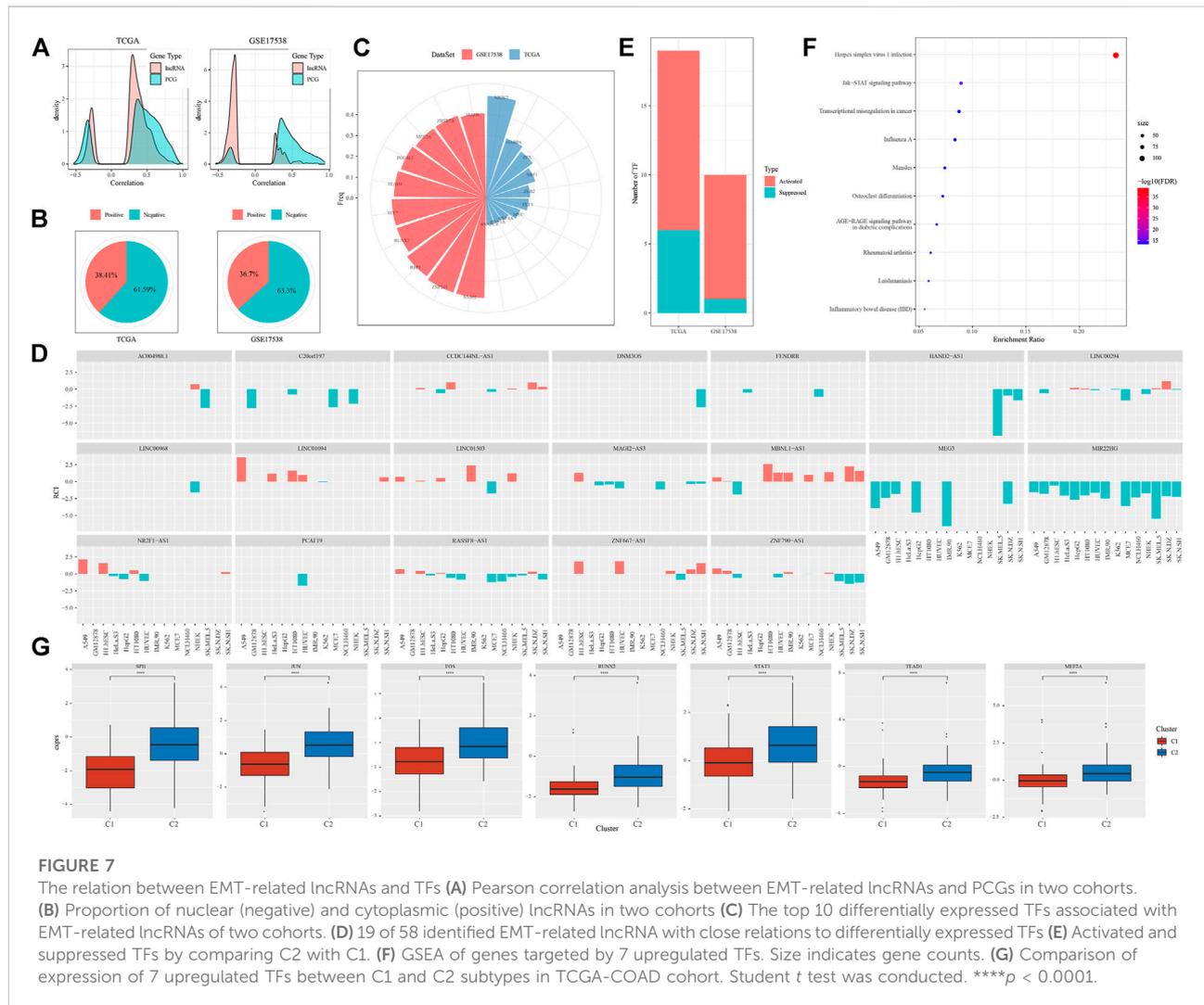
The crosstalk between epithelial-mesenchymal transition-related lncRNAs and transcription factors

Given that we have illustrated distinct molecular features of different subtypes, we implicated that EMT-related lncRNAs were greatly involved in modulating oncogenic pathways or the expression of immune-related genes. Actually, close associations (both negative and positive) were observed between EMT-related lncRNAs and protein-coding genes (PCGs) (Figure 7A). It is known that the function of lncRNAs is highly associated with their subcellular locations. Therefore, to evaluate the possible mechanism of the regulation, we assessed the locations of these 58 identified EMT-related lncRNAs. We found over a half of lncRNAs localized in the nuclear with 61.59% and 63.30% in TCGA-COAD and GSE17538 cohorts respectively

(Figure 7B). As most of EMT-related lncRNAs localized in the nuclear, we supposed that they possibly regulated gene expression by interacting with TFs. We then analyzed the TF activity and screened differentially expressed TFs between two subtypes (131 TFs in TCGA-COAD and 106 TFs in GSE17538). Correlation analysis between 58 EMT-related lncRNAs and dysregulated TFs discovered a group of important TFs and lncRNAs that may closely interact with each other. The top 10 identified TFs in two cohorts were listed (Figure 7C), and 19 EMT-related lncRNAs were screened to have a close relation to differentially expressed TFs (Figure 7D). By comparing C2 with C1 subtype, we found that a majority of TFs were upregulated in C2 with 7 same TFs upregulated in both two cohorts (Figure 7E). Functional analysis on the genes targeting by these 7 upregulated TFs showed that tumor-related pathways of Jak-STAT signaling and transcriptional misregulation in cancer were significantly enriched (Figures 7F,G). The above results indicated that the 19 identified EMT-related lncRNAs may regulate oncogenic pathways through interacting with the 7 upregulated TFs.

Six epithelial-mesenchymal transition-related lncRNAs were identified to serve as prognostic biomarkers

In the previous section, we identified 19 EMT-related lncRNAs with a relation to TFs. To understand which



lncRNAs among them acted a key role between EMT-related genes and EMT activity, we applied first order partial correlation analysis on them. When eliminating some of lncRNAs, the correlation between EMT activity and EMT-related genes greatly decreased. As a result, six EMT-related lncRNAs were identified, including ZNF667-AS1, CCDC144NL-AS1, MAGI2-AS3, HAND2-AS1, LINC01094 and PCAT19 (Figure 8A). Not surprisingly, GSEA on EMT-related genes associated with these 6 lncRNAs dug out that tumor- and immune-related pathways including proteoglycans in cancer, leukocyte transendothelial migration and focal adhesion were significantly enriched (Figure 8B). Finally, based on the expression of the 6 lncRNAs, we constructed a prognostic model. Samples in two cohorts could be both clearly stratified into high-risk and low-risk groups ($p = 0.0058$ and $p = 0.00052$ in TCGA-COAD and GSE17538 respectively,

Figures 8C,D), implicating robust performance of the prognostic model.

Discussion

A number of studies have found the regulatory role of lncRNAs in EMT and thus promotes tumor development and metastasis. Usually, the function of lncRNAs can be divided into two classifications, EMT promoters or suppressors (Cheng et al., 2019). However, some of lncRNAs are controversial that they act different roles varied by cancer types, indicating the complication of tumor development. To further understand the role of lncRNAs in EMT in colon cancer with early stages (I and II), we identified 58 lncRNAs that were possibly involved in EMT process, and constructed two molecular subtypes based on these EMT-related lncRNAs.

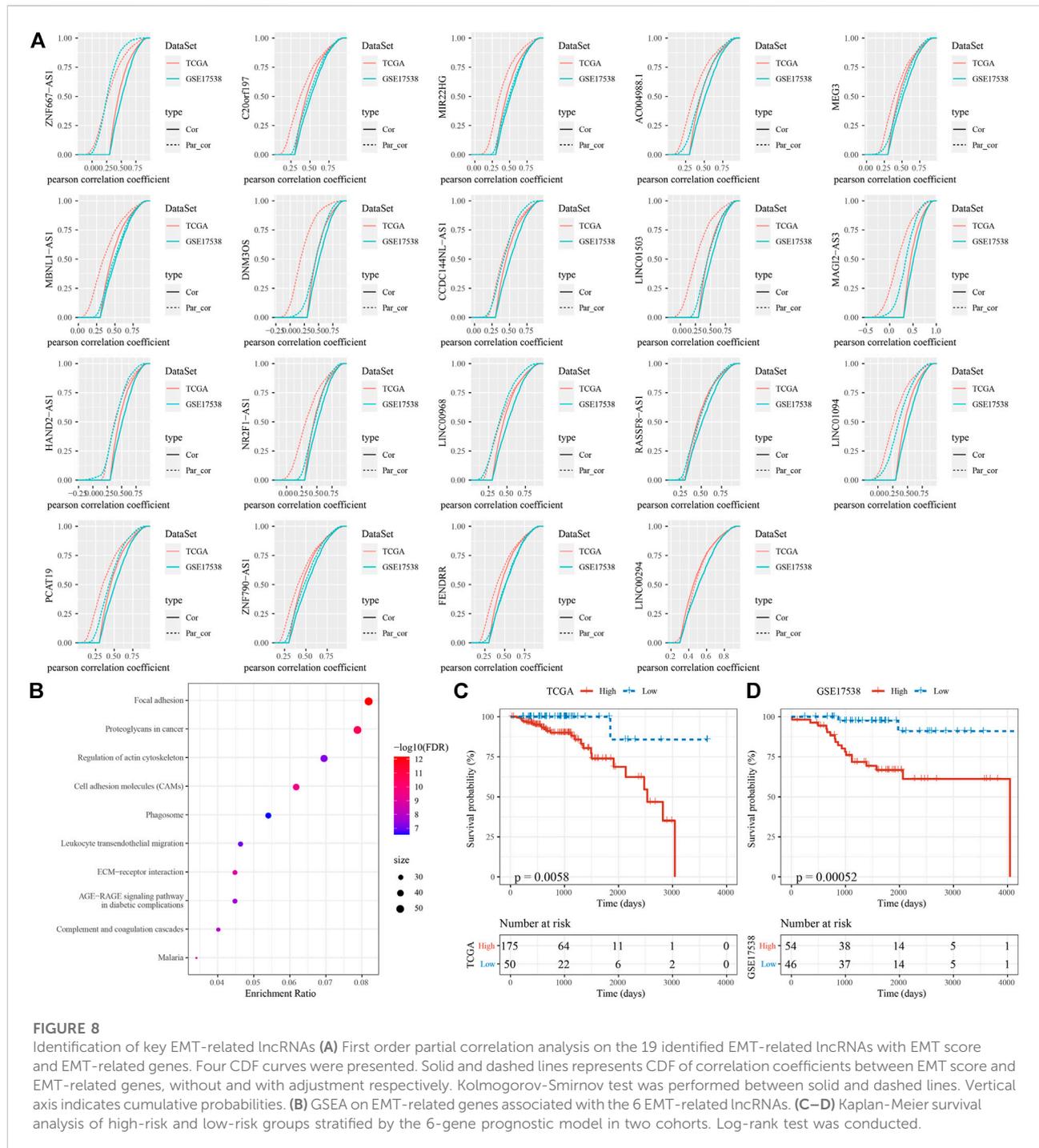


FIGURE 8

Identification of key EMT-related lncRNAs (A) First order partial correlation analysis on the 19 identified EMT-related lncRNAs with EMT score and EMT-related genes. Four CDF curves were presented. Solid and dashed lines represents CDF of correlation coefficients between EMT score and EMT-related genes, without and with adjustment respectively. Kolmogorov-Smirnov test was performed between solid and dashed lines. Vertical axis indicates cumulative probabilities. (B) GSEA on EMT-related genes associated with the 6 EMT-related lncRNAs. (C–D) Kaplan-Meier survival analysis of high-risk and low-risk groups stratified by the 6-gene prognostic model in two cohorts. Log-rank test was conducted.

In the comparison between two subtypes, it was reasonable that C1 subtype had more favorable OS, with lower EMT activity than C2. However, two subtypes displayed no obvious difference on genomic features. Notably, pathways activated in cancer development were observed to be more enriched in C2 subtype, indicating higher invasive activity of tumor cells leading to

migration in C2 subtype. There was no doubt in the results that EMT pathway was the most enriched in C2 subtype compared with other enriched tumor-related pathways such as angiogenesis, hypoxia, TNF- α signaling and TGF- β signaling. These pathways were all reported to be associated with EMT process under the regulation of lncRNAs or miRNAs.

Vascular endothelial growth factor (VEGF) contributing for angiogenesis is demonstrated to be linked with Twist2 expression and the reduction of E-cadherin levels (Rojas-Puentes et al., 2016). It has been reported that VEGF upregulation in solid tumors with hypoxia can stimulate the transformation of endothelial cells to mesenchymal cells (Holderfield and Hughes, 2008). Multiple studies on different cancer types demonstrated that VEGF administration induced the detection of EMT markers (Yang et al., 2006; Gonzalez-Moreno et al., 2010; Desai et al., 2013). Hypoxia is a critical hallmark of solid tumors, and are illustrated to induce EMT through hypoxia-inducible factor-1 α (HIF-1 α) activating SNAI1 (Zhang et al., 2013; Tam et al., 2020). Notably, HIF-1 α is a mediator of E-cadherin expression, a major epithelial tumor suppressor, whose reduction can promote angiogenesis (Evans et al., 2007). Links among hypoxia, HIF-1 α , E-cadherin, angiogenesis and EMT are shown to promote tumor invasion.

Techasen et al. have demonstrated that tumor necrosis factor- α (TNF- α), an inflammatory cytokine largely secreted from tumor stromal cells, stimulates EMT activation and significantly upregulated Snail expression in cholangiocarcinoma tissues (Techasen et al., 2012). Another critical cytokine tumor growth factor- β (TGF- β) expressing by tumor-infiltrating immune cells also serves as an inducer of EMT through forming EMT-permissive microenvironment, which creates a linkage between EMT and TME (Fuxe and Karlsson, 2012). Moreover, we found that JAK-STAT signaling controlling immune response was highly enriched in C2 subtype. JAK-STAT signaling mediates immune cells in response to cytokines and growth factors, which is involved in metastasis and EMT (Jin, 2020). Xue et al. have revealed that lncRNA-AB073614 promotes EMT through JAK-STAT3 signaling pathway in colorectal cancer cells (Xue et al., 2018). These associations support the reliability of EMT-related lncRNAs for classifying COAD patients into two subtypes where EMT-related pathways were highly enriched in C2 subtype.

Besides enriched pathways, two subtypes also showed distinct TME with high immune infiltration in C2 subtype. A series of cytokines, chemokines and immune checkpoints secreted by tumor cells and tumor-infiltrated immune cells contribute EMT-permissive TME (Jung et al., 2015). Although C2 subtype was highly infiltrated with higher proportion of cytotoxic immune cells, immunosuppressive cells such as macrophages, regulatory T cells, and MDSCs were also highly enriched. In addition, higher expression of immune checkpoints was shown in C2, further supporting the close relation between TME and EMT. In lung adenocarcinoma, Lou et al. found that multiple immune checkpoints associating with increased regulatory T cells including PD-L1, PD-1, TIM-3, B7-H3, BTLA and CTLA-4 displayed EMT phenotype (Lou et al., 2016). Similarly, C2 subtype with high EMT activity also

shown increased expression of BTLA, CD274 (PD-L1) and CTLA-4, suggesting that these immune checkpoints may be involved in EMT process.

To understand the possible mechanism of EMT-related lncRNAs regulating EMT-related genes, we investigated the relation between lncRNAs and TFs. We discovered that 7 TFs were significantly upregulated in C2 subtype, and JAK-STAT signaling pathway was found to be enriched in these TFs, which was consistent with the previous findings. Among the 19 identified lncRNAs correlated with TF expression, MEG3 and MIR22HG were massively expressed in the nuclear. MEG3, considering as an oncogenic lncRNA (Al-Rugeebah et al., 2019), is indicated to affect EMT in various cancer types such as breast cancer (Zhang et al., 2017), ovarian cancer (Wang et al., 2019), lung cancer (Terashima et al., 2017) and gastric cancer (Xu et al., 2018). However, it has not been reported in colon cancer, which may serve as a novel direction for further characterizing the mechanism of EMT. MIR22HG is demonstrated as a tumor suppressor through TGF- β /SMAD signaling in colorectal cancer, whose depletion can promote EMT process and tumor metastasis (Xu et al., 2020). We speculated that these 19 lncRNAs are highly involved in regulating EMT through interacting with TFs or other regulators.

Furthermore, to identify key EMT-related lncRNAs, we used first-order partial correlation analysis and dug out six key lncRNAs including ZNF667-AS1, CCDC144NL-AS1, MAGI2-AS3, HAND2-AS1, LINC01094, and PCAT19. These six lncRNAs were all reported to be involved in cancer progression. ZNF667-AS1 was reported to be involved in cancer progression and migration in laryngeal squamous cell carcinoma and cervical cancer (Li et al., 2019; Meng et al., 2019). CCDC144NL-AS1 was identified as a prognostic biomarker in non-small cell lung cancer and it could also promote hepatocellular carcinoma development (Zhang et al., 2021a; Zhang et al., 2021b). MAGI2-AS3 and HAND2-AS1 were dysregulated in many cancer types and were suggested as potential biomarkers for cancer prognosis (Gu et al., 2021; Kai-Xin et al., 2021). LINC01094 could promote the progression of ovarian cancer (Chen et al., 2021), breast cancer (Wu et al., 2021), pancreatic cancer (Luo et al., 2021), and other cancers. PCAT19 could activate cell-cycle genes thereby promoting cancer cell growth and cancer metastasis in pancreatic cancer (Hua et al., 2018).

In conclusion, this study proposed two novel molecular subtypes based on EMT-related lncRNAs. The distinct features of enriched pathways and TME between two subtypes supported that EMT-related lncRNAs played important roles in EMT process through regulating TFs involved in JAK-STAT signaling. Moreover, we identified six key EMT-related lncRNAs associated overall survival and the six lncRNAs could serve as a prognostic signature for COAD patients. As the study focused on COAD samples with stage I and II, we expanded the fundamental research on the early stages of COAD and the six lncRNAs may

serve as biomarkers for early diagnose of COAD. However, the limitation was that only pure bioinformatics analysis was applied in the present study. In addition, we did not distinguish left-sided and right-sided colon cancer, which may lower the accuracy of our results. In the future work, the role and prognostic value of six key EMT-related lncRNAs needed to be further explored and validated in more clinical patients.

Data availability statement

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/[Supplementary Material](#).

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by KL, KC, DL, SZ, and ZZ. The first draft of the manuscript was written by HL, YZ, and YX, DW and YP revised the manuscript. All authors read and approved the final manuscript.

References

- Al-Rugeebah, A., Alanazi, M., and Parine, N. R. (2019). MEG3: An oncogenic long non-coding RNA in different cancers. *Pathol. Oncol. Res.* 25, 859–874. doi:10.1007/s12253-019-00614-3
- Argilés, G., Tabernero, J., Labianca, R., Hochhauser, D., Salazar, R., Iveson, T., et al. (2020). Localised colon cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 31, 1291–1305. doi:10.1016/j.annonc.2020.06.022
- Chen, H., Liu, Y., Liu, P., Dai, Q., and Wang, P. (2021). LINC01094 promotes the invasion of ovarian cancer cells and regulates the Wnt/ β -catenin signaling pathway by targeting miR-532-3p. *Exp. Ther. Med.* 22, 1228. doi:10.3892/etm.2021.10662
- Cheng, J. T., Wang, L., Wang, H., Tang, F. R., Cai, W. Q., Sethi, G., et al. (2019). Insights into biological role of lncRNAs in epithelial-mesenchymal transition, *Cells*, 8, E1178. doi:10.3390/cells8101178
- Desai, S., Laskar, S., and Pandey, B. N. (2013). Autocrine IL-8 and VEGF mediate epithelial-mesenchymal transition and invasiveness via p38/JNK-ATF-2 signalling in A549 lung cancer cells. *Cell. Signal.* 25, 1780–1791. doi:10.1016/j.cellsig.2013.05.025
- Dongre, A., and Weinberg, R. A. (2019). New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* 20, 69–84. doi:10.1038/s41580-018-0080-4
- Evans, A. J., Russell, R. C., Roche, O., Burry, T. N., Fish, J. E., Chow, V. W., et al. (2007). VHL promotes E2 box-dependent E-cadherin transcription by HIF-mediated regulation of SIP1 and snail. *Mol. Cell. Biol.* 27, 157–169. doi:10.1128/MCB.00892-06
- Fuxe, J., and Karlsson, M. C. (2012). TGF- β -induced epithelial-mesenchymal transition: A link between cancer and inflammation. *Semin. Cancer Biol.* 22, 455–461. doi:10.1016/j.semcancer.2012.05.004
- Garcia-Alonso, L., Iorio, F., Matchan, A., Fonseca, N., Jaaks, P., Peat, G., et al. (2018). Transcription factor Activities enhance markers of drug sensitivity in cancer. *Cancer Res.* 78, 769–780. doi:10.1158/0008-5472.CAN-17-1679
- Gonzalez-Moreno, O., Lecanda, J., Green, J. E., Segura, V., Catena, R., Serrano, D., et al. (2010). VEGF elicits epithelial-mesenchymal transition (EMT) in prostate

Conflict of interest

Authors YX, ZZ, and DW were employed by the company YuceBio Technology Co., Ltd.

The remaining 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.997739/full#supplementary-material>

intraepithelial neoplasia (PIN)-like cells via an autocrine loop. *Exp. Cell Res.* 316, 554–567. doi:10.1016/j.yexcr.2009.11.020

Gu, X., Zheng, Q., Chu, Q., and Zhu, H. (2021). HAND2-AS1: A functional cancer-related long non-coding rna. *Biomed. Pharmacother.* 137, 111317. doi:10.1016/j.biopha.2021.111317

HänZELMANN, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7

Holderfield, M. T., and Hughes, C. C. (2008). Crosstalk between vascular endothelial growth factor, notch, and transforming growth factor-beta in vascular morphogenesis. *Circ. Res.* 102, 637–652. doi:10.1161/CIRCRESAHA.107.167171

Hua, J. T., Ahmed, M., Guo, H., Zhang, Y., Chen, S., Soares, F., et al. (2018). Risk SNP-mediated promoter-enhancer switching drives prostate cancer through lncRNA PCAT19. *Cell* 174, 564–575. doi:10.1016/j.cell.2018.06.014

Jin, W. (2020). Role of JAK/STAT3 signaling in the regulation of metastasis, the transition of cancer stem cells, and chemoresistance of cancer by epithelial-mesenchymal transition. *Cells* 9, E217. doi:10.3390/cells9010217

Jung, H. Y., Fattet, L., and Yang, J. (2015). Molecular pathways: Linking tumor microenvironment to epithelial-mesenchymal transition in metastasis. *Clin. Cancer Res.* 21, 962–968. doi:10.1158/1078-0432.CCR-13-3173

Kai-Xin, L., Cheng, C., Rui, L., Zheng-Wei, S., Wen-Wen, T., and Peng, X. (2021). Roles of lncRNA MAGI2-AS3 in human cancers. *Biomed. Pharmacother.* 141, 111812. doi:10.1016/j.biopha.2021.111812

Lee, G. H., Malietz, G., Askari, A., Bernardo, D., Al-Hassi, H. O., and Clark, S. K. (2015). Is right-sided colon cancer different to left-sided colorectal cancer? - a systematic review. *Eur. J. Surg. Oncol.* 41, 300–308. doi:10.1016/j.ejso.2014.11.001

Li, Y. J., Yang, Z., Wang, Y. Y., and Wang, Y. (2019). Long noncoding RNA ZNF667-AS1 reduces tumor invasion and metastasis in cervical cancer by counteracting microRNA-93-3p-dependent PEG3 downregulation. *Mol. Oncol.* 13, 2375–2392. doi:10.1002/1878-0261.12565

- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425. doi:10.1016/j.cels.2015.12.004
- Liu, Y., He, M., Wang, D., Diao, L., Liu, J., Tang, L., et al. (2017). HisGAtlas 1.0: A human immunosuppression gene database. *Database.* 2017, bax094. doi:10.1093/database/bax094
- Lou, Y., Diao, L., Cuentas, E. R., Denning, W. L., Chen, L., Fan, Y. H., et al. (2016). Epithelial-mesenchymal transition is associated with a distinct tumor microenvironment including elevation of inflammatory signals and multiple immune checkpoints in lung adenocarcinoma. *Clin. Cancer Res.* 22, 3630–3642. doi:10.1158/1078-0432.CCR-15-1434
- Luo, C., Lin, K., Hu, C., Zhu, X., Zhu, J., and Zhu, Z. (2021). LINC01094 promotes pancreatic cancer progression by sponging miR-577 to regulate LIN28B expression and the PI3K/AKT pathway. *Mol. Ther. Nucleic Acids* 26, 523–535. doi:10.1016/j.omtn.2021.08.024
- Mas-Ponte, D., Carlevaro-Fita, J., Palumbo, E., Hermoso Pulido, T., Guigo, R., and Johnson, R. (2017). LncAtlas database for subcellular localization of long noncoding RNAs. *Rna* 23, 1080–1087. doi:10.1261/rna.060814.117
- Meng, W., Cui, W., Zhao, L., Chi, W., Cao, H., and Wang, B. (2019). Aberrant methylation and downregulation of ZNF667-AS1 and ZNF667 promote the malignant progression of laryngeal squamous cell carcinoma. *J. Biomed. Sci.* 26, 13. doi:10.1186/s12929-019-0506-0
- O'Brien, S. J., Bishop, C., Hallion, J., Fiechter, C., Scheurle, K., Paas, M., et al. (2020). Long non-coding RNA (lncRNA) and epithelial-mesenchymal transition (EMT) in colorectal cancer: A systematic review. *Cancer Biol. Ther.* 21, 769–781. doi:10.1080/15384047.2020.1794239
- Petrelli, F., Tomasello, G., Borronovo, K., Ghidini, M., Turati, L., Dallera, P., et al. (2017). Prognostic survival associated with left-sided vs right-sided colon cancer: A systematic review and meta-analysis. *JAMA Oncol.* 3, 211–219. doi:10.1001/jamaoncol.2016.4227
- Reverter, A., and Chan, E. K. (2008). Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. *Bioinformatics* 24, 2491–2497. doi:10.1093/bioinformatics/btn482
- Rojas-Puentes, L., Cardona, A. F., Carranza, H., Vargas, C., Jaramillo, L. F., Zea, D., et al. (2016). Epithelial-mesenchymal transition, proliferation, and angiogenesis in locally advanced cervical cancer treated with chemoradiotherapy. *Cancer Med.* 5, 1989–1999. doi:10.1002/cam4.751
- Şenbabaoglu, Y., Gejman, R. S., Winer, A. G., Liu, M., Van Allen, E. M., De Velasco, G., et al. (2016). Erratum to: Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* 17, 46. doi:10.1186/s13059-017-1180-8
- Shen, W., Song, Z., Xiao, Z., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 3, e36. doi:10.1002/imt2.36
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi:10.1073/pnas.0506580102
- Sun, Z., Yang, S., Zhou, Q., Wang, G., Song, J., Li, Z., et al. (2018). Emerging role of exosome-derived long non-coding RNAs in tumor microenvironment. *Mol. Cancer* 17, 82. doi:10.1186/s12943-018-0831-z
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660
- Tam, S. Y., Wu, V. W. C., and Law, H. K. W. (2020). Hypoxia-induced epithelial-mesenchymal transition in cancers: HIF-1 α and beyond. *Front. Oncol.* 10, 486. doi:10.3389/fgene.2020.00486
- Techasen, A., Namwat, N., Loilome, W., Bungkanjana, P., Khuntikeo, N., Puapairoj, A., et al. (2012). Tumor necrosis factor- α (TNF- α) stimulates the epithelial-mesenchymal transition regulator Snail in cholangiocarcinoma. *Med. Oncol.* 29, 3083–3091. doi:10.1007/s12032-012-0305-x
- Terashima, M., Tange, S., Ishimura, A., and Suzuki, T. (2017). MEG3 long noncoding RNA contributes to the epigenetic regulation of epithelial-mesenchymal transition in lung cancer cell lines. *J. Biol. Chem.* 292, 82–99. doi:10.1074/jbc.M116.750950
- Ulanja, M. B., Rishi, M., Beutler, B. D., Sharma, M., Patterson, D. R., Gullapalli, N., et al. (2019). Colon cancer sidedness, presentation, and survival at different stages. *J. Oncol.* 2019, 4315032. doi:10.1155/2019/4315032
- Wang, L., Yu, M., and Zhao, S. (2019). lncRNA MEG3 modified epithelial-mesenchymal transition of ovarian cancer cells by sponging miR-219a-5p and regulating EGFR. *J. Cell. Biochem.* 120, 17709–17722. doi:10.1002/jcb.29037
- Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26, 1572–1573. doi:10.1093/bioinformatics/btq170
- Wu, Z. H., Wang, X. L., Tang, H. M., Jiang, T., Chen, J., Lu, S., et al. (2014). Long non-coding RNA HOTAIR is a powerful predictor of metastasis and poor prognosis and is associated with epithelial-mesenchymal transition in colon cancer. *Oncol. Rep.* 32, 395–402. doi:10.3892/or.2014.3186
- Wu, X., Kong, C., and Wu, Y. (2021). Long intergenic non-protein coding RNA 1094 (LINC01094) promotes the progression of breast cancer (BC) by regulating the microRNA-340-5p (miR-340-5p)/E2F transcription factor 3 (E2F3) axis. *Bioengineered* 12, 9046–9057. doi:10.1080/21655979.2021.1993715
- Xu, G., Meng, L., Yuan, D., Li, K., Zhang, Y., Dang, C., et al. (2018). MEG3/miR-21 axis affects cell mobility by suppressing epithelial-mesenchymal transition in gastric cancer. *Oncol. Rep.* 40, 39–48. doi:10.3892/or.2018.6424
- Xu, J., Shao, T., Song, M., Xie, Y., Zhou, J., Yin, J., et al. (2020). MIR22HG acts as a tumor suppressor via TGF β /SMAD signaling and facilitates immunotherapy in colorectal cancer. *Mol. Cancer* 19, 51. doi:10.1186/s12943-020-01174-w
- Xue, J., Liao, L., Yin, F., Kuang, H., Zhou, X., and Wang, Y. (2018). lncRNA AB073614 induces epithelial-mesenchymal transition of colorectal cancer cells via regulating the JAK/STAT3 pathway. *Cancer Biomark.* 21, 849–858. doi:10.3233/CBM-170780
- Yang, A. D., Camp, E. R., Fan, F., Shen, L., Gray, M. J., Liu, W., et al. (2006). Vascular endothelial growth factor receptor-1 activation mediates epithelial to mesenchymal transition in human pancreatic carcinoma cells. *Cancer Res.* 66, 46–51. doi:10.1158/0008-5472.CAN-05-3086
- Yoshihara, K., Shahmoradgol, M., MartiNEZ, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic* 16, 284–287. doi:10.1089/omi.2011.0118
- Zhang, L., Huang, G., Li, X., Zhang, Y., Jiang, Y., Shen, J., et al. (2013). Hypoxia induces epithelial-mesenchymal transition via activation of SNAIL1 by hypoxia-inducible factor -1 α in hepatocellular carcinoma. *BMC Cancer* 13, 108. doi:10.1186/1471-2407-13-108
- Zhang, W., Shi, S., Jiang, J., Li, X., Lu, H., and Ren, F. (2017). lncRNA MEG3 inhibits cell epithelial-mesenchymal transition by sponging miR-421 targeting E-cadherin in breast cancer. *Biomed. Pharmacother.* 91, 312–319. doi:10.1016/j.biopha.2017.04.085
- Zhang, L., Chi, B., Chai, J., Qin, L., Zhang, G., Hua, P., et al. (2021a). lncRNA ccdc144nl-AS1 serves as a prognosis biomarker for non-small cell lung cancer and promotes cellular function by targeting miR-490-3p. *Mol. Biotechnol.* 63, 933–940. doi:10.1007/s12033-021-00351-6
- Zhang, Y., Zhang, H., and Wu, S. (2021b). lncRNA-CCDC144NL-AS1 promotes the development of hepatocellular carcinoma by inducing WDR5 expression via sponging miR-940. *J. Hepatocell. Carcinoma* 8, 333–348. doi:10.2147/JHC.S306484