Biochemistry and Biophysics Reports 41 (2025) 101938

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

Identification and multi-omics analysis of essential coding and long non-coding genes in colorectal cancer

Yanguo Li [b,1], Zixing Meng [a,1], Chengjiang Fan [a], Hao Rong [a], Yang Xi [a,**], Qi Liao [a,*]

[a] Department of Biochemistry and Molecular Biology and Zhejiang Key Laboratory of Pathophysiology, Health Science Center, Ningbo University, Ningbo, Zhejiang, China
[b] Institute of Drug Discovery Technology, Ningbo University, Ningbo, Zhejiang, China

ARTICLE INFO

Keywords:
Essential genes
Essential lncRNAs
Single-cell RNA sequencing
CRISPR/Cas9
Tumor microenvironment
Multi-omics

ABSTRACT

Essential genes are indispensable for the survival of cancer cell. CRISPR/Cas9-based pooled genetic screens have distinguished the essential genes and their functions in distinct cellular processes. Nevertheless, the landscape of essential genes at the single cell levels and the effect on the tumor microenvironment (TME) remains limited. Here, we identified 396 essential protein-coding genes (ESPs) by integration of 8 genome-wide CRISPR loss-of-function screen datasets of colorectal cancer (CRC) cell lines and single-cell RNA sequencing (scRNA-seq) data of CRC tissues. Then, 29 essential long non-coding genes (ESLs) were predicted using Hypergeometric Test (HT) and Personalized PageRank (PPR) algorithms based on ESPs and co-expressed network constructed from scRNA-seq. CRISPR/Cas9 knockout experiment verified the effect of several ESPs and ESLs on the survival of CRC cell line. Furthermore, multi-omics features of ESPs and ESLs were illustrated by examining their expression patterns and transcription factor (TF) regulatory network at the single cell level, as well as DNA mutation and DNA methylation events at bulk level. Finally, through integrating multiple intracellular regulatory networks with cell-cell communication network (CCN), we elucidated that *CD47* and *MIF* are regulated by multiple CRC essential genes, and the anti-cancer drugs sunitinib can interfere the expression of them potentially. Our findings provide a comprehensive asset of CRC ESPs and ESLs, sheding light on the mining of potential therapy targets for CRC.

1. Introduction

Colorectal cancer is one of the most frequently diagnosed cancers worldwide. There were estimated 1.88 million new cases and 0.92 million deaths of CRC in 2020 [1]. Although novel therapies and drugs have approximately doubled the average survival time over the past decade, patients with advanced CRC often succumb to the disease within three years [2].

Gene is defined as essential if its complete loss of function will lead to impaired cell development, gene activation, and cellular reprogramming [3–5]. The pooled CRISPR loss-of-function screen is a widely used method for identifying essential protein-coding genes involved in cancer-related biological processes including drug resistance [6], immune suppression [7], and immunotherapy response [8]. A number of studies have already obtained multiple sets of pan-cancer core ESPs and non-ESPs using human cancer cell lines, including Hart's ESPs and Hart's non-ESPs [9,10], Achilles' ESPs [11], and Zhang's ESPs [12]. Notably, the Dependency Map (DepMap) portal and BioGRID ORCS have provided an index of CRISPR screens research focused on identifying and understanding the landscape of pan-cancer ESPs [13,14]. Although these studies have highlighted the functions of ESPs and their regulatory networks in cancer tissues, how their impacts on the TME are still unknown and required further exploration and establishment. Besides, the genotypes and phenotypes of various human cancer cell lines and tissues were heterogeneous. It was reported that ESPs in breast, pancreatic, and ovarian cancer cell lines were partially overlapped, indicating certain ESPs may function in a tumor-specific fashion [15]. Nowadays, few studies were focused on specific ESPs in CRC.

During the past several years, scRNA-seq has become an important technique for resolving critical cancer genes and illustrating the composition of TME. Numerous studies have been conducted for a comprehensive investigation of the cellular and intercellular

* Corresponding author.
** Corresponding author.
E-mail addresses: xiyang@nbu.edu.cn (Y. Xi), liaoqi@nbu.edu.cn (Q. Liao).
[1] Co-first author.

https://doi.org/10.1016/j.bbrep.2025.101938
Received 5 December 2024; Received in revised form 19 January 2025; Accepted 28 January 2025

interactions in CRC, demonstrating the dynamic changes in tumor immunity and providing guidance for clinical treatment strategies [16,17]. Recently, the technology that combined scRNA-seq with CRISPR - activation [5], CRISPR - knockout screen [18], and CRISPR - interference [19] have been emerged to identify genetic regulators and lineage traces. For example, Vishnubalaji et al. integrated the scRNA-seq with CRISPR functional screen data from Achilles project [11] and identified potential therapeutic targets for different subtypes of Breast cancer [20]. These studies suggest that combined analysis of scRNA-seq and CRISPR functional screen data can more accurately identify essential genes in CRC and will broaden our understanding of their role in the high-dimensional TME.

In this study, we first integrated genome-wide CRISPR/Cas9 loss-of-function screen data and scRNA-seq data to identify ESPs in CRC. Then CRC ESLs were predicted using HT and PPR algorithms based on gene co-expressed network constructed from scRNA-seq data of tumor epithelial cells. To further illustrate the characteristics of ESPs and ESLs, we examined their single cell expression patterns, pathway enrichment, DNA mutation, DNA methylation events, and transcription factor (TF) regulatory network. Finally, we integrated co-expressed network, TF regulatory network, and CCN to provide valuable insights into the impact of ESPs and ESLs on the remodeling of TME. The potential drugs that can interfere key factors of target signal process were also predicted. Collectively, our study presents novel insights into the essential roles of CRC ESPs and ELSs in tumor cell development, TF regulation, and TME remodeling, thus establishing a foundation for identification of novel

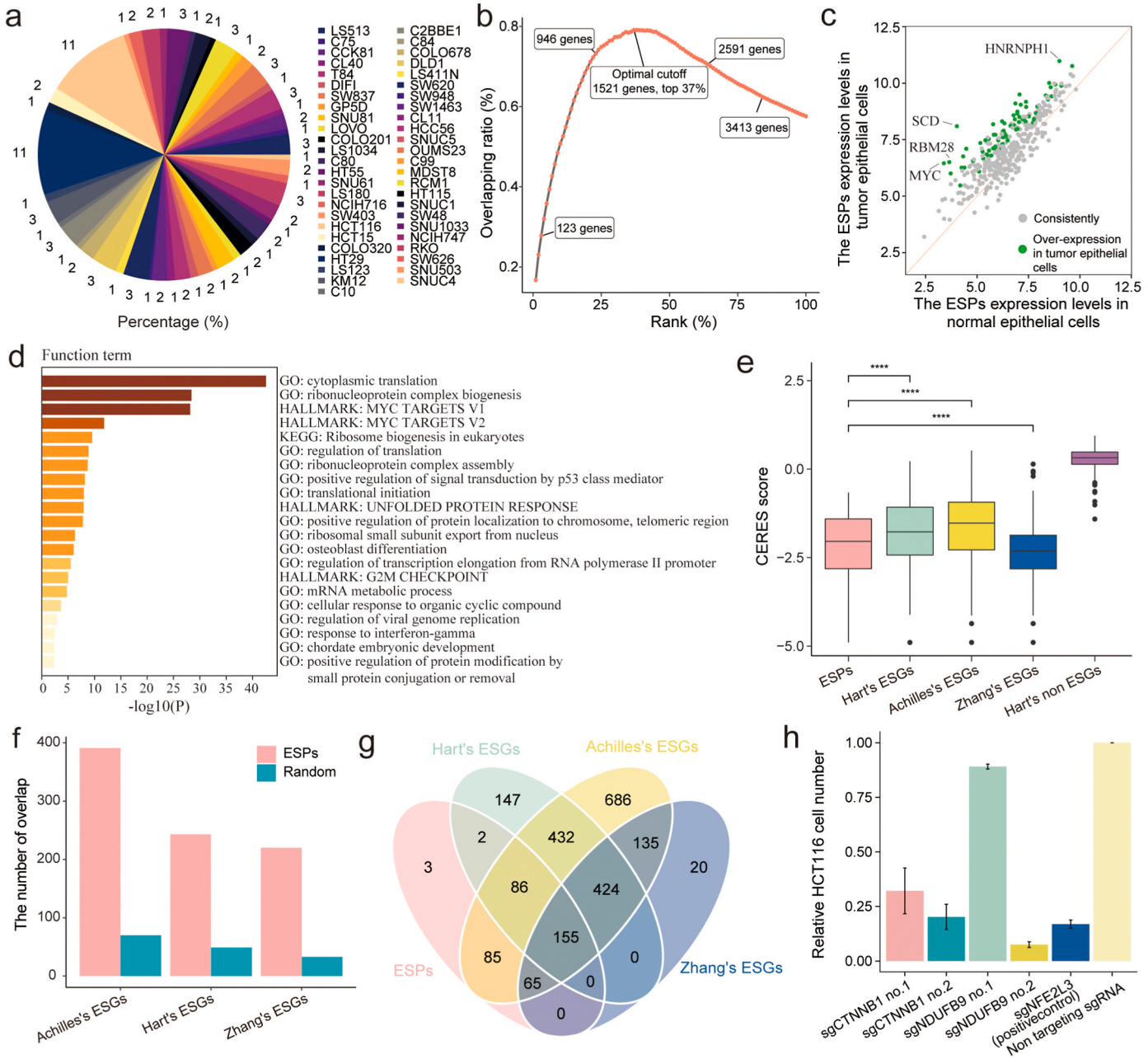

Fig. 1. Identification of CRC ESPs using CRISPR loss-of-function screen and scRNA-seq datasets. **a.** The pie chart displays the CRC cell line types in the CRISPR loss-of-function screen samples. **b.** Overlapping ratios between previously identified pan-cancer ESP sets and our results with different ranking proportion cutoffs. **c.** Average expression levels of CRC ESPs in tumor and normal epithelial cells. **d.** Functional enrichment analysis of 74 overexpressed CRC ESPs. **e.** Comparison of CERES scores among different ESP sets. **f.** Overlap between CRC ESPs or randomly selected genes and three pan-cancer ESP sets. **g.** The venn diagram illustrates the intersection of different ESP sets. **h.** The effect of *CTNNB1* and *NDUFB9* knockout by CRISPR/Cas9 on the survival of HCT116 cell lines.

therapeutic targets.

2. Results

2.1. Identification of CRC ESPs based on integration of CRISPR screen and scRNA-seq datasets

Eight genome-wide CRISPR/Cas9 loss-of-function screen datasets of CRC were collected, including a total of 49 CRC cell lines and 301 samples with replicates (Fig. 1a, Supplementary Table 1). To evaluate the essentiality of each gene, the sgRNAs depletion score for each gene in each sample was calculated by CERES algorithm and defined as CERES score [11]. A lower CERES score indicates a higher level of gene essentiality as all the screen datasets are negative selection in this study. Considering multiple datasets were involved, we integrated CERES scores for each gene across various datasets by utilizing Robust Rank Aggregation (RRA) algorithm [21]. The optimal cutoff to obtain ESPs was determined by calculating the degree of overlap between RRA results and other pan-cancer ESP sets previously identified, including Achilles' ESPs, Hart's ESPs, and Zhang's ESPs. As a result, 1521 candidates of CRC ESPs were obtained (Fig. 1b). We further collected scRNA-seq data of CRC (smart-seq, GSE81861) to analyze the detection fraction of candidate ESPs in tumor epithelial cells [22]. We found 396 ESPs were expressed in more than 50 % of tumor epithelial cells and were thus considered as final CRC ESPs (Supplementary Table 2). Among them, 74 ESPs were found to be overexpressed in tumor epithelial cells (Fig. 1c, Supplementary Table 2). These ESPs were principally linked to cancer-related functions, such as MYC TARGETS V1/V2, positive regulation of signal transduction by p53 class mediator, and response to interferon-gamma (Fig. 1d). Notably, we found ESPs exhibit higher expression than other genes in both tumor cells and normal cells (Supplementary Fig. a), which corroborated the findings from a previous study [9].

Then we compared the CERES scores of CRC ESPs with those of pan-cancer ESPs and Hart's non-ESPs to assess the reliability of our results. As anticipated, Hart's non-ESPs showed the highest CERES scores and indicated the lowest necessity, while the CRC ESPs were lower, implying their essential roles in CRC (Fig. 1e). Additionally, we observed larger overlap between CRC ESPs and other pan-cancer ESP sets than that of randomly selected genes, further reflecting the high quality of CRC ESPs (Fig. 1f). Among 396 ESPs, 155 ESPs co-exist in all ESP sets, indicating the homogeneity of different tumors (Fig. 1g). Notably, we found three novel CRC ESPs, *CTNNB1*, *NDUFB9*, and *SCNM1* that were not found in any other ESP sets (Fig. 1g). To determine whether these novel CRC ESPs are functionally important in the survival of CRC cancer cell lines, we detected the effect of *CTNNB1* and *NDUFB9* knockout on HCT116 cell line by CRISPR/Cas9. Remarkably, HCT116 cell line lost viability while *CTNNB1* and *NDUFB9* were knockout, respectively (Fig. 1h). *CTNNB1* was reported as a driver gene in CRC according to the Pan-Cancer Analysis of Whole Genomes (PCAWG) study [23]. *NDUFB9*, an important gene involved in mitochondrial function, has been identified as a prognostic biomarker for endometrial cancer [24]. However, the role of *SCNM1* in CRC is needed to further investigate.

2.2. Prediction of CRC ESLs

LncRNAs play an important regulatory role in various cellular processes of CRC, including cell apoptosis, proliferation, and epithelial-mesenchymal transition [25]. To elucidate a comprehensive landscape of essential genes in CRC, we employed HT and PPR methods to predict essential lncRNAs [12] (**see methods**). These methods required the construction of a co-expression network comprising of both protein-coding genes and lncRNAs, where lncRNAs essentiality were predicted based on their closeness to ESPs [12]. We built four different co-expression networks to assess the best network, including an unweighted network and a topological overlap weighted network based on the RNA sequencing profiles of CRC from The Cancer Genome Atlas (TCGA, referred to as unWCN-TCGA and WCN-TCGA, respectively), as well as an unweighted network and a topological overlap weighted network based on expression profiles of tumor epithelial cells from scRNA-seq dataset (GSE81861, referred to as unWCN-TEC and WCN-TEC, respectively) (Fig. 2a, **see method**). Then we calculated the adjusted *p*-value of each gene by HT to evaluate whether the directly related genes of a query gene are enriched in ESPs based on unWCN-TCGA and unWCN-TEC (Fig. 2a). We observed that the adjusted *p*-values of CRC ESPs, Achilles' ESPs, Zhang's ESPs, and Hart's ESPs showed significantly smaller than those of Hart's non-ESPs (Fig. 2b). Importantly, the adjusted *p*-values of ESPs based on unWCN-TEC were significantly lower than those based on unWCN-TCGA (Fig. 2b), suggesting the higher quality of co-expression network constructed from scRNA-seq than RNA-seq. Consequently, we identified 26 lncRNAs as potential candidates of CRC ESLs by using the cutoff of adjusted *p*-value lower than 0.01 based on unWCN-TEC.

Next, we utilized CRC ESPs as original seeds to calculate the score of all genes by PPR method based on WCN-TCGA and WCN-TEC, respectively (Fig. 2a). The higher score represents the higher probability of a gene being essential. As expected, the PPR scores of CRC ESPs, Achilles' ESPs, Zhang's ESPs, and Hart's ESPs were higher, and WCN-TEC also showed greater advantage than WCN-TCGA as higher PPR scores of ESPs in WCN-TEC were observed (Fig. 2c). We then selected the top 5 % lncRNAs in PPR results based on WCN-TEC as CRC ESL candidates (n = 39). Results yielded 41 ESLs after combining the results of the HT and PPR methods, and 24 lncRNAs were found to be common between both methods (Fig. 2a). Finally, we retained the ESLs detected in more than 30 % of CRC tumor epithelial cells and defined them as CRC ESLs (n = 29, Supplementary Table 3). Expression analysis revealed that 11 CRC ESLs were significantly overexpressed in CRC tumor epithelial cells compared to normal (Fig. 2d, Supplementary Table 3). Among the 11 CRC ESLs, CRISPR/Cas9 knockout experiments demonstrated that knocking out *H19*, *MIF-AS1*, *NORAD*, *SNHG17*, *SNHG6*, and *SNHG8* in the HCT116 cell line resulted in a decrease in the number of surviving cells (Fig. 2e), implying the dependent role of those ESLs in CRC.

In our results, some ESLs were known as CRC associated lncRNAs according to previous reports, such as *PVT1*, *H19*, and *GAS5* (Fig. 2f, Supplementary Table 3). By analyzing the functions and clinical association of CRC ESLs in Lnc2Cancer 3.0 database [26], we found that most CRC ESLs are involved in cell apoptosis, growth, survival, and metastasis (Fig. 2f). Among these ESLs, 11 lncRNAs were previously identified as part of Zhang's ESLs (n = 97) and are associated with CRC, while 13 genes were identified as potential novel CRC-related lncRNAs (Fig. 2f). Next, to identify potential Gene Ontology (GO) biological process functions of CRC ESLs, we conducted enrichment analysis on the top 200 associated genes of each ESL based on WCN-TEC. The results revealed that a significant number of CRC ESLs are involved in ribosome/ncRNA/rRNA/DNA processing, regulation of translation, telomere maintenance, and signal transduction by p53 class mediator (Supplementary Table 4), providing additional evidence for the crucial roles of CRC ESLs in fundamental metabolic processes of tumor cells.

2.3. Expression, DNA mutation, and DNA methylation patterns of CRC essential genes

To examine the expression characteristics of essential genes in different cell types of CRC, we utilized CRC ESPs and ESLs as gene sets to calculate enrichment scores in each cell type based on the 10x genomics scRNA-seq dataset of CRC (GSE132465) [17]. As anticipated, both CRC ESPs and ESLs gene set enrichment scores were higher in epithelial cells of tumor samples compared with normal samples (Fig. 3a). Interestingly, the gene set scores of CRC ESPs were particularly elevated in tumor epithelial cells compared to other pan-cancer ESP sets (Fig. 3a), indicating the specificity of CRC ESPs for CRC tumor epithelial cells. We also observed the expressions of ESPs and ESLs in other cell lineages, such as

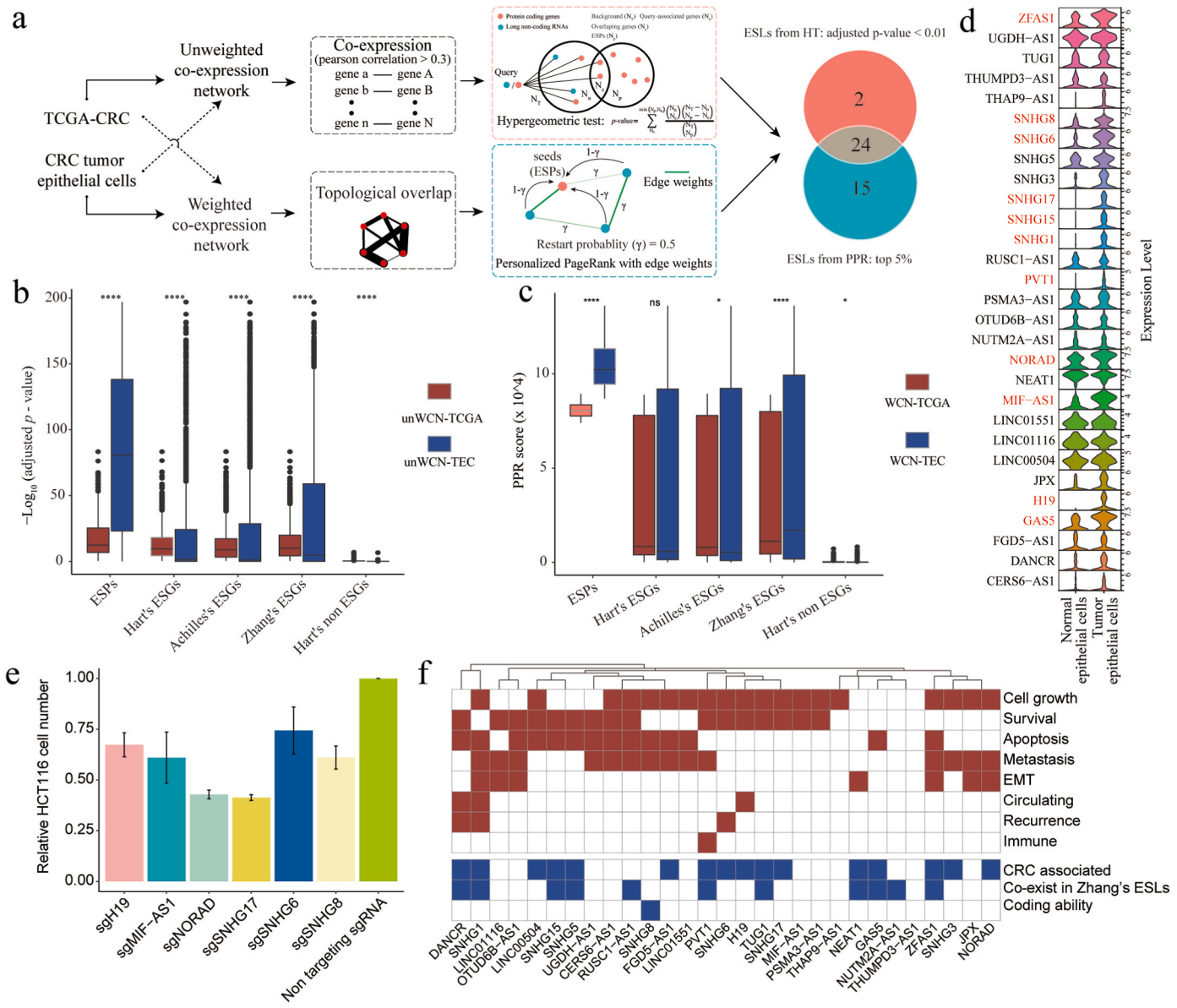

Fig. 2. Prediction of CRC ESLs using co-expression networks. **a.** The workflow of predicting CRC ESLs. The Venn diagram displays the intersection of predicted CRC ESLs from the two methods. **b.** The distributions of adjusted *p*-value for different ESP sets by HT method based on unWCN-TCGA and unWCN-TEC. **c.** The distributions of PPR rank scores for different ESP sets based on WCN-TCGA and WCN-TEC. **d.** The expression levels of CRC ESLs in CRC tumor and normal epithelial cells. Genes marked in red indicate significantly overexpression in tumor cells. **e.** The effect of ESLs knockout by CRISPR/Cas9 on survival of HCT116 cell line. **f.** The functional annotation of ESLs in CRC. Red block indicates publication-reported functions of lncRNAs in pan-cancer, blue block represents publication-reported properties of lncRNAs, and white block signifies the absence of publication reports.

T cells and B cells (Fig. 3a), implying that target essential genes may have side effect on cell lineages in TME. Notably, 4 ESPs (*KRT8*, *MYC*, *PRELID3B*, and *SCD*) were specifically expressed in tumor epithelial cells (Fig. 3b, **see methods**). Three of them, *KRT8*, *MYC*, and *SCD* are well-known biomarkers or therapeutic targets for cancer [27–29], while few research has performed on *PRELID3B* in cancer. CRISPR/Cas9 knockout experiments confirmed that function loss of these genes affects the survival of HCT116 cell line (Fig. 3c).

Mutation of essential genes has been shown to disrupt normal cell development and promote cancer, such as *MYC*. We conducted mutation analysis using whole genome sequencing (WGS) datasets of CRC in PCAWG and whole exome sequencing (WES) datasets of CRC in TCGA to explore the genomic alteration of CRC ESPs and ESLs. In general, the mutation rates of ESPs and ESLs in WGS dataset ranged from 0 % to 25 % and 0 %–33 %, respectively (Fig. 3d–e), while the mutation rates of ESPs were relatively low in WES dataset, ranging from 0 % to 8 % (Supplementary Fig. b). Among the mutated genes in WGS study, *DYNC1H1* (25 %), *CHD4* (23 %), *CNOT1* (21 %), and *CTNNB1* (19 %) exhibited the highest mutation frequency (Fig. 3d). Similarly, these genes also showed relatively high frequencies and ranked within the top 5 mutated genes in the TCGA CRC dataset (Supplementary Fig. b). Importantly, most CRC ESPs mutated in a co-occurrence manner in both the WGS and WES datasets (Supplementary Fig. c-d), suggesting that the mutation of ESPs typically activates synergistic oncogenic pathways.

DNA methylation also plays a crucial role in influencing gene expression. Through analysis of differentially methylated positions (DMPs) in ESPs and ESLs in TCGA-COAD, we identified a total of 71 DMPs with 45 unique genes (adjusted *p*-value <0.05, absolute $\log_2FC > 0.2$, Supplementary Fig. e). Of these DMPs, 26 showed DNA hypermethylation ($\log_2FC > 0.2$), while 45 exhibited DNA hypomethylation ($\log_2FC < -0.2$). Subsequently, we calculated the correlation coefficients between the expression and DNA methylation levels of each

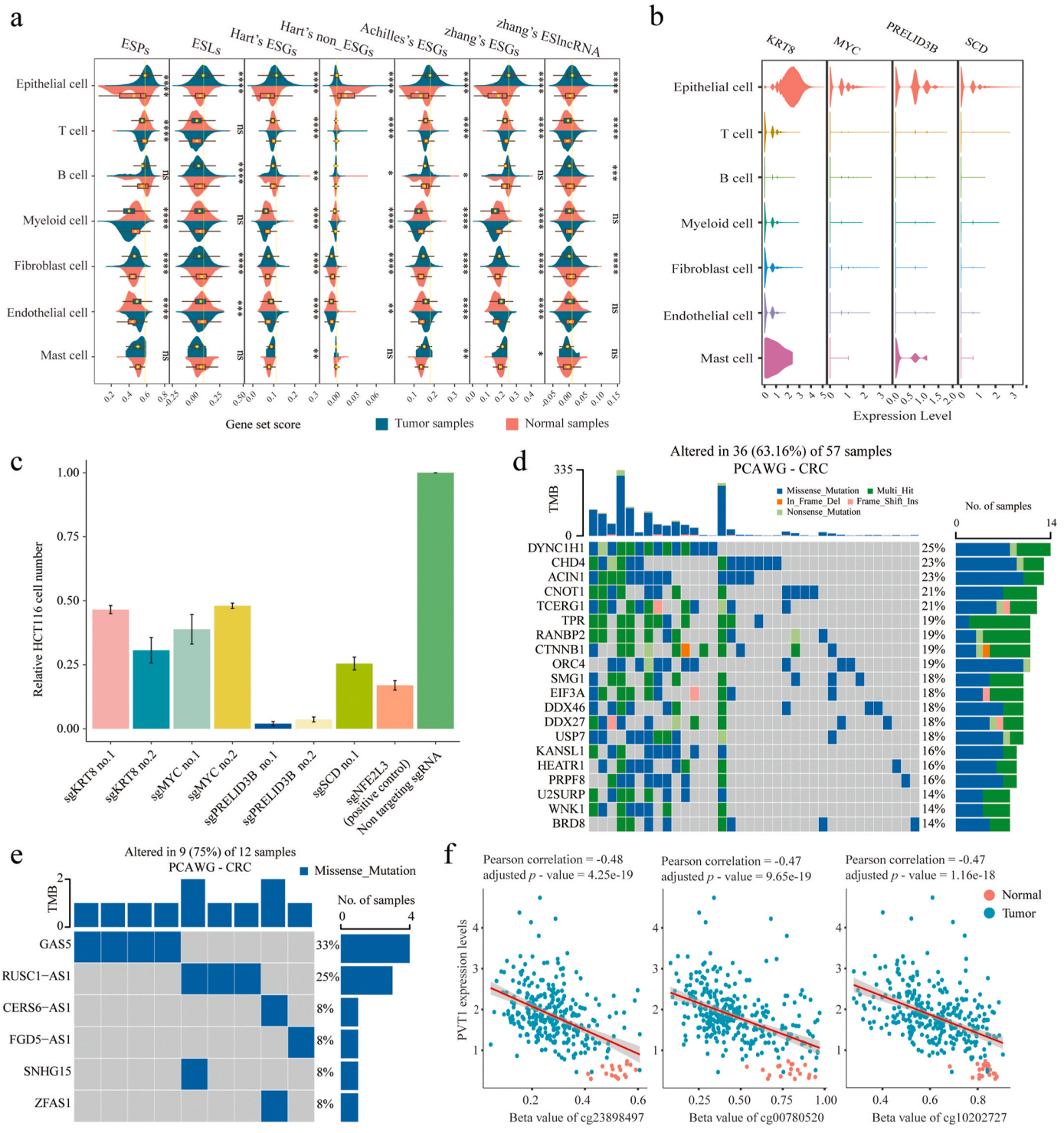

Fig. 3. Expression, DNA mutation, and DNA methylation patterns of CRC essential genes. **a.** The violin plot shows the gene sets enrichment scores of CRC ESPs, ESLs, other pan-cancer ESP sets, and Hart's non-ESPs in different cell lineages between tumor and normal samples. The yellow line represents the mean score of CRC ESPs in tumor epithelial cells. **b.** Four CRC ESPs were specifically expressed in tumor epithelial cells compared with other cell lineages. **c.** The effect of four CRC ESPs knockout by CRISPR/Cas9 on survival of HCT116 cell lines. **d-e.** The top mutated CRC ESPs (d) and ESLs (e) based on WGS datasets of PCAWG. **f.** The correlation between *PVT1* expression and DNA methylation levels of three probes.

gene. The results revealed three ESPs (*MYC*, *SCD*, *EIF6*) and one ESLs (*PVT1*) had negative correlation coefficients (Supplementary Table 5). We observed hypomethylation of two probes cg08526705 and cg00163372 in the gene body may activate *MYC* expression in CRC (Supplementary Fig. f-g). These two probes were found to be related with chemotherapy drug resistance previously [30]. In *PVT1*, three probes cg23898497, cg00780520, and cg10202727 were found to be associated with *PVT1* expression in CRC (Fig. 3f, Supplementary Table 5). *PVT1* is an adjacent gene of *MYC* on chromosome 8q24.21 and has been shown to regulate *MYC* expression and contribute to cancer development [31]. Our findings indicated that the hypomethylation of *MYC* and *PVT1* lead to the upregulation of their expression, and thus

promote the development of CRC.

2.4. TF regulatory network of essential genes in CRC tumor epithelial cells

Based on scRNA-seq dataset of CRC (GSE132465), we generated a TF regulatory network by using SCENIC algorithm [32]. A total of 30,743 TF-target relationships associated with CRC ESPs and ESLs were established, including 177 TFs and 9837 genes (Supplementary Table 6). These relationships were further categorized into high- or low-confidence annotations [32] (**see methods**). Interestingly, 9 TFs from ESPs related to 80.63 % (n = 24,788) of identified TF-target relationships. Among them, 3 TFs had a highest number of high-confidence relationships compared with others essential TFs, including *BCLAF1*, *MYC*, and *YY1* (Fig. 4a). Area under the curve (AUC) scores from SCENIC were used to quantify the activity of TF targets (regulon). Differential AUC scores of regulons revealed the significant activation of 79 regulons in tumor epithelial cells, including *BCLAF1* regulon, *YY1* regulon, and *MYC* regulon (Fig. 4b–c, Supplementary Table 7). Furthermore, three novel CRC ESPs (*CTNNB1*, *NDUFB9* and *SCNM1*) that mentioned above were also regulated by *BCLAF1*, *YY1*, and *MYC* (Fig. 4d), suggesting that these TFs may act as oncogenes in CRC. Previously, knockdown of the L isoform of *BCLAF1* in mouse tumor model inhibited the tumor growth, suggesting the carcinogenic characteristics of *BCLAF1* [33]. Above all, our findings demonstrated the intricate regulatory networks between TFs and CRC essential genes.

In the regulatory network, 424 relationships were linked between TFs (n = 108) and ESLs (n = 28, Fig. 4e, Supplementary Table 6). Notably, the lncRNAs such as *FGD5-AS1*, *GAS5*, and *PVT1* were regulated by more than 20 TFs. Among these, *FGD5-AS1*, has been emerged as a crucial regulator in CRC and other types of cancer via promoting cell proliferation, drug resistance, and epithelial-mesenchymal transition [34,35]. In our study, 14 TFs were found to regulate the expression of *FGD5 - AS1* (Fig. 4f), some of them were known well associated with CRC, such as *EGR1* [36] and *XBP1* [37]. Collectively, our findings provided a reliably regulatory network connecting TFs and CRC essential genes based on scRNA-seq data.

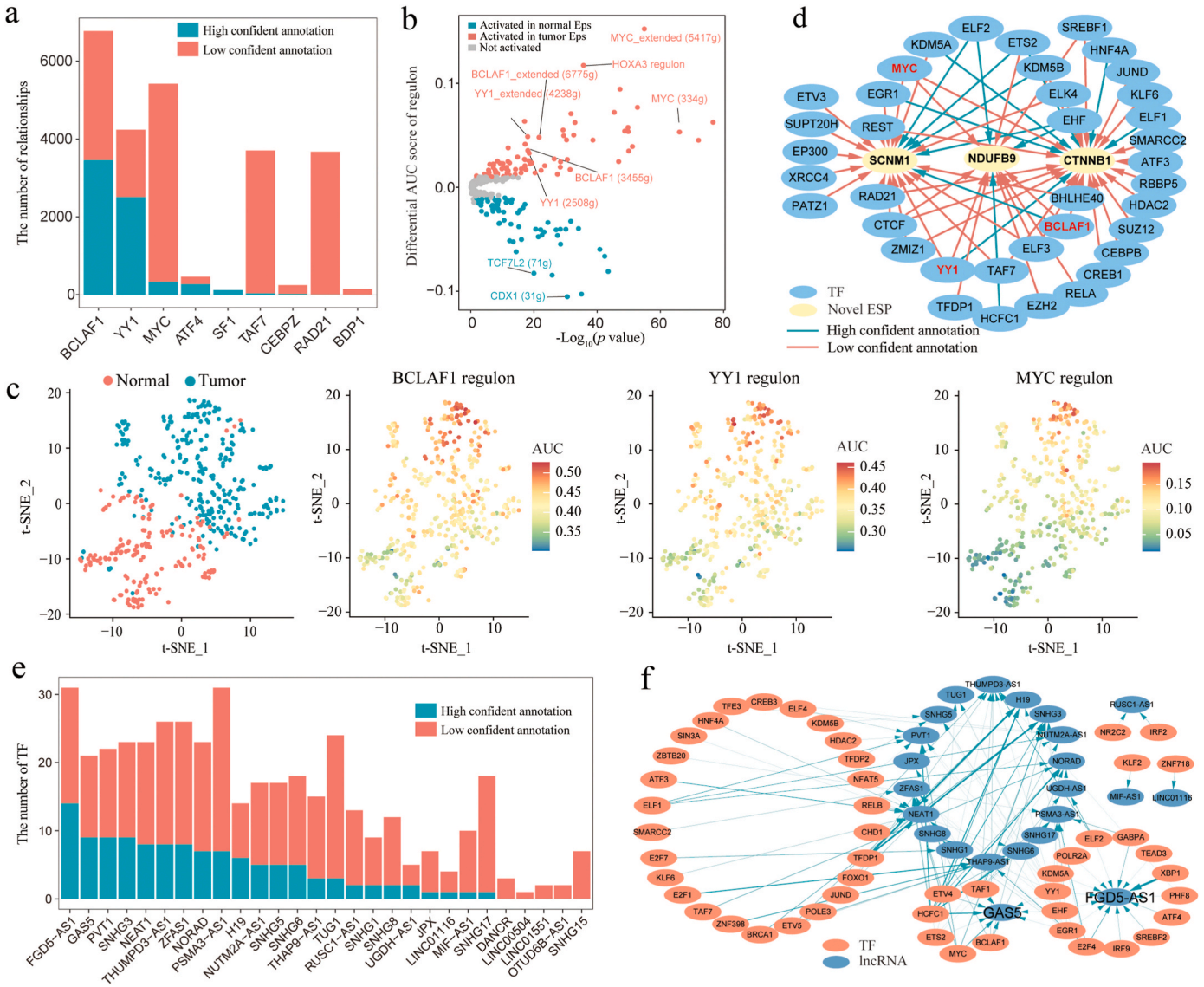

Fig. 4. Regulatory network between CRC essential genes and TFs. **a.** The number of target genes regulated by 9 TFs from ESPs. **b.** The t-SNE plot displays the distribution of tumor and normal epithelial cells, along with the AUC scores of three regulons. **c.** The differential AUC scores of regulons between tumor and normal epithelial cells. **d.** The network visualization of relationships between TFs and three novel CRC ESPs. **e.** The number of TFs regulating ESLs. **f.** The network visualization of TF-ESL relationships with high confidence. The width of the line represents the weight of associations.

2.5. Crosstalk between CRC essential genes and TME

The CCN between cancer cells and cells in TME mediated by ligand-receptor interactions, plays a crucial role in shaping tumor behavior and TME remodeling. In our study, only six CRC ESPs, including *COPA*, *RPS19*, *RACK1*, *PPIA*, *TMED2*, and *TFRC*, were known to act as ligands or receptors [38]. The majority of CRC essential genes may participate in the CCN by regulating the expression of ligands or receptors in tumor epithelial cells (Fig. 5a). To verify such relationships, we first constructed a CCN associated with tumor epithelial cells using CellphoneDB [38]. The CCN consists of 169 connections targeting tumor epithelial cells and 141 signals source from tumor epithelial cells (Supplementary Table 8). We found CRC tumor epithelial cells interacted more frequently with endothelial cells, fibroblast cells and myeloid cells (Fig. 5b), which are known to play important roles in angiogenesis, tumor invasion, and immune response.

We then integrated the WCN-TEC, TF regulatory network, and CCN, resulting in a total of 1734 relationships (Supplementary Table 9), where 356 ESPs (including 9 TFs) and 28 ESLs affect the expression of 85 ligands/receptors of tumor epithelial cells, suggesting that CRC ESPs and TFs have a widespread influence on CCN (Supplementary Fig. h). We observed that six ligands/receptors of tumor epithelial cells, including *GPI*, *CDH1*, *MIF*, *RPS19*, *CXADR*, and *CD47*, are associated with more than 100 essential genes (Supplementary Table 9), indicating these ligands/receptors may act as the critical mediators between essential genes and TME. Among them, *MIF* secreted from tumor epithelial cells could interact with *CD74*, TNFRSF10D, and *TNFRSF14* (Fig. 5c), the relationships may responsible for the suppression of anti-tumor immune response [39]. Additionally, it has been reported that the CD47-SIRPα/γ interaction protects tumor cells from killing by suppressing both macrophage phagocytosis and antigen presentation of dendritic cells [40]. Consistent with these findings, we observed a significant activation of CD47-SIRPα interactions between tumor epithelial cells and myeloid cells (Fig. 5c). Clinically, CRC patients from TCGA with high expression levels of *CD47* and *MIF* exhibited a significantly lower overall survival rate compared to those patients with low expression levels of *CD47* and *MIF* (Fig. 5d). Our results supported the notion that essential genes can facilitate tumor cells evasion from immune system by regulating the expression of receptors/ligands such as *CD47* and *MIF* in CRC.

Next, we aimed to identify existing drugs that could potentially repress CRC essential genes and TFs ($n = 304$) associated with *CD47* and *MIF* (Fig. 5a). We utilized an integrative web platform, iLINCS [41], to analyze the expression pattern of 304 targets in pre-computed signatures of 10 anti-cancer agents in 5 CRC cell lines (GSE116439) [42]. We found that drugs such as sunitinib, dasatinib, topotecan, and lapatinib were sensitive to cancer cells, as they can effectively inhibit the activity of target genes at appropriate concentration and treatment time (Fig. 5e, Supplementary Table 10). For instance, sunitinib acts as an inhibitor of *CSF1R*, *VEGF* receptor, c-kit, *PDGF* receptor, and *RET*, and has been widely used in cancer therapy [43]. Treatment with sunitinib resulted in the down-regulation of *MYC*, *KRT8*, *CDC37*, and other genes in the CRC cell line KM12 (Fig. 5e, signature ID: PG_4150). In addition, the patients with high *CD47* expression benefit from sunitinib monotherapy in clear cell renal cell carcinoma [44], further supporting sunitinib may act as CD47 inhibitor in CRC. On the other hand, CRC cell lines treated with cisplatin or vorinostat exhibited over-expression of target genes such as *MYC* and *TAF7* (Fig. 5e), indicating that CRC cells may develop resistance to these drugs. In conclusion, our results suggest that CRC essential genes may impact the TME through CCN and provides valuable insights into potential therapeutic strategies for CRC.

3. Discussion

Pan-cancer essential genes are widely recognized as therapeutic targets in cancer. However, the failure to account for tumor heterogeneity when targeting pan-essential genes likely contributes to the clinical failure [15,45]. In this study, we combined genome-wide CRISPR screen, scRNA-seq, co-expression networks, and CRISPR/Cas9 knockout experiments to identify and validate ESPs and ESLs associated with CRC. Our results reveal that CRC ESPs partially overlap with well-known pan-cancer ESPs, exhibit high expression in CRC tumor epithelial cells, and play critical roles in fundamental pathways, transcription factor regulation, and cell-cell communication networks. Collectively, our findings provide a comprehensive landscape of CRC ESPs and ESLs, offering new avenues for therapeutic strategies.

We streamlined the candidate ESPs from 1521 to 396 by prioritizing those expressed in over 50 % of tumor epithelial cells based on scRNA-seq. Consistent with previous studies, single-cell expression pattern of ESPs and ESLs showed higher expression in tumor than normal samples [12]. We found 155 CRC ESPs overlap with various sets of pan-cancer ESPs, highlighting both the homogeneity and heterogeneity of different tumors. Hart et al. also found that essential genes across different tumor cell lines exhibit 52 % consistency, and knocking out these genes results in cellular dysfunction and cell death [4,46]. A novel CRC ESP, *CTNNB1*, has been reported to drive CRC evolution, with mutations occurring primarily in early-stage of CRC [23]. As expected, knockout of *CTNNB1* significantly inhibited CRC cell proliferation, indicating its role in CRC tumor cells survival. Additionally, we found that most pan-cancer ESPs and CRC ESPs are expressed not only in tumor epithelial cells but also immune cells and stromal cells, suggesting targeting essential genes may lead to low efficiency, side effect, and even therapeutic failure [45]. Notably, we identified four ESPs that can truly serve as drug targets with specificity for CRC epithelial cells, lacking expression in immune cells and stromal cells. Our results yield new insights into the biological characteristic of CRC ESPs.

Genome-wide CRISPR/Cas9 screen are not ideal for identifying lncRNAs [47]. Instead, co-expression networks can be can be used to predict essential lncRNAs [48]. We constructed specific co-expression networks using scRNA-seq data (Smart-seq) from CRC tumor epithelial cells, which provided low noise and redundancy compared to bulk transcriptomics. By leveraging this approach, we accurately predicted 29 CRC ESLs using HT and PPR algorithms, including 9 previously reported ESLs and 13 novel CRC-related lncRNAs [12]. Similar to ESPs, these ESLs were highly expressed in tumor epithelial cells and were involved in tumor-promoting and tumor-suppressing functions, as well as fundamental metabolic processes. Furthermore, we observed that the activity of ESLs also influenced by DNA mutations and methylations. For instance, *GAS5*, a known tumor suppressor [49,50], exhibited a high mutation rate in CRC patients from PCAWG and elevated expression in CRC. Three CpG sites of *PVT1* identified in our study were significantly hypomethylated in CRC and negatively correlated with *PVT1* expression. Targeting to methylation sites may serve as a potential strategy for regulating *PVT1* expression.

Essential genes involved in critical signaling and metabolic pathways can also receive extracellular signals and transmit to the TME via ligands/receptors within CCN. In the signal outgoing axis, essential genes may influenced by transcription factors, DNA mutations, and DNA methylation. These essential genes can impact the ligand/receptor expression on tumor cells, which in turn transmit signals to the TME. Through integrative network analysis, we identified six ligands/receptors, including *MIF* and *CD47*, potentially regulated by over 100 essential genes and TFs. Elevated expression levels of *MIF* and *CD47* in CRC is linked to poor prognosis, possibly due to their role in immune evasion [51,52]. We also identified drugs, such as sunitinib, that modulate MIF and CD47 expression by targeting their associated essential genes and TFs. Sunitinib has been shown to regulate CD47 expression in clear cell renal cell carcinoma [44]. This approach represents an efficient strategy to interfere essential genes expression and CCN ligands/receptors for cancer therapy.

Although this study provides a comprehensive analysis, there are some limitations. First, due to the low sequencing depth in 10X genomic

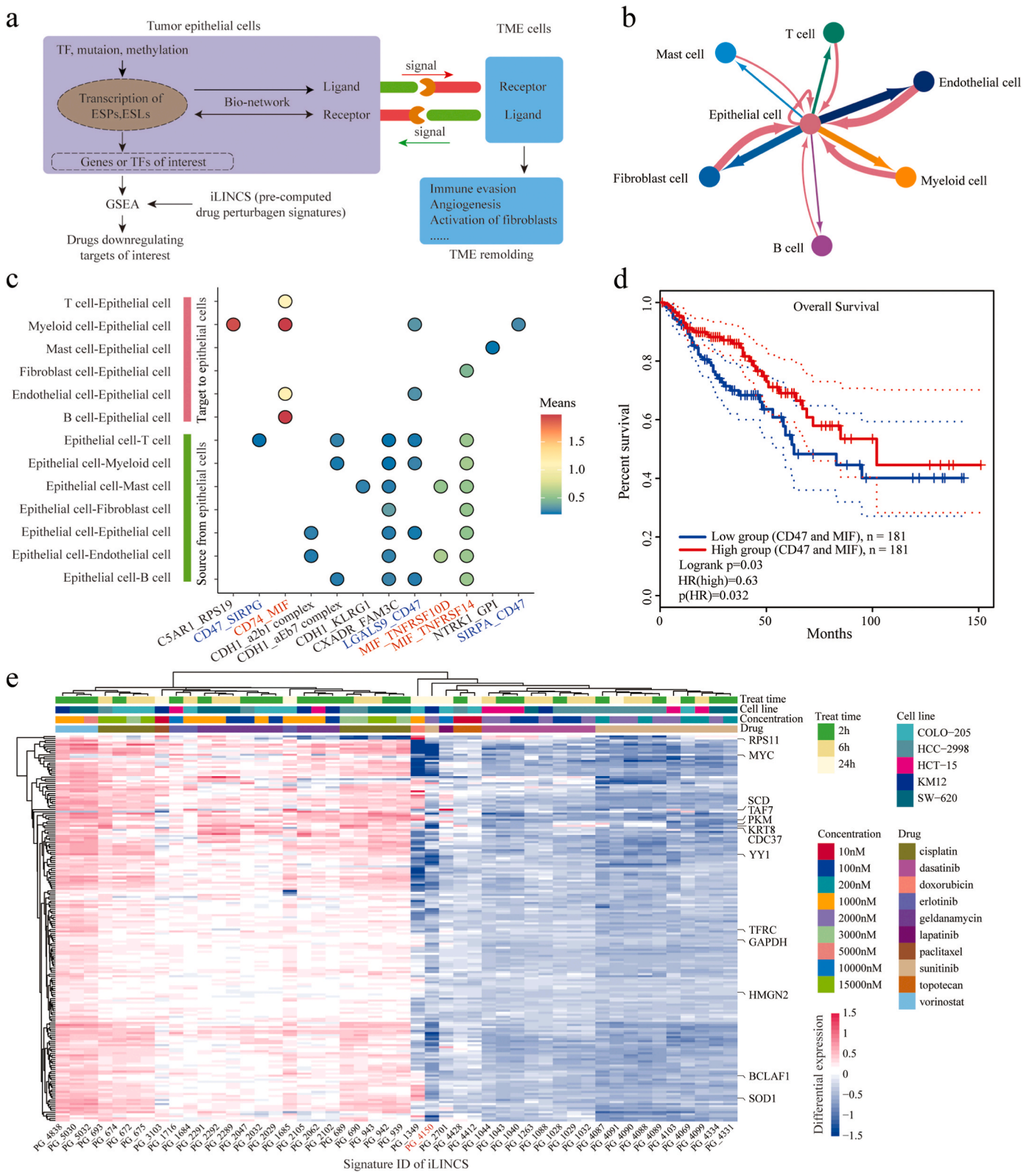

Fig. 5. CRC essential genes involved in TME remodeling through the regulation of ligands/receptors expression. **a.** The relationships between CRC essential genes and TME. The targets of interest can be used to identify candidate drugs through iLINCS analysis. **b.** The CCN associated with tumor epithelial cells. The arrow indicates the direction of signal transduction, and the line width represents the strength of interactions. **c.** An overview of selected ligands or receptors interactions between tumor epithelial cells and cells in TME. The values of means indicate the average expression levels of the interacting pairs. **d.** Overall survival rates of CRC patients with high or low expression levels of *CD47* and *MIF*. **e.** The expression pattern of target genes in 5 CRC cell lines after treatment with different drugs.

scRNA-seq, full-length scRNA-seq data of CRC is necessary for constructing a co-expression network. However, only 270 tumor cells were available, so further validation with larger datasets is required. Additionally, while we elucidated patterns of gene expression, DNA mutations, methylation, TF regulation, and TME remodeling for essential genes, more work is needed to confirm how essential genes perform irreplaceable functions at the protein level. Finally, further experiments are necessary to verify the key role of CRC essential genes in tumorigenesis.

4. Materials and methods

4.1. Genome-scale CRISPR data collection and processing

Eight genome-scale CRISPR loss-of-function screen datasets of CRC were downloaded from articles and the DepMap database (https://depmap.org/portal/). Then CERES algorithm, a computational method that estimates gene dependency levels from CRISPR-Cas9 essentiality screens while accounting for copy number-specific effect, was then applied. CERES provided the sgRNA depletion score for each gene across all cell lines. To integrate the gene list by CERES scores across all cell lines, we used the Robust Rank Aggregation algorithm (RRA) in R ("RobustRankAggreg" package), and adjusted the *p*-value using "Bonferroni" method. Once the integrated ranking list was obtained, we explored various proportions, ranging from 1 % to 23 % (adjusted *p*-value <0.05), to determine the optimized threshold to identify candidate ESPs. The optimized threshold was determined when the maximum value of the overlapping ratios between genes obtained from different ranking proportions and the union pan-cancer essential gene set of Hart's ESGs (1246 genes), Achilles' ESGs (2149 genes), and Zhang's ESGs (799 genes).

4.2. Single-cell RNA sequencing data processing

We obtained scRNA-seq data (smart-seq, GSE81861) from Gene Expression Omnibus database (GEO), consisting of 272 CRC tumor epithelial cells and 160 normal epithelial cells. The gene expression matrices of tumor and normal epithelial cells were combined and converted into a Seurat object using the Seurat R package for downstream analysis [53].

We also obtained 10x Genomics scRNA-seq data (GSE132465) from GEO, consisting of 56,465 cells from 23 patients with primary CRC samples and 10 matched normal mucosae samples. We performed quality filtering to remove cells with less than 500 expressed genes, less than 500 unique molecular identifiers (UMIs), more than 20 % UMIs derived from mitochondrial genes, and $\log_{10}$(expressed genes/UMIs) greater than 0.78. Then, we normalized and scaled the gene expression matrices using SCTransform. Subsequently, harmony algorithm was employed to integrate scRNA-seq data across different patients [54]. For cell type identification, we first applied a graph-based clustering approach by using "FindNeighbors" and "FindClusters" function. Then the clusters were further annotated into seven major cell types: epithelial cells (*EPCAM*), myeloid cells (*CD14*), T cells (*CD3D*), B cells (*CD79A*), fibroblast cells (*COL1A1*), endothelial cells (*PECAM1*), and mast cells (*TPSAB1*).

The differentially expressed essential genes in different group or cell lineages were calculated using the Seurat FindMarkers function with "MAST" algorithm. The genes with default parameters of $\log_2FC > 0.5$ and adjusted *p*-value <0.05 were considered as overexpressed genes. Specifically expressed genes in each cell type were defined as those detected in certain cell type at a percentage three times higher than other cell types. To assess the gene set score fore each cell type, "AddModuleScore" function in Seurat was applied.

4.3. Cell culture and CRISPR/Cas9 knockout

HCT116 (human colon carcinoma cell line, Ubigene, YC-C004) were cultured in RPMI-1640 (Thermo Fisher Scientific) with 10 % fetal bovine serum (FBS; VISTECH) and 1 % penicillin-streptomycin (Thermo Fisher Scientific). HEK293T cells (ATCC CRL3216) were cultured in DMEM (VISTECH) with 10 % FBS and 1 % penicillin-streptomycin. All cell lines were maintained at 37 °C in a humidified incubator with 5 % CO2.

To clone individual gene targeting single-guide RNAs (sgRNAs), the lentiviral vector (pSLQ1373) was digested with BlpI and BstXI, and gel-purified [55]. For each gene, two sgRNA were designed on the CHOP-CHOP website [56] (Supplementary Table 11). SgRNA fragments were synthesized by Tsingke (Tsingke Biotechnology Co., Ltd.) as forward and reverse primers (Table S1), which were then annealed, gel-purified, and ligated to the linearized pSLQ1373 vector by homologous recombination (ClonExpress Ultra One Step Cloning Kit, Vazyme).

To produce lentivirus, HEK293T cells were transiently transfected with polyethylenimine (PEI MAX, Polysciences) and packaging plasmids psPAX2 and pMD2.0G. The ratio of target plasmid to psPAX2 and pMD2.0G is 5:4:1. The ratio of the total mass of the added plasmid (μg) to the PEI MAX (μl) is 1:3. Lentivirus was collected by filtering the supernatant through a 0.45-μm filter 48 h after transfection. The lentiviruses were collected and stored at −80 °C.

HCT116 cells were treated with 15 μg/μl BSD for 5 days to select Cas9-expressing cells, after being infected with EFS-spCas9-BSD lentivirus for 48 h. The purified cells were sorted into 96-well plates using a flow cytometry sorter CytoFLEX SRT (BECKMAN COULTER). After 12–14 days, the well-growing clonal cell population in the 96-well plate was picked out, expanded, and frozen.

HCT116-Cas9 was infected with a lentivirus construct expressing individual sgRNA, or two sgRNAs simultaneously for lncRNA knock-out. After 72 h of infection, the infected cells were treated with 2 μg/ml puromycin for 24 h and then recovered in fresh culture medium without puromycin for 48 h.

Gene knock-out cells were counted and seeded to a 12-well plate at 0.1 million cells/well. After 72 h of growth, cells were uniformly passaged at a 1:10 and grown for another 72 h. Cells were then collected, washed, and resuspended in phosphate-buffered saline (PBS). Ten microliters of the cell suspension were mixed with 0.4 % trypan blue solution (Solarbio) at a 1:1 ratio and counted on a hemocytometer (Hirschmann). Each sample was counted three times to obtain a mean cell number.

4.4. Collection and processing of gene expression, DNA mutation, and DNA methylation data

We collected and combined the expression profiles, DNA mutation data, and DNA methylation data of TCGA-COAD and TCGA-READ from the UCSC data portal. The PCAWG mutation data of CRC was collected from International Cancer Genome Consortium (ICGC). The mutation data was processed, analyzed, and visualized using Maftools package in R. The DNA methylation data were processed and visualized by using ChAMP package in R. Differentially methylated positions (DMPs) were identified by "champ.DMP" function in ChAMP package with adjusted *p*-value <0.05 and absolute $\log_2FC > 0.2$. The Pearson correlation coefficients between the expression and DNA methylation levels of each gene were calculated to highlight the association of them.

4.5. Co-expression network construction

Co-expression network with both protein-coding genes and lncRNAs were constructed based on bulk RNA-seq from TCGA without normal samples and scRNA-seq data of tumor epithelial cells from GSE81861, respectively. For the construction of an unweighted co-expression network, the gene Pearson correlation coefficients were calculated. Fisher's asymptotic test from WGCNA packages [57] and FDR correction

were used to screen reliable relationships. Eventually, only the pairs with adjusted *p*-value lower than 0.01 and absolute value of correlation coefficient greater than 0.3 were reserved and further converted to undirected network. We also constructed topological overlap weighted co-expression network by using "TOMsimilarityFromExpr" function from WGCNA with parameter as follow: soft-thresholding power of 6, Pearson correlation coefficient, and TOMDenom specifying as mean. Finally, the pairs with weight greater than 0 were exported and further converted to undirected network.

4.6. Hypergeometrics test and Personalized PageRank

We applied the HT algorithm (Fig. 2a) to evaluate the degree of overlap between ESPs and neighboring genes of a lncRNA within a co-expression network [12]. For each lncRNA, *p*-value were calculated and adjusted using FDR correction, which were then used to evaluate the potential of the lncRNA as an ESL. Candidate ESLs were defined as lncRNAs with adjusted *p*-values lower than 0.01. The HT algorithm is as follows:

$$\boldsymbol{p}\text{-value} = \sum_{N_c}^{\min(N_p,N_n)} \frac{\binom{N_n}{N_c}\binom{N_T - N_n}{N_p - N_c}}{\binom{N_T}{N_p}}$$

N_T represents the total number of protein coding genes and lncRNAs obtained from co-expression network; N_p denotes the number of ESPs; N_n indicates the neighboring genes of a query lncRNA within the co-expression network; N_c represents the number of overlaping genes between N_n and N_p.

PPR is a well-established method derived from random walk with restart, which computes ranking score for each lncRNA using multiple seeds in the network. In this study, we used the Python implementation for random walk with restart (PyRWR, https://github.com/jinhongjung/pyrwr) to compute the ranking score of each lncRNA. We set the restart probability to 0.5, with ESPs designated as the seeds, and all other parameters were configured to the software's defaults. Only the top 5 % of lncRNAs in the PPR results were considered potential ESL candidates.

4.7. Inferring TF regulatory network of CRC scRNA-seq data

We utilized the R implementation of the SCENIC pipeline to infer TF regulatory network between transcription factors and essential genes. Two gene-motif rankings databases were employed to determine the *cis*-acting element around the transcription start site (TSS), including 10 kb around the TSS and 500 bp upstream of the TSS. The gene-motif annotations from cisTarget databases can be categorized as either high-confidence or low-confidence. The high-confidence annotations are "direct annotation" and "inferred by orthology" in the annotation source of gene-motif, while low-confidence annotations are "inferred by motif similarity". The regulon named with the sufix "-extended" indicate lower confidence annotation. To assess the differential AUC score of the regulon between tumor and normal epithelial cells, *t*-test was conducted.

4.8. CCN network construction and integration

We employed CellPhoneDB V3 to construct the CCN network of CRC based on scRNA-seq data (GSE132465). For downstream analysis, we only considered the ligand-receptor pairs with *p*-value lower than 0.05. We then integrating multiple intracellular regulatory networks, including co-expressed network, TF regulatory network, and CCN of tumor epithelial cells. First, we filtered the top 100 co-expressed pairs of each ESP and the pairs with weight great than 0.001 in WCN-TEC. Second, we filtered relationships associated with ESPs and ESLs from TF regulatory network. Finally, we integrated those two intracellular regulatory networks with CCN, the source node of integrated network are ESPs, ESLs, or TFs, while the target node are ligands or receptors of tumor epithelial cells.

4.9. Function enrichment and survival analysis

GO, KEGG, and HALLMARK function analyses of protein-coding genes were carried out by Metascape [58]. To determine the functions of lncRNAs, we first checked the existing reports from Lnc2Cancer 3.0. Second, top 200 associated genes for each ESL based on weighted co-expression network of tumor epithelial cells were selected. Then, GO biological process enrichment analysis was performed using "ClusterProfiler" R package. Additionally, we conducted survival analysis using GEPIA [59], an interactive web server to analyze RNA sequencing data from TGCA.

CRediT authorship contribution statement

Yanguo Li: Methodology, Formal analysis, Software, Data curation, Visualization, Writing – original draft. **Zixing Meng:** Investigation, Formal analysis, Visualization. **Chengjiang Fan:** Investigation. **Hao Rong:** Validation, Software. **Yang Xi:** Resources, Writing – review & editing, Funding acquisition. **Qi Liao:** Conceptualization, Project administration, Methodology, Supervision, Writing – review & editing, Funding acquisition.

Code availability

All scripts used for data analysis are available from GitHub (https://github.com/bioyanguo/CRC_ESP_ESL).

Declaration of AI and AI-assisted technologies in the writing process

ChatGPT was utilized during the writing process to enhance the readability and language of the manuscript.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgements

This work was supported by the Zhejiang Provincial Natural Science Foundation of China (No. LY21C060002, No. LY21C060001); the National Natural Science Foundation of China (No. 31970630); the Ningbo Natural Science Foundation of China (No. 2021J124); the Fundamental Research Funds for the Provincial Universities of Zhejiang (No. SJLZ2021001); the Keynote Research Project of Ningbo City (No. 2023Z171, No. 2023Z226); the Research Project of Zhejiang (No. Y202352483).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrep.2025.101938.

Data availability

The data that has been used is confidential.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality

worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 71 (3) (2021) 209–249.
[2] R.M. McQuade, V. Stojanovska, J.C. Bornstein, K. Nurgali, Colorectal cancer chemotherapy: the evolution of treatment and new approaches, Curr. Med. Chem. 24 (15) (2017) 1537–1557.
[3] T. Wang, K. Birsoy, N.W. Hughes, K.M. Krupczak, Y. Post, J.J. Wei, E.S. Lander, D. M. Sabatini, Identification and characterization of essential genes in the human genome, Science (New York, NY) 350 (6264) (2015) 1096–1101.
[4] O. Shalem, N.E. Sanjana, E. Hartenian, X. Shi, D.A. Scott, T. Mikkelson, D. Heckl, B. L. Ebert, D.E. Root, J.G. Doench, et al., Genome-scale CRISPR-Cas9 knockout screening in human cells, Science (New York, NY) 343 (6166) (2014) 84–87.
[5] C. Alda-Catalinas, D. Bredikhin, I. Hernando-Herraez, F. Santos, O. Kubinyecz, M. A. Eckersley-Maslin, O. Stegle, W. Reik, A single-cell transcriptomics CRISPR-activation screen identifies epigenetic regulators of the zygotic genome activation program, Cell Syst 11 (1) (2020) 25–41, e29.
[6] K.H. Lin, J.C. Rutter, A. Xie, B. Pardieu, E.T. Winn, R.D. Bello, A. Forget, R. Itzykson, Y.R. Ahn, Z. Dai, et al., Using antagonistic pleiotropy to design a chemotherapy-induced evolutionary trap to target drug resistance in cancer, Nat. Genet. 52 (4) (2020) 408–417.
[7] J.T. Cortez, E. Montauti, E. Shifrut, J. Gatchalian, Y. Zhang, O. Shaked, Y. Xu, T. L. Roth, D.R. Simeonov, Y. Zhang, et al., CRISPR screen in regulatory T cells reveals modulators of Foxp3, Nature 582 (7812) (2020) 416–420.
[8] R.T. Manguso, H.W. Pope, M.D. Zimmer, F.D. Brown, K.B. Yates, B.C. Miller, N. B. Collins, K. Bi, M.W. LaFleur, V.R. Juneja, et al., In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target, Nature 547 (7664) (2017) 413–418.
[9] T. Hart, K.R. Brown, F. Sircoulomb, R. Rottapel, J. Moffat, Measuring error rates in genomic perturbation screens: gold standards for human functional genomics, Mol. Syst. Biol. 10 (7) (2014) 733.
[10] T. Hart, A.H.Y. Tong, K. Chan, J. Van Leeuwen, A. Seetharaman, M. Aregger, M. Chandrashekhar, N. Hustedt, S. Seth, A. Noonan, et al., Evaluation and design of genome-wide CRISPR/SpCas9 knockout screens, G3 (Bethesda) 7 (8) (2017) 2719–2727.
[11] R.M. Meyers, J.G. Bryan, J.M. McFarland, B.A. Weir, A.E. Sizemore, H. Xu, N. V. Dharia, P.G. Montgomery, G.S. Cowley, S. Pantel, et al., Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells, Nat. Genet. 49 (12) (2017) 1779–1784.
[12] Y. Zhang, Y. Tao, H. Ji, W. Li, X. Guo, D.M. Ng, M. Haleem, Y. Xi, C. Dong, J. Zhao, et al., Genome-wide identification of the essential protein-coding genes and long non-coding RNAs for human pan-cancer, Bioinformatics 35 (21) (2019) 4344–4349.
[13] J.M. Dempster, J. Rossen, M. Kazachkova, J. Pan, G. Kugener, D.E. Root, A. Tsherniak, Extracting Biological Insights from the Project Achilles Genome-Scale CRISPR Screens in Cancer Cell Lines, 2019 720243.
[14] R. Oughtred, J. Rust, C. Chang, B.J. Breitkreutz, C. Stark, A. Willems, L. Boucher, G. Leung, N. Kolas, F. Zhang, et al., The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions, Protein Sci. 30 (1) (2021) 187–200.
[15] R. Marcotte, K.R. Brown, F. Suarez, A. Sayad, K. Karamboulas, P.M. Krzyzanowski, F. Sircoulomb, M. Medrano, Y. Fedyshyn, J.L.Y. Koh, et al., Essential gene profiles in breast, pancreatic, and ovarian cancer cells, Cancer Discov. 2 (2) (2012) 172–189.
[16] L. Zhang, Z. Li, K.M. Skrzypczynska, Q. Fang, W. Zhang, S.A. O'Brien, Y. He, L. Wang, Q. Zhang, A. Kim, et al., Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer, Cell 181 (2) (2020) 442–459, e429.
[17] H.O. Lee, Y. Hong, H.E. Etlioglu, Y.B. Cho, V. Pomella, B. Van den Bosch, J. Vanhecke, S. Verbandt, H. Hong, J.W. Min, et al., Lineage-dependent gene expression programs influence the immune landscape of colorectal cancer, Nat. Genet. 52 (6) (2020) 594–603.
[18] J.M. Replogle, T.M. Norman, A. Xu, J.A. Hussmann, J. Chen, J.Z. Cogan, E.J. Meer, J.M. Terry, D.P. Riordan, N. Srinivas, et al., Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing, Nat. Biotechnol. 38 (8) (2020) 954–961.
[19] R.M.J. Genga, E.M. Kernfeld, K.M. Parsi, T.J. Parsons, M.J. Ziller, R. Maehr, Single-cell RNA-sequencing-based CRISPRi screening resolves molecular drivers of early human endoderm development, Cell Rep. 27 (3) (2019) 708–718, e710.
[20] R. Vishnubalaji, N.M. Alajez, Single-cell transcriptome analysis revealed heterogeneity and identified novel therapeutic targets for breast cancer subtypes, Cells 12 (8) (2023).
[21] R. Kolde, S. Laur, P. Adler, J. Vilo, Robust rank aggregation for gene list integration and meta-analysis, Bioinformatics 28 (4) (2012) 573–580.
[22] H. Li, E.T. Courtois, D. Sengupta, Y. Tan, K.H. Chen, J.J.L. Goh, S.L. Kong, C. Chua, L.K. Hon, W.S. Tan, et al., Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors, Nat. Genet. 49 (5) (2017) 708–718.
[23] M. Gerstung, C. Jolly, I. Leshchiner, S.C. Dentro, S. Gonzalez, D. Rosebrock, T. J. Mitchell, Y. Rubanova, P. Anur, K. Yu, et al., The evolutionary history of 2,658 cancers, Nature 578 (7793) (2020) 122–128.
[24] J. Wang, F. Meng, F. Mao, Single cell sequencing analysis and transcriptome analysis constructed the liquid-liquid phase separation(LLPS)-related prognostic model for endometrial cancer, Front. Oncol. 12 (2022) 1005472.
[25] S. Chen, X. Shen, Long noncoding RNAs: functions and mechanisms in colon cancer, Mol. Cancer 19 (1) (2020) 167.
[26] Y. Gao, S. Shang, S. Guo, X. Li, H. Zhou, H. Liu, Y. Sun, J. Wang, P. Wang, H. Zhi, et al., Lnc2Cancer 3.0: an updated resource for experimentally supported lncRNA/circRNA cancer associations and web tools based on RNA-seq and scRNA-seq data, Nucleic Acids Res. 49 (D1) (2021) D1251–D1258.
[27] M.K.D. Scott, M. Limaye, S. Schaffert, R. West, M.G. Ozawa, P. Chu, V.S. Nair, A. C. Koong, P. Khatri, A multi-scale integrated analysis identifies KRT8 as a pan-cancer early biomarker, Pac Symp Biocomput 26 (2021) 297–308.
[28] S. Trop-Steinberg, Y. Azar, Is myc an important biomarker? Myc expression in immune disorders and cancer, Am. J. Med. Sci. 355 (1) (2018) 67–75.
[29] K. Vriens, S. Christen, S. Parik, D. Broekaert, K. Yoshinaga, A. Talebi, J. Dehairs, C. Escalona-Noguero, R. Schmieder, T. Cornfield, et al., Evidence for an alternative fatty acid desaturation pathway increasing cancer plasticity, Nature 566 (7744) (2019) 403–406.
[30] J. Krushkal, T. Silvers, W.C. Reinhold, D. Sonkin, S. Vural, J. Connelly, S. Varma, P. S. Meltzer, M. Kunkel, A. Rapisarda, et al., Epigenome-wide DNA methylation analysis of small cell lung cancer cell lines suggests potential chemotherapy targets, Clin. Epigenet. 12 (1) (2020) 93.
[31] Y.Y. Tseng, B.S. Moriarity, W. Gong, R. Akiyama, A. Tiwari, H. Kawakami, P. Ronning, B. Reuland, K. Guenther, T.C. Beadnell, et al., PVT1 dependence in cancer with MYC copy-number increase, Nature 512 (7512) (2014) 82–86.
[32] S. Aibar, C.B. Gonzalez-Blas, T. Moerman, V.A. Huynh-Thu, H. Imrichova, G. Hulselmans, F. Rambow, J.C. Marine, P. Geurts, J. Aerts, et al., SCENIC: single-cell regulatory network inference and clustering, Nat. Methods 14 (11) (2017) 1083–1086.
[33] X. Zhou, X. Li, Y. Cheng, W. Wu, Z. Xie, Q. Xi, J. Han, G. Wu, J. Fang, Y. Feng, BCLAF1 and its splicing regulator SRSF10 regulate the tumorigenic potential of colon cancer cells, Nat. Commun. 5 (2014) 4581.
[34] L. Zhu, X. Shi, X. Yu, Z. Wang, M. Zhang, Y. He, W. Guo, Expression and crucial role of long non-coding RNA FGD5-AS1 in human cancers, Am J Transl Res 13 (10) (2021) 10964–10976.
[35] D. Li, X. Jiang, X. Zhang, G. Cao, D. Wang, Z. Chen, Long noncoding RNA FGD5-AS1 promotes colorectal cancer cell proliferation, migration, and invasion through upregulating CDCA7 via sponging miR-302e, In Vitro Cell. Dev. Biol. Anim. 55 (8) (2019) 577–585.
[36] S.H. Kim, Y.Y. Park, S.N. Cho, O. Margalit, D. Wang, R.N. DuBois, Kruppel-like factor 12 promotes colorectal cancer growth through early growth response protein 1, PLoS One 11 (7) (2016) e0159899.
[37] X. Jiang, D. Li, G. Wang, J. Liu, X. Su, W. Yu, Y. Wang, C. Zhai, Y. Liu, Z. Zhao, Thapsigargin promotes colorectal cancer cell migration through upregulation of lncRNA MALAT1, Oncol. Rep. 43 (4) (2020) 1245–1255.
[38] M. Efremova, M. Vento-Tormo, S.A. Teichmann, R. Vento-Tormo, CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes, Nat. Protoc. 15 (4) (2020) 1484–1506.
[39] C.R. Figueiredo, R.A. Azevedo, S. Mousdell, P.T. Resende-Lara, L. Ireland, A. Santos, N. Girola, R. Cunha, M.C. Schmid, L. Polonelli, et al., Blockade of MIF-CD74 signalling on macrophages and dendritic cells restores the antitumour immune response against metastatic melanoma, Front. Immunol. 9 (2018) 1132.
[40] S. Kaur, J.S. Isenberg, D.D. Roberts, CD47 (cluster of differentiation 47), Atlas Genet Cytogenet Oncol Haematol 25 (2) (2021) 83–102.
[41] M. Pilarczyk, M. Fazel-Najafabadi, M. Kouril, B. Shamsaei, J. Vasiliauskas, W. Niu, N. Mahi, L. Zhang, N.A. Clark, Y. Ren, et al., Connecting omics signatures and revealing biological mechanisms with iLINCS, Nat. Commun. 13 (1) (2022) 4678.
[42] A. Monks, Y. Zhao, C. Hose, H. Hamed, J. Krushkal, J. Fang, D. Sonkin, A. Palmisano, E.C. Polley, L.K. Fogli, et al., The NCI transcriptional pharmacodynamics workbench: a tool to examine dynamic expression profiling of therapeutic response in the NCI-60 cell line panel, Cancer Res. 78 (24) (2018) 6807–6817.
[43] G.D. Demetri, A.T. van Oosterom, C.R. Garrett, M.E. Blackstein, M.H. Shah, J. Verweij, G. McArthur, I.R. Judson, M.C. Heinrich, J.A. Morgan, et al., Efficacy and safety of sunitinib in patients with advanced gastrointestinal stromal tumour after failure of imatinib: a randomised controlled trial, Lancet 368 (9544) (2006) 1329–1338.
[44] W. Jiang, H. Zeng, Z. Liu, K. Jin, B. Hu, Y. Chang, L. Liu, Y. Zhu, L. Xu, Z. Wang, et al., Immune inactivation by CD47 expression predicts clinical outcomes and therapeutic responses in clear cell renal cell carcinoma patients, Urol. Oncol. 40 (4) (2022), 166 e115-166 e125.
[45] L. Chang, P. Ruiz, T. Ito, W.R. Sellers, Targeting pan-essential genes in cancer: challenges and opportunities, Cancer Cell 39 (4) (2021) 466–479.
[46] T. Hart, M. Chandrashekhar, M. Aregger, Z. Steinhart, K.R. Brown, G. MacLeod, M. Mis, M. Zimmermann, A. Fradet-Turcotte, S. Sun, et al., High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities, Cell 163 (6) (2015) 1515–1526.
[47] S.J. Liu, M.A. Horlbeck, S.W. Cho, H.S. Birk, M. Malatesta, D. He, F.J. Attenello, J. E. Villalta, M.Y. Cho, Y. Chen, et al., CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells, Science (New York, NY) 355 (6320) (2017).
[48] X. Li, W. Li, M. Zeng, R. Zheng, M. Li, Network-based methods for predicting essential genes or proteins: a survey, Briefings Bioinf. 21 (2) (2020) 566–583.
[49] L. Liu, H.J. Wang, T. Meng, C. Lei, X.H. Yang, Q.S. Wang, B. Jin, J.F. Zhu, lncRNA GAS5 inhibits cell migration and invasion and promotes autophagy by targeting miR-222-3p via the GAS5/PTEN-signaling pathway in CRC, Mol. Ther. Nucleic Acids 17 (2019) 644–656.
[50] W. Ni, S. Yao, Y. Zhou, Y. Liu, P. Huang, A. Zhou, J. Liu, L. Che, J. Li, Long noncoding RNA GAS5 inhibits progression of colorectal cancer by interacting with and triggering YAP phosphorylation and degradation and is negatively regulated by the m(6)A reader YTHDF3, Mol. Cancer 18 (1) (2019) 143.
[51] M.E.W. Logtenberg, F.A. Scheeren, T.N. Schumacher, The CD47-SIRPalpha immune checkpoint, Immunity 52 (5) (2020) 742–752.

[52] B. Otvos, D.J. Silver, E.E. Mulkearns-Hubert, A.G. Alvarado, S.M. Turaga, M. D. Sorensen, P. Rayman, W.A. Flavahan, J.S. Hale, K. Stoltz, et al., Cancer stem cell-secreted macrophage migration inhibitory factor stimulates myeloid derived suppressor cell function and facilitates glioblastoma immune evasion, Stem Cell. 34 (8) (2016) 2026–2039.
[53] T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W.M. Mauck 3rd, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive integration of single-cell data, Cell 177 (7) (2019) 1888–1902, e1821.
[54] I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P.R. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony, Nat. Methods 16 (12) (2019) 1289–1296.
[55] Y. Liu, C. Yu, T.P. Daley, F. Wang, W.S. Cao, S. Bhate, X. Lin, C. Still 2nd, H. Liu, D. Zhao, et al., CRISPR activation screens systematically identify factors that drive neuronal fate and reprogramming, Cell Stem Cell 23 (5) (2018) 758–771, e758.
[56] K. Labun, T.G. Montague, M. Krause, Y.N. Torres Cleuren, H. Tjeldnes, E. Valen, CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing, Nucleic Acids Res. 47 (W1) (2019) W171–W174.
[57] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008) 559.
[58] Y. Zhou, B. Zhou, L. Pache, M. Chang, A.H. Khodabakhshi, O. Tanaseichuk, C. Benner, S.K. Chanda, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets, Nat. Commun. 10 (1) (2019) 1523.
[59] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, Z. Zhang, GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses, Nucleic Acids Res. 45 (W1) (2017) W98–W102.