

PRIMARY RESEARCH

Open Access



# Network analysis of *KLF5* targets showing the potential oncogenic role of *SNHG12* in colorectal cancer

Qi Liao<sup>1†</sup> , Linbo Chen<sup>4†</sup>, Ning Zhang<sup>6</sup>, Yang Xi<sup>3</sup>, Shiyun Hu<sup>1</sup>, Derry Minyao Ng<sup>1</sup>, Fatma Yislam Hadi Ahmed<sup>1</sup>, Guofang Zhao<sup>5</sup>, Xiaoxiang Fan<sup>5</sup>, Yangyang Xie<sup>5</sup>, Xiaoyu Dai<sup>5</sup>, Yanping Jin<sup>2</sup>, Jiaxin Ge<sup>2</sup>, Changzheng Dong<sup>1</sup>, Xinjun Zhang<sup>2\*</sup> and Junming Guo<sup>3\*</sup>

## Abstract

**Background:** *KLF5* is a member of the Kruppel-like factor, subfamily of zinc finger proteins that are involved in cancers. *KLF5* functions as a transcription factor and regulates the diverse protein-coding genes (PCGs) in colorectal cancer (CRC). However, the long non-coding RNAs (lncRNAs) regulated by *KLF5* in CRC are currently unknown.

**Methods:** In this study, we first designed a computational pipeline to determine the PCG and lncRNA targets of *KLF5* in CRC. Then we analyzed the motif pattern of the binding regions for the lncRNA targets. The regulatory co-factors of *KLF5* were then searched for through bioinformatics analysis. We also constructed a regulatory network for *KLF5* and annotated its functions. Finally, one of the *KLF5* lncRNA targets, *SNHG12*, was selected to further explore its expression pattern and functions in CRC.

**Results:** We were able to identify 19 lncRNA targets of *KLF5* and found that the motifs of the lncRNA binding sites were GC-enriched. Next, we pinpointed the transcription factors *AR* and *HSF1* as the regulatory co-factors of *KLF5* through bioinformatics analysis. Then, through the analysis of the regulatory network, we found that *KLF5* may be involved in DNA replication, DNA repair, and the cell cycle. Furthermore, in the cell cycle module, the *SNHG12* up-regulating expression pattern was verified in the CRC cell lines and tissues, associating it to CRC invasion and distal metastasis. This indicates that *SNHG12* may play a critical part in CRC tumorigenesis and progression. Additionally, expression of *SNHG12* was found to be down-regulated in CRC cell lines when *KLF5* expression was knocked-down by siRNA; and a strong correlation was observed between the expression levels of *SNHG12* and *KLF5*, further alluding to their regulatory relationship.

**Conclusions:** In conclusion, the network analysis of *KLF5* targets indicates that *SNHG12* may be a significant lncRNA in CRC.

**Keywords:** Long non-coding RNA (lncRNA), *KLF5*, *SNHG12*, Colorectal cancer, Bioinformatics

## Background

Colorectal cancer (CRC) is one of the most common cancers in digestive system. CRC is the third most frequent malignancy in males and the second most frequent in females worldwide [1]. Approximately 1.4 million new cases of CRC and 700,000 CRC-related deaths are reported worldwide each year [1], suggesting that it is a major public health problem. Unfortunately, treatments

\*Correspondence: zxj14@163.com; guojunming@nbu.edu.cn

†Qi Liao and Linbo Chen contributed equally to this work

<sup>2</sup>The Affiliated Hospital of School of Medicine, Ningbo University, Ningbo 315020, China

<sup>3</sup>Department of Biochemistry and Molecular Biology, Zhejiang Provincial Key Laboratory of Pathophysiology, Ningbo University School of Medicine, Ningbo 315211, Zhejiang, China

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



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

such as aggressive surgical management and chemotherapy have a limited overall impact on cure rates and long-term survival rates, partially due to the complicated pathogenesis and underlying regulatory mechanisms of CRC [2].

Kruppel-like factor 5 (*KLF5*) is one of the 17 known members of the Kruppel-like factor subfamily of zinc finger proteins [3]. Increasing evidences have revealed that *KLF5* is involved in various cancers, such as prostate cancer [4], non-small-cell lung cancer [5], and bladder cancer [6], as well as other diseases [7, 8]. Interestingly, *KLF5* is critical for intestinal development and homeostasis. Furthermore, it has been detected to have a positive role in intestinal tumorigenesis [9]. *KLF5* can also activate beta-catenin with lysophosphatidic acid to stimulate the proliferation of colon cancer cells [10]. *KLF5* also participates in the cell cycle, inducing the expression of several cell cycle-related genes including cyclin D1 and cyclin B [11, 12]. Owing to the important role of *KLF5* in CRC, a number of studies have aimed to identify novel small molecule compounds that can inhibit the function of *KLF5*, yielding potential therapeutic targets for the treatment of CRC [13, 14]. For example, ML264, a novel small molecule compound, was able to inhibit the proliferation of CRC in vitro by suppressing the expression of *KLF5* [14].

Long non-coding RNAs (lncRNAs) are transcripts that consist of more than 200 nucleotides and have limited to no protein-coding capacity [15]. lncRNAs are considered to be a relatively new type of regulatory RNAs. Emerging evidences have shown that lncRNAs have prominent roles in human diseases, owing to its participation in a large range of biological processes [16, 17] as well as in cancer development [18, 19]. Some lncRNAs are up-regulated in cancers and act as oncogenes, such as *LINC00941* in gastric cancer [20], HOX transcript antisense RNA (*HOTAIR*) in hepatocellular carcinoma and CRC [21, 22], and *H19* in bladder and prostate cancer [23, 24]. Other lncRNAs, including Growth Arrest Specific 5 (*GAS5*), Maternally Expressed Gene 3 (*MEG3*), and *LOC285194* have been reported to function as tumor suppressors in cancers [25–27]. *KLF5* has been shown to be regulated by *CASC15* and *PVT1* in breast cancer [28, 29] along with *MALAT1* in pulmonary artery hypertension [30]. However, the role of lncRNAs in CRC remains unknown.

As previous reports, we developed several bioinformatics tools [31, 32] and constructed a comprehensive regulatory network of lncRNAs in CRC [2]. Acknowledging the critical role of *KLF5* and lncRNAs, we hypothesize that *KLF5* may regulate several specific pivotal lncRNAs in CRC [33]. In this research, we initially designed a computational pipeline to investigate the lncRNA and

protein-coding gene (PCG) targets of *KLF5* in CRC. Next, we analyzed the DNA motif pattern of the binding regions on the lncRNA targets and predicted for the regulatory co-factors associated with *KLF5*. Furthermore, we established a *KLF5* regulatory network based on co-expression relationships and protein–protein interactions (PPIs) to identify the functional modules related to *KLF5*. Finally, one of the lncRNA targets, termed *SNHG12*, which has been shown to be increased in osteosarcoma tissues in contrast with normal tissue [34], was selected to explore its expression patterns in CRC tumors and cell lines, potential functions, and association with clinical–pathological factors of patients with CRC. In addition, expression of *SNHG12* was observed after knocking-down expression of *KLF5* in CRC cell line with siRNA and the relationship between the expression levels of *SNHG12* and *KLF5* in CRC tissues or adjacent non-tumorous tissues was reviewed to verify their regulatory relationship.

## Materials and methods

### Identification of CRC associated lncRNAs and PCGs

The expression datasets for lncRNAs and PCGs in primary cancer, normal colon tissue, and metastasized CRC were collected from the GEO database (GSE50760) [35]. Differential expression analysis was performed between the normal colon and primary cancer, as well as normal colon and metastasized cancer, respectively, using a *t*-test with False Discovery Rate (FDR)-adjusted *p*-value less than 0.05. To obtain more accurate results, the differentially expressed genes were obtained based on the intersection with the Cancer RNA sequencing (RNA-Seq) Nexus database [36], which includes expression profiles from The Cancer Genome Atlas (TCGA).

### The potential PCG and lncRNA targets of *KLF5* in CRC

First, using the Cistrome database (<https://cistrome.org/>) [37], we collected the target genes identified from the chromatin immunoprecipitation DNA sequencing (ChIP-Seq) datasets of *KLF5* in two colon adenocarcinoma cell lines: the GP5d cell line with transfecting and non-transfecting RAD21 short interfering RNA (GSM1240834, GSM1240820) [38], as well as the LoVo cell line, which was blocked by double thymidine to cause cell cycle arrest in the early S phase, was cultured in a medium containing nocodazole in M-phase synchronization (GSM1208642, GSM1242268, GSM1242274) [38]. Overall, five samples of *KLF5* ChIP-Seq datasets were obtained, plus PCGs and lncRNAs with a *KLF5* binding site in at least two samples were selected for the first step. Second, the differently expressed lncRNAs and PCGs were obtained from the intersection results of the RNA-Seq dataset (GSE50760) and Cancer RNA-Seq Nexus database. Third, the potential

lncRNA and PCG targets of *KLF5* were picked up from the intersection of the results obtained in the first and second steps. Finally, the lncRNA targets of *KLF5* with higher confidence were further selected based on the lncRNA PCG co-expression relationships.

### Motif analysis of *KLF5* binding sites

We downloaded the *KLF5* binding peaks of ChIP-seq datasets from the Cistrome database [37], and then mapped them to the genome regions of lncRNA and PCG targets in GENCODE v25 database [39] with 2 kb extended at the transcription start site. Furthermore, MEME software was used to analyze the motif pattern of *KLF5* binding sites [40, 41].

### Prediction of *KLF5* regulatory co-factors in CRC

We first downloaded the ChIP-Seq profiles of all other transcription factors (TFs) from the Cistrome database [37] and obtained the target genes of the TFs detected from each ChIP-Seq profile. Then we proposed a two-step enrichment analysis method to identify the TFs that may co-regulate with *KLF5* (Additional file 1: Figure S1).

In the first step, we calculated the significance of enrichment by overlapping, between the *KLF5*'s targets and the target genes of each ChIP-Seq profile collected, using a hypergeometric test. Then the ChIP-Seq profiles with *p*-values less than 0.01 ranked in the top 50 were selected in a descending order, denoted as "sigChIP-Seq-profiles". The formula of the hypergeometric test was as follows:

$$p\text{-value} = \sum_{N_s}^{\min(N_c, N_k)} \frac{\binom{N_k}{N_s} \binom{N_T - N_k}{N_c - N_s}}{\binom{N_T}{N_c}},$$

where  $N_T$  represents the total number of all genes (lncRNAs or PCGs) obtained from the GENCODE database [42];  $N_c$  represents the number of other TFs' targets from each ChIP-Seq profile;  $N_k$  represents the number of *KLF5*'s targets; and  $N_s$  represents the number of common targets between  $N_k$  and  $N_c$ .

In the second step, for each TF involved in the dataset of "sigChIP-Seq-profiles" constructed above, we compared the number of ChIP-Seq profiles for this TF in this dataset and the pool of all ChIP-Seq profiles, using the following hypergeometric test:

$$p\text{-value} = \sum_{N_s}^{\min(N_e, N_{tf})} \frac{\binom{N_{tf}}{N_s} \binom{N_T - N_{tf}}{N_e - N_s}}{\binom{N_T}{N_e}},$$

where  $N_T$  represents the total number of ChIP-Seq profiles downloaded from the Cistrome database [37];  $N_e$

represents the number of "sigChIP-Seq-profiles" selected in the first step (that is, 50 or less);  $N_{tf}$  represents the number of ChIP-Seq profiles corresponding to the same TF in the pool of all ChIP-Seq profiles;  $N_s$  represents the number of significant ChIP-Seq profiles corresponding to the same TF in "sigChIP-Seq-profiles". Finally, TFs with *p*-values less than 0.01 and the number of significant ChIP-Seq profiles not less than 3 were determined to be the co-regulatory TFs of *KLF5*.

### *KLF5* regulation network construction and module analysis

We constructed a *KLF5* regulatory network by combining the targets of *KLF5*, the co-expression relationships in CRC among *KLF5* with its lncRNA and PCG targets, as well as the PPIs among *KLF5* and protein-coding targets. Next, we downloaded three datasets from the Co-LncRNA database, obtaining the co-expression relationships in CRC based on Spearman correlation [43], the final co-expression links consisted of relationships observed in at least two datasets along with the PPIs from the HPRD database [44]. Additionally, we used the MCL algorithm with default parameters to identify the modules [45]. Then, we obtained calculations of the enriched gene ontology biological process using the hypergeometric test with FDR-adjusted *p*-values less than 0.05.

### Function prediction of *SNHG12* based on co-expression network

The co-expressed PCG partners of *SNHG12* were obtained in the same way as stated above. We organized a lncRNA PCG co-expression network for CRC based on three CRC datasets obtained from the Co-LncRNA database [43], requiring that the co-expressed links should be observed in at least two datasets.

### Patients and tissue samples

We obtained CRC tissues and their paired adjacent-normal tissues between the years 2012 to 2015 from surgical patients at the Affiliated Hospital of Ningbo University School of Medicine and the First Affiliated Hospital of Sun Yat-sen University, China. All tissue specimens were immediately placed in an RNA-fixer Reagent (Bioteke, Beijing, China) after surgical resections and then stored at  $-80^{\circ}\text{C}$  until use. The diagnoses of all 111 CRC patients were based on histopathological evaluation using the clinical staging performed according to the 7th edition of the American Joint Committee on Cancer (AJCC) Cancer Staging Manual. It is important to mention that no patients received any treatment before surgery and all patients were anonymous with written informed consent and that all aspects of this study were approved by the Human Research Ethics Committee of Ningbo University

with the methods performed in accordance with the relevant guidelines and regulations.

#### RNA extraction, reverse transcription, and qRT-PCR detection

According to the manufacturer's protocol, TRIzol reagent (Ambion, Carlsbad, CA, USA) was used to extract total RNA from tissues and cultured cells. We reverse-transcribed 2 µg total RNA into cDNA using the PromegaGo-Script Reverse Transcription System (Promega, Madison, WI, USA). The *SNHG12* expression level was normalized to  $\beta$ -actin. Furthermore, qRT-PCR was performed using GoTaq qPCR master mix (Promega) on the Mx3005P QPCR System (Stratagene, La Jolla, CA, USA), the 25 µl PCR reaction mix contained 12.5 µl GoTaq qPCR Master Mix, 5 µl cDNA product, 1 µl forward primer, 1 µl reverse primer, and 5.5 µl nuclease-free water. Overall, a total of 45 cycles of amplification were performed after 10 min of preheating. Each cycle consisting of 95 °C for 15 s, 60 °C for 30 s, and 72 °C for 30 s. The primer sequences were as follow: 5'-TCTGGTGATCGAGGACTTCC-3' (forward) and 5'-ACCTCCTCAGTATCACACACT-3' (reverse) for *SNHG12*, 5'-CCTGGTCCAGACAAGATG TGA-3' (forward) and 5'-GAACTGGTCTACGACTGA GGC-3' (reverse) for *KLF5* and 5'-CTCCTTAATGTC ACGCACGAT-3' for  $\beta$ -actin. The expression levels of *SNHG12* were calculated using the  $\Delta C_t$  method and relative expression levels of *KLF5* and *SHNG12* were calculated using the  $2^{-\Delta\Delta C_t}$  method [46]. Each experiment was performed in triplicate; and all results were expressed as the means  $\pm$  SD.

#### Cell culture

The human colon epithelial cell line (NCM460) and the four CRC cell lines (HCT116, HT29, SW620 and COLO-205) were purchased from the Chinese Academy of Sciences Cell Bank (Shanghai, China). The cells were maintained by being supplemented with 10% fetal bovine serum (FBS) and grown in humidified air containing 5% CO<sub>2</sub> at 37 °C in RPMI 1640 or McCoy's 5A (Invitrogen, USA) medium.

#### siRNA and transfection

The Lipofectamine<sup>®</sup> RNAiMAX Transfection Reagent (Invitrogen, Germany) was used for the knock-down of *KLF5* gene following the manufacturer's instructions. The sequence of the *KLF5* siRNA (siKLF5) was: 5'-AAAGTA TAGACGAGACAGTGC-3', which was collected from a previous study [47]. Besides, the negative control siRNA (siNC) was: 5'-CUUACGCUGAGUACUUCGATT-3'. The siRNAs were synthesized by Genechem company (Shanghai, China). The siRNA knock-down experiment was performed for 72 h, then proteins were extracted.

#### Western blot analysis

Total proteins were extracted and separated by SDS-PAGE. Primary antibodies against *KLF5* (#DF7135, Affinity, Cincinnati, OH, USA) and *GAPDH* (#AP0063, Bioworld Antibodies, Bloomington, MN, USA) were used.

#### Statistical analysis

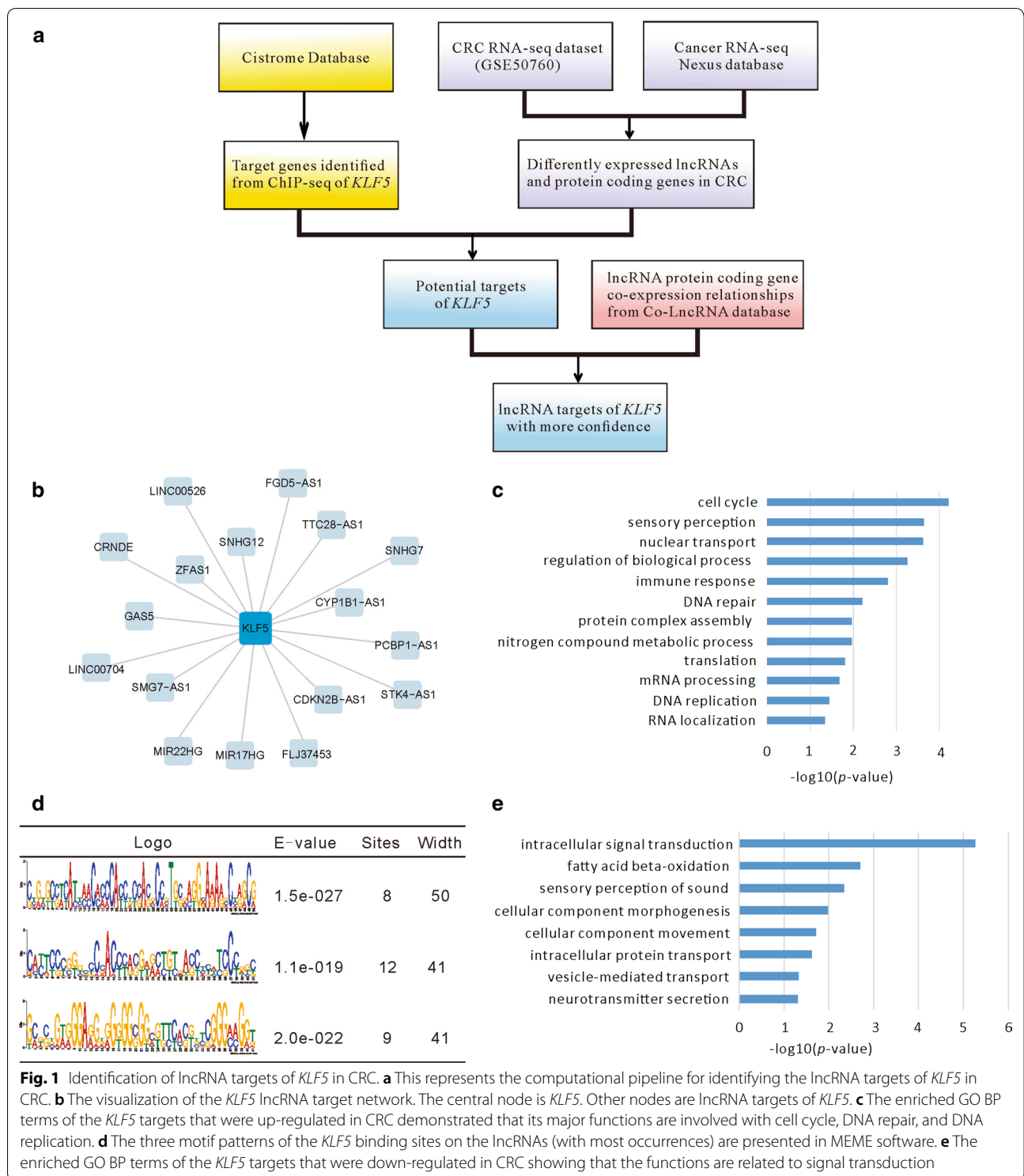
All statistical analyses were performed using SPSS version 13.0 (SPSS, Chicago, IL, USA), R and GraphPad Prism 6.0 (GraphPad Software, La Jolla, CA, USA). The expression differences between the CRC tissues and adjacent non-cancer tissues were evaluated using paired sample *t*-tests. We also used the rank-sum test to evaluate the expression between the CRC cell lines and NCM460. Furthermore, independent sample *t*-tests and One-way Analysis of Variance (ANOVA) test was assisted in analyzing the correlation between the expression levels and clinical-pathological characteristics of CRC patients and the comparison of expression between siNC and siKLF5 group. To further illustrate the diagnostic values, we used the Receiver Operating Characteristic (ROC) curve for assessment. A two-sided *p*-value < 0.05 was considered to indicate statistical significance.

## Results

#### Identification of PCG and lncRNA targets of *KLF5* in CRC

We designed a computational pipeline to obtain the PCG and lncRNA targets of *KLF5* in CRC (Fig. 1a). First, we collected the differently expressed PCGs and lncRNAs based on the intersection between the CRC RNA-Seq dataset GSE50760 [35] and the cancer RNA-Seq Nexus database [36]. Among these differentially expressed genes, 71 lncRNAs and 3,464 PCGs were down-regulated in CRC, while 73 lncRNAs and 3786 PCGs were up-regulated. Some of the lncRNAs, such as *UCA1*, *H19*, *GASS*, and *PVT1*, have previously been reported to be associated with CRC [48–51]. Second, we required that the lncRNA targets of *KLF5* should have the *KLF5* binding sites at the corresponding promoter regions (from – 2000 to 200 bp of TSS), based on ChIP-Seq datasets for *KLF5* [38]. As a result, 19 lncRNA and 1744 PCG targets were identified, among which 10 lncRNAs and 956 PCGs were up-regulated while 9 lncRNAs and 788 PCGs were down-regulated in CRC, respectively (Fig. 1b, Additional file 2: Table S1 and Additional file 3: Table S2).

Function enrichment analysis established that the up-regulated targets of *KLF5* in CRC were correlated with cell cycle, immune response, DNA repair, and DNA replication, while the down-regulated targets were linked to intracellular signal transduction, fatty acid beta-oxidation, and sensory perception of sound (Fig. 1c, e). Several lncRNAs, such as *GASS*, have shown to be involved in



CRC development and progression, reflecting the accuracy of our analysis. *KLF5* may also be able to regulate key lncRNAs in CRC. Specifically, 11 lncRNAs, including *SNHG12*, were co-expressed with *KLF5* in CRC. This

finding enhanced our confidence that there are probable lncRNA targets from our established coding non-coding gene co-expression network [32] (Additional file 2: Table S1).

### Motif analysis of *KLF5* binding sites on lncRNA targets

*KLF5* is considered to be a GC-box binding factor [52]. According to the transcription assays in vitro, *KLF5* level of enhanced transcription depends on the presence of GC boxes in the promoter region of the target gene [53]. However, this conclusion was based on PCGs. We further investigated the binding patterns of *KLF5* on lncRNA targets by analyzing the ChIP-Seq datasets from the Cistrome database and identified 62 and 4180 peaks for lncRNA and PCG targets of *KLF5*, respectively [37]. Although there were far fewer peaks for lncRNAs than for PCGs, the lengths (mean = 258 bp and 275 bp, respectively,  $p > 0.05$ ) and the signal levels (mean = 12.9 and 13.7, respectively,  $p > 0.05$ ) were similar in both cases. Furthermore, MEME software was used to demonstrate the motif patterns of the peaks for lncRNAs, resulting in the detection of GC sequence-enriched motifs, suggesting the presence of active transcriptional regulation [54] (Fig. 1d). In a previous study, *KLF5* was also found to regulate target genes with a GC-rich motif in a sequence-specific manner in lymphoid cells and tissues [55], indicating that the transcription regulation mechanism of *KLF5* to lncRNAs is identical to that of PCGs.

### Identification of regulatory co-factors for *KLF5*

Some TFs have been shown to co-operate with other TFs to regulate the expression of downstream genes. In a bioinformatics analysis, *KLF5* was discovered to interact with several TFs including *CREB1*, *Sp1*, *MYC*, *ER*, and *AR* to regulate gene expression [56]. To illustrate the regulatory mechanism of *KLF5* in CRC, we obtained the target genes of other TFs based on 5500 ChIP-Seq profiles from the Cistrome database [37] and proposed a two-step enrichment analysis method to identify candidate TFs that may co-operate with *KLF5* (Additional file 1: Figure S1). First, we analyzed the enrichment values of the overlapping genes between the lncRNA or PCG targets of *KLF5* and each target set detected from 5500 ChIP-Seq profiles of other TFs in diverse cell types, using a hypergeometric test. Second, according to the results of the first step, we can select the significant ChIP-Seq profiles with a  $p$ -value less than 0.01 and ranked the top 50 in descending order, denoted as “sigChIP-Seq-profiles”.

Third, for each TF, we calculated the significance of the proportion of this TF involved in the “sigChIP-Seq-profiles” compared with the background, using the same hypergeometric test. Finally, the TFs with a number of significant ChIP-Seq profiles not less than 3 and a  $p$ -value less than 0.01 were determined to be the regulatory co-factors.

Consequently, we found that the *AR* was a co-regulatory TF of *KLF5* with respect to the lncRNA targets (Fig. 2a). In an initial study, *KLF5* was revealed to be a target of *AR* and to promote cancer progression in prostate cancer [57]. *AR* was also predicted to be a co-factor of *KLF5* in a bioinformatics analysis [56]. These results suggest the presence of a relationship between *AR* and *KLF5*. Additionally, *HSF1* was established to be a co-regulator TF of *KLF5* for the PCG target (Fig. 2b).

Thus, we explored the common targets of *KLF5* and its co-regulatory TFs. For *AR*, the same lncRNA targets are *TTC28-AS1*, *CYP11B1-AS1*, *ZFAS1*, *MIR17HG*, *CDKN2B-AS1*, and *SNHG7* (Fig. 2c), whereas the common PCG targets of *HSF1* and *KLF5* included *HSPE1*, *HSP90AB1*, *STIP1*, *CACYBP*, *FXR1*, and *CKS1B* (Fig. 2d). A visualization of the network for *KLF5* and its co-regulatory TFs is presented in Fig. 2e. Some of these targets have been confirmed to be associated with CRC such as *MIR17HG* [58] and *SNHG7* [59]. *KLF5* may co-operate or form a transcription complex with *AR* or *HSF1* to promote tumorigenesis in CRC.

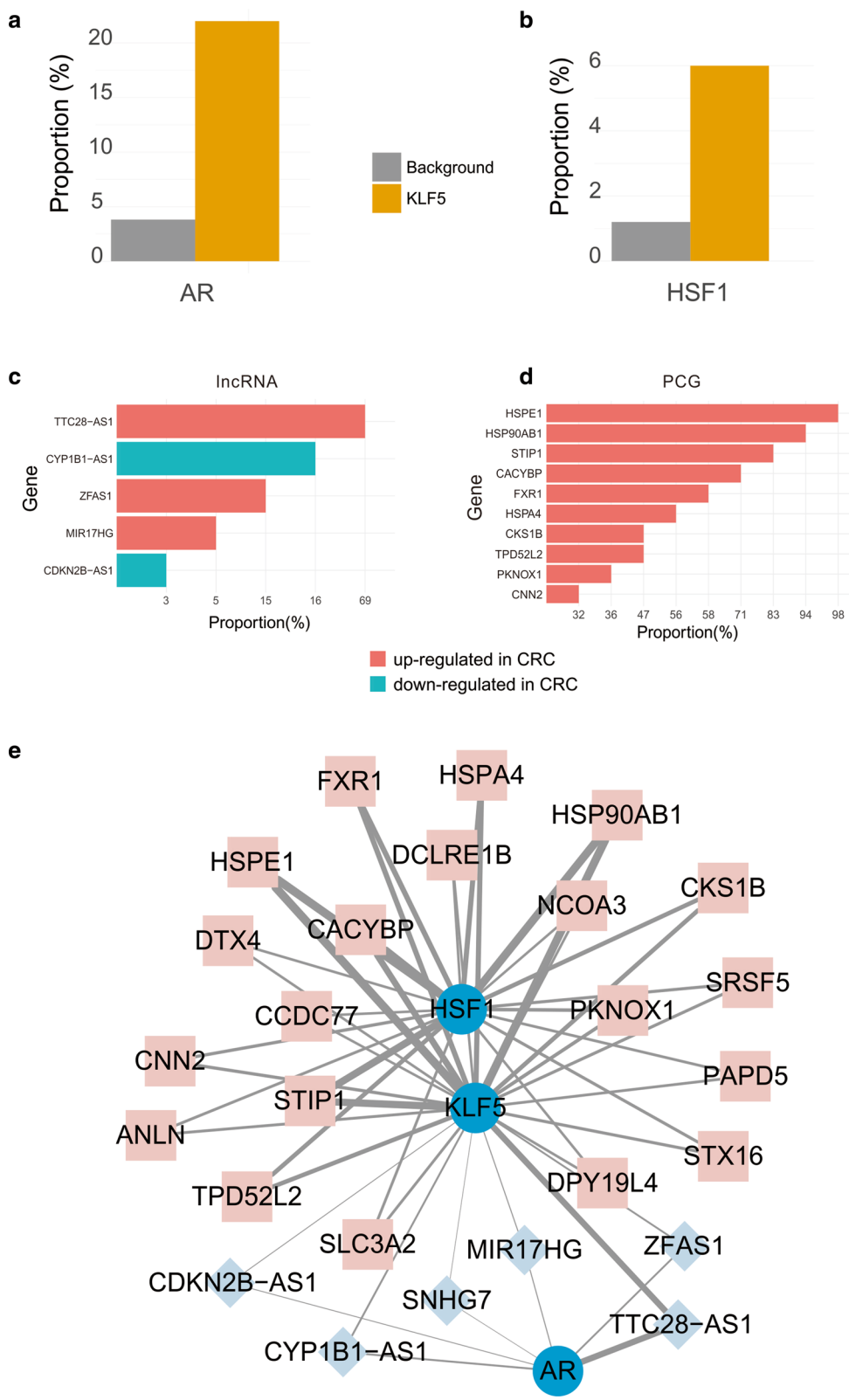
### Module analysis of *KLF5* regulatory network

The *KLF5* regulatory network was constructed based on the relationships between targets together with *KLF5*, PPIs among PCGs, and lncRNA-PCG co-expression. Finally, 14,823 lncRNA–protein co-expression relationships and 934 PPIs were obtained, with eight up-regulated and seven down-regulated lncRNAs, plus 927 up-regulated and 726 down-regulated PCGs (Fig. 3a). Among the genes in the network, *EP300* and *RARA* have previously been shown to be associated with *KLF5* in CRC [60, 61].

Next, we performed module analysis using the MCL algorithm with its default parameters [45], resulting in 15 modules with at least 25 genes. It is also important to

(See figure on next page.)

**Fig. 2** Regulatory co-factors of *KLF5* in CRC. **a, b** The orange bar represents the proportion of the ChIP-Seq profiles of *AR* (**a**) and *HSF1* (**b**) whose targets were enriched with *KLF5*'s, while the grey bar represents the proportion at random. **c** Common lncRNA targets of *KLF5* and *AR*. The bar represents the percentage of the gene that was detected by ChIP-Seq profiles corresponding to *AR*. The red bar represents the up-regulated genes in CRC while the green bar represents the down-regulated genes in CRC. **d** The common PCG targets of *KLF5* and *HSF1*. The bar represents the percentage of the gene that was detected in the ChIP-Seq profiles corresponding to *HSF1*. The red bar represents the up-regulated genes in CRC. The picture shows the top 10 PCGs. **e** Network visualization of the co-regulatory network for *KLF5* in CRC. The squares nodes represents the PCG, the circle nodes represents TF while the diamond nodes represents lncRNA. Furthermore, the width of the edge represents the percentage of ChIP-Seq profiles that detected the gene as the target of the corresponding TF (*AR* or *HSF1*)



(See figure on next page.)

**Fig. 3** The visualization of the regulatory network within *KLF5* targets. **a** The 15 small circles are the 15 modules with gene numbers larger than 25 identified by the MCL algorithm, while the genes in the biggest circle are the other genes that did not belong to these modules. The main enriched GO BP terms are shown in each module. **b** The subnetwork visualization of the DNA replication module. The circle nodes represent PCGs while the triangle nodes represent lncRNAs. The red nodes represent the genes up-regulated in CRC while the green nodes represent the genes down-regulated in CRC. **c** The enriched GO BP terms of PCGs in the DNA replication module. **d** The subnetwork visualization of the cell cycle module. The meaning of the nodes with different colors are the same as (b). **e** The enriched GO BP terms of PCGs in the cell cycle module

note that the functional enrichment analysis showed that these modules were involved in DNA replication (module 1), DNA repair (module 2), stimulus-response (module 3), immune system process (module 5), apoptosis and cell death (module 6), metabolic and biosynthetic processes (modules 7, 8, and 10), development (module 11), and the cell cycle (module 14) (Fig. 3a). For example, module 1 included one up-regulated lncRNA, *SNHG7*, one down-regulated lncRNA, *FGD5-AS1*, and 138 PCGs, which has major roles in DNA replication and other biogenesis or biosynthetic process (Fig. 3b, c). On the other hand, module 14 included one up-regulated lncRNA, *SNHG12*, and 37 PCGs, whose functions are mainly restricted to the cell cycle (Fig. 3d, e). The enriched functions of the modules suggest that *KLF5* and its targets may participate in similar biological processes.

#### ***SNHG12* in cell cycle module is up-regulated in CRC tissues and cell lines**

Considering the crucial role of *KLF5* in the cell cycle [11, 12], we focused on the module 14, with the enriched cell cycle function. *SNHG12*, the only lncRNA in this module, promotes tumorigenesis and metastasis in human osteosarcoma cells [34] and hepatocellular carcinoma [62]. In particular, we determined that there were *KLF5* binding sites in *SNHG12* promoter regions (Fig. 4a). Function prediction of *SNHG12* based on the co-expression network showed that *SNHG12* might be involved in metabolic processes and the cell cycle (Fig. 4b, c). For validation, we first examined the expression pattern of *SNHG12* in 111 matched CRC tissues and its corresponding adjacent normal tissues. The results showed that *SNHG12* ( $p < 0.01$ , Fig. 4d) was significantly up-regulated in CRC tissues. Next, the expression levels were evaluated in four CRC cell lines (HCT116, HT29, SW620, and COLO-205) and one human colon epithelial cell line (NCM460). After normalization to NCM460, a substantial induction of *SNHG12* was observed in all detected CRC cell lines ( $p < 0.05$ , Fig. 4e).

#### ***SNHG12* expression is regulated by *KLF5* in CRC cell line**

To further determine the relationship between *SNHG12* and *KLF5*, we then studied the expression levels of *KLF5* in the same CRC tissue samples. As hypothesized, *KLF5* expression was positively correlated with that of

*SNHG12* in both cancer and normal tissues (Fig. 5a, b). As a comparison, two other lncRNAs studied previously, *LINC00472* [63] and *LINC00473* (manuscript submitted), displayed a much lower PCCs (Fig. 5c). Furthermore, to verify the regulatory relationship of the two genes, biological experiments were further performed. First, Western blot experiment showed the success of knocking-down expression of *KLF5* in HCT116 cell line with siRNA (Fig. 5d). Second, using qRT-PCR, *SNHG12* expression was observed to be down-regulated significantly after knocking-down expression of *KLF5* with siRNA (Fig. 5e). Our results show *SNHG12* may be regulated by *KLF5*.

#### **Association between *SNHG12* expression and clinical-pathological characteristics of patients with CRC**

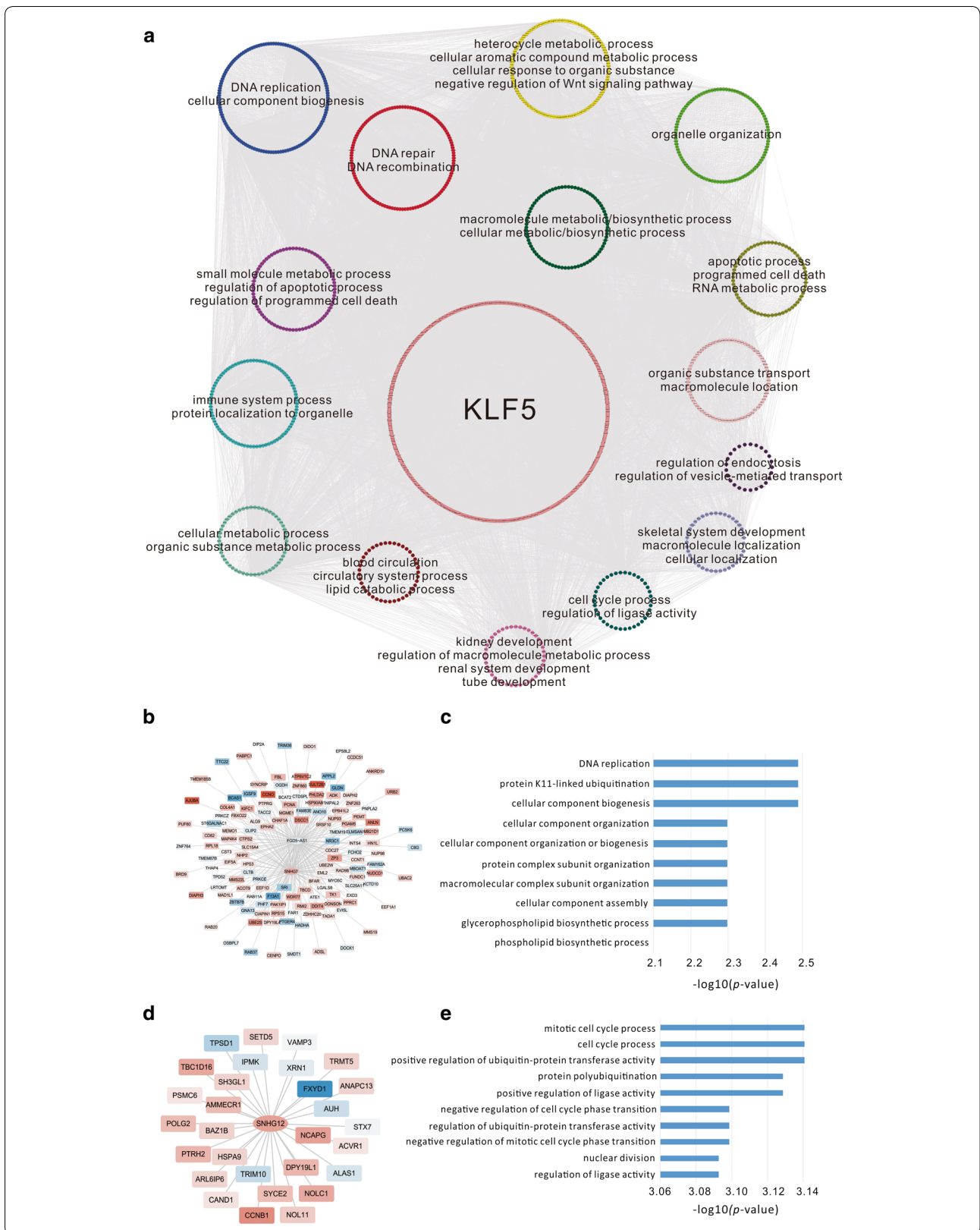
To investigate the oncogenic role of *SNHG12* in CRC progression, we assessed the correlation between *SNHG12* expression levels and the clinical-pathological characteristics of CRC patients. As illustrated in Table 1, *SNHG12* expression was positively associated with invasion ( $p = 0.047$ ) and distal metastasis ( $p = 0.03$ ). However, we did not encounter any association between *SNHG12* expression and other clinical-pathological parameters such as age, gender, tumor location, tumor size, differentiation, lymph node metastasis, or TNM stage (Table 1).

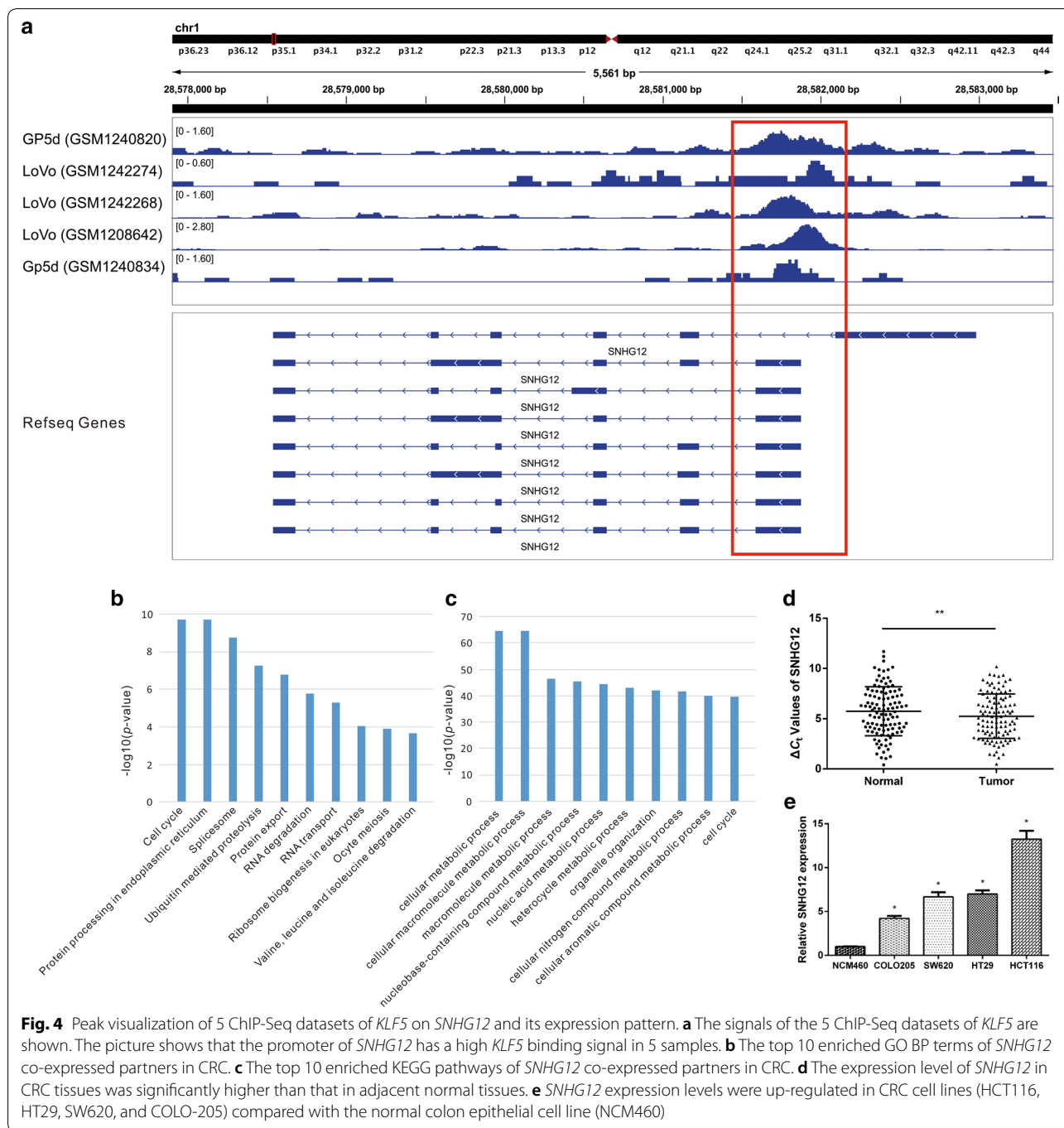
Lastly, we reviewed whether *SNHG12* could be utilized as a diagnostic biomarker of CRC. Therefore, we constructed a Receiver Operating Characteristic (ROC) curves using adjacent normal tissues as a control, which resulted in an area of 0.56 under the ROC curve for *SNHG12* (Fig. 5f).

#### **Discussion**

CRC is one of the most common malignancies, leading to high mortality rates worldwide [1]. The 5-year survival rate of patients with distant metastases is only 11% for colon cancer and 12% for rectal cancer [64]. Although some molecular markers, such as carcinoembryonic antigen, have been used for CRC diagnosis, their diagnostic values are not satisfactory, partially owing to different gene mutations and the complex protein–RNA regulatory network [2, 65]. In this study, a computational pipeline was designed to investigate the functional modules



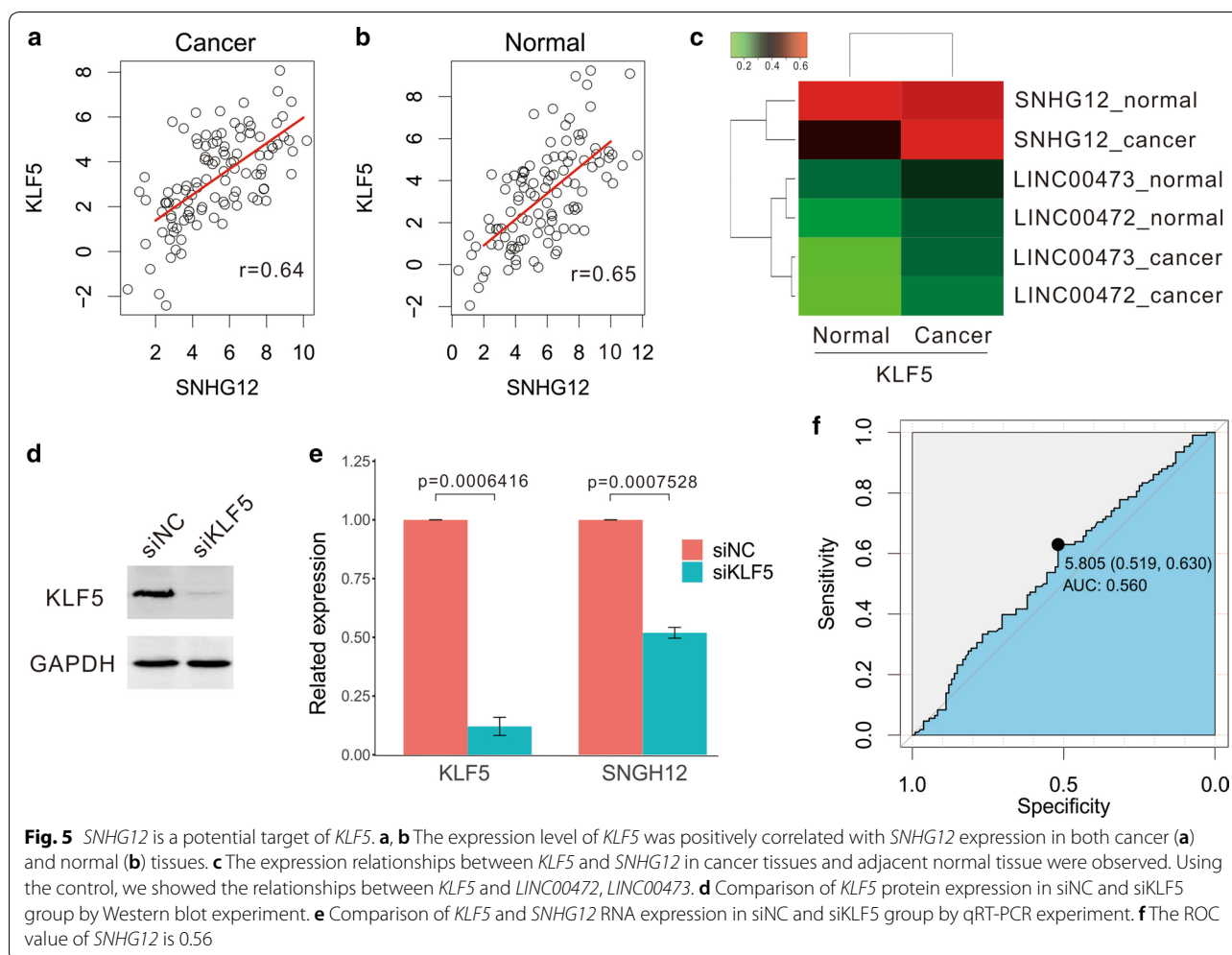




of *KLF5* in CRC and its downstream target lncRNA *SNHG12*, highlighting their oncogenic properties.

Recently, overwhelming evidence has indicated that any dysregulation in the expression of lncRNAs is associated with carcinogenesis and cancer metastasis [66]. As regulatory RNA molecules, lncRNAs have an essential role in the epigenetic, transcriptional, and post-transcriptional regulation of gene expression [67]. Additionally,

many results have indicated that the abnormal expression of several lncRNAs is involved in the carcinogenesis of CRC [68, 69]. For instance, *HOTAIR* participates in CRC invasion and metastasis by stimulating chromatin modifications [22]. Taurine-upregulated gene 1 (*TUG1*), a lncRNA, functions as an oncogene in CRC by regulating cell proliferation, migration, and invasion [70]. Further experiments have demonstrated that *TUG1* knockdown



could impair the migratory and invasive ability of CRC cells by regulating epithelial–mesenchymal transition [70]. Ge et al. [71] also found that prostate cancer-associated transcript 1 (*PCAT-1*) was up-regulated in CRC. Furthermore, research has shown that the overall survival of CRC patients with high *PCAT-1* expression was significantly lower than that of those with low expression [71]. *MEG3* expression was also significantly lower in CRC tissues compared to matched normal tissues, and *MEG3* down-regulation predicted poor prognosis in patients with CRC [72].

The biological network is a complex network that is involved in multiple types of molecular structures such as DNA, protein, and ncRNA. Acting as a TF, *KLF5* may co-operate with other TF to regulate targets together. In this study, we found *HSF1* is a potential co-factor of *KLF5* to regulate protein-coding genes. *HSF1* has been reported to have a central role in the heat shock response to maintain protein homeostasis in all eukaryotic cells [73]. In contrast, accumulating evidence revealed that

*HSF1* has multiple additional functions, including roles in autophagy, apoptosis, immune response, cell growth arrest, and even cancer development [74]. The prominent role of *HSF1* has been observed in several cancers, including gastric cancer [75], osteosarcoma [76], breast cancer [77], and esophageal squamous cell carcinoma [78]. The up-regulation of *HSF1* promotes the proliferation, migration, and invasion of cancer cells and can serve as a prognostic marker in cancer [75]. In CRC, the expression of *HSF1* has been shown to be associated with metastasis, with respect to both RNA and protein levels [79, 80], indicating the key role of *HSF1* in CRC.

Regulatory ncRNAs such as miRNAs and lncRNAs have been shown to regulate the expression, stability, and subcellular location of PCGs in a various number of biological processes [81, 82]. Nowadays, several studies about the association between miRNAs and *KLF5* have been published [83–85]. For example, miR-5195-3p inhibits the proliferation and invasion of human bladder cancer cells by directly targeting *KLF5* [83]. Interestingly,

**Table 1 Association between *SNHG12* expression levels ( $\Delta C_t$ ) and clinical-pathological factors**

Characteristics	No. of patients (%)	<i>SNHG12</i> (Mean $\pm$ SD)	<i>p</i> -value
Age (year)			
≥ 60	67 (60.4)	5.25 $\pm$ 2.19	0.953
< 60	44 (39.6)	5.23 $\pm$ 2.27	
Gender			
Male	66 (59.5)	5.27 $\pm$ 1.97	0.886
Female	45 (40.5)	5.21 $\pm$ 2.54	
Tumor location			
Colon	54 (48.6)	4.95 $\pm$ 2.22	0.169
Rectal	57 (51.4)	5.53 $\pm$ 2.18	
Diameter (cm)			
> 5	28 (25.2)	5.12 $\pm$ 2.19	0.731
≤ 5	83 (74.8)	5.29 $\pm$ 2.23	
Differentiation			
Well	4 (3.6)	4.49 $\pm$ 2.04	0.692
Moderate	87 (78.4)	5.22 $\pm$ 2.25	
Poor	20 (18.0)	5.50 $\pm$ 2.11	
Invasion			
T1	7 (6.3)	6.79 $\pm$ 2.49	0.047*
T2	23 (20.7)	5.41 $\pm$ 2.24	
T3	16 (14.4)	5.91 $\pm$ 2.61	
T4	65 (58.6)	4.80 $\pm$ 1.92	
Lymph node metastasis			
N0	57 (51.4)	5.35 $\pm$ 2.25	0.253
N1	31 (27.9)	5.55 $\pm$ 2.26	
N2	23 (20.7)	4.58 $\pm$ 1.97	
Distal metastasis			
M 0	102 (91.9)	5.36 $\pm$ 2.20	0.030*
M 1	9 (8.1)	3.70 $\pm$ 1.68	
TNM stage			
I	21 (18.9)	5.56 $\pm$ 2.21	0.256
II	36 (32.4)	5.21 $\pm$ 2.30	
III	44 (39.6)	5.41 $\pm$ 2.18	
IV	10 (9.0)	3.96 $\pm$ 1.78	

A smaller  $\Delta C_t$  value indicates higher expression

\* *p* values are significant under 0.05

*KLF5* acts as a transcription factor, which can regulate the expression of miRNA such as miR-200, to maintain epithelial characteristics and preventing epithelial-mesenchymal transition in epithelial cells [85]. Regarding lncRNAs, *KLF5* has been reported to regulate *LINC0346* in gastric cancer [86] and enhance the expression of *RPI* in breast cancer [87]. However, few studies were about lncRNAs and *KLF5* in CRC. The relationship between lncRNAs and TFs is mutual. On one side, lncRNAs can regulate TFs. On the other side, TFs can also regulate lncRNAs through the binding of its corresponding promoter regions just like PCGs.

In this research, we aimed to explore the PCG and lncRNA targets of *KLF5* and analyze the biological regulation network of *KLF5* in CRC. Through bioinformatics analysis, TFs *AR* and *HSF1* were found to be co-regulatory TFs with *KLF5* for lncRNAs and PCGs, respectively (Fig. 2a, b). Next, we showed that *SNHG12* function in the cell cycle module of *KLF5* may be vital in CRC (Fig. 3d, e). We verified the up-regulated expression patterns of *SNHG12* in CRC tissues and cell lines, and further examined the association with the clinical-pathological characteristics of patients with CRC (Fig. 4d, e, Table 1), proving that *SNHG12* is positively involved in CRC invasion and distal metastasis by having an oncogenic role in CRC. Furthermore, we observed *SNHG12* was down-regulated when expression of *KLF5* was knocked-down in CRC cell line (Fig. 5d, e), and the high correlation between *SNHG12* and *KLF5* expression in cancerous and normal tissues (Fig. 5a, b), suggesting that *SNHG12* is a potential target of *KLF5* in CRC. However, whether *KLF5* binding promoter of *SNHG12*, further biological experiments such as Chromatin immunoprecipitation (ChIP) should be conducted in future.

## Conclusion

In conclusion, our results provided several potential PCG and lncRNA targets for *KLF5* in CRC and further demonstrated that one of the targets, *SNHG12*, might function as an oncogene in CRC. Whether targeting *KLF5*–*SNHG12* will produce any therapeutic benefits will require further investigation.

## Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12935-020-01527-x>.

**Additional file 1: Figure S1.** The pipeline of identifying regulatory co-factors of *KLF5*.

**Additional file 2: Table S1.** Potential lncRNA targets of *KLF5* in CRC.

**Additional file 3: Table S2.** Potential protein-coding gene targets of *KLF5* in CRC.

## Abbreviations

*KLF5*: Kruppel like factor 5; PCGs: Protein coding genes; CRC: Colorectal cancer; lncRNAs: Long non-coding RNAs; *SNHG12*: Small nucleolar RNA host gene 12; LPA: Lysophosphatidic acid; *HOTAIR*: HOX transcript antisense RNA; GASS: Growth arrest specific 5; *MEG3*: Maternally expressed gene 3; PPIs: Protein–protein interactions; CNC: Coding-non-coding gene co-expression network; HSR: Heat shock response; CEA: Carcinoembryonic antigen; *TUG1*: Taurine-upregulated gene 1; EMT: Epithelial–mesenchymal transition; *PCAT-1*: Prostate cancer-associated transcript 1; TCGA: The Cancer Genome Atlas; GO: Gene ontology; BP: Biological process; AJCC: American Joint Committee on Cancer; ANOVA: Analysis of variance; ROC: Receiver operating characteristic.

## Acknowledgements

We thank Dr. Weiling He from the Department of Gastrointestinal Surgery, The First Affiliated Hospital, Sun Yat-sen University, who provided suggestions for this study.

### Authors' contributions

QL and LC wrote the manuscript. QL, JG and XZ designed the idea of this study. QL and CD did the bioinformatics analysis. LC, YX, SH, YJ, and JG performed the biology experiments. Derry, Fatma improved English writing in manuscript. NZ and XZ provided the clinical samples and edited the manuscript. GZ, XF, YX, and XD provided some suggestions for the study. All authors read and approved the final manuscript.

### Funding

This work was supported by National Natural Science Foundation of China (Grant No. 31970630), the Fundamental Research Funds for the Provincial Universities of Zhejiang (No. SJLZ2021001), the Natural Science Foundation of Ningbo (Grant No. 2017A610154 and No. 2016A610121), Zhejiang Key Laboratory of Pathophysiology (No. 201812), The Scientific Innovation Team Project of Ningbo (No. 2016C51001 and 2017C110019) and K. C. Wong Magna Fund in Ningbo University.

### Data availability statement

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

### Ethics approval and consent to participate

All patients were anonymous with written informed consent, also all aspects of this study were approved by the Human Research Ethics Committee of Ningbo University with the methods performed in accordance with the relevant guidelines and regulations.

### Consent for publication

Written informed consent for publication was obtained from all participants.

### Competing interests

The authors declare that they have no conflict of interest.

### Author details

<sup>1</sup> Department of Preventative Medicine, Zhejiang Provincial Key Laboratory of Pathophysiology, Ningbo University School of Medicine, Ningbo 315211, Zhejiang, China. <sup>2</sup> The Affiliated Hospital of School of Medicine, Ningbo University, Ningbo 315020, China. <sup>3</sup> Department of Biochemistry and Molecular Biology, Zhejiang Provincial Key Laboratory of Pathophysiology, Ningbo University School of Medicine, Ningbo 315211, Zhejiang, China. <sup>4</sup> Department of Gastroenterology, The Affiliated People's Hospital of Ningbo University, Ningbo 315040, Zhejiang, China. <sup>5</sup> Hua Mei Hospital, University of Chinese Academy of Science, Ningbo 315000, China. <sup>6</sup> Department of Gastrointestinal Surgery, The First Affiliated Hospital, Sun Yat-Sen University, Guangzhou 510080, China.

Received: 4 May 2020 Accepted: 29 August 2020

Published online: 07 September 2020

### References

- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65:87–108.
- Zhang Y, Tao Y, Li Y, Zhao J, Zhang L, Zhang X, Dong C, Xie Y, Dai X, Zhang X, Liao Q. The regulatory network analysis of long noncoding RNAs in human colorectal cancer. *Funct Integr Genom*. 2018. <https://doi.org/10.1007/s10142-017-0588-2>.
- McConnell BB, Yang VW. Mammalian Kruppel-like factors in health and diseases. *Physiol Rev*. 2010;90:1337–811.
- Li X, Zhang B, Wu Q, Ci X, Zhao R, Zhang Z, Xia S, Su D, Chen J, Ma G, Fu L, Dong JT. Interruption of KLF5 acetylation converts its function from tumor suppressor to tumor promoter in prostate cancer cells. *Int J Cancer*. 2015;136:536–46.
- Li X, Liu X, Xu Y, Liu J, Xie M, Ni W, Chen S. KLF5 promotes hypoxia-induced survival and inhibits apoptosis in non-small cell lung cancer cells via HIF-1 $\alpha$ . *Int J Oncol*. 2014;45:1507–14.
- Pattison JM, Posternak V, Cole MD. Transcription factor KLF5 binds a cyclin E1 polymorphic intronic enhancer to confer increased bladder cancer Risk. *Mol Cancer Res*. 2016;14:1078–86.
- Meng G, Xiao Y, Ma Y, Tang X, Xie L, Liu J, Gu Y, Yu Y, Park CM, Xian M, Wang X, Ferro A, Wang R, Moore PK, Zhang Z, Wang H, Han Y, Ji Y. Hydrogen sulfide regulates Kruppel-like factor 5 transcription activity via specificity protein 1 S-sulphydration at Cys664 to prevent myocardial hypertrophy. *J Am Heart Assoc*. 2016;5:e004160.
- Yan Y, Tan MW, Xue X, Ding XY, Wang GK, Xu ZY. Involvement of Oct4 in the pathogenesis of thoracic aortic dissection via inducing the dedifferentiated phenotype of human aortic smooth muscle cells by directly upregulating KLF5. *J Thorac Cardiovasc Surg*. 2016;152(820–829):e4.
- Nandan MO, Ghaleb AM, Liu Y, Bialkowska AB, McConnell BB, Shroyer KR, Robine S, Yang VW. Inducible intestine-specific deletion of Kruppel-like factor 5 is characterized by a regenerative response in adult mouse colon. *Dev Biol*. 2014;387:191–202.
- Guo L, He P, No YR, Yun CC. Kruppel-like factor 5 incorporates into the beta-catenin/TCF complex in response to LPA in colon cancer cells. *Cell Signal*. 2015;27:961–8.
- Nandan MO, Chanchevalap S, Dalton WB, Yang VW. Kruppel-like factor 5 promotes mitosis by activating the cyclin B1/Cdc2 complex during oncogenic Ras-mediated transformation. *FEBS Lett*. 2005;579:4757–62.
- Nandan MO, Yoon HS, Zhao W, Ouko LA, Chanchevalap S, Yang VW. Kruppel-like factor 5 mediates the transforming activity of oncogenic H-Ras. *Oncogene*. 2004;23:3404–13.
- Bialkowska AB, Du Y, Fu H, Yang VW. Identification of novel small-molecule compounds that inhibit the proliferative Kruppel-like factor 5 in colorectal cancer cells by high-throughput screening. *Mol Cancer Ther*. 2009;8:563–70.
- Ruiz de Sabando A, Wang C, He Y, Garcia-Barros M, Kim J, Shroyer KR, Bannister TD, Yang VW, Bialkowska AB. ML264, a novel small-molecule compound that potently inhibits growth of colorectal cancer. *Mol Cancer Ther*. 2016;15:72–83.
- Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell*. 2009;136:629–41.
- Wapinski O, Chang HY. Long noncoding RNAs and human disease. *Trends Cell Biol*. 2011;21:354–61.
- Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell*. 2011;43:904–14.
- Zhang H, Chen Z, Wang X, Huang Z, He Z, Chen Y. Long non-coding RNA: a new player in cancer. *J Hematol Oncol*. 2013;6:37.
- Bartonicek N, Maag JL, Dinger ME. Long noncoding RNAs in cancer: mechanisms of action and technological advancements. *Mol Cancer*. 2016;15:43.
- Luo C, Tao Y, Zhang Y, Zhu Y, Minyao DN, Haleem M, Dong C, Zhang L, Zhang X, Zhao J, Liao Q. Regulatory network analysis of high expressed long non-coding RNA LINC00941 in gastric cancer. *Gene*. 2018;662:103–9.
- Yang Z, Zhou L, Wu LM, Lai MC, Xie HY, Zhang F, Zheng SS. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. *Ann Surg Oncol*. 2011;18:1243–50.
- Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, Shibata K, Suzuki A, Komune S, Miyano S, Mori M. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. *Cancer Res*. 2011;71:6320–6.
- Luo M, Li Z, Wang W, Zeng Y, Liu Z, Qiu J. Long non-coding RNA H19 increases bladder cancer metastasis by associating with EZH2 and inhibiting E-cadherin expression. *Cancer Lett*. 2013;333:213–21.
- Zhu M, Chen Q, Liu X, Sun Q, Zhao X, Deng R, Wang Y, Huang J, Xu M, Yan J, Yu J. lncRNA H19/miR-675 axis represses prostate cancer metastasis by targeting TGFBI. *FEBS J*. 2014;281:3766–75.
- Pickard MR, Mourtada-Maarabouni M, Williams GT. Long non-coding RNA GAS5 regulates apoptosis in prostate cancer cell lines. *Biochim Biophys Acta*. 2013;1832:1613–23.
- Zhou Y, Zhong Y, Wang Y, Zhang X, Batista DL, Gejman R, Ansell PJ, Zhao J, Weng C, Klibanski A. Activation of p53 by MEG3 non-coding RNA. *J Biol Chem*. 2007;282:24731–42.
- Liu Q, Huang J, Zhou N, Zhang Z, Zhang A, Lu Z, Wu F, Mo YY. lncRNA loc285194 is a p53-regulated tumor suppressor. *Nucleic Acids Res*. 2013;41:4976–87.
- Yu L, Xu Q, Yu W, Duan J, Dai G. lncRNA cancer susceptibility candidate 15 accelerates the breast cancer cells progression via miR-153-3p/KLF5 positive feedback loop. *Biochem Biophys Res Commun*. 2018;506:819–25.

29. Tang J, Li Y, Sang Y, Yu B, Lv D, Zhang W, Feng H. LncRNA PVT1 regulates triple-negative breast cancer through KLF5/beta-catenin signaling. *Oncogene*. 2018;37:4723–34.
30. Liao D, Xu H, Wu B, Jiang S, Pan H, Wang R, Chen J. Long noncoding RNA MALAT1 sponges miR1243p.1/KLF5 to promote pulmonary vascular remodeling and cell cycle progression of pulmonary artery hypertension. *Int J Mol Med*. 2019;44:871–84.
31. Liao Q, Xiao H, Bu D, Xie C, Miao R, Luo H, Zhao G, Yu K, Zhao H, Skogerbo G, Chen R, Wu Z, Liu C, Zhao Y. ncFANS: a web server for functional annotation of long non-coding RNAs. *Nucleic Acids Res*. 2011;39:W118–W124.
32. Liao Q, Liu C, Yuan X, Kang S, Miao R, Xiao H, Zhao G, Luo H, Bu D, Zhao H, Skogerbo G, Wu Z, Zhao Y. Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. *Nucleic Acids Res*. 2011;39:3864–78.
33. Zhang Y, Tao Y, Liao Q. Long noncoding RNA: a crosslink in biological regulatory network. *Brief Bioinform*. 2017. <https://doi.org/10.1093/bib/bbx042>.
34. Ruan W, Wang P, Feng S, Xue Y, Li Y. Long non-coding RNA small nucleolar RNA host gene 12 (SNHG12) promotes cell proliferation and migration by upregulating angiominin gene expression in human osteosarcoma cells. *Tumour Biol*. 2016;37:4065–73.
35. Kim SK, Kim SY, Kim JH, Roh SA, Cho DH, Kim YS, Kim JC. A nineteen gene-based risk score classifier predicts prognosis of colorectal cancer patients. *Mol Oncol*. 2014;8:1653–66.
36. Li JR, Sun CH, Li W, Chao RF, Huang CC, Zhou XJ, Liu CC. Cancer RNA-Seq Nexus: a database of phenotype-specific transcriptome profiling in cancer cells. *Nucleic Acids Res*. 2016;44:D944–D951.
37. Mei S, Qin Q, Wu Q, Sun H, Zheng R, Zang C, Zhu M, Wu J, Shi X, Taing L, Liu T, Brown M, Meyer CA, Liu XS. Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. *Nucleic Acids Res*. 2017;45:D658–D662.
38. Yan J, Enge M, Whittington T, Dave K, Liu J, Sur I, Schmierer B, Jolma A, Kivioja T, Taipale M, Taipale J. Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell*. 2013;154:801–13.
39. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kocicinski F, Aken BL, Barrell D, Zadissa A, Searle S, Barnes I, Bignell A, Boychenko V, Hunt T, Kay M, Mukherjee G, Rajan J, Despacio-Reyes G, Saunders G, Steward C, Harte R, Lin M, Howald C, Tanzer A, Derrien T, Chrast J, Walters N, Balasubramanian S, Pei B, Tress M, Rodriguez JM, Ezkurdia I, van Baren J, Brent M, Haussler D, Kellis M, Valencia A, Reymond A, Gerstein M, Guigo R, Hubbard TJ. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*. 2012;22:1760–74.
40. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol*. 1994;2:28–36.
41. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 2009;37:W202–W208.
42. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J, Barnes I, Berry A, Bignell A, Carbonell Sala S, Chrast J, Cunningham F, Di Domenico T, Donaldson S, Fiddes IT, Garcia Giron C, Gonzalez JM, Grego T, Hardy M, Hourlier T, Hunt T, Izuogu OG, Lagarde J, Martin FJ, Martinez L, Mohanan S, Muir P, Navarro FCP, Parker A, Pei B, Pozo F, Ruffier M, Schmitt BM, Stapleton E, Suner MM, Sycheva I, Szczynska-Ratajczak B, Xu J, Yates A, Zerbino D, Zhang Y, Aken B, Choudhary JS, Gerstein M, Guigo R, Hubbard TJP, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res*. 2018. <https://doi.org/10.1093/nar/gky955>.
43. Zhao Z, Bai J, Wu A, Wang Y, Zhang J, Wang Z, Li Y, Xu J, Li X. Co-LncRNA: investigating the lncRNA combinatorial effects in GO annotations and KEGG pathways based on human RNA-Seq data. *Database*. 2015. <https://doi.org/10.1093/database/bav082>.
44. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shaheen B, Venugopal A, Balakrishnan L, Marimuthu A, Banerjee S, Somanathan DS, Sebastian A, Rani S, Ray S, Harshy Kishore CJ, Kanth S, Ahmed M, Kashyap MK, Mohmood R, Ramachandra YL, Krishna V, Rahiman BA, Mohan S, Ranganathan P, Ramabadrans S, Chaerkady R, Pandey A. Human protein reference database—2009 update. *Nucleic Acids Res*. 2009;37:D767–D772.
45. Enright AJ, Van Dongen S, Ouzounis CA. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res*. 2002;30:1575–84.
46. Li P, Chen S, Chen H, Mo X, Li T, Shao Y, Xiao B, Guo J. Using circular RNA as a novel type of biomarker in the screening of gastric cancer. *Clin Chim Acta*. 2015;444:132–6.
47. Chen C, Benjamin MS, Sun X, Otto KB, Guo P, Dong XY, Bao Y, Zhou Z, Cheng X, Simons JW, Dong JT. KLF5 promotes cell proliferation and tumorigenesis through gene regulation and the TSU-Pr1 human bladder cancer cell line. *Int J Cancer*. 2006;118:1346–55.
48. Ni B, Yu X, Guo X, Fan X, Yang Z, Wu P, Yuan Z, Deng Y, Wang J, Chen D, Wang L. Increased urothelial cancer associated 1 is associated with tumor proliferation and metastasis and predicts poor prognosis in colorectal cancer. *Int J Oncol*. 2015;47:1329–38.
49. Li S, Hua Y, Jin J, Wang H, Du M, Zhu L, Chu H, Zhang Z, Wang M. Association of genetic variants in lncRNA H19 with risk of colorectal cancer in a Chinese population. *Oncotarget*. 2016;7:25470–7.
50. Zheng Y, Song D, Xiao K, Yang C, Ding Y, Deng W, Tong S. LncRNA GAS5 contributes to lymphatic metastasis in colorectal cancer. *Oncotarget*. 2016;7:83727–34.
51. Takahashi Y, Sawada G, Kurashige J, Uchi R, Matsumura T, Ueo H, Takano Y, Eguchi H, Sudo T, Sugimachi K, Yamamoto H, Doki Y, Mori M, Mimori K. Amplification of PVT-1 is involved in poor prognosis via apoptosis inhibition in colorectal cancers. *Br J Cancer*. 2014;110:164–71.
52. Sogawa K, Imataka H, Yamasaki Y, Kusume H, Abe H, Fujii-Kuriyama Y. cDNA cloning and transcriptional properties of a novel GC box-binding protein, BTEB2. *Nucleic Acids Res*. 1993;21:1527–32.
53. Kojima S, Kobayashi A, Gotoh O, Ohkuma Y, Fujii-Kuriyama Y, Sogawa K. Transcriptional activation domain of human BTEB2, a GC box-binding factor. *J Biochem*. 1997;121:389–96.
54. Boeva V. Analysis of genomic sequence motifs for deciphering transcription factor binding and transcriptional regulation in eukaryotic cells. *Front Genet*. 2016;7:24.
55. Yang XO, Doty RT, Hicks JS, Willerford DM. Regulation of T-cell receptor D beta 1 promoter by KLF5 through reiterated GC-rich motifs. *Blood*. 2003;101:4492–9.
56. Xing C, Ci X, Sun X, Fu X, Zhang Z, Dong EN, Hao ZZ, Dong JT. Klf5 deletion promotes Pten deletion-initiated luminal-type mouse prostate tumors through multiple oncogenic signaling pathways. *Neoplasia*. 2014;16:883–99.
57. Frigo DE, SHERK AB, Wittmann BM, Norris JD, Wang Q, Joseph JD, Toner AP, Brown M, McDonnell DP. Induction of Kruppel-like factor 5 expression by androgens results in increased CXCR4-dependent migration of prostate cancer cells in vitro. *Mol Endocrinol*. 2009;23:1385–96.
58. Xu J, Meng Q, Li X, Yang H, Xu J, Gao N, Sun H, Wu S, Familiari G, Relucenti M, Zhu H, Wu J, Chen R. Long noncoding RNA MIR17HG promotes colorectal cancer progression via miR-17-5p. *Cancer Res*. 2019;79:4882–95.
59. Hu Y, Wang L, Li Z, Wan Z, Shao M, Wu S, Wang G. Potential prognostic and diagnostic values of CDC6, CDC45, ORC6 and SNHG7 in colorectal cancer. *Onco Targets Ther*. 2019;12:11609–21.
60. Zhou J, Zhan S, Tan W, Cheng R, Gong H, Zhu Q. P300 binds to and acetylates MTA2 to promote colorectal cancer cells growth. *Biochem Biophys Res Commun*. 2014;444:387–90.
61. Zhang XH, Zheng B, Han M, Miao SB, Wen JK. Synthetic retinoid Am 80 inhibits interaction of KLF5 with RAR alpha through inducing KLF5 dephosphorylation mediated by the PI3K/Akt signaling in vascular smooth muscle cells. *FEBS Lett*. 2009;583:1231–6.
62. Lan T, Ma W, Hong Z, Wu L, Chen X, Yuan Y. Long non-coding RNA small nucleolar RNA host gene 12 (SNHG12) promotes tumorigenesis and metastasis by targeting miR-199a/b-5p in hepatocellular carcinoma. *J Exp Clin Cancer Res*. 2017;36:11.
63. Chen L, Zhang W, Li DY, Wang X, Tao Y, Zhang Y, Dong C, Zhao J, Zhang L, Zhang X, Guo J, Zhang X, Liao Q. Regulatory network analysis of LINC00472, a long noncoding RNA downregulated by DNA hypermethylation in colorectal cancer. *Clin Genet*. 2018. <https://doi.org/10.1111/cge.13245>.
64. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin*. 2015;65:5–29.
65. Lin SH, Raju GS, Huff C, Ye Y, Gu J, Chen JS, Hildebrandt MAT, Liang H, Menter DG, Morris J, Hawk E, Stroehlein JR, Futreal A, Kopetz S, Mishra L, Wu X. The somatic mutation landscape of premalignant colorectal adenoma. *Gut*. 2018;67:1299–305.

66. Shen XH, Qi P, Du X. Long non-coding RNAs in cancer invasion and metastasis. *Mod Pathol*. 2015;28:4–13.
67. Yang L, Froberg JE, Lee JT. Long noncoding RNAs: fresh perspectives into the RNA world. *Trends Biochem Sci*. 2014;39:35–433.
68. Xie X, Tang B, Xiao YF, Xie R, Li BS, Dong H, Zhou JY, Yang SM. Long non-coding RNAs in colorectal cancer. *Oncotarget*. 2016;7:5226–399.
69. Han D, Wang M, Ma N, Xu Y, Jiang Y, Gao X. Long noncoding RNAs: novel players in colorectal cancer. *Cancer Lett*. 2015;361:13–211.
70. Wang L, Zhao Z, Feng W, Ye Z, Dai W, Zhang C, Peng J, Wu K. Long non-coding RNA TUG1 promotes colorectal cancer metastasis via EMT pathway. *Oncotarget*. 2016. <https://doi.org/10.18632/oncotarget.10563>.
71. Ge X, Chen Y, Liao X, Liu D, Li F, Ruan H, Jia W. Overexpression of long non-coding RNA PCAT-1 is a novel biomarker of poor prognosis in patients with colorectal cancer. *Med Oncol*. 2013;30:588.
72. Yin DD, Liu ZJ, Zhang E, Kong R, Zhang ZH, Guo RH. Decreased expression of long noncoding RNA MEG3 affects cell proliferation and predicts a poor prognosis in patients with colorectal cancer. *Tumour Biol*. 2015;36:4851–9.
73. Akerfelt M, Morimoto RI, Sistonen L. Heat shock factors: integrators of cell stress, development and lifespan. *Nat Rev Mol Cell Biol*. 2010;11:545–55.
74. Barna J, Csermely P, Vellai T. Roles of heat shock factor 1 beyond the heat shock response. *Cell Mol Life Sci*. 2018;75:2897–916.
75. Kim SJ, Lee SC, Kang HG, Gim J, Lee KH, Lee SH, Chun KH. Heat shock factor 1 predicts poor prognosis of gastric cancer. *Yonsei Med J*. 2018;59:1041–8.
76. Zhou Z, Li Y, Jia Q, Wang Z, Wang X, Hu J, Xiao J. Heat shock transcription factor 1 promotes the proliferation, migration and invasion of osteosarcoma cells. *Cell Prolif*. 2017;50:e12346.
77. Gokmen-Polar Y, Badve S. Upregulation of HSF1 in estrogen receptor positive breast cancer. *Oncotarget*. 2016;7:84239–45.
78. Liao Y, Xue Y, Zhang L, Feng X, Liu W, Zhang G. Higher heat shock factor 1 expression in tumor stroma predicts poor prognosis in esophageal squamous cell carcinoma patients. *J Transl Med*. 2015;13:338.
79. Cen H, Zheng S, Fang YM, Tang XP, Dong Q. Induction of HSF1 expression is associated with sporadic colorectal cancer. *World J Gastroenterol*. 2004;10:3122–6.
80. Zhou B, Guo R. Genomic and regulatory characteristics of significant transcription factors in colorectal cancer metastasis. *Sci Rep*. 2018;8:17836.
81. Zhang Y, Tao Y, Liao Q. Long noncoding RNA: a crosslink in biological regulatory network. *Brief Bioinform*. 2018;19:930–45.
82. Pradhan AK, Emdad L, Das SK, Sarkar D, Fisher PB. The enigma of miRNA regulation in cancer. *Adv Cancer Res*. 2017;135:25–522.
83. Jiang Z, Zhang Y, Cao R, Li L, Zhong K, Chen Q, Xiao J. MiR-5195-3p inhibits proliferation and invasion of human bladder cancer cells by directly targeting oncogene KLF5. *Oncol Res*. 2017. <https://doi.org/10.3727/096504016X14831120463349>.
84. Morimoto Y, Mizushima T, Wu X, Okuzaki D, Yokoyama Y, Inoue A, Hata T, Hirose H, Qian Y, Wang J, Miyoshi N, Takahashi H, Haraguchi N, Matsuda C, Doki Y, Mori M, Yamamoto H. miR-4711-5p regulates cancer stemness and cell cycle progression via KLF5, MDM2 and TFDP1 in colon cancer cells. *Br J Cancer*. 2020;122:1037–49.
85. Zhang B, Zhang Z, Xia S, Xing C, Ci X, Li X, Zhao R, Tian S, Ma G, Zhu Z, Fu L, Dong JT. KLF5 activates microRNA 200 transcription to maintain epithelial characteristics and prevent induced epithelial-mesenchymal transition in epithelial cells. *Mol Cell Biol*. 2013;33:4919–35.
86. Xu TP, Ma P, Wang WY, Shuai Y, Wang YF, Yu T, Xia R, Shu YQ. KLF5 and MYC modulated LINC00346 contributes to gastric cancer progression through acting as a competing endogenous RNA and indicates poor outcome. *Cell Death Differ*. 2019;26:2179–93.
87. Jia X, Shi L, Wang X, Luo L, Ling L, Yin J, Song Y, Zhang Z, Qiu N, Liu H, Deng M, He Z, Li H, Zheng G. KLF5 regulated lncRNA RP1 promotes the growth and metastasis of breast cancer via repressing p27kip1 translation. *Cell Death Dis*. 2019;10:373.

## Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

