RESEARCH

Identification of fatty acid metabolism-related lncRNAs in the prognosis and immune microenvironment of colon adenocarcinoma

Shuang Wu[1†], Yuzhu Gong[1†], Jianfang Chen[1†], Xiang Zhao[1], Huimin Qing[1], Yan Dong[1], Sisi Li[2], Jianjun Li[1*] and Zhe Wang[1*]

Abstract

Background: Cancer metabolism is largely altered compared to normal cells. This study aims to explore critical metabolism pathways in colon adenocarcinoma (COAD), and reveal the possible mechanism of their role in cancer progression.

Methods: Expression data and sequencing data of COAD samples were obtained from The Cancer Genome Atlas and Gene Expression Omnibus databases. The expression profiles between tumor and normal samples were compared to identify differential metabolism pathways through single sample gene set enrichment analysis.

Results: Fatty acid synthesis was identified as a key metabolism pathway in COAD. Based on fatty acid-related lncRNAs, two molecular subtypes (C1 and C2) were defined. C2 subtype with worse prognosis had higher immune infiltration and higher expression of immune checkpoints. Five transcription factors (TFs) including FOS, JUN, HIF1A, STAT3 and STAT2 were highly expressed in C2 subtype. Five fatty acid-related lncRNAs were identified to be biomarkers for predicting COAD prognosis. Finally, further experients showed that knockdown of lncRNA PAXIP1-AS1 decreased the triglyceride content and the fatty acid synthase and acetyl-CoA carboxylase 1 expressions, which suggested that lncRNA PAXIP1-AS1 plays an important role in fatty acid metabolism of COAD.

Conclusions: This study demonstrated that fatty acid synthesis was greatly altered in COAD. Fatty acid-related lncRNAs were speculated to be involved in cancer progression through associating with TFs. The five screened TFs may serve as new drug targets for treating COAD.

Keywords: Colon adenocarcinoma, Fatty acid metabolism, Long non-coding RNAs, Molecular subtypes, Prognosis, Immune infiltration

[†]Shuang Wu, Yuzhu Gong and Jianfang Chen contributed equally to this work.

*Correspondence: leejjun2007@163.com; wangzhe.com@hotmail.com

[1] Department of Oncology and Southwest Cancer Center, Southwest Hospital, Third Military Medical University (Army Medical University), Chongqing 400038, China
Full list of author information is available at the end of the article

Introduction

Colon adenocarcinoma (COAD) is one of the most common cancers within digestive system, whose death rate ranking in the fifth of all cancer types [1]. American Joint Committee on Cancer (AJCC) classifies COAD patients into different stages with different clinical features including histological features, metastasis and lymph node. With the development of

Wu *et al. Biology Direct*   (2022) 17:19

Page 2 of 14

new therapies for cancer treatment, especially immunotherapy and other molecular-targeted therapies, improved overall survival of cancer patients is shown [2, 3]. However, not all cancer patients can benefit from these new therapeutics. To address this issue, a number of studies also explored a series of molecular features such as genomic features [4] and biomarkers [5–7] for predicting the response to immunotherapy or chemotherapy, thus guiding personalized treatments.

Cancer metabolism is one of important aspects in cancer development, such as activated glycolysis contributing a lot in cancer metastasis [8]. The complexity of cancer metabolism has been underscoring in the previous studies, but comprehensive understanding on cancer metabolism can largely facilitate the development of new targeted drugs. ACLY is the first-step rate-controlling enzyme in lipid synthesis, which is found to be upregulated in colon cancer and to contribute metastasis [9]. Evidences have demonstrated that the remodeling of tumor microenvironment (TME) can be modulated by metabolism-related pathways [10]. Conversely, TME can also alter metabolic features through cellular metabolites of cells in TME [11]. It has been suggested that targeting cancer metabolism is a promising strategy for cancer treatment [12–15]. Lipid metabolism-related signature is also developed for predict colon cancer survival [16].

In the regulation of metabolism-related pathways, long non-coding RNAs (lncRNAs) play a critical role in metabolic reprogramming [17–19]. For example, lncRNA-GLCC1 was found to reprogram glycolytic metabolism and was associated with prognosis in colorectal cancer [20]. CCAT1/FABP5 promotes tumour progression through mediating fatty acid metabolism in lung adenocarcinoma [21]. LINC01606 protects colon cancer cells from ferroptotic cell death and functions as an oncogene to facilitate tumor cell stemness [22]. CDKN2B-AS1 is considered as a promising biomarker or therapeutic target in various cancer types [23].

Therefore, in this study, we introduced a series of bioinformatics analysis to further explore the role of metabolism and metabolism-related lncRNAs in COAD development. We identified fatty acid biosynthesis as a critical pathway that was greatly dysregulated in COAD. Based on lncRNAs related to fatty acid, we constructed two novel molecular subtypes with different molecular features. Furthermore, we explored the relation between lncRNAs and transcription factors (TFs), and identified key fatty acid-related lncRNAs for predicting COAD prognosis. Fatty acid was considered as a critical metabolism pathway in COAD progression.

Methods

Additional file 1: Fig. S1A shows the workflow of this study.

Information of COAD samples and data preprocessing

TCGA-COAD dataset containing RNA-seq data (FPKM) and clinical information was downloaded from The Cancer Genome Atlas (TCGA) database in October 01, 2021. GSE17538 dataset was obtained from Gene Expression Omnibus (GEO) database in October 01, 2021. Samples without clinical information, survival time and survival status were removed. Only samples with stage II, III and IV were remained. Ensembl ID and probes were converted to gene symbol. The fpkm of gene expression was converted into TPM expression profile according to the method of Li et al. [24]. Expression data in two datasets were identified into mRNAs and lncRNAs according to GTF file downloaded from GENCODE (https://www.gencodegenes.org/).

Identification of fatty acid-related lncRNAs

Pearson correlation analysis was conducted between fatty acid activity score and lncRNAs. Fatty acid-related lncRNAs were identified under conditions of |correlation coefficient|> 0.25 and $P < 0.05$. The lncRNAs that were identified both in TCGA-COAD and GSE17538 datasets were considered as important fatty acid-related lncRNAs for further analysis.

Unsupervised consensus clustering based on fatty acid-related lncRNAs

Unsupervised consensus clustering is a widely used method for identifying molecular subtypes with different molecular features based on expression profiles. We used ConsensusClusterPlus R package to conduct unsupervised consensus clustering based on the expression of fatty acid-related lncRNAs [25]. KM algorithm was used and Euclidean was selected as distance. 500 bootstraps were repeated with each bootstrap containing 80% samples in TCGA-COAD dataset. Cluster number k was set from 2 to 10, and cumulative distribution function (CDF) and consensus matrix were used to confirm the optimal cluster number.

Gene set enrichment analysis (GSEA)

GSEA is a powerful method for delineating biological processes based on gene expression data [26]. We applied GSEA to calculate the enrichment score of hallmark pathways for molecular subtypes. Hallmark pathways with a series of gene sets were obtained from Molecular Signature database (MSigDB) [27]. ClusterProfiler R

Wu *et al. Biology Direct*   (2022) 17:19

Page 3 of 14

package was used to annotate gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways on fatty acid-related lncRNAs and TFs [28].

Characterization of tumor microenvironment

The proportion of different immune cells was calculated by GSEA based on a series of gene signatures. Estimation of STromal and Immune cells in MAlignant Tumours using Expression data (ESTIMATE) algorithm was conducted to evaluate immune infiltration from two aspects including stromal score and immune score [29]. ESTIMATE score is the combined score of stromal score and immune score. Tumor Immune Dysfunction and Exclusion (TIDE) analysis was implemented for assess T cell function [30].

Assessing the function of fatty acid-related lncRNAs

To analyze the localization of fatty acid-related lncRNAs, relative concentration index (RCI) was calculated for each lncRNA based on LncATLAS database [31]. $RCI < 0$ indicates nucleus-localized lncRNAs and $RCI < 0$ indicates cytoplasm-localized lncRNAs. The TF activity was assessed in accordance with the algorithm from Garcia-Alonso et al. [32]. Pearson correlation analysis was conducted to analyze the association between lncRNAs and TFs.

The first-order partial correlation analysis

The first order partial correlation analysis was an efficient method for analyzing one variable in the relation between another two variables. It was used to investigate the association among fatty acid-related lncRNAs, fatty acid score and fatty acid-related genes [33]. The association between two variables was largely weakened when removing the effect of another variable. Of the case, the variable was considered as key fatty acid-related lncRNA.

Construction of a prognostic model based on key fatty acid-related lncRNAs

Key fatty acid-related lncRNAs were identified by the first order partial correlation analysis. Univariate Cox regression analysis was employed on the key lncRNAs to obtain coefficients. The prognostic model was defined as risk score = beta i * expression i, where beta represents coefficients and i represents lncRNAs. Risk score was calculated for each sample in two datasets. Median value of risk score was selected as a cut-off to divide samples into high-risk and low-risk groups.

Cell culture

NCM460 cell line (normal human colonic epithelial cell line) and all the four human colorectal cancer (COAD) cell lines (LoVo, HCT116, SW480, and HT-29) were purchased from the Cell Bank of Type Culture Collection of the Chinese Academy of Sciences (CBTCCCAS, Shanghai, China). The cells were cultured in RPMI 1640 medium (Gibco, United States) or McCoy's 5A medium (Gibco, United States) supplemented with 10% fetal bovine serum (Gibco, United States), 100 U/ml penicillin, and 100 mg/ml streptomycin at 37 °C in a humidified incubator with a 5% CO2 concentration.

Quantitative real-time PCR

Total RNA was extracted from NCM460 cells and COAD cells by TRIzol Reagent (Beijing Solarbio Technology Co., Ltd., Beijing, China) according on the manufacturer's instruction. PrimeScript™ RT-PCR kit (TaKaRa, Mountain View, CA) was used to perform reverse transcription after quality validation. qRT-PCR was conducted by using the SYBR Premix Ex Taq™ (TaKaRa). The expression of β-actin was used as an internal control. The fatty acid-related lncRNA expression was measured via the $2^{-\Delta\Delta CT}$ method. All primer sequences used in this research are listed in Additional file 2: Table S1.

Short interference RNAs (siRNAs), and transfection

Small interfering RNA (siRNA) of PAXIP1-AS1 were synthesized by RiboBio (Guangzhou, China). HCT116 and HT-29 cells were transfected with si-PAXIP1-AS1for 48 h by using Lipofectamine 2000 (Invitrogen, CA, USA). qRT-PCR was used to evaluate the knockdown efficiency of si-PAXIP1-AS1.

Intracellular triglyceride assay

Intracellular triglyceride was determined using an E1687Ge General TGC/Triglyceride ELISA Kit (EIAab Science Inc., Wuhan, China) at 450 nm according to the manufacturer's protocol.

Western blot

HCT116 and HT-29 cells transfected with si-PAXIP1-AS1 were lysed with RIPA lysis buffer (R0010, Solarbio, China) supplemented with protease inhibitors (Roche). BCA Kit (Pierce, Rockford, IL) was used to determine the protein concentration. The total protein was added into loading buffer and then was separated by 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), then the separated protein was transferred to polyvinylidene fluoride (PVDF) membrane (Merck Millipore, Billerica, MA). The membrane was blocked with 5% skimmed milk for 1 h, the primary anti-fatty acid synthase (FASN) and acetyl-CoA carboxylase 1 (ACC1) antibodies (Cell Signaling Technology, Danvers, MA) was used to block protein overnight at 4 °C. The PVDF membrane was then washed with TBST and incubated with the corresponding secondary antibody for 2 h. The

Wu *et al. Biology Direct*   (2022) 17:19

Page 4 of 14

proteins were identified by using Pierce SuperSignal West Pico Chemiluminescent Substrate (Termo Fisher, Waltham, MA) according to the manufacturer's protocol. β-actin antibody was used as the internal reference.

Statistical analysis
Parameters of packages and tools were default if without indication. Statistical analysis was performed in R (4.1.1) platform. Statistical methods were indicated in the corresponding sections. $P < 0.05$ was considered as significant. ns, no sifnificance. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$.

Results
Identification of important metabolism-related pathways
Firstly, single sample gene set enrichment analysis (ssGSEA) was used to calculate enrichment score for each sample in two datasets. Then Wilcoxon test was conducted to identify differential metabolism-related pathways between normal and tumor samples in TCGA-COAD dataset. Within the identified differential pathways, univariate Cox regression analysis was performed to identify metabolism-related pathways that were associated with prognosis in two datasets. Finally, fatty acid biosynthesis pathway was identified in both two datasets.

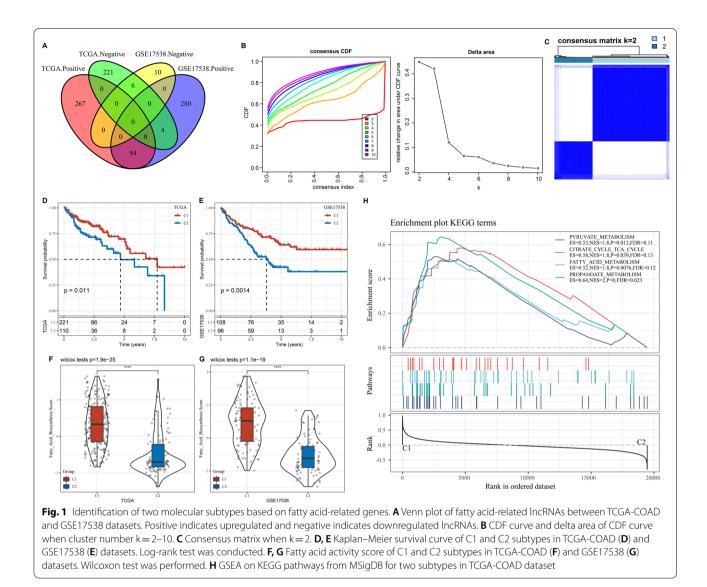### Constructing molecular subtypes based on lncRNAs related to fatty acid
To identify lncRNAs related to fatty acid, Pearson correlation analysis was performed between activity score of fatty acid and lncRNA expression. In TCGA-COAD and GSE17538 datasets, we identified 592 and 394 lncRNAs related to fatty acid respectively. Among these lncRNAs, 6 lncRNAs were negatively correlated with fatty acid in both two datasets, and 94 lncRNAs were positivity correlated with fatty acid (Fig. 1A). Then univariate Cox regression analysis was conducted within 100 lncRNAs, and 15 lncRNAs related to prognosis were identified. By using unsupervised consensus clustering, samples in TCGA-COAD dataset were classified into different groups. According to CDF and area under CDF curve in different cluster numbers, cluster number k = 2 was determined to clearly classify samples into two groups (Fig. 1B, C). As a consequence, cluster 1 (C1) and cluster 2 (C2) were defined as two molecular subtypes. Survival analysis on two subtypes showed that C1 and C2 had differential overall survival with C1 had more favorable prognosis in both TCGA-COAD and GSE17538 datasets ($P = 0.011$ and $P = 0.0014$ respectively, Fig. 1D, E). Moreover, C1 subtype had extremely higher fatty acid activity than C2 subtype in both TCGA-COAD and GSE17538 datasets ($P = 1.9e-25$ and $P = 1.1e-19$ respectively, Fig. 1F, G), indicating that high fatty acid activity may have a

protective role to against tumor progression. To understand if there was a difference on pathways between C1 and C2 subtypes, we used GSEA to calculate enrichment score of hallmark gene sets from MSigDB [27]. As a result, we observed that metabolism-related pathways including pyruvate metabolism, fatty acid metabolism, and citrate cycle TCA cycle were significantly enriched in C1 subtype ($P < 0.05$, FDR < 0.25, Fig. 1H).

Different genomic and mutation features between C1 and C2 subtypes
To compare genomic features of two subtypes, we obtained genomic data of TCGA-COAD dataset from previous research [34], including aneuploidy score, homologous recombination defects, number of segments, fraction altered and tumor mutation burden. Except for number of segments, C1 and C2 subtypes had a significant difference on other four genomic features ($P < 0.0001$, Fig. 2A). C1 subtype showed obviously higher enrichment score of aneuploidy, homologous recombination defects and fraction altered, while C2 subtype had higher score of tumor mutation burden. Among five genomic features, aneuploidy score and fraction altered were significantly correlated with fatty acid score (R = 0.223, $P = 8.61e-05$ and R = 0.216, $P = 0.00012$ respectively, Fig. 2B). Furthermore, analysis on single nucleotide variations revealed the top 16 significantly mutated genes ($P < 0.05$, Fig. 2C). In C1 subtype, APC and TP53 were highly mutated with mutation frequencies of 84% and 65% respectively, while they were 49% and 39% in C2 subtype. Compared with C1 subtypes, mutations in C2 subtype were more scattered. The different genomic features and gene mutations can alter gene expression and thus contribute to dysregulated pathways.

Differentially enriched pathways between C1 and C2 subtypes
Given that C2 subtype had worse prognosis than C1 subtype, we tried to find out whether there were differentially enriched pathways that may affect prognosis between them. For this purpose, hallmark pathways from MSigDB were included for conducting GSEA in TCGA-COAD and GSE17538 datasets. By comparing C2 to C1, we observed that two subtypes exhibited differential enrichment in a number of pathways, with C2 subtype showing lower enrichment of many metabolism-related pathways. Within these differential pathways between two subtypes, 15 pathways such as VEGF signaling, ECM receptor interaction and Toll-like receptor signaling pathways were upregulated in C2 subtype in both two datasets, and 9 pathways such as retinol metabolism, P450, starch and sucrose metabolism were downregulated (Fig. 3A). Within 15 upregulated pathways in C2 subtype, Rader

Wu *et al. Biology Direct*   (2022) 17:19

Page 5 of 14

Fig. 1 Identification of two molecular subtypes based on fatty acid-related genes. **A** Venn plot of fatty acid-related lncRNAs between TCGA-COAD and GSE17538 datasets. Positive indicates upregulated and negative indicates downregulated lncRNAs. **B** CDF curve and delta area of CDF curve when cluster number k = 2–10. **C** Consensus matrix when k = 2. **D, E** Kaplan–Meier survival curve of C1 and C2 subtypes in TCGA-COAD (**D**) and GSE17538 (**E**) datasets. Log-rank test was conducted. **F, G** Fatty acid activity score of C1 and C2 subtypes in TCGA-COAD (**F**) and GSE17538 (**G**) datasets. Wilcoxon test was performed. **H** GSEA on KEGG pathways from MSigDB for two subtypes in TCGA-COAD dataset
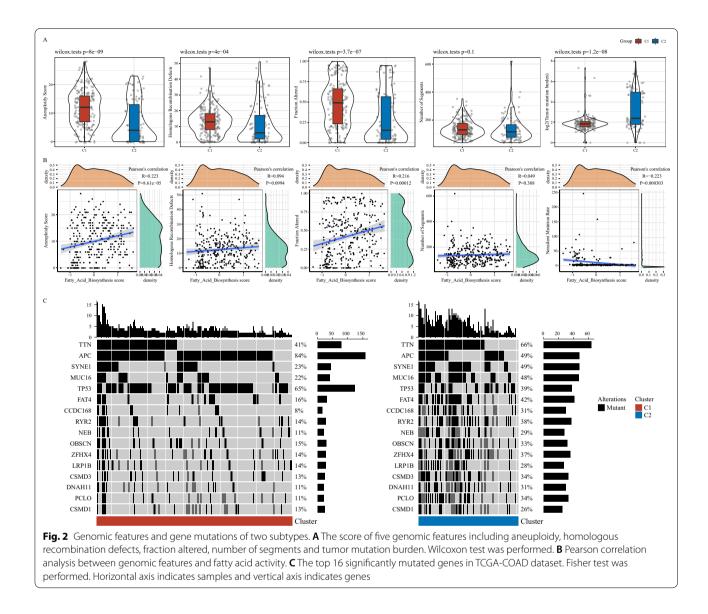
plots presented similar normalized enrichment score (NES) in two datasets (Fig. 3B, C), suggesting that the subtyping based on fatty acid-related lncRNA expression was reliable, and the these lncRNAs may be involved in regulating these activated pathways in C2 subtype.

High immune infiltration and T cell dysfunction were shown in C2 subtype

The component of TME is a critical hallmark in cancer development. Therefore, we evaluate the distribution of different types of immune cells in two subtypes by using a series of gene signatures. As a result, among 24 immune cells, we found that 19 of them were differentially enriched in two subtypes in TCGA-COAD dataset (Fig. 4A). Notably, CD8 T cells, cytotoxic cells and dendritic cells (DCs) were more enriched in C2 subtype, while immunosuppressive cells such as macrophages

and regulatory T (Treg) cells were also highly distributed ($P < 0.05$). ESTIMATE analysis further supported the above result that C2 subtype had significantly higher stromal score and immune score than C1 subtype, indicating higher immune infiltration of C2 subtype ($P < 0.01$, Fig. 4B). In GSE17538 dataset, we observed similar results (Additional file 1: Fig. S1A, B).

Furthermore, we investigated the expression of immune checkpoints that can affect the function of anti-tumor immune cells. It was observed that most of immune checkpoints were more highly expressed in C2 subtype than C1 (Fig. 4C). Especially, some important checkpoints whose high expression were reported to inhibit T cell function were significantly higher expressed in C2 subtype, such as LAG3, CTLA4, PDCD1, IDO1, and CD274 ($P < 0.0001$). High expression of these immunosuppressive checkpoints in

Wu *et al. Biology Direct*   (2022) 17:19

Page 6 of 14

Fig. 2 Genomic features and gene mutations of two subtypes. **A** The score of five genomic features including aneuploidy, homologous recombination defects, fraction altered, number of segments and tumor mutation burden. Wilcoxon test was performed. **B** Pearson correlation analysis between genomic features and fatty acid activity. **C** The top 16 significantly mutated genes in TCGA-COAD dataset. Fisher test was performed. Horizontal axis indicates samples and vertical axis indicates genes
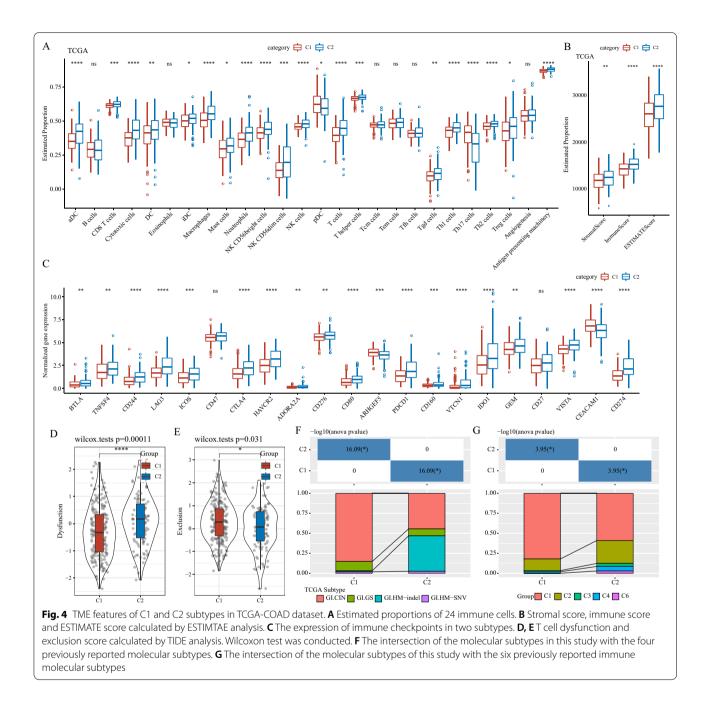
C2 subtype may explain its worse outcome. In addition, TIDE analysis further demonstrated that C2 subtype had more severe T cell dysfunction than C1 ($P = 0.00011$, Fig. 4D), which was consistent with the above results. Simultaneously, T cell exclusion was shown to be a bit more severe in C1 ($P = 0.031$, Fig. 4E). In GSE17538 dataset, the expression of immune checkpoints were not obviously differential between two subtypes, but CD274 and CTLA4 were still overexpressed in C2 subtype ($P < 0.01$, Additional file 1: Fig. S1C). TIDE analysis in GSE17538 dataset displayed similar result with the result in TCGA-COAD dataset (Additional file 1: Fig. S1D and E), supporting that the unfavorable prognosis of C2 subtype may result from the dysfunction of cytotoxic T cells. We also compared

the relationship between the two molecular subtypes and the four molecular subtypes previously reported by Liu et al. [35]. It can be observed that there are significant differences in the distribution of the corresponding four old molecular subtypes in the two subtypes, and the C1 subtype is mainly GI CIN subtype and C2 subtype are rich in GI Patients with HM-indel subtype (Fig. 4F). Thorsson et al. [34] reported the 6 immune subtypes in pan cancers. We compared these 6 subtypes with our C1 and C2 subtypes. Similarly, we can observe that there are significant differences in the distribution of 6 different immune subtypes in C1 and C2 subtypes in this study (Fig. 4G). Moreover, in the modulation of TME, fatty acid-related lncRNAs played a nonnegligible role.

Wu *et al. Biology Direct*   (2022) 17:19

Page 7 of 14

Fig. 3 The enrichment of hallmark pathways in two subtypes. **A** Differential enriched pathways between C1 and C2 subtypes. NES was shown by comparing C2 with C1. Orange indicates relatively activated pathways and purple indicates relatively suppressed pathways in C2 subtype. **B, C** Rader plots of 15 pathways that were activated in C2 subtype in both TCGA-COAD (**B**) and GSE17538 (**C**) datasets. NES was shown between -4 to 4 from inner to outside circles

Fatty acid-related lncRNAs were involved in regulating oncogenic pathways

Given that two molecular subtypes based on fatty acid-related lncRNAs displayed distinct molecular features

and prognosis, we suspected that these lncRNAs may be highly involved in delivering tumor-promotive signals in COAD. LncRNAs are considered as important regulators in controlling gene expression. No surprisingly, we

Wu *et al. Biology Direct* (2022) 17:19

Page 8 of 14

Fig. 4 TME features of C1 and C2 subtypes in TCGA-COAD dataset. **A** Estimated proportions of 24 immune cells. **B** Stromal score, immune score and ESTIMATE score calculated by ESTIMTAE analysis. **C** The expression of immune checkpoints in two subtypes. **D, E** T cell dysfunction and exclusion score calculated by TIDE analysis. Wilcoxon test was conducted. **F** The intersection of the molecular subtypes in this study with the four previously reported molecular subtypes. **G** The intersection of the molecular subtypes of this study with the six previously reported immune molecular subtypes

observed that the expression of protein-coding genes (PCGs) was highly consistent with the expression of fatty acid-related lncRNAs, with a higher proportion of positive correlations (Fig. 5A). To understand the localization of fatty acid-related lncRNAs which largely decided their function, LncATLAS database was utilized for calculating their RCI [31]. The result showed that fatty acid-related lncRNAs were localized more in the nucleus (RCI < 0) in both TCGA-COAD and GSE17538 datasets, with a proportion of 58.98% and 63.50% respectively (Fig. 5B).

LncRNAs regulating gene expression through suppressing or activating transcription factors (TFs) is one of the main mechanism. Next we analyzed TF activity of C1 and C2 subtypes according to the algorithm developed by Garcia-Alonso et al. [32], and identified 108 and 76 dysregulated TFs in TCGA-COAD and GSE17538 datasets respectively. We found that these dysregulated TFs were mostly activated (Fig. 5C), and five TFs were found to be upregulated in C2 subtype in both two datasets including FOS, HIF1A, JUN, STAT3 and STAT2 ($P < 0.0001$,
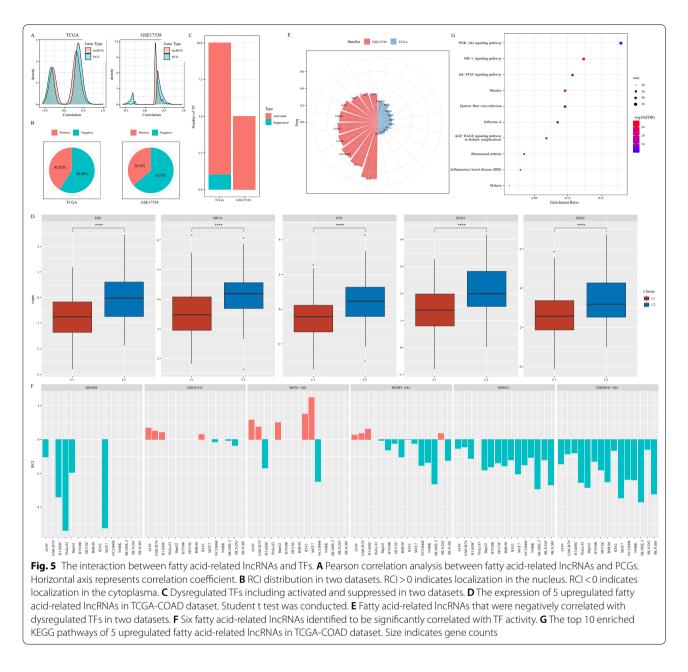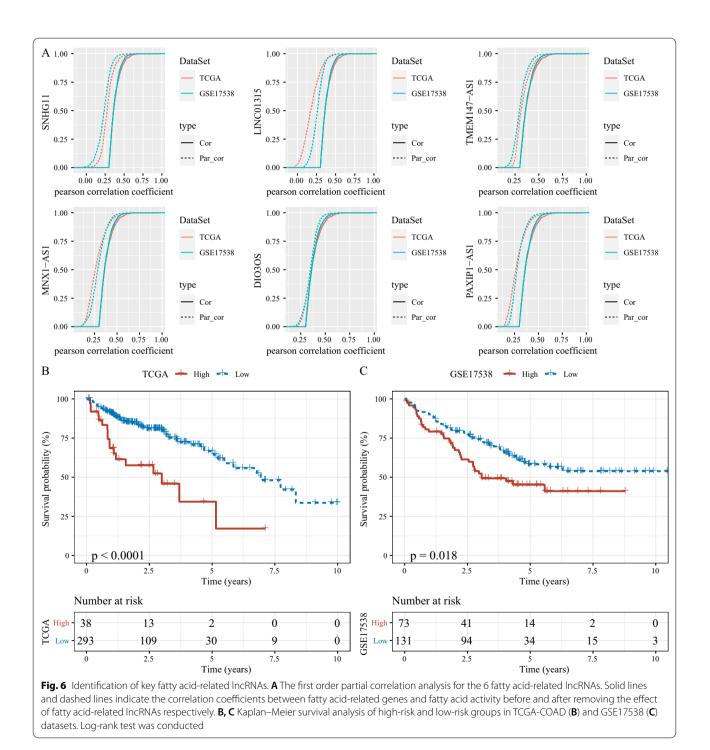
Wu *et al. Biology Direct*   (2022) 17:19

Page 9 of 14

Fig. 5 The interaction between fatty acid-related lncRNAs and TFs. **A** Pearson correlation analysis between fatty acid-related lncRNAs and PCGs. Horizontal axis represents correlation coefficient. **B** RCI distribution in two datasets. RCI > 0 indicates localization in the nucleus. RCI < 0 indicates localization in the cytoplasma. **C** Dysregulated TFs including activated and suppressed in two datasets. **D** The expression of 5 upregulated fatty acid-related lncRNAs in TCGA-COAD dataset. Student t test was conducted. **E** Fatty acid-related lncRNAs that were negatively correlated with dysregulated TFs in two datasets. **F** Six fatty acid-related lncRNAs identified to be significantly correlated with TF activity. **G** The top 10 enriched KEGG pathways of 5 upregulated fatty acid-related lncRNAs in TCGA-COAD dataset. Size indicates gene counts

Fig. 5D). Based on Pearson correlation analysis between these dysregulated TFs and nucleus-localized fatty acid-related lncRNAs, a group of TFs negatively associated with the lncRNAs were identified ($P < 0.05$, Fig. 5E). Within 100 fatty acid-related lncRNAs, 6 lncRNAs in the nucleus were identified to be significantly associated with TF activity, including DIO3OS, LINC01315, MNX1-AS1, PAXIP1-AS1, SNHG11, and TMEM147-AS1 (Fig. 5F). We inferred that these 6 lncRNAs may be greatly involved in regulating activated TFs observed in C2 subtype. Moreover, to better understand the function of the 5 upregulated TFs in C2 subtype, KEGG pathways were

annotated. Consequently, three tumor-related pathways were identified within the top 10 significantly enriched pathways, including PI3K-Akt signaling, HIF-1 signaling and JAK-STAT signaling pathways (Fig. 5G).

Identification of key fatty acid-related lncRNAs

To identify key lncRNAs that were involved in regulating fatty acid metabolism, we applied the first order partial correlation analysis among fatty acid-related genes, fatty acid activity and the identified 6 fatty acid-related lncRNAs that were associated with TF activity. As a consequence, 5 fatty acid-related lncRNAs were screened

in two datasets (Fig. 6A). The correlation between fatty acid activity and fatty acid-related genes was significantly weakened when removing these fatty acid-related lncRNAs, indicating that these lncRNAs played important roles in the signals related to fatty acid. Functional analysis on these 5 lncRNAs revealed that metabolism-related pathways such as purine metabolism, pyrimidine

metabolism and other glycan degradation were enriched (Additional file 3: Fig. S2). Based on the expression of 5 lncRNAs, high-risk and low-risk groups were defined according to the Optimal cut-off of risk score. Two groups manifested distinct overall survival in both two datasets (Fig. 6B, C), indicating that these 5 lncRNAs
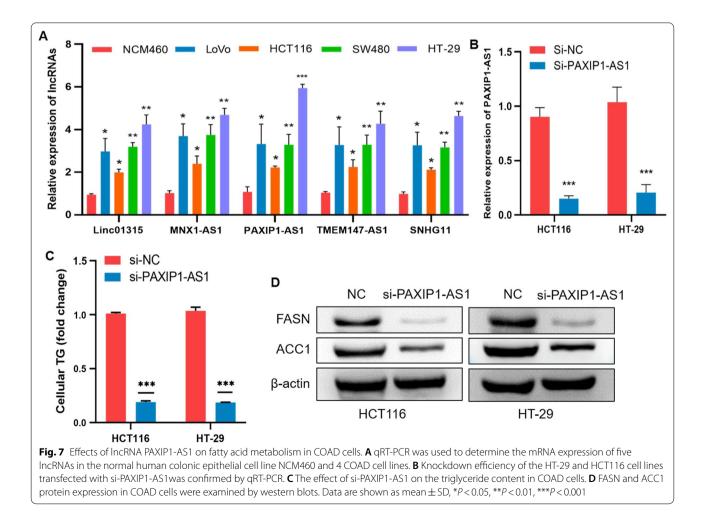

Fig. 6 Identification of key fatty acid-related lncRNAs. **A** The first order partial correlation analysis for the 6 fatty acid-related lncRNAs. Solid lines and dashed lines indicate the correlation coefficients between fatty acid-related genes and fatty acid activity before and after removing the effect of fatty acid-related lncRNAs respectively. **B, C** Kaplan–Meier survival analysis of high-risk and low-risk groups in TCGA-COAD (**B**) and GSE17538 (**C**) datasets. Log-rank test was conducted

could serve as biomarkers to predict prognosis for COAD patients.

Effects of lncRNA PAXIP1-AS1 on fatty acid metabolism of colorectal cancer

To furthermore explore the role of lncRNAs on fatty acid metabolism of colorectal cancer, we determined the expression of five lncRNAs (linc01315, MNX1-AS1, PAXIP1-AS1, SNHG11, TMEM147-AS1) in 4 human COAD cell lines (LoVo, HCT116, SW480, and HT-29) and the normal human colonic epithelial cell line NCM460, the results showed that compared with NCM460 cell line, the mRNA expression levels of five lncRNAs in the 4 COAD cell lines were all up-regulated, of which, HT-29 cell line exhibited the highest expression of five lncRNAs, while HCT116 cell line exhibited the lowest expression of five lncRNAs (Fig. 7A). Therefore, HT-29 and HCT116 cell lines were chosen as candidate cell lines for the following experiment. Among the five lncRNAs, lncRNA PAXIP1-AS1

showed the highest expression, then we explored the role of PAXIP1-AS1 on fatty acid metabolism in COAD cells. qRT-PCR showed the expression of PAXIP1-AS1 were significantly decreased in the si-PAXIP1-AS1 HT-29 and HCT116 cell lines (Fig. 7B). Afterward, intracellular triglyceride content was determined by the triglyceride assay kit, the result showed that knockdown of PAXIP1-AS1 reduced the triglyceride content in COAD cells (Fig. 7C). Consistently, FASN and ACC1, the key enzymes of fatty acid synthesis metabolism, were significantly downregulated in si-PAXIP1-AS1 COAD cells (Fig. 7D). Together, these results suggest that lncRNA PAXIP1-AS1 accelerates the fatty acid metabolism of COAD.

Discussion

In the comparison of metabolism-related pathways between normal and tumor samples, we identified that fatty acid synthesis pathway was significantly dysregulated in COAD samples. Then based on the expression of lncRNAs related to fatty acid, we constructed two

Fig. 7 Effects of lncRNA PAXIP1-AS1 on fatty acid metabolism in COAD cells. **A** qRT-PCR was used to determine the mRNA expression of five lncRNAs in the normal human colonic epithelial cell line NCM460 and 4 COAD cell lines. **B** Knockdown efficiency of the HT-29 and HCT116 cell lines transfected with si-PAXIP1-AS1was confirmed by qRT-PCR. **C** The effect of si-PAXIP1-AS1 on the triglyceride content in COAD cells. **D** FASN and ACC1 protein expression in COAD cells were examined by western blots. Data are shown as mean ± SD, *P < 0.05, **P < 0.01, ***P < 0.001

Wu *et al. Biology Direct*　　(2022) 17:19

Page 12 of 14

molecular subtypes that had differential prognosis. C2 subtype had unfavorable prognosis and lower the enrichment of fatty acid synthesis, compared with C1 subtype. High activity score of fatty acid seemed to play a protective role in COAD.

Fatty acid synthesis is an essential part of lipid metabolism, which can convert nutrients into metabolic components for synthesizing biological membranes, storing energy and generating signaling molecules [36]. Fatty acid synthesis has been demonstrated to be involved in tumorigenesis and tumor progression in various cancer types. We found that some tumor-related pathways such as VEGF signaling, Toll-like receptor signaling and ECM receptor signaling were associated with high activity of fatty acid synthesis. VEGF is a key mediator of angiogenesis in cancer, which can promote metastasis by activating VEGF signaling. In colorectal cancer, fatty acid was observed to be higher expressed in VEGF-positive tumor tissues, compared to VEGF-negative tumor tissues [37], which was accordant with our study. Inhibiting fatty acid synthase can suppress angiogenesis through modulating VEGF-A expression in glioma cells [38]. Zaytseva et al. also support that knock-down of fatty acid synthase decreases the expression of VEGF and thus suppress angiogenesis in colorectal cancer [39].

Toll-like receptors (TLRs) contributes to activate immune response in inflammation-induced diseases and cancers, which is stimulated by saturated fatty acids [40, 41]. Overexpression of TLR4 indicates progression and poor prognosis in colon cancer [42]. The previous findings suggest that high expression of fatty acid can promote cancer progression through activating VEGF signaling and Toll-like receptor signaling, however, which seems inconsistent to our observation that C1 subtype had higher activity of fatty acid synthesis and higher enrichment of these pathways.

We put our sight into TME, one of critical hallmarks of cancer, which determines the prognosis to some extent. C2 subtype had significantly higher immune infiltration but also higher expression of immune checkpoints than C1 subtype, which may result in its worse prognosis. Especially, high expression of LAG3, CTLA4, PDCD1 (PD-1), IDO1, and CD274 (PD-L1) were widely reported to be associated with suppressive immune response and suppressed T cell function in cancer. Immune checkpoint blockade such as anti-CTLA-4 and anti-PD-1/PD-L1 have been considered as promising strategies for treating various cancer types also including COAD [43].

An interaction between fatty acid metabolism and TME has been reported in a number of studies. In glioblastoma, Tregs obtain energy relying on fatty acid oxidation (FAO) under hypoxic conditions [44]. Within the process, hypoxia-inducible factor 1α (HIF-1α) is an inducer to drive Tregs depending on lipid oxidation in mitochondrial metabolism. M2 macrophages, which are associated with cancer progression and invasion secret increased IL-1β under conditions of FAO dependence and HIF-1α upregulation [45]. Therefore, fatty acid metabolism plays an important role in TME alternation.

To further reveal the mechanism of fatty acid metabolism in cancer development, the association between fatty acid-related lncRNAs and TFs were analyzed. We identified six fatty acid-related lncRNAs that had close association with dysregulated TFs, including DIO3OS, LINC01315, MNX1-AS1, PAXIP1-AS1, SNHG11, and TMEM147-AS1. Furthermore, five TFs were found to be significantly upregulated in C2 subtype in both two datasets, including FOS, HIF1A, JUN, STAT3 and STAT2. HIF1A gene expressing HIF-1α contributes to hypoxia-induced FAO dependence, which is consistent with the previous findings [44, 45]. STAT3 and STAT2 belong to STAT family and are involved in STAT signaling, which are highly associated with TME and cancer progression [46]. Especially, STAT3 signaling induces the upregulation of immunosuppressive factors such as TGF-β and PD-1/PD-L1, and promotes angiogenesis through upregulating HIF-1α and VEGF [47]. Higher expression of STAT2 and STAT3 may lead to the worse prognosis of C2 subtype. However, the relation between fatty acid-related lncRNAs and these TFs except for HIF1A remains to be unclear, and it needs exploring in the further study.

To validate the function of the lncRNAs in the fatty acid metabolism-related signature we constructed, lncRNA PAXIP1-AS1 was selected to conduct functional experiments in COAD cells. The results showed that knockdown of PAXIP1-AS1 decreased the content of triglyceride in COAD cells. As we know, FASN and ACC1were the key enzymes involving in the fatty acid synthesis metabolism, both of them modulate the lipogenesis, growth, and apoptosis of COAD cells [39, 48]. Besides, the upregulation of FASN is correlated with the metastasis in COAD [49, 50]. We then explored the expression of FASN and ACC1 in si-PAXIP1-AS1 COAD cells, the result showed the expression of FASN and ACC1 were decreased when PAXIP1-AS1 expression was downregulated. Taken together, this study demonstrates that PAXIP1-AS1 plays a key role in the fatty acid metabolism of COAD.

Conclusions

In conclusion, based on comparing the molecular features between C1 and C2 subtypes, we considered that fatty acid-related lncRNAs were involved in modulating fatty acid synthesis and TME, and thus contributes

Wu *et al. Biology Direct*   (2022) 17:19

Page 13 of 14

to cancer progression. A possible relation was identified between fatty acid-related lncRNAs and immunosuppressive factors such as HIF-1α and STAT. The five identified TFs may be new targets for exploiting new molecular drugs. Moreover, five fatty acid-related lncRNAs were screened to establish a prognostic signature for predicting COAD prognosis.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13062-022-00332-y.

Additional file 1. Supplementary Fig. S1: Workflow and TME features of C1 and C2 subtypes in GSE17538 dataset. **A** Work flow chart. **B** Estimated proportions of 24 immune cells. **C** Stromal score, immune score and ESTIMATE score calculated by ESTIMTAE analysis. **D** The expression of immune checkpoints in two subtypes. **E, F** T cell dysfunction and exclusion score calculated by TIDE analysis. Wilcoxon test was conducted.

Additional file 2. Supplementary Table S1: Primer Sequences used in qRT-PCR.

Additional file 3. Supplementary Fig. S2: Functional analysis of 5 key fatty acid-related lncRNAs.

Author contributions
ZW and JJL designed the research. SW, YZG, JFC, XZ, HMQ, YD and SSL collected and analyzed data. SW, YZG and JFC drafted the manuscript. ZW and JJL reviewed and revised the manuscript. All authors read and approved the final manuscript.

Availability of data and materials
TCGA-COAD dataset containing RNA-seq data (FPKM) and clinical information was downloaded from The Cancer Genome Atlas (TCGA) database in October 01, 2021. GSE17538 dataset was obtained from Gene Expression Omnibus (GEO) database in October 01, 2021.

Declarations

Ethics approval and consent to participate
For analyses of de-identifed data from the TCGA and GEO databases, institutional review board approval and informed consent were not required.

Consent for publication
Not applicable.

Competing interests
The authors have no conflicts of interest to declare.

Author details
[1]Department of Oncology and Southwest Cancer Center, Southwest Hospital, Third Military Medical University (Army Medical University), Chongqing 400038, China. [2]Key Laboratory of Tumor Immunopathology of Ministry of Education of China, Institute of Pathology and Southwest Cancer Center, Southwest Hospital, Third Military Medical University (Army Medical University), Chongqing 400038, China.

References
1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Stein A, Folprecht G. Immunotherapy of colon cancer. Oncol Res Treat. 2018;41(5):282–5.
3. Wang ZX, Cao JX, Liu ZP, Cui YX, Li CY, Li D, et al. Combination of chemotherapy and immunotherapy for colon cancer in China: a meta-analysis. World J Gastroenterol. 2014;20(4):1095–106.
4. Gavin PG, Colangelo LH, Fumagalli D, Tanaka N, Remillard MY, Yothers G, et al. Mutation profiling and microsatellite instability in stage II and III colon cancer: an assessment of their prognostic and oxaliplatin predictive value. Clin Cancer Res. 2012;18(23):6531–41.
5. Dienstmann R, Mason MJ, Sinicrope FA, Phipps AI, Tejpar S, Nesbakken A, et al. Prediction of overall survival in stage II and III colon cancer beyond TNM system: a retrospective, pooled biomarker study. Ann Oncol. 2017;28(5):1023–31.
6. Sadeghi A, Roudi R, Mirzaei A, Zare Mirzaei A, Madjd Z, Abolhasani M. CD44 epithelial isoform inversely associates with invasive characteristics of colorectal cancer. Biomark Med. 2019;13(6):419–26.
7. Roudi R, Barodabi M, Madjd Z, Roviello G, Corona SP, Panahei M. Expression patterns and clinical significance of the potential cancer stem cell markers OCT4 and NANOG in colorectal cancer patients. Mol Cell Oncol. 2020;7(5):1788366.
8. Ganapathy-Kanniappan S. Molecular intricacies of aerobic glycolysis in cancer: current insights into the classic metabolic phenotype. Crit Rev Biochem Mol Biol. 2018;53(6):667–82.
9. Wen J, Min X, Shen M, Hua Q, Han Y, Zhao L, et al. ACLY facilitates colon cancer cell metastasis by CTNNB1. J Exp Clin Cancer Res. 2019;38(1):401.
10. Lyssiotis CA, Kimmelman AC. Metabolic interactions in the tumor microenvironment. Trends Cell Biol. 2017;27(11):863–75.
11. Anastasiou D. Tumour microenvironment factors shaping the cancer metabolism landscape. Br J Cancer. 2017;116(3):277–86.
12. Stine ZE, Schug ZT, Salvino JM, Dang CV. Targeting cancer metabolism in the era of precision oncology. Nat Rev Drug Discov. 2022;21(2):141–62.
13. Machala M, Procházková J, Hofmanová J, Králiková L, Slavík J, Tylichová Z, et al. Colon cancer and perturbations of the sphingolipid metabolism. Int J Mol Sci. 2019;20(23):6051.
14. Piccinin E, Cariello M, Moschetta A. Lipid metabolism in colon cancer: role of liver X receptor (LXR) and stearoyl-CoA desaturase 1 (SCD1). Mol Asp Med. 2021;78:100933.
15. Zeng H, Umar S, Rust B, Lazarova D, Bordonaro M. Secondary bile acids and short chain fatty acids in the colon: a focus on colonic microbiome, cell proliferation, inflammation, and cancer. Int J Mol Sci. 2019;20(5):1214.
16. Jiang C, Liu Y, Wen S, Xu C, Gu L. In silico development and clinical validation of novel 8 gene signature based on lipid metabolism related genes in colon adenocarcinoma. Pharmacol Res. 2021;169:105644.
17. Liu H, Luo J, Luan S, He C, Li Z. Long non-coding RNAs involved in cancer metabolic reprogramming. Cell Mol Life Sci. 2019;76(3):495–504.
18. Ghafouri-Fard S, Shoorei H, Taheri M. The role of long non-coding RNAs in cancer metabolism: a concise review. Front Oncol. 2020;10:555825.
19. Sobhani N, Chahwan R, Roudi R, Morris R, Volinia S, Chai D, et al. Predictive and prognostic value of non-coding RNA in breast cancer. Cancers (Basel). 2022;14(12):2952.
20. Tang J, Yan T, Bao Y, Shen C, Yu C, Zhu X, et al. LncRNA GLCC1 promotes colorectal carcinogenesis and glucose metabolism by stabilizing c-Myc. Nat Commun. 2019;10(1):3499.
21. Chen J, Alduais Y, Zhang K, Zhu X, Chen B. CCAT1/FABP5 promotes tumour progression through mediating fatty acid metabolism and stabilizing PI3K/AKT/mTOR signalling in lung adenocarcinoma. J Cell Mol Med. 2021;25(19):9199–213.
22. Luo Y, Huang S, Wei J, Zhou H, Wang W, Yang J, et al. Long noncoding RNA LINC01606 protects colon cancer cells from ferroptotic cell death and promotes stemness by SCD1-Wnt/β-catenin-TFE3 feedback loop signalling. Clin Transl Med. 2022;12(4):e752.

Wu *et al. Biology Direct*   (2022) 17:19

Page 14 of 14

23. Song C, Qi Y, Zhang J, Guo C, Yuan C. CDKN2B-AS1: an indispensable long non-coding RNA in multiple diseases. Curr Pharm Des. 2020;26(41):5335–46.
24. Li B, Ruotti V, Stewart RM, Thomson JA, Dewey CN. RNA-seq gene expression estimation with read mapping uncertainty. Bioinformatics. 2010;26(4):493–500.
25. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26(12):1572–3.
26. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.
27. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417–25.
28. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–7.
29. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.
30. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24(10):1550–8.
31. Mas-Ponte D, Carlevaro-Fita J, Palumbo E, Hermoso Pulido T, Guigo R, Johnson R. LncATLAS database for subcellular localization of long noncoding RNAs. RNA. 2017;23(7):1080–7.
32. Garcia-Alonso L, Iorio F, Matchan A, Fonseca N, Jaaks P, Peat G, et al. Transcription factor activities enhance markers of drug sensitivity in cancer. Cancer Res. 2018;78(3):769–80.
33. Reverter A, Chan EK. Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. Bioinformatics. 2008;24(21):2491–7.
34. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. Immunity. 2018;48(4):812-30.e14.
35. Liu Y, Sethi NS, Hinoue T, Schneider BG, Cherniack AD, Sanchez-Vega F, et al. Comparative molecular analysis of gastrointestinal adenocarcinomas. Cancer Cell. 2018;33(4):721-35.e8.
36. Röhrig F, Schulze A. The multifaceted roles of fatty acid synthesis in cancer. Nat Rev Cancer. 2016;16(11):732–49.
37. Yang K, Li H, Dong J, Dong Y, Wang CZ. Expression profile of polyunsaturated fatty acids in colorectal cancer. World J Gastroenterol. 2015;21(8):2405–12.
38. Zhou Y, Jin G, Mi R, Zhang J, Zhang J, Xu H, et al. Inhibition of fatty acid synthase suppresses neovascularization via regulating the expression of VEGF-A in glioma. J Cancer Res Clin Oncol. 2016;142(12):2447–59.
39. Zaytseva YY, Elliott VA, Rychahou P, Mustain WC, Kim JT, Valentino J, et al. Cancer cell-associated fatty acid synthase activates endothelial cells and promotes angiogenesis in colorectal cancer. Carcinogenesis. 2014;35(6):1341–51.
40. Li TT, Ogino S, Qian ZR. Toll-like receptor signaling in colorectal cancer: carcinogenesis to cancer therapy. World J Gastroenterol. 2014;20(47):17699–708.
41. Hwang DH, Kim JA, Lee JY. Mechanisms for the activation of Toll-like receptor 2/4 by saturated fatty acids and inhibition by docosahexaenoic acid. Eur J Pharmacol. 2016;785:24–35.
42. Cammarota R, Bertolini V, Pennesi G, Bucci EO, Gottardi O, Garlanda C, et al. The tumor microenvironment of colorectal cancer: stromal TLR-4 expression as a potential prognostic marker. J Transl Med. 2010;8:112.
43. Passardi A, Canale M, Valgiusti M, Ulivi P. Immune checkpoints as a target for colorectal cancer treatment. Int J Mol Sci. 2017;18(6):1324.
44. Miska J, Lee-Chang C, Rashidi A, Muroski ME, Chang AL, Lopez-Rosas A, et al. HIF-1α is a metabolic switch between glycolytic-driven migration and oxidative phosphorylation-driven immunosuppression of Tregs in glioblastoma. Cell Rep. 2019;27(1):226-237.e4.
45. Zhang Q, Wang H, Mao C, Sun M, Dominah G, Chen L, et al. Fatty acid oxidation contributes to IL-1β secretion in M2 macrophages and promotes macrophage-mediated tumor cell migration. Mol Immunol. 2018;94:27–35.
46. Verhoeven Y, Tilborghs S, Jacobs J, De Waele J, Quatannens D, Deben C, et al. The potential and controversy of targeting STAT family members in cancer. Semin Cancer Biol. 2020;60:41–56.
47. Zou S, Tong Q, Liu B, Huang W, Tian Y, Fu X. Targeting STAT3 in cancer immunotherapy. Mol Cancer. 2020;19(1):145.
48. Zhan Y, Ginanni N, Tota MR, Wu M, Bays NW, Richon VM, et al. Control of cell growth and survival by enzymes of the fatty acid synthesis pathway in HCT-116 colon cancer cells. Clin Cancer Res. 2008;14(18):5735–42.
49. Zaytseva YY, Rychahou PG, Gulhati P, Elliott VA, Mustain WC, O'Connor K, et al. Inhibition of fatty acid synthase attenuates CD44-associated signaling and reduces metastasis in colorectal cancer. Cancer Res. 2012;72(6):1504–17.
50. Zaytseva YY, Harris JW, Mitov MI, Kim JT, Butterfield DA, Lee EY, et al. Increased expression of fatty acid synthase provides a survival advantage to colorectal cancer cells via upregulation of cellular respiration. Oncotarget. 2015;6(22):18891–904.