The tsRNAs (tRFdb-3013a and tRFdb-3013b) Serve as Novel Biomarkers for Colon Adenocarcinomas

Xiaoling Wu
Chongqing Medical University  https://orcid.org/0000-0002-2534-8375

Lihong Tan
Chongqing Medical and Pharmaceutical College

Zhurong Tang
Chongqing Medical University

Chunjie Wen
Chongqing Medical University

Huan Chen
Chongqing Medical University

Weiguo Cao
Chongqing Medical University

Hecun Zou (✉ zouhecun@cqmu.edu.cn)
Chongqing Medical University  https://orcid.org/0000-0001-7902-7385

Research

Keywords: Postoperative ileus, Ileostomy closure, Diverting loop ileostomy, Risk factor, Rectal cancer

DOI: https://doi.org/10.21203/rs.3.rs-812336/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background: Postoperative ileus is one of the most common complications after diverting loop ileostomy closure. Some reports have investigated the risk factors for postoperative complications or ileus after ileostomy closure; however, these studies did not evaluate the index surgery sufficiently. In this study, we evaluated the risk factors, including the details of the index surgery, for ileus after diverting ileostomy closure.

Methods: This was a retrospective study of patients who underwent ileostomy closure following index surgery for rectal cancer. Patients who developed postoperative ileus (POI (+)) and patients who did not (POI (−)) after ileostomy closure were compared.

Results: Fifty-nine patients were evaluated and were divided into two groups: POI (+) (n=9) and POI (−) (n=50), and the groups were compared. There were no significant differences in the details of the index surgery, operative procedure, transanal total mesorectal excision, lateral lymph node dissection, operating time, or blood loss. The incidence of Clavien–Dindo grade ≥ III complications after the index surgery was significantly higher in the POI (+) group.

Conclusions: The incidence of Clavien–Dindo grade ≥ III complications after the index surgery may increase the risk of postoperative ileus after ileostomy closure.

Introduction

The tsRNAs (tRNA-derived small RNAs) are novel class of small non-coding RNAs derived from transfer-RNAs (tRNAs), and that also known as tRNA-derived fragments (tRFs) [1]. Identification and characterization of tsRNAs are accelerated greatly in the last few years, due in large part to the emergence of next generation sequencing (NGS) systems-level approaches to molecular genetics [2]. The tsRNA fragments are produced through cleavage of mature or premature tRNAs at various positions, the fragments are short and range in length from 16 to 35 nucleotides [3]. A general classification of tsRNAs is based on their position in relation to the parental tRNA sequence, which can be classified into several distinct categories, including 5′-tRFs, 3′-tRFs, and internal tRFs, these tsRNAs originate precisely from the span of mature tRNAs [3, 4]. Two remaining tsRNAs: 5′U-tRFs, which derived from the 5′ leader sequence of primary tRNA genes; and tRF-1 that comprise part of the 3′ trailer sequence [3]. Numerous researches suggest that the production of tRNAs into specific small fragments is a regimented process, do not result from degradation [5]. These cleavages of tRNAs are context-dependent, including sex, ethnicity, disease, and tissue type of an individual. The dependencies increase an urgency to understand the regulatory roles of tsRNAs. Such efforts are gaining momentum, and comprise experimental and computational approaches [2, 6].

Analysis of the datasets in the Cancer Genome Atlas (TCGA) have discriminated more than thousand microRNA isoforms [7], it is worth mentioning that many of the tsRNAs are sequenced at an abundance comparable to that of many abundant miRNAs. Subsequent efforts attempted to classify and tabulate
the tsRNAs that were being discovered in some higher organisms into a curated repository, and structured the tRFdb database [8]. Several studies have linked the tsRNAs to different human cancers, including colorectal [9], liver, and breast cancer [10], etc. Growing evidences demonstrated that many tsRNAs were differentially expressed in the tissues of multiple cancer patients, these dysregulated tsRNA fragments might serve as potential candidates for diagnostic and prognostic biomarkers, and even contribute to various biological processes of cancer development and progression [11]. Colon adenocarcinoma (COAD) are the most common type of primary tumor originating in the intestinal tract [12]. According to cancer statistics, mortality due to colon adenocarcinomas is the highest among gastrointestinal tumors [13]. The pathogenesis of colon adenocarcinomas is extremely complicated, which involves the aberrant activation of proto-oncogenes and inactivation of anti-oncogenes [12, 14]. However, it is not clear about whether the tsRNAs would serve as novel molecular biomarkers for diagnosis, prognosis and target therapy of colon adenocarcinomas.

In the present study, we aimed to identify the tsRNAs profiles in the sncRNA-seq datasets of human colon adenocarcinomas samples via a deterministic and exhaustive tsRNAs mining pipeline, then investigate the expression pattern and biological function of tRFdb-3013a and tRFdb-3013b within colon adenocarcinomas. Additionally, the relationship between tRFdb-3013a/b and its target genes are also studied in order to reveal the underlying mechanisms.

Materials And Methods

Data collection

The sequences of human tsRNAs were taken from tRFdb database (http://genome.bioch.virginia.edu/trfdb/) and tRFexplorer database (https://trfexplorer.cloud/ ), including different types of tsRNA fragments: 5′-tRFs, 3′-tRFs, 5′U-tRFs and tRF-1 [8, 15]. Then a custom annotation of the reference human genome (hg19) containing only known tsRNAs was assembled. The hg19 gene annotations and corresponding sequences of tRNAs in humans were taken from GtRNAdb (genomic tRNA database, http://gtm adb.ucsc.edu/) [16]. The small non-coding RNA sequencing (sncRNA-seq) datasets of colon adenocarcinoma (COAD) samples on TCGA, the RNA sequencing data of TCGA-COAD samples, and the corresponding patient clinical information were retrieved from GDC data portal (https://portal.gdc.cancer.gov/ ) [17]. The primary clinical and molecular pathology characteristics of COAD patients were listed in Table S4. The gene expression profiles of GSE39582 were downloaded from the Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/ ) database, the gene expression microarrays are based on Affymetrix Human Genome U133 Plus 2.0 Array platform (Affymetrix, USA) [18].

Identification pipeline

The data processing and flow chart for tsRNAs identification are shown in Fig. 1A. a conservative pipeline was implemented to get an accurate estimation of tsRNAs expression as described previously [15]. The adaptor trimming and quality filtering of raw FASTQ files on sncRNA-seq data were pre-processed via
using Trim Galore (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/), which is a wrapper tool for Cutadapt and FastQC method. Then, clean reads of small RNA sequencing were re-mapped to the reference human genome (GRCh37 /hg19) and the sequences of our tsRNAs annotation file by applying TopHat2 tool [19]. After alignment, only the mapped reads could be quantified to count the number of reads belonging to each of the candidate tsRNAs with HTSeq software [20], other ambiguous reads were removed for a more accurate and conservative analysis. Finally, the expression value of tsRNAs were calculated and normalized as reads per million (RPM) of total raw counts [21], and the tsRNAs with average expression level less than one log2 RPM were filtered to eliminate random degradation sequences. Data analysis was performed with R statistical software (https://www.r-project.org/).

**Bioinformatic analysis and visualization**

The RNA expression profiles of TCGA datasets were downloaded and prepared using TCGAbioliinks package of Bioconductor R [22]. The differential expression analysis between colon adenocarcinoma samples and non-tumor controls was performed with the Limma package [23]. An absolute log2 fold-change (llogFC)] more than one and \(P\)-value less than 0.05 were set as cut-off point. Hierarchical clustering of differentially expressed tsRNA genes was performed based on gene expression values of each sample to verifying the difference between colon adenocarcinomas and non-tumors [24]. Visualization of the identified tsRNAs including volcano plots and heat-map was performed with the ggplot2 [25] and pheatmap packages of R (https://cran.r-project.org/web/packages/pheatmap/), respectively. Survival analysis with Kaplan-Meier curves plus the log-rank test were performed and visualized via the R survival and survminer package [26]. Hierarchical cluster analysis was performed on tsRNAs expression and primary pathology characteristics, its corresponding heat-map was visualized by pheatmap package. *Pearson* correlation was estimated for each pair of tsRNA and candidate mRNA, its scatter diagram was plotted by ggplot2 and ggpubr package in R software [25]. Enrichment analyses were performed to provide biological insights concerning tRFdb-3013a/b-related genes, the molecular signatures database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/) were used as functional gene sets for gene set enrichment analysis (GSEA) [27]. Gene ontology (GO), KEGG pathway, and GSEA analysis were performed and visualized via clusterProfiler and enrichplot package (https://yulab-smu.github.io/clusterProfiler-book/) [28]. The computational binding interactions between tsRNA and target genes were predicted via tRFtarget database with RNAhybrid and IntaRNA tools [29]. The interaction relationships of top 50 targets and tRFdb-3013a/b was organized as a network using Cytoscape (v3.7.2).

**Cell culture and transfections**

Human colon adenocarcinoma cell lines SW480 and Caco-2 were purchased from the Cell Bank of Chinese Academy of Sciences (Shanghai, China). The cells were cultured in Leibovitz’s L-15 medium or Eagle’s Minimum Essential Medium (MEM; Gibco) supplemented with 10% or 20% fetal bovine serum (FBS; Gibco) at 37°C in a humidified incubator with 5 % CO2. The cell lines at 50–60% confluences were transfected with micromolar either mimic for tRFdb-3013a/b (tRF-3013a mimic, tRF-3013b mimic), or
inhibitor for tRFdb-3013a/b (tRF-3013a inhibitor), or negative controls (NC mimic, NC inhibitor) using the Lipofectamine RNAiMAX reagent (Invitrogen, USA). The mimic and inhibitor for tRFdb-3013a/b and negative control were purchased from RiboBio (Guangzhou, China).

**RNA extraction and qRT-PCR analysis**

Total RNA was extracted from tissues or cultured cells using the TRIzol reagent (Invitrogen, USA). One microgram total RNA of each sample was reverse transcribed into cDNA under standard conditions by using PrimeScript RT reagent Kit with gDNA Eraser (Takara, Japan). Quantitative real-time polymerase chain reaction (qRT-PCR) was performed with the SYBR® Premix DimerEraser™ (Takara, Japan) on the LightCycler® 480 system (Roche Diagnostics, Switzerland). The Bulge-Loop miRNA qRT-PCR Stater Kit (Ribobio, China) with specific stem-loop RT primers was used to quantify the expression of tsRNAs, following the manufacturer's protocol. RNU6 (U6) was used as the internal control for tsRNAs template normalization and ACTB (β-actin) for mRNA template normalization. Relative quantification of gene expression was calculated by the comparative cycle-threshold (CT) (2-ΔΔCT) method [24]. Stem-loop primers for tRFdb-3013a, tRFdb-3013b, and RNU6 were designed and synthesized by Ribobio Biotech (Guangzhou, China). The general primers for ST3GAL1 and ACTB were designed and synthesized by Sangon Biotech (Shanghai, China).

**Protein extraction and western blot**

Total protein was extracted from cultured cells using RIPA buffer (Beyotime, China) and the protein concentration was determined using a BCA Protein Assay Kit (Bio-Rad). The specific primary antibody (anti-ST3GAL1 antibody diluted at 1:500, Santa Cruz, USA; anti-β-actin antibody diluted at 1:1000, Sigma-Aldrich, USA). The β-actin served as an endogenous protein for normalization. The protein bands were visualized using ChemiDocTM XRS + imaging system and Image Lab software (Bio-Rad, California, USA).

**Cell proliferation, migration and invasion assays**

Cell proliferation assays were performed with Cell Counting Kit-8 (CCK-8; Dojindo Laboratories, Japan), the absorbance (A) in each well was measured at 450 nm with a Microplate Reader (BioTek, VT, USA). Cell migration and invasion ability were tested via the scratch and Transwell assays. The images of scratch gap and invasion cells were recorded and photographed using a Leica microscope (Heidelberg, Germany).

**Luciferase Reporter Assay**

The wild-type or mutant sequences of ST3GAL1 3’-UTR containing putative tRFdb-3013a and tRFdb-3013b binding sites were subcloned into pmir-RB-Report vector (RiboBio, China). SW480 cells were co-transfected with pmir-RB-Report vector plus or without tRFdb-3013a/b. The luciferase activity was measured using the Dual-luciferase Reporter Assay System (Promega, USA).

**Statistical analysis**
The statistical analyses were performed with the R software (https://www.r-project.org/), GraphPad Prism version 8.0 (GraphPad Software, Inc., USA) was used for graphing and analysis. Data were exhibited as means ± standard deviation (SD). The χ² test was used to compare the categorical variables. Regarding the numerical variables, statistical significance of differences between two groups was assessed using two-sided Student's t test; and comparisons of multiple groups were made by one-way analysis of variance (ANOVA). All experiments were performed in triplicate and P values less than 0.05 were considered a statistically significant difference.

Results

The identification of tsRNAs in colon adenocarcinomas datasets.

First of all, we collected more than thousands human tsRNA signatures data from tRFdb (http://genome.bioch.virginia.edu/trfdb/) [8] and tRFexplorer (https://trfexplorer.cloud/), all tsRNA sequences have mapped within the specific regions of tRNA (transfer-RNA) as described previously [15]. As statistical analysis, these tsRNA fragments are mainly derived from 5′ end (5′-tRFs) and 3′ end (3′-tRFs) of mature tRNA, as well as the 5′ leader (5′U-tRF) and 3′ trailer (tRF-1) regions of primary tRNA genes.

With computational approaches, we determined lots of tsRNAs and assessed their expression profile within the sncRNA-seq (small non-coding RNA sequence) data of TCGA-COAD dataset, the summary of our identification pipeline is shown in Fig. 1A. After filtering, a total of 192 tsRNAs with available expression abundance were identified within COAD samples (Table S1), of which about thirty-two percent derived from the 3′ trailer (tRF-1) regions of primary tRNA genes, and thirty percent tsRNAs from the 3′ end (3′-tRFs) of mature tRNA (Fig. 1C). As for the chromosomes, more than thirty-five percent of tsRNAs are originated from the human chr1 region, and nearly twenty percent originated from the chr6 region (Fig. 1B).

Base on the differential expression analysis, we discriminated a total of forty-two differentially expressed tsRNA genes (DEGs), of which fifteen were down-regulated and twenty-seven were up-regulated (Fig. 2A) between colon adenocarcinomas and non-tumor controls. Hierarchical cluster analysis and heat-map of differentially expressed tsRNA genes were performed and exhibited in Fig. 2B. The expression patterns of tsRNAs were significantly distinguished between colon cancers and non-tumor groups. As shown in the blue dotted line frame of Fig. 2B, we identified several tsRNAs (including ts-43, ts-44, tRFdb-3013a and tRFdb-3013b) that derived from tRNA-His-GUG, which is a histidyl-transfer-RNA [30].

Significance of tRFdb-3013a and tRFdb-3013b within colon adenocarcinomas

Among the identified tsRNA fragments, we are interested in the fragments derived from tRNA-His-GUG or tRNA-His-GTG gene. As shown in Fig. 3A, we have identified seven tsRNA fragments in the sncRNA-seq
datasets of colon adenocarcinomas samples, there are two 3′-tRFs fragments (tRFdb-3013a and tRFdb-3013b) can re-map to the cleavage of 3′ end of several different mature tRNA-His-GUG (Fig. 3A and 4A), and five tRF-1 fragments (ts-30, ts-43, ts-44, ts-1 and ts-88) that can re-mapping to the down-stream regions (35 bases) of primary tRNA-His-GTG genes. To reveal the implication of these tsRNAs in colon adenocarcinomas, the expression patterns of above tsRNAs were analyzed in the samples of TCGA-COAD datasets. It showed that the expressions of four tsRNAs (ts-43, ts-44, tRFdb-3013a and tRFdb-3013b) were significantly decreased in colon adenocarcinoma samples compared to paired non-tumor controls (Fig. 3 all P < 0.05), however, no significant difference for other tsRNAs was observed between colon adenocarcinomas and controls (both P > 0.05).

Next, we investigated the relationship between tsRNAs expression and overall survival of colon adenocarcinoma samples using Kaplan–Meier curve analysis with a log-rank comparison. Base on the median value of tsRNAs expression, colon adenocarcinoma samples were divided into low expression group and high expression group. As shown in Fig. 4B, the clinical survival of colon adenocarcinoma patients with low tRFdb-3013a expression was significantly worse than that of the high expression patients (P = 0.029); the survival of patients with low tRFdb-3013b expression was also worse than that of high expression patients (P = 0.026). These results suggested that down-regulated tRFdb-3013a and tRFdb-3013b (known simply as tRFdb-3013a/b) were associated with poor survival prognosis of colon adenocarcinoma patients.

The primary clinical and molecular pathology characteristics of colon adenocarcinoma patients were listed in Table S4. The correlations of tRFdb-3013a/b expression with pathology characteristics were analyzed using Chi-square test, and that were visualized by Hierarchical clustering heatmap. As shown in Fig. 5A-5B, the expression of tRFdb-3013a and tRFdb-3013b was significantly down-regulated in colon adenocarcinoma samples with four stages (both P < 0.05) compared to non-tumor controls. Moreover, tRFdb-3013a/b may tend to down-regulated in the patients with lymphatic or vascular invasion present (Fig. 5A). However, no significant relationship was found between tRFdb-3013a/b expression and other pathology characteristics including gender, grade, MLH1 silencing, methylation subtype, colon polyps and MSI status, etc. (all P > 0.05). These findings implied that tRFdb-3013a and tRFdb-3013b may serve as novel biomarkers for diagnosis and prognosis of colon adenocarcinomas.

**Biological insights of tRFdb-3013a/b on colon adenocarcinoma cells.**

To further explore the effects of tRFdb-3013a/b on colon adenocarcinomas. The mimic or inhibitor of tRFdb-3013a/b was used to over-expressed or inhibit the tRFdb-3013a/b in cells. Cell proliferation of colon adenocarcinomas was determined by CCK-8 assays, it showed that cell proliferation was significantly suppressed in tRFdb-3013a mimic transfected cells compared to negative controls (both P < 0.05; Fig. 6A). Transwell and scratch assays were performed to test the effect of tRFdb-3013a on invasion and migration ability of colon adenocarcinoma cells. The results displayed that migratory rate of cells transfected with tRFdb-3013a mimic was significantly reduced in comparison with negative controls (Fig. 6B); the numbers of invasive cells with tRFdb-3013a mimic were significantly reduced and
that cells with tRFdb-3013a inhibitor were increased compared with those of negative controls (Fig. 6C). These indicated that tRFdb-3013a may be effective to regulate the cell proliferation, cell migration and invasion of colon adenocarcinomas.

Moreover, some enrichment analyses were performed to provide biological insights concerning the role of tRFdb-3013a/b. TCGA-COAD mRNA expression profiling data were used to find some correlated genes of tRFdb-3013a/b via Spearman’s correlation method. As shown in Fig. 7A, for gene ontology (GO) analysis, some tRFdb-3013a-related genes were enriched in the specific GO terms, including extracellular matrix organization, collagen-containing extracellular matrix, and extracellular matrix structural constituent (EMILIN1, ADAT2, FBLN1). The dot plots (Fig. 7B) showed that tRFdb-3013a-related genes were enriched in several KEGG pathway, such as phagosom (C1R). In the GSEA analyses (Fig. 7C), a molecular signature (LMTK3 target genes) was found to associated with tRFdb-3013a, such as FBLN1 and ILVBL-AS1 are both the tRFdb-3013a correlated-genes. The enrichment analyses of tRFdb-3013b were also conducted and presented in supplemental material (Figure S1). These data suggested that tRFdb-3013a/b may play an important role in tumor progression of colon adenocarcinomas.

The tRFdb-3013a/b target and regulate ST3GAL1 expression in colon adenocarcinomas.

By using bioinformatics database and tools, many potential targets of tRFdb-3013a/b were predicted and screened out. The interaction relationship of tRFdb-3013a/b and their target genes were presented in Fig. 8A. Among these targets, we picked out a molecular with strong binding score, ST3GAL1, which is a type unity membrane protein with sialyl-transferase activity [31]. The positions in 3’ untranslated region (3’UTR) of ST3GAL1 mRNA contained a putative binding site of tRFdb-3013a/b (Fig. 8B). Dual-luciferase reporter assay was performed to validate their direct binding affinity, it showed that relative luciferase activity was reduced in the co-transfection of pmirRB-ST3GAL1 wild-type reporter and tRFdb-3013a/b mimic (both P < 0.05, Fig. 9B-left), but no significant difference in that mutant reporter. Moreover, the analysis of TCGA-COAD and GSE39582 datasets showed that ST3GAL1 was highly expressed in colon adenocarcinomas compare to non-tumor controls (Fig. 9A), as was mentioned above, tRFdb-3013a and tRFdb-3013b were decreased in colon adenocarcinomas (Fig. 5B). In addition, the cell experiments demonstrated that expression levels of ST3GAL1 mRNA and protein were down-regulated in colon adenocarcinoma cells transfected with tRFdb-3013a/b mimic (both P < 0.01, Fig. 9B-right and 9C). These results suggested tRFdb-3013a and tRFdb-3013b might directly target ST3GAL1 and regulate ST3GAL1 expression in colon adenocarcinomas.

Discussion

The tsRNAs are novel class of small non-coding RNAs derived from transfer-RNAs (tRNAs), also known as tRNA-derived fragments (tRFs) [1]. Identification and characterization of tsRNAs are accelerated greatly in the last few years, due in large part to the emergence of next generation sequencing (NGS) systems-level approaches to molecular genetics [2]. In the present study, we firstly identified some tsRNA fragments within the sncRNA-seq datasets of colon adenocarcinoma samples, most of tsRNAs were
derived from the 3′ end (3′-tRFs) of mature tRNA or 3′ trailer (tRF-1) regions of primary tRNA genes. Furthermore, a total of forty-two differentially expressed tsRNA genes were discriminated by bioinformatics analysis, of which fifteen were down-regulated and twenty-seven were up-regulated between colon adenocarcinomas and non-tumor controls. Among them, we picked out seven tsRNA fragments derived from tRNA-His-GUG, which is a histidyl-transfer-RNA [30]. There are two 3′-tRFs fragments (tRFdb-3013a, tRFdb-3013b) can re-map to the cleavage of 3′ end of mature tRNA-His-GUG, the expressions of tRFdb-3013a and tRFdb-3013b were significantly decreased in colon adenocarcinomas compared to non-tumor controls. Kaplan–Meier curve analyses demonstrated that clinical survival of colon adenocarcinoma patients with low tRFdb-3013a/b expressions were significantly worse than that of the high expression patients, it suggested that down-regulated tRFdb-3013a and tRFdb-3013b were associated with poor survival prognosis of colon adenocarcinoma patients. Further analyses revealed that tRFdb-3013a/b expression may tend to down-regulated in the patients with lymphatic or vascular invasion present. These findings implied that tRFdb-3013a and tRFdb-3013b serve as novel biomarkers for diagnosis and prognosis of colon adenocarcinomas.

Several studies have linked the tsRNA fragments to different human cancers, including colorectal [9], liver, and breast cancer [10], etc. Growing evidences demonstrated that many tsRNAs were differentially expressed in the tissues or plasma of multiple cancer patients, these dysregulated tsRNA fragments might serve as potential candidates for diagnostic and prognostic biomarkers, and even contribute to various biological processes of cancer development and progression [11]. For example, Tao et.al. have revealed the expression profile of tsRNAs in human colorectal cancer (CRC) specimens via sequencing analyses, and identified a specific tsRNA derived from the internal of mature tRNA-His-GTG, 5′tiRNA-His-GTG (its base sequence is GCCGTGATCGTATAGTGGTTAGTACTCTGCGTTGT), which is up-regulated in CRC tissues [9]. Then, in vitro and in vivo experiments have revealed that inhibition of 5′tiRNA-His-GTG suppressed cancer cell proliferation and colony formation, induced cell apoptosis, and affected tumor growth in nude mice, these indicate that 5′tiRNA-His-GTG may act an oncogenic role in the progression of colorectal cancers [9]. In breast cancer, Mo et.al. have determined the differential expressed tsRNAs between tumor tissues with matched non-tumor controls, and found a tsRNA AS-tDR-001430 was significantly decreased in tissues and serum from breast cancer patients [32]. Moreover, lower expression of AS-tDR-001430 is associated with higher stage progression and lymph node metastasis in breast cancers. AS-tDR-001430 is a 5′-half fragment derived from the internal of mature tRNA-Val-CAC and mapped to chromosome 6p22. In breast cancer cells, overexpression of this tsRNA suppress cells malignant activities, these suggest that tsRNA AS-tDR-001430 may serve as a potential tumor-suppressor in breast cancer [32]. In present study, certain cyto-biology assays were performed to provide biological insights concerning the role of tRFdb-3013a on colon adenocarcinoma cells. It demonstrated that tRFdb-3013a over-expression suppressed cell proliferation, cell migration and invasion of colon adenocarcinomas in vitro. Moreover, the enrichment analyses showed tRFdb-3013a was associated with specific GO terms, such as extracellular matrix organization, collagen-containing extracellular matrix, and extracellular matrix structural constituent (EMILIN1, ADAT2, FBLN1); the tRFdb-3013a-related genes may refer to phagosome (C1R), as well as the GSEA molecular signature genes (FBLN1 and ILVBL-AS1). These
results indicated that down-regulated tRFdb-3013a might act as a suppressive role in cell migration and invasion of colon adenocarcinomas via specific signaling pathway.

Though the biological functions of tsRNAs are complex and require further elucidation, current knowledge of their function have been a key focus on cancer research. Previous researches have revealed that tsRNA fragments could bind to argonaute (AGO) proteins and play a critical role in gene silencing by directly targeting the 3’ untranslated regions (3’UTRs) of some mRNAs in a manner similar to miRNAs [2]. For example, Goodarzi et al. have identified the new class of tsRNAs derived from tRNA(Glu), tRNA(Asp), tRNA(Gly), and tRNA(Tyr) through hypoxic stress induction, this tsRNA fragments can bind and displace the 3’UTRs of YBX1 gene, which encoded a RNA-binding protein, then suppress the stability of multiple oncogenic transcripts such as EIF4EBP1 and AKT1, results in tumor-suppressive and metastasis-suppressive activity in breast cancers [10]. Furthermore, in colorectal cancer cells, the upregulated 5’tiRNA-His-GTG is found to directly target LATS2, which is a key component of the hippo signaling pathway, and inhibits cell proliferation and promotes cell apoptosis by regulating YAP protein activity [9]. Our bioinformatics analyses displayed the interaction relationship of tRFdb-3013a/b and their predicted target genes, and found tRFdb-3013a/b could bind the 3’UTR of ST3GAL1, which is highly expressed in colon adenocarcinomas. ST3GAL1 is a membrane protein with sialyl-transferase activity, which has been reported to promote tumor metastasis and invasion by previous study [31]. Dual-luciferase reporter assay showed that tRFdb-3013a/b could directly target 3’UTR of ST3GAL1 and reduce its relative luciferase activity. In addition, over-expression of tRFdb-3013a/b could down-regulated the expression levels of ST3GAL1 protein and mRNA in colon adenocarcinoma cells. These results suggest tRFdb-3013a and tRFdb-3013b might directly target the 3’UTR of ST3GAL1 and regulate ST3GAL1 expression in colon adenocarcinomas.

Conclusion

Taken together, our investigation identified the tsRNAs profiles, especially several tsRNA fragments derived from tRNA-His-GUG, within the sncRNA-seq datasets of colon adenocarcinoma samples. Among them, tRFdb-3013a and tRFdb-3013b were significantly down-regulated in colon adenocarcinomas, the expressions of tRFdb-3013a/b were associated with clinical survival prognosis of colon adenocarcinoma patients. Over-expression of tRFdb-3013a suppressed cell proliferation, cell migration and invasion of colon adenocarcinomas. Down-regulated tRFdb-3013a/b might act as tumor-suppressive role in colon adenocarcinomas via specific signaling pathway. Additionally, tRFdb-3013a/b could directly target the 3’UTR of ST3GAL1 and regulate ST3GAL1 expression in colon adenocarcinomas. These implied that tRFdb-3013a and tRFdb-3013b may play important role in tumor progression of colon adenocarcinomas, and could be explored as novel diagnostic and prognostic biomarkers for colon adenocarcinomas.

Abbreviations

tRNA
Declarations

Ethics approval and consent to participate

This study was approved by the Ethics Committee of Tokushima University (approval number. 3215-1), and the study was conducted in accordance with the provisions of the Declaration of Helsinki. All patients provided informed consent for the use of their data.

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The authors declare that they have no financial interests.

Authors' contributions

TN, MS and KY were involved in study design and data interpretation. TT, MN, HK, CT, YW, SY and YI were involved in the data analysis. All authors revised the manuscript, approved the manuscript to be published, and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

We thank Jane Charbonneau, DVM, from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.
References

1. Kim HK, Fuchs G, Wang S, Wei W, Zhang Y, Park H, Roy-Chaudhuri B, Li P, Xu J, Chu K, et al. A transfer-RNA-derived small RNA regulates ribosome biogenesis. Nature. 2017;552(7683):57–62.

2. Xie Y, Yao L, Yu X, Ruan Y, Li Z, Guo J. Action mechanisms and research methods of tRNA-derived small RNAs. Signal Transduct Target Ther. 2020;5(1):109.

3. Shi J, Zhang Y, Zhou T, Chen Q. tsRNAs: The Swiss Army Knife for Translational Regulation. Trends Biochem Sci. 2019;44(3):185–9.

4. Schorn AJ, Gutbrod MJ, LeBlanc C, Martienssen R. LTR-Retrotransposon Control by tRNA-Derived Small RNAs. Cell. 2017;170(1):61–71 e11.

5. Su Z, Wilson B, Kumar P, Dutta A. Noncanonical Roles of tRNAs: tRNA Fragments and Beyond. Annu Rev Genet. 2020;54:47–69.

6. Zou H, Wu LX, Tan L, Shang FF, Zhou HH. Significance of Single-Nucleotide Variants in Long Intergenic Non-protein Coding RNAs. Front Cell Dev Biol. 2020;8:347.

7. Telonis AG, Magee R, Loher P, Chervoneva I, Londin E, Rigoutsos I. Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. Nucleic Acids Res. 2017;45(6):2973–85.

8. Kumar P, Mudunuri SB, Anaya J, Dutta A. tRFdb: a database for transfer RNA fragments. Nucleic Acids Res. 2015;43(Database issue):D141–5.

9. Tao EW, Wang HL, Cheng WY, Liu QQ, Chen YX, Gao QY. A specific tRNA half, 5'tiRNA-His-GTG, responds to hypoxia via the HIF1alpha/ANG axis and promotes colorectal cancer progression by regulating LATS2. J Exp Clin Cancer Res. 2021;40(1):67.

10. Goodarzi H, Liu X, Nguyen HC, Zhang S, Fish L, Tavazoie SF. Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. Cell. 2015;161(4):790–802.

11. Yu M, Lu B, Zhang J, Ding J, Liu P, Lu Y. tRNA-derived RNA fragments in cancer: current status and future perspectives. J Hematol Oncol. 2020;13(1):121.

12. Bien J, Lin A. A Review of the Diagnosis and Treatment of Metastatic Colorectal Cancer. JAMA. 2021;325(23):2404–5.

13. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2020. CA Cancer J Clin. 2020;70(3):145–64.

14. Zou H, Wen C, Peng Z, Shao Y, Hu L, Li S, Li C, Zhou HH. P4HB and PDIA3 are associated with tumor progression and therapeutic outcome of diffuse gliomas. Oncol Rep. 2018;39(2):501–10.

15. La Ferlita A, Alaimo S, Veneziano D, Nigita G, Balatti V, Croce CM, Ferro A, Pulvirenti A: Identification of tRNA-derived ncRNAs in TCGA and NCI-60 panel cell lines and development of the public database tRFexplorer. Database (Oxford) 2019, 2019.
16. Chan PP, Lowe TM. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. Nucleic Acids Res. 2016;44(D1):D184–9.

17. Hutter C, Zenklusen JC. The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. Cell. 2018;173(2):283–5.

18. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, et al. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41(Database issue):D991–5.

19. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013;14(4):R36.

20. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166–9.

21. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.

22. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44(8):e71.

23. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.

24. Zou H, Wu LX, Yang Y, Li S, Mei Y, Liu YB, Zhang L, Cheng Y, Zhou HH. IncRNAs PVT1 and HAR1A are prognosis biomarkers and indicate therapy outcome for diffuse glioma patients. Oncotarget. 2017;8(45):78767–80.

25. Wickham H: ggplot2: Elegant Graphics for Data Analysis. In: Use R!, 2nd edn. Cham: Springer International Publishing: Imprint: Springer; 2016: 1 online resource (XVI, 260 pages 232 illustrations, 140 illustrations in color.

26. Peng Z, Chen Y, Cao H, Zou H, Wan X, Zeng W, Liu Y, Hu J, Zhang N, Xia Z, et al. Protein disulfide isomerases are promising targets for predicting the survival and tumor progression in glioma patients. Aging. 2020;12(3):2347–72.

27. Liberzon A, Birger C, Thorvaldsdottir 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. Li N, Shan N, Lu L, Wang Z. tRFtarget: a database for transfer RNA-derived fragment targets. Nucleic Acids Res. 2021;49(D1):D254–60.

30. Veneziano D, Tomasello L, Balatti V, Palamarchuk A, Rassenti LZ, Kipps TJ, Pekarsky Y, Croce CM. Dysregulation of different classes of tRNA fragments in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 2019;116(48):24252–8.
31. Wu X, Zhao J, Ruan Y, Sun L, Xu C, Jiang H. Sialyltransferase ST3GAL1 promotes cell migration, invasion, and TGF-beta1-induced EMT and confers paclitaxel resistance in ovarian cancer. Cell Death Dis. 2018;9(11):1102.

32. Mo D, Jiang P, Yang Y, Mao X, Tan X, Tang X, Wei D, Li B, Wang X, Tang L, et al. A tRNA fragment, 5’-tRNA(Val), suppresses the Wnt/beta-catenin signaling pathway by targeting FZD3 in breast cancer. Cancer Lett. 2019;457:60–73.

**Figures**
Figure 1

A summary of tsRNAs identification and characterization. (A) the flow chart of data processing and tsRNAs identification pipeline. (C) The characterization of tsRNA corresponding derived tRNA gene sources and (B) the chromosome locations statistical analysis.
Figure 2

Differential expression analysis between colon adenocarcinomas and non-tumor controls. (A) Volcano plots of the differentially expressed tsRNAs, including 15 up-regulated and 27 down-regulated tsRNAs. (B) Hierarchical clustering and heat-map of differentially expressed tsRNAs, the up-regulated tsRNAs are clustered in the magenta-shaded areas, and the green-shaded areas indicate down-regulation.
Figure 3

(A) Representative images of the sequence alignments between tsRNA fragments and candidate tRNA sources. (B) The expression patterns of several tRNA-His-GUG derived tsRNA fragments (including tRFdb-3013a, tRFdb-3013b, ts-1, ts-43 and ts-44) in colon adenocarcinomas and paired non-tumor controls of TCGA-COAD datasets.
Figure 4

The clinical analysis of tRFdb-3013a and tRFdb-3013b. (A) Secondary Structure of tRNA-His-GUG and its derived fragments (tRFdb-3013a and tRFdb-3013b). (B) Kaplan–Meier survival curve analysis with a log-rank test based on tRFdb-3013a and tRFdb-3013b expression within colon adenocarcinoma samples (n=587) of TCGA-COAD datasets.
Figure 5

The correlation analysis of tRFdb-3013a/b expression with the primary pathology characteristics. (A) Hierarchical clustering heatmap of tRFdb-3013a/b expression and the pathology characteristics parameters. (B) The scatter plus box plots of tRFdb-3013a and tRFdb-3013b expression between non-tumor controls and colon adenocarcinomas with four stages.
The effects of tRFdb-3013a on cell proliferation, cell migration and invasion of colon adenocarcinomas. (A) CCK-8 assays were performed to measure the cell proliferation of colon adenocarcinoma SW480 cell. (B) Relative gap analysis scratch assay was used to assess the migration ability of SW480 cell. (C) Transwell assays were performed to assess cell invasion of colon adenocarcinomas. ** P < 0.01, * P < 0.05. The tRF-3013a mean as tRFdb-3013a, the NC mean as negative control.
Figure 7

The enrichment analyses of tRFdb-3013a-related genes within TCGA-COAD datasets. (A) The top gene ontology (GO) terms, including biological process, cellular component and molecular function. (B) The top KEGG pathway for tRFdb-3013a-related genes. The correlation scatter-plots of tRFdb-3013a and its correlated-genes (ADAT2, EMILIN1, and C1R). (C) GSEA (gene set enrichment analysis) plots of the molecular signatures, the scatter plots for tRFdb-3013a and its correlated-genes (ILVBL-AS1, FBLN1).
Figure 8

(A) The interaction networks that involved tRFdb-3013a/b and their target genes based on tRFtarget database, the nodes represent target genes and lines represent the interaction relationships. (B) the putative binding sites of tRFdb-3013a/b in the 3'-UTR regions of ST3GAL1 based on bioinformatics prediction.
Figure 9

ST3GAL1 is directly regulated by tRFdb-3013a and tRFdb-3013b in colon adenocarcinomas. (A) The expression patterns of ST3GAL1 within samples of TCGA-COAD and GSE39582 datasets, COAD: colon adenocarcinomas, CRC: colorectal cancers. (B-left) The luciferase reporter assay of colon adenocarcinoma cell co-transfected with pmiR-RB-ST3GAL1 wild-type or mutant report vector, together with tRFdb-3013a/b or NC mimics. (C) Expression levels of ST3GAL1 mRNA (B-right) and protein in colon
adenocarcinomas cells transfected with tRFdb-3013a/b mimics or controls. ** P < 0.01, * P < 0.05. The tRF-3013a/b means as tRFdb-3013a/b, NC means as negative control.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.docx
- TableS1S2.xlsx
- TableS3S4.xlsx