Identification the role of mono-ADP-ribosylation in colorectal cancer by integrated Transcriptome analysis

Shuxian Zhang  
chongqing medical university

Jiale Duan  
Chongqing Medical University

Yanping Yang  
Chongqing Medical University

Hanjuan Gong  
chongqing medical university

Yi Tang  
Chongqing Medical University

Ming Xiao  
Chongqing Medical University

Ming Li  
Chongqing Medical University

Qingshu Li  
Chongqing Medical University

Ya-Lan Wang ( wangyalan@cqmu.edu.cn )  
Chongqing Medical University  https://orcid.org/0000-0001-5252-7259

Research Article

Keywords: colorectal cancer, mono ADP ribosylation, transcriptome sequencing, endoplasmic reticulum, alternative splicing

Posted Date: February 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-212814/v1

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

Purpose

Our previous study has clarified the carcinogenic properties of arginine-specific mono-ADP ribosyltransferase 1 (ART1), which is considered to be a critical post-translational modification that changes the structure and function of proteins and is widely involved in important processes. This study provides, for the first time, a comprehensive insight of transcriptomic analysis for colorectal cancer cells interfered with ART1 silencing by Illumina RNA-Seq and related verification experiments.

Methods

Lentiviral infection was used to construct a CT-26 cell line that stably knocks down the ART1 gene, a whole transcriptome sequencing technique was performed to identify differentially expressed genes (DEGs). GO and KEGG classification/enrichment analysis and verification experiments were performed to determine the role of ART1 in the progression of colorectal cancer.

Results

A total of 5552 DEGs, GO function and KEGG pathway with highest enrichment, forms of SNP and diverse splicing patterns were able to be identified. Importantly, knockdown of ART1 affected the occurrence of the splicing of certain key genes related to tumor cell growth, also down-regulated expression of the key gene PTBP1 for alternative splicing. The overall attenuation of the endoplasmic reticulum unfolded protein response (UPR) signaling pathway caused by ART1 inhibition would unbalance UPR signaling, leading to the occurrence of apoptosis to impede tumorigenesis.

Conclusion

ART1, which clustered in organelles, may promote the development of colorectal cancer by participating in a variety of new mechanisms including endoplasmic reticulum stress regulation, metabolic process or alternative splicing, which may provide a good clinical drug candidate closer to targeted therapy of CRC.

Introduction

Colorectal cancer (CRC) is the second leading cause of cancer-related deaths worldwide. In both sexes combined, CRC ranks third (6.1%) in terms of incidence and second (9.2%) in terms of mortality according to the status report on the global burden of cancer issued by the International Agency for Research on Cancer (IARC) based on the GLOBOCAN[1]. The 5-year survival rate is almost less than 10% when the tumor has already diagnosed at advanced stages. Targeted therapy is currently a vital treatment that can prolong the survival of patients with advanced stages. Regrettably, there are still patients who have no alternative targeted drugs, also, adverse reactions of targeted drugs limit their long-term therapeutic effect, some new breakthroughs are expected. Therefore, it is urgently needed to explore more pivotal and specific key molecule from more aspects as therapeutic targets for this malignancy, also to elucidate the molecular mechanisms of CRC by expanding basic research.

ADP-ribosylation is a common modification of macromolecules, such as proteins, nucleic acids etc, such a widespread modification is of increasing interest because of its association with may key biological function and cellular process, including transcription, DNA damage repair, stress responses, cell signaling[2]. ADP-ribosylation is the enzymatic transfer of a single or multiple ADP-ribose from NAD⁺ to acceptor molecule, termed MARylation and PARylation respectively[3, 4]. In comparison to PARylation, the research on functions of MARylation is not particularly thorough, especially the lack of the identification of mono-ADP-ribosylated proteins. However, as time passed, MARylation have gradually been attached importance due to an increasing number of evidence that MARylation also serves as a reversible post-translational modification that can be read by readers and removed by MAR-specific erasers, furthermore, mono-ADP-ribosyltransferase-catalyzed ribosylation of protein substrate usually leads to protein inactivation, providing a mechanism to inhibit protein functions in both physiological and pathological conditions. ART1, which is widely anchored in the cell membrane and localized in the cytoplasm, is the only enzyme reported to catalyze the arginine-specific mono ADP ribosylation in humans and mouse. Besides, in past five years, our research group has conducted extensive research on mono-ADP-ribosylation and found evidence to support that ART1 and PARP-1 have a synergistic effect on CDDP-induced apoptosis of CT26 cells, and the mechanism may be related to ART1 regulating RhoA/ROCK pathway and thereby affecting the expression of NF-kB and PARP-1[5], suggesting poly-ADP ribosyltransferase is regulated by mono ADP ribosyltransferase to a certain extent in CRC[6], mono-ADP ribosylation may be the basis of poly-ADP ribosylation to play an important role. In addition, roles in the regulation of cell proliferation, apoptosis, tumor angiogenesis, DNA repair of ART1 have been described in our past research, among them, the results showed that ART1 is a vital cancer-promoting factor, which is significantly increased in colorectal cancer[7, 8]. While it is known that ART1 is a dangerous carcinogen in colorectal cancer, we hope to find more differential expression genes through high-

Page 2/14
throughput RNA sequencing, and collect functional enrichment and signal pathway enrichment information to explore the molecular mechanism of ART1 in the development of colorectal cancer, which is conducive to the explore of targeted therapy for colorectal cancer.

CT-26 cells share molecular features with undifferentiated and refractory human colorectal carcinoma cells[9]. Therefore, as a widely used clinical and basic research mouse tumor cell line, it is the best alternative for biological mechanism research or preclinical evaluation of targeting and immunotherapy. Here, we used lentiviral transfection to establish a CT-26 cell line with stable knockdown of ART1 to carry out high-throughput sequencing and related enrichment analysis.

Material And Methods

Cell culture and lentivirus cell lines establishment. Mouse colorectal cancer cell line CT26 isolated from Balb/c mouse was provided by Professor Wei YQ of Sichuan University. The construction and packaging of lentivirus were provided by GeneChem cooperation (Shanghai, China). Non-transfection, GFP-vector, GFP-shART1 cell lines establishment steps were described previously[6]. Briefly, CT-26 cells were infected with lentivirus for 48h and selected in 2ug/ml puromycin in culture medium. All group of CT26 cells were grown in the recommended RPMI-1640 medium (Hyclone, Logan, UT, USA), supplemented with 10% fetal bovine serum and antibiotic solution (100ug/ml streptomycin and 100U/ml penicillin, Hyclone), at 37°C and an atmosphere of 5% CO2. The cells in the three groups were lysed with RNAiso Plus (Takara Bio, Inc.) and sent to the Sangon BiotechCo., Ltd.

RNA isolation. Total RNAs were extracted using a Total RNA Extractor (Trizol) kit (B511311, Sangon, China) according to its manual. DNA contamination was cleaned by DNase I treatment.

RNA-Seq reads mapping. Clean reads were processed and aligned to mouse reference genome (NCBIM37) using HISAT2 (http://ccb.jhu.edu/software/hisat2/manual.shtml, version 2.0) with default parameters. Hierarchical indexing and multiple alignment strategies were used to solve the short spliced-reads alignment problem. Then, RSeQC (http://rseqc.sourceforge.net/; version 2.6.1) was used to statistics the alignment results.

Analysis of gene structure and chromosome distribution. The homogeneity distribution and the genome structure were checked by Qualimap (http://www.qualimap.org; version 2.2.1). BEDTools (http://samtools.sourceforge.net; version 2.26.0) was used to statistical analysis the gene coverage ratio.

Genetic structural analysis. In the latest version of BCFTools (http://samtools.sourceforge.net; version 1.5), SNP calling can be achieved by using BCFTools instead of Samtools and BCFTools at the same time. SNP (Single Nucleotide Polymorphisms) and Indel (insertion or deletion) in each sample were extracted by SNP/InDel calling, and Snp Eff was used to count the distribution of mutation sites on the genome structure. Also quality value>20 and coverage cutoffs >8 was used as filtering criteria in this process. In this study, ASprofile software (download form http://ccb.jhu.edu/software/ASprofile/; version 2.36) was used to extract and quantify and compare alternative splicing (AS) events from RNA sequence data. We can directly compare multiple transcripts of the same gene through asport software to identify different alternative splicing events.

Differential expression genes (DEGs) analysis. We measure the relative abundances of transcripts through standardized RNA-Seq fragment counts with StringTie (http://ccb.jhu.edu/software/stringtie; version 1.3.3b). The unit of measurement is transcripts per million (TPM) which eliminates the influence of gene lengths and sequencing discrepancies to enable direct comparison of gene expression between samples. The calculation formula of TPM was as follows:

$$TPM_i = \frac{X_i}{L_i} \times \frac{1}{\sum_j \frac{X_j}{L_j}} \times 10^6$$

$$X_i = \text{total exon fragment/reads} \quad L_i = \frac{\text{exon length}}{\text{KB}}$$

The Venn Diagram package in R language (Venn Diagram R package; https://CRAN.R-project.org/package=Venn Diagram; version1.6.17) was used to visually showed the similarity and overlap of the number of expressed genes between samples. DESeq2 (R package; https://github.com/mikelove/DESeq2; version 1.12.4) is an R package and was used to estimate variance-mean dependence in count data from RNA sequencing assays and determine differentially expressed genes (DEGs) between two samples based on a model using the negative binomial distribution. Genes were considered as significant differentially expressed if q-value <0.001 and |Fold Change| >2. Gene expression differences were visualized by scatter plot, MA plot and volcano plot.

Functional analysis of differential expression genes (DEGs). Functional enrichment analysis including Gene Ontology (GO) and KEGG was performed to identify which DEGs, proteins or other molecules were significantly enriched in GO terms or pathways. DAVID v 6.7 was used to map
DEGs to the GO terms in the database followed by calculation of the number of genes in every term, and a hypergeometric test is performed to identify significantly enriched GO terms in the gene list out of the background of the reference gene list. cluster Profiler was used for functional enrichment analysis. GO terms annotation were visualized by Histogram. GO enrichment analysis result was displayed using bubble chart. Kobas software was used for enrichment analysis of KEGG pathway of differentially expressed genes. In general, GO terms and KEGG pathway with Q-value <0.05 were considered as significantly altered.

**Reagents and antibodies.** Antibody against PERK (#5683), IRE1α (#3294), ATF6 (#65880) and p-eif2α (#3398) were purchased from Cell Signaling Technology (Beverly, MA, USA). Anti-XBP1 (ab220783) antibody was obtained from Abcam (Cambridge, UK), anti-GRP78 (sc-13539) antibody was obtained from Santa Cruz biotechnology (CA, USA), ART1 (66958-1-ig) and β-actin (60008-1-ig) antibody was obtained from Proteintech Technology(Wuhan, Hubei, PR.C).anti-rabbit, antimouse and anti-rat secondary antibodies (Proteintech, Wuhan, Hubei, PR.C) were also obtained. Total RNA from CT-26 cells in each group were isolated with RNaseq Plus reagent (Takara Bio, Inc.). Total RNA was transcribed reversely to cDNA using an PrimeScriptTM RT regent kit with gDNA Eraser (Takara Bio, Inc.).

**Western blotting analysis.** Total protein was extracted with RIPA lysis buffer and cocktail (1:100). Equal protein in each group were loaded into sodium dodecyl sulfate (SDS) polyacrylamide gel at 80 V for 0.5 h, then 100 V for 1hand finally transferred to polyvinylidene fluoride (PVDF) blotting membranes. The membrane is blocked by 5% milk or bovine serum albumin (BSA) for 1h at room temperature, and then incubated with the primary antibody overnight at 4℃, all antibody concentrations are strictly in accordance with the instructions. The membrane is taken out the next day, washed thrice and then incubated with the secondary antibody for 1h at room temperature. Protein bands were visualized by luminal chemiluminescence (Bio-Rad, Hercules, CA, USA). The gray value is detected and analyzed by image J software.

**Real-time polymerase chain reaction analysis.** Briefly, quantitative real-time PCR of PERK, IRE1α, ATF6 and GRP78was performed using SYBR Green Master mix (Takara Bio, Inc.) on a 7500 system (Applied Biosystems, Foster City, CA). The primers are shown as follows:

<table>
<thead>
<tr>
<th>Name</th>
<th>Sense Strand/Sense Primer (5’-3’)</th>
<th>Sense Strand/Sense Primer (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PERK</td>
<td>AGTCCCTGCTGAATCTCTTCC</td>
<td>TCCCAAGGCAGAACAGATATACC</td>
</tr>
<tr>
<td>IRE1α</td>
<td>ACACTGCTGAGACCCTTGTG</td>
<td>GGAGCCGTCTCTTGCTA</td>
</tr>
<tr>
<td>ATF6</td>
<td>GACTCACCACCATCCAGGTGT</td>
<td>CTCCAGTCTATCATCTGGTCC</td>
</tr>
<tr>
<td>GRP78</td>
<td>ACTGGGGGACACCATTTCT</td>
<td>ATCGCAATACAGACGCTCC</td>
</tr>
<tr>
<td>β-actin</td>
<td>AGACTACGAGCTGCTGAC</td>
<td>AGCACTGTGGTGGCGTGAC</td>
</tr>
</tbody>
</table>

**Results**

**RNA-seq evaluation.** High-quality reads were aligned to mouse reference genome (NCBIM37) using HISAT2 software with higher accuracy and faster running speed. Between 96%-98% of reads were aligned to the reference genome, about 88%-91% of reads aligned to the reference genome in the unique method, all multiple mapped ratios in different three group <10%, the difference between total mapped ratio with unique mapped ratio<10% (Table.1), suggesting that the sequencing quality and mapping results all meet the requirements, also DNA or ribosome contamination is eliminated. Besides, we can also intuitively illustrate that there is no problem in the process of RNA extraction and cDNA library construction through statistical analysis of genome structure distribution and density distribution of sequences on chromosome. Statistics show that the percentage of sequencing sequences located in the exon region is the highest, while the location of introns and intergenic regions is relatively low (Fig 1A), thereby eliminating the contamination of immature mRNA in the three samples. Located the sequence to each chromosome on the genome and plot it, statistics reveal that there are no abnormal coverage positions in chromosomes (Fig 1B).

**mRNA structural variation.** SNP analysis results showed that there are 37896 and 42592 SNP sites in Non-transfection and GFP-vector respectively, the ART1 knockdown cells possessed 29418 SNP sites. The incidence of insertion or deletion (Indel) of small fragments at a certain position in the genome in three group were much lower than that of SNP (Fig 1C). Among this SNP sites, the number of transition mutations(A>G, C>T, G>A, T>C) was far larger than that of transversion (A>C, A>T, C>A, C>G, G>C, G>T, T>A, T>G). Also, in this detection, the majority of SNP sites occurred in CDS, intronic, ncRNA and 3’UTR regions, and less happened in splice, intergenic and 5’UTR regions (Fig 1D).

RNA-Seq and ASprole software together can be used to predict different alternative splicing patterns and splicing sites. Among the ten possible splice forms, skipped exon (SKIP), alternative 5’first exon (transcription start site, TSS) and alternative 3’last exon (transcription terminal site, TTS) had the highest incidence (Fig 1F). Researches have demonstrated that many types of factors including polyypyrimidine tract binding protein (PTB), CELF family (CUG-BP, Elav-like family), CALD1, can participate in the regulation of alternative splicing, and splicing variants of variable splicing factors which affect tumor cell growth and angiogenesis[10-12]. Interestingly, in this study, the prediction of possible alternative splicing events
found that the knockdown of ART1 affected the occurrence of the splicing of SLC39a14, VEGF, Met, and the number of alternative splicing events in multiple alternative splicing factors (Mbnl1, Celf1, Cald1 and Ptbp3) has also changed significantly. (Supplement Table1&2.)

**Identification of the DEGs.** TPM is currently recognized as the most accurate transcript quantification formula. The number of genes uniquely expressed in each group and the number of genes shared between three groups are shown in a Venn diagram (TPM>0). (Fig 2A). In the subsequent DEGs and enrichment analysis, GFP-vector was selected as the control group. DEGseq(version 1.12.4), an R package to identify differentially expressed genes or subtypes from RNA-seq data in different samples. The DEGs were screened with criterions: Q-Value<0.05 and |Fold Change|>2. In the GFP-shART1 vs. GFP-vector, 1052 genes highly expressed in GFP-shART1 group, and 4500 down-regulated genes were identified (Fig 2B). In addition, the top 10 up-and down-regulated genes in GFP-shART1 group were listed in Tables.2. It should be pointed out that after knocking down ART1, PTBP1, which plays an important role in pre-mRNA splicing and in the regulation of alternative splicing events was significantly down-regulated, log2FoldChange=-20.604763, qValue<0.01. This also suggested that ART1 has a certain regulatory effect on alternative splicing.

**GO term enrichment analysis of DEGs between GFP-vector and GFP-shART1 groups.** Enrichment analysis can be performed on the significance of GO functions of differential genes, which can help us to identify the biological pathways most relevant to biological phenomena in differential gene sets. It is considered that the function is enriched when the corrected Q-value <0.05. The GO enriched analysis indicated that the most enriched GO terms in DEGs between GFP-vector and GFP-shART1 were "intracellular" (GO: 0005622) and "intracellular part" (GO: 0044424), which belongs to the cellular component, followed by "intracellular organelle" (GO: 0043229) and "organellar" (GO: 0043226) (Q-value=0.05) (Fig 2C). It means that the active position of DEGs caused by ART1 knockdown are mainly located in intracellular, especially in organelles. For molecular function, DEGs were significantly enriched in "binding" (GO: 0005488), followed by "heterocyclic compound binding" (GO: 0005488) (Q-value<0.05) (Fig 2D), besides, more than half of the top 30 enriched molecular functions listed in the bubble chart are related to various forms of bindings; For biological process, the highest enrichment of DEGs was "cellular metabolic process" (GO: 0044237), followed by "metabolic process" (GO:0008152) (Q-value<0.05) (Fig 2E). All of the above are in line with our previous research that ART1 is abundantly expressed in the endoplasmic reticulum, may increase glycolysis and energy metabolism in CT26 cells under high glucose conditions by regulating the Akt/c-Myc signaling pathway and the expression of glycolytic enzymes[13].

**KEGG enrichment analysis.** The KEGG pathway enrichment analysis of differentially expressed genes was performed using Kobas software. The DEGs were enriched significantly in 33 KEGG pathways (Q-value<0.05), and the top 30 signal pathways with the highest enrichment were shown in the bubble chart. Ubiquitin mediated proteolysis(ko04120), Protein processing in endoplasmic reticulum(ko04141), Ribosome biogenesis in eukaryotes(ko03008), RNA transport(ko03013), Endocytosis(ko04144) were ranked in the top five (Fig 2F).

Among them, what arouses our great interest is the mitogen activated protein kinases (MAPK) signaling pathway with 87 differential genes enriched (Q-value: 0.0002466). The extracellular-signal-regulated kinases (ERK1/2), the jun N-terminal kinase or stress activated protein kinases (JUK or SAPK), p38/MAPK and ERK5 are four major subfamilies of MAPK signaling[14]. In particular, researchers have recognized that the main regulator of colorectal cancer cell proliferation is the ERK MAPK signaling pathway. Moreover, the mutation of the KRAS proto-oncogene is an early event in the development of CRC. The high activity of Ras is accompanied by the increase of ERK activity, which causes colorectal cancer cells to escape from normal growth and differentiation control and has the ability to invade surrounding tissues and organs[15]. Kras also hemizygously mutated at p.G12D in CT26 cells, however, KEGG enrichment map showed that the knockdown of ART1 in CT26 cells reduced the expression of MEK1(log2FoldChange=-1.85), ERK (log2FoldChange=-2.17) and its downstream c-myc (log2FoldChange=-1.66). Also, knocking down ART1 significantly down-regulated mitogen-activated protein kinase 4 and 7(MKK4 and MKK7) (log2FoldChange=-2.36= and -3.34) which can activate the JNK double phosphorylation site to prevent the activation of the JUK signaling pathway, in addition, the transcription level of JNK also decreased (Supplement Figure1). All of the above indicated that ART1 silencing can prevent the proliferation of colorectal cancer cells to a certain extent. It needs to be added that the previous research which published in 2017 also suggested that ART1 serves a facilitatory role in the proliferation and migration of CT26 cells, and this effect may associate with the factors downstream of FAK and RhoA, c-myc, c-fos and COX2[16, 17], which is also consistent with the results obtained by this RNA sequencing. However, the potential biological mechanism of the down-regulation of various key signaling molecules caused by ART1 inhibition needs further research.

ART1 is a mono-ADP-ribosyltransferase that is stably anchored to the cell membrane or abundantly expressed in the cytoplasm and specifically catalyzes arginine residues. In our previous research, ART1 was also found to be specifically expressed in the endoplasmic reticulum, which is combined with the result that more differential expressed genes obtained in the GO classification and enrichment analysis can be enriched into the cellular component of organelle, let us have a strong interest in protein processing in endoplasmic reticulum pathway (ko04141) obtained from KEGG enrichment analysis (Fig 3A). Sequencing results showed that ART1 silencing significantly downregulated the expression of its substrate protein glucose-regulated protein 78(GRPs78/BIP) (log2FoldChange=-1.98), also inhibited the expression of endoplasmic reticulum transmembrane protein (PKP like ER protein kinase(PERK) (log2FoldChange=-1.26), activating transcription factor 6(ATF6) (log2FoldChange=-2.81) and inositol-enquiring enzyme 1(IRE1a) (log2FoldChange=-4.6) in the three important unfolded protein reaction(UPR) signal transduction pathways, suggesting that ART1 is likely to play a key role in the regulation of unmatured protein flow and the relief of endoplasmic reticulum load in endoplasmic reticulum stress environment. In addition, we also selected the top 10 functional categories with the highest enrichment degree and the differential genes related to this function, drew the network diagram of the significant enrichment function and gene interaction (Figure3B).
**ART1 Inhibition induced the apoptosis of CT-26 cells by overall repression of UPR reaction.** Endoplasmic reticulum stress promotes the processing of misfolded or unfolded proteins accumulated in the cavity of the endoplasmic reticulum, in this way, the normal function of the cell is better maintained, which is conducive to cell survival, and the subsequent response triggered by this is the unfolded protein response. However, the unfolded protein response is characterized by two-way regulation in the process of processing cell growth and death. On the one hand, it can maintain cell survival. On the other hand, cell apoptosis will occur immediately when severe endoplasmic reticulum stress exceeds the load of the cell. More importantly, recent studies have found that if the unfolded protein response is fully inhibited, the cell will lose the ability to maintain the homeostasis of the endoplasmic reticulum and die[18, 19]. GRP78/Bip plays an important role in this adjustment process. Under normal physiological conditions, the N-terminals of the three receptor proteins, PERK (PEK like ER kinase), IRE1α (inositol requiring 1), and ATF6α (activating transcription factor 6), combine with GRP78/Bip to form a complex, thereby being in an inactive state and inhibiting signal transduction. When stress occurs, misfolded or unfolded proteins will accumulate to cause the complex to disaggregate, release PERK, IRE1α, ATF6α, and activate related pathways to promote the degradation of misfolded proteins and the correct folding of unfolded proteins. Emerging evidence indicated that UPR signaling components may play important roles in cancer since the upregulation of GRP78, PERK, IRE1α and ATF6 were detected in various clinical and tumor cell samples. The inhibition of the signal pathway mediated by a single component will cause the activation of the compensation mechanism of other components. Therefore, it is necessary to find the key points that can directly and comprehensively destroy the three related signal pathways in the UPR response of tumor cells, which can inhibit the survival of tumor cells.

It has been established that ART1 is widely distributed in endoplasmic reticulum, and its role in the endoplasmic reticulum is mainly to modify GRP78 by mono-ADP ribosylation to inactivate it[20, 4]. Recent studies demonstrated that ART1 was transiently up-regulated during ER stress induction with the treatment of dithiothreitol or thapsigargin[21]. Also, compared with normal tissues, the significantly higher expression of ART1 in colorectal cancer has been clarified in previous experiments There, combined with the results obtained by RNA sequencing, we speculated that the increase of ART1 in colorectal cancer may be also associated to the presence of more endoplasmic reticulum stress stimuli in tumor cells.

Our previous study confirmed that ART1 was abundantly expressed in endoplasmic reticulum. In this study, we detected the activation of the signaling pathways mediated by the three interactors proteins of GRP78. PERK, IRE1α and ATF6 all decreased significantly after the knockdown of ART1 by lentiviral transfection, in addition, XBP-1s and p-eif2α were downregulated by ART1 inhibition, suggesting that inhibiting the expression of ART1 can completely block the activation of the three signal pathways. Besides, the expression of GRP78 also decreases after the UPR signaling pathways were blocked, which consistent with the result obtained in RNA-seq analysis (Fig 4A). At the same time, real-time PCR were performed to detect the mRNA level PERK, IRE1α, ATF6 and GRP78, the degree of PERK, IRE1α, ATF6 and GRP78 all lower in GFP-shART1 group (Fig 4B). Interestingly, it was found that the expression of Cleaved-caspase3 increased and the expression of Bcl2 decreased (Fig 4C), indicating that knocking down ART1 may hinder cell growth by completely inhibiting the UPR signaling pathway in the endoplasmic reticulum, and at the same time, it affects the level of Bcl2 and makes the cells tend to apoptosis. Through the results of this experiment, it is found that tumor cell apoptosis not only occurs when the endoplasmic reticulum stress is overloaded, but also when the UPR response, an important regulatory signaling pathway for endoplasmic reticulum stress, is fully inhibited.

**Discussion**

Our previous research found that ART1 plays an important role in the proliferation, apoptosis, adhesion, movement, metastasis and angiogenesis of mouse colon cancer cells. First, ART1 serves a facilitatory role in the proliferation and migration of CT26 cell, and the potential mechanism associated with downstream factors FAK, RhoA, c-Myc, c-fos and COX-2 [10], also under high-glucose conditions, ART1 may increase the generation of ATP and lactic acid by upregulating the AKT/mTOR/c-Myc pathway, thus stimulating the proliferation and inhibiting apoptosis of CT-26 cells by increasing the expression of PKD1 and LDHA[22]. Secondly, ART1 could influence apoptosis rate through regulation of mRNA and protein expression of a landmark infector Tubb3[23]. In addition, ART1 has a positive correlation with VEGF and integrin αVβ3 expression and associated with the angiogenesis of colorectal carcinoma or upregulates HIF-1α via PI3K/AKT signaling pathway, thus promote the expression of angiogenic factors[8]. Therefore, ART1 is highly overexpressed in colorectal carcinoma tissues and plays an accelerated role in many aspects in its development. Simultaneously, research on targeted protein modified by ART1 also received lots of attention. Integra7, human neutrophil peptide (HNP-1), platelet derived growth factor-BB(PDGF-BB), fibroblast growth factor2 (FGF-2) and GRP78/Bip are receptor protein directly modified by ART1, and are related to tumor immunity, angiogenesis, cell proliferation, etc[24-26, 21]. More important, the above previous research results are roughly consistent with the results of this transcriptome analysis, RhoA(log2FoldChange=-1.04), c-Myc(log2FoldChange=-1.66), COX-2(log2FoldChange=-3.81), PKD1(log2FoldChange=-4.53) all decreased significantly. Therefore, we hope to find more molecular mechanisms to consolidate the carcinogenic properties of ART1 in colorectal cancer. At the same time, since ART1 mainly plays a role by modifying its substrate protein, we hope to find a signaling pathway related to the substrate protein directly modified by ART1, so as to play a positive role in the targeted or combined treatment of colorectal cancer.

This study provides the first comprehensive insight into the transcriptome of colorectal cancer cells CT-26 affected by ART1 silencing. We selected CT-26 cells which share molecular features with aggressive, undifferentiated, refractory human colorectal carcinoma cells, strictly adhered to high quality control standards of sequencing data, and then performed accurate reference sequence alignment. Using a whole transcriptome sequencing technique (RNA-Seq), we were able to identify the levels of differentially expressed genes with TPM value and performed subsequently GO and KEGG classification and enrichment analysis. Moreover, the numbers and forms of SNP and diverse splicing patterns have also been subjected to
diversified statistical analysis. The above provides clues for us to study the effect of ART1 in colorectal cancer from multiple directions and explore the potential specific molecular mechanism.

TPM was used as a standardized value to standardize the number of read counts in transcripts and genes. However, compared with the classic FPKM (Fragments Per Kilobase Million) and RPKM (Trans Per Million) algorithms, TPM has a different processing order, that is, the gene length is first considered, and then the sequencing depth is considered, and the calculated value can be directly used for comparison between samples. Generally, it is a data standardization algorithm superior to FPKM and RPKM[27]. 1052 highly expressed genes and 4500 down-regulated genes in GFP-shART1 group were identified. Among the top ten genes listed in this study that increased and decreased after ART1 knockdown, sorted by log2 FoldChange, We noticed that ART1 silencing can significantly increase the expression of protein arginine transferase prmt1. Prmt1, mainly distributed in the cytoplasm and nucleus, is associated with a variety of cancerous diseases and participates in various cellular processes such as cell signal transduction, DNA damage repair, and transcription regulation. In particular, studies have suggested that PRMT1 is positively correlated with the proliferation and migration ability of colorectal cancer cells, and can inhibit the apoptosis of some colorectal cancer cells. So, whether the high expression of prmt1 caused by ART1 silence shows the same effect in colorectal cancer is worthy of further investigation. Moreover, ART1 silence can also seriously increase the expression of GAPDH, which also provides new ideas for the selection of internal controls in our daily research.

GO functional annotation and KEGG pathway enrichment analysis were to screen the biological functions, cell localization and signaling pathways that ART1 may affect. The differential genes obtained by knocking down ART1 were enriched into 33 signaling pathways when Q-value is less than 0.05 as the standard, and KEGG enrichment analysis screened 73 pathways when P-value< 0.05 as the standard. MAPK signaling pathway(ko04010) and Protein processing in endoplasmic reticulum(ko04141) are the two signaling pathway we are going to focus on. Fabrizio Get al found ADP-ribosylated GRP78/Bip on arginine residues largely distributed on cellular plasma membrane, in the lumen of ER, which is the result of the modification of enzyme ART1[21], also, the information from Uniprot (https://www.uniprot.org) and our previous results indicated that there is a large amount of ART1 distributed in the endoplasmic reticulum. Researchers also provided evidence that GRP78/Bip undergoes ADP ribosylation, and that ADP ribosylation as a rapid posttranslational mechanism for reversible inactivation of GRP78/Bip through interfering with allosteric coupling of GRP78/Bip' two domains may play an important role in fast regulation of unfolded protein load [28, 29]. Afterwards, more direct evidence has been reported that ADP-ribosylated GRP78/Bip provides a buffering system that balanced the protein procession rates with those of protein synthesis. All of the above suggested that ART1 is abundantly expressed in the endoplasmic reticulum and modified GRP78 to inactivate it, which makes us very interested in this signaling pathway in this transcriptome analysis. Therefore we speculated that ADP-ribosylated GRP78/Bip will definitely affect its downstream UPR response and the regulation process of endoplasmic reticulum homeostasis, Consistent with the sequencing result, our Western blot and RT-PCR result showed that ART1 silencing significantly downregulated the expression of itssubstrateproteinGRP78/Bip also inhibited the expression of endoplasmic reticulum transmembrane protein PERK, ATF6, IRE1α. The downstream key signaling molecules XBP-1sand p-eif2ai were also significantly inhibited, indicating that ART1 inhibition can completely block the activation of the unfolded protein response signaling pathway in the endoplasmic reticulum. More importantly, and consistent with the sequencing results, BCL-2 (log2FoldChange=-1.95) decreased significantly, and cleaved caspase3 increased significantly after intervention in ART1, suggesting the occurrence of apoptosis events. Since members of all classes of the Bcl-2 family localize to the ER membrane and have been confirmed to influence ER homeostasis, we speculate that the unfolded protein response is fully inhibited, which increases the release of Ca 2+ in the ER and inhibits the expression of Bcl-2, induces the occurrence of apoptosis. However, the deep molecular mechanism is still unclear, and it remains should be defined whether increased ADP-ribosylated GRP78/Bip can affect the release of GRP78/Bip interactors and then determine the downstream signaling pathways that can participate in it.

Through analysis, we thus found that the different number of variable shear events in each group, but ART1 knockdown can affect many types of factors including PTBP, CELF family, CALD1, etc, which can participate in the regulation of alternative splicing. PTBP1, which binds to the pyrimidine-rich region in the RNA sequence, interferes with the expression of the corresponding protein by affecting the positioning of the mRNA variable splicing, and ultimately triggers the change of the biological effect in the cell[30, 31]. In the sequencing analysis results of this study, PTBP1 was just affected by ART1 knockdown, which showed a significant decrease in expression. Therefore, How the alternative splicing molecule PTBP1 regulated by ART1 affects the development of colorectal cancer is still the subject of future research.

In this study, CT26 cells with ART1 intervention were successfully sequenced. A total of 5552 differential genes were found and annotated. Enrichment analysis helps us screen out ART1 may involve in the development of colorectal cancer by participating in a variety of new mechanisms including endoplasmic reticulum stress regulation and variable shear regulation. All of the above provide us with new ideas for understanding the carcinogenic properties of arginine mono-ADP ribosylation in colorectal cancer.

**Declarations**

**Declaration**

Not applicable

**Funding**
The research was supported by Innovation Project of Graduate Student in Chongqing (CYB19160); National Nature Science Foundation of China (30870946); Science and technology Research Foundation of Chongqing Municipal Education Commission(KJQN201800435).

Conflicts of interest

The author declares that there have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and materials

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

Code availability

Not applicable

Authors' contributions

SX Z and JL D processed the cell culture, RNA extraction, analyze of the sequencing result, SX Z wrote the manuscript; YP Y and HJ G were responsible for the process of validation experiments, YT, MX designed and conducted the experiments, QS L, ML and also participated in analysis and interpretation of data of RNA-seq. YL W were responsible for funding acquisition and manuscript modification.

Ethics approval

Not applicable

References


Tables

Table 1: Mapping results statistics
Non-transfection | GFP-vector | GFP-shART1
---|---|---
Total reads | 54002444(100.00%) | 54798870(100.00%) | 72651342(100.00%)
Total mapped | 52459772(97.14%) | 53392798(97.43%) | 70220553(96.65%)
Multiple mapped | 3626433(6.72%) | 3952311(7.21%) | 6106597(8.41%)
Uniquely mapped | 48833399(90.43%) | 49440487(90.22%) | 64113956(88.25%)
Read-1 mapped | 24468106(45.31%) | 24755948(45.18%) | 32088437(44.17%)
Read-2 mapped | 24365233(45.12%) | 24684539(45.05%) | 32025519(44.08%)
Reads map to '+' | 24422904(45.23%) | 24730789(45.13%) | 32315123(44.48%)
Reads map to '-' | 24410435(45.20%) | 24709698(45.09%) | 31798833(43.77%)
Non-splice reads | 24870606(46.05%) | 26003232(47.45%) | 35164099(48.40%)
Splice reads | 23962733(44.37%) | 23437255(42.77%) | 28949857(39.85%)
Reads mapped in proper pairs | 46457178(86.03%) | 47163648(86.07%) | 58918434(81.10%)

Table 2. Top ten up-regulated genes in GFP-shART1 CT26 cells

<table>
<thead>
<tr>
<th>Transcript ID</th>
<th>MeanTPM (GFP-shRNA)</th>
<th>MeanTPM (GFP-vector)</th>
<th>log2Fold Change</th>
<th>pValue</th>
<th>qValue</th>
<th>Result</th>
<th>GenelD</th>
<th>GeneName</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSMUST00000073605</td>
<td>1915.21</td>
<td>0.0001</td>
<td>24.1909993</td>
<td>0</td>
<td>0</td>
<td>up</td>
<td>ENSMUSG00000057666</td>
<td>Gapdh</td>
</tr>
<tr>
<td>ENSMUST0000066401</td>
<td>199.03</td>
<td>0.0001</td>
<td>20.9245545</td>
<td>0</td>
<td>0</td>
<td>up</td>
<td>ENSMUSG0000007891</td>
<td>Ctsd</td>
</tr>
<tr>
<td>ENSMUST0000122845</td>
<td>111.97</td>
<td>0.0001</td>
<td>19.6197042</td>
<td>0</td>
<td>0</td>
<td>up</td>
<td>ENSMUSG00000026553</td>
<td>Copa</td>
</tr>
<tr>
<td>ENSMUST0000128787</td>
<td>80.56</td>
<td>0.0001</td>
<td>18.9450666</td>
<td>1.3396E-206</td>
<td>2.0004E-205</td>
<td>up</td>
<td>ENSMUSG00000025260</td>
<td>Ndufv1</td>
</tr>
<tr>
<td>ENSMUST0000112617</td>
<td>80.51</td>
<td>0.0001</td>
<td>17.5745E-194</td>
<td>1.7574E-194</td>
<td>2.5053E-193</td>
<td>up</td>
<td>ENSMUSG00000074884</td>
<td>Serf2</td>
</tr>
<tr>
<td>ENSMUST000044923</td>
<td>59.7</td>
<td>0.0001</td>
<td>16.1973714</td>
<td>0</td>
<td>0</td>
<td>up</td>
<td>ENSMUSG00000041645</td>
<td>Ddx24</td>
</tr>
<tr>
<td>ENSMUST000117099</td>
<td>54.43</td>
<td>0.0001</td>
<td>15.0540425</td>
<td>0</td>
<td>0</td>
<td>up</td>
<td>ENSMUSG00000022945</td>
<td>Chaf1b</td>
</tr>
<tr>
<td>ENSMUST000148575</td>
<td>50.47</td>
<td>0.0001</td>
<td>14.9506666</td>
<td>1.3396E-206</td>
<td>2.0004E-205</td>
<td>up</td>
<td>ENSMUSG00000019324</td>
<td>Prmt1</td>
</tr>
<tr>
<td>ENSMUST000119865</td>
<td>0.0001</td>
<td>121.31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG00000027601</td>
<td>Mtf1</td>
</tr>
<tr>
<td>ENSMUST000176485</td>
<td>0.0001</td>
<td>136.14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG00000021595</td>
<td>Nsun2</td>
</tr>
<tr>
<td>ENSMUST000111878</td>
<td>0.0001</td>
<td>139.86</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG00000015766</td>
<td>Eps8</td>
</tr>
<tr>
<td>ENSMUST000156623</td>
<td>0.0001</td>
<td>150.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG00000027523</td>
<td>Gnas</td>
</tr>
<tr>
<td>ENSMUST000095457</td>
<td>0.0001</td>
<td>159.46</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG0000006498</td>
<td>Ptbp1</td>
</tr>
<tr>
<td>ENSMUST000174096</td>
<td>0.0001</td>
<td>172.72</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG00000029767</td>
<td>Calu</td>
</tr>
<tr>
<td>ENSMUST000179478</td>
<td>0.0001</td>
<td>191.52</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG00000024079</td>
<td>Hnmpf</td>
</tr>
<tr>
<td>ENSMUST000073308</td>
<td>0.0001</td>
<td>192.22</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG0000009079</td>
<td>Ewsr1</td>
</tr>
<tr>
<td>ENSMUST000127252</td>
<td>0.0001</td>
<td>204.95</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG00000027642</td>
<td>Rpn2</td>
</tr>
<tr>
<td>ENSMUST000021698</td>
<td>0.0001</td>
<td>239.03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>down</td>
<td>ENSMUSG00000021270</td>
<td>Hsp90aa1</td>
</tr>
</tbody>
</table>

Figures
Figure 1:

A. Statistical pie chart of genome structure distribution in each sample. The area occupied by each color corresponds to the percentage of the number of reads aligned to the genome structure and the number of reads aligned to the reference genome. Blue: exonic, orange: intronic, gray: intergenic.

B. Density distribution map of sequence on chromosome. Horizontal axis: the length of the chromosome (in millions of bases), vertical axis: the median of the RNA-Seq sequence density (logarithm to base 2).

C. Number of SNPs. Horizontal axis: sample name, vertical axis: number of variations. Red: SNP, blue: InDel.

D. Type of AS events statistics in each sample. Horizontal axis represents classification of AS forms, and the vertical axis is the number of various AS events. AE, alternative exon ends; IR, retention of single; MIR, multiple introns; MSKIP, cassette exons; SKIP, exon skipping; TSS, alternative transcription start site; TTS, alternative transcription termination site. XAE: approximate AE; XIR: approximate IR; XMIR: approximate MIR; XMSKIP: approximate MSKIP; XSKIP: approximate SKIP.
Figure 2

DEGs, GO and KEGG enrichment analysis. (A). Venn diagram of shared genes. Different samples are represented by different colors, and the numbers in the figure represent the number of specific or shared expressed genes. The overlapping area shows the number of expressed genes shared by different samples, whereas the non-overlapping area indicates the number of unique expressed genes between different samples. (B). Scatter plot of expression differences. The horizontal and vertical axes are the log (TPM) values of the two groups of samples. In the figure, each dot represents a gene, and the closer to the origin, the lower the expression level. Among them, red indicates up-regulated genes, green indicates down-regulated genes, black indicates non-differential genes, and up/down-regulation are both vertical axis samples relative to horizontal axis samples. (C, D, E). Scatterplot of significantly enriched function. The vertical axis represents the function annotation information, and the horizontal axis represents the rich factor corresponding to the function (the number of differential genes annotated to the function divided by the number of genes annotated to the function). Q-value is represented by the color of the dot. The number of differential genes contained in each function is expressed by the size of the dot. (Only the top 30 gene function with the highest enrichment were drew in each scatterplot). (F). Scatterplot of significantly enriched KEGG pathway. The vertical axis represents the KEGG pathway classification, and the horizontal axis represents the rich factor corresponding to the pathway (the number of differential genes annotated to the function divided by the number of genes enriched in the pathway). Q-value is represented by the color of the dot. The number of differential genes contained in each function is expressed by the size of the dot. (Only the top 30 KEGG pathway with the highest enrichment were drew in each scatterplot).
Figure 3

Processing in endoplasmic reticulum pathway maps of DEGs and KOG enrichment network KEGG enrichment analysis. (A). All the highlighted gene products in the figure belong to the genes annotated in this transcriptome. The rectangular nodes represent gene products (such as enzymes or some RNA regulators), and the circular nodes represent compounds (that is, substrates or products). The white background circle the rectangular rectangle indicates other pathways associated with this pathway. The colors and positive and negative indicate the relative up-and-down relationship of genes in the two comparison samples. Red indicates up-regulated genes. Reddish colors indicate higher gene scheduling, green indicates down-regulated genes, and green colors indicate higher gene scheduling. Yellow indicates non-differential genes. The color block size represents the proportion of the corresponding gene in the gene product. (B). The square represents the functional classification enriched by DEGs, the circle represents the names of related genes, and the connecting lines indicate the correlation between functions or genes.
Figure 4

ART1 inhibition induced the apoptosis of CT-26 cells. (A). Relative protein levels of key molecules in the UPR signaling pathway were detected in cells after infected with vector and shART1 lentivirus using western blot. (B). Real-time PCR amplification of PERK, IRE1α, ATF6 and GRP78. (C). Relative protein levels of Cleaved-caspase3 and Bcl2 and statistical result were shown.

Supplementary Files

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

- sup.doc