

Exploration of the typical features of tubulovillous adenoma using in-depth quantitative proteomics analysis

Yin Zhang^{a*}, Meng Pan^{b*}, Chun-Yuan Li^{b*}, Jing-ying Li^b, Wei Ge^b, Lai Xu^a, and Yi Xiao^a

^aDepartment of General Surgery, Division of Colorectal Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; ^bState Key Laboratory of Medical Molecular Biology & Department of Immunology, Institute of Basic Medical Sciences Chinese Academy of Medical Sciences, School of Basic Medicine Peking Union Medical College, Beijing, China

ABSTRACT

This present study aimed to explore the typical protein features of tubulovillous adenoma (TVA) using proteomic and bioinformatic analyses. Tandem mass tag (TMT)-based quantitative proteomic analyses were conducted on normal mucosa, tubular adenoma, TVA and adenocarcinoma tissues. We identified 5,665 proteins categorized into seven clusters based on Pearson's correlation analysis. The bioinformatic analysis showed mitochondrial and metabolism-related events were typical characteristics of TVA and mitochondrial-, ribosome- and matrisome-related biological processes may contribute to carcinogenesis. PLOD3 was identified as a key protein associated with the malignant potential of TVA and promoted the viability of adenoma organoids. The Cancer Genome Atlas (TCGA) analysis revealed PLOD3 as a risk factor for disease-free and overall survival. Furthermore, the PLOD3 expression correlated negatively with the abundance of B cells, CD8 + T cells, CD4 + T cells, neutrophils, macrophages and myeloid dendritic cells. In conclusion, enhanced metabolic and mitochondrial reprogramming are typical features of TVA, and PLOD3 might be related to the "immune desert" phenotype and contribute to TVA tumorigenesis and colorectal cancer development.

ARTICLE HISTORY

Received 16 June 2021
Revised 17 August 2021
Accepted 17 August 2021

KEYWORDS

Tubulovillous adenoma; tumorigenesis; PLOD3; quantitative proteomics; metabolic reprogramming; tumor microenvironment

Introduction


Colorectal cancer (CRC) is the third most commonly diagnosed cancer in the world and the second leading cause of cancer deaths [1]. A comprehensive understanding of the evolution of the CRC is of critical importance in improving screening and prevention. In the canonical carcinogenesis of CRC, the conventional adenoma-to-carcinoma (A-C) sequence accounts for approximately 85% of the cases [2]. However, only 5% of the conventional adenomas could finally become carcinoma [3]. Therefore, classification of conventional adenoma based on the risk of carcinogenesis is essential for accurate surveillance. Patients with advanced adenoma (tumor size >1 cm, villous component (VC) >25% or/and high-grade intraepithelial neoplasia) have a higher risk of developing cancer and cancer mortality compared with

non-adenoma patients, while the patients with non-advanced adenoma do not [4].

At the histological level, adenoma with VC is the representative type of advanced adenoma. According to the proportion of the VC, VC-type is categorized as tubulovillous adenoma (TVA, 25%–75% VC) and villous adenoma (VA, >75% VC) [5]. Using tubular adenoma (TA) as a reference, the relative risks of malignant transition of TVA and VA are 1.51 and 2.53, respectively, with a higher proportion of VC correlated with a higher risk of malignant transition [6]. Hence, it is important to determine the features that characterize the high-risk malignant transition of TVA/VA. At the genetic level, VC is significantly associated with *Kras* mutations, but without correlation to adenoma size or the number of chromosomal abnormalities [3]. Activating

CONTACT Lai Xu  drxulai@pumch.cn; Yi Xiao  xiaoy@pumch.cn  Division of Colorectal Surgery, Department of General Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, No.1 Shuaifuyuan Road, Wangfujing, Dongcheng District, Beijing 100730, China

*These authors are contributed equally to this article

 Supplemental data for this article can be accessed [here](#).

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

GNAS mutations occur frequently in VA, but these mutations are irrelevant to carcinogenesis of VA [7,8]. At the epigenetic level, the CpG island methylator phenotype-positive status is frequently identified in TVA/VA and the degree of MGMT and RASSF2 methylation is significantly higher in TVA/VA than in TA [9]. At the gene expression level, AMN and PTGDR were found to be down-regulated, while osteopontin and osteonectin were upregulated in the VA-carcinoma sequence [10].

In this study, we aimed to provide a comprehensive description of the typical protein features of TVA and identify key proteins that reflect the potential malignant risk of TVA with proteomic strategy and bioinformatic analysis. Our results could contribute to the further understanding of the features and carcinogenesis potential of TVA at the protein level.

Methods

Patients and tissue samples

Normal mucosa (NM), TA, TVA and AC tissues were obtained from 24 patients recruited at the Division of Colorectal Surgery, Department of General Surgery, Peking Union Medical College Hospital (China). Details of the patients and tissues are listed in Table S1. After removal, all tissues were stored temporarily on dry ice and then transferred to -80°C .

This study was approved by the Ethics Committee of Peking Union Medical College Hospital (Number: JS-2094). Written informed consent was obtained from each patient prior to the study commencement.

Protein extraction and tandem mass tag-labeling (TMT-labeling)

Lysate proteins (100 μg) in 8 M urea were alkylated with dithiothreitol (DTT) and iodoacetamide (IAA) before protein digestion. The proteins were digested with Trypsin/Lys-C and protein labeling was performed using the TMT-labeling kit according to the manufacturer's instructions. The TMT-labeled peptides were analyzed by high performance liquid chromatography with a directly interfaced Thermo Orbitrap Fusion mass spectrometer (Thermo Scientific, USA). Protein identification was performed using Proteome Discoverer 2.2 software

(Thermo Scientific, USA) with the SEQUEST search engine. Proteins with a false discovery rate (FDR) <0.01 and unique peptides ≥ 2 qualified for further analysis. Relative protein quantification was performed using the TMT-6plex method. All the methods were performed as previously described [11]. The mass spectrometry proteomics data have been deposited with the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) in the iProX partner repository with the dataset identifier: PXD023899 [12].

Bioinformatic analysis

Heatmaps were generated using the HCE 3.5 software based on the standardized proteomic profilings of NM, TVA and AC [13]. Gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Wiki-pathway analyses were performed using the ClueGO and CluePedia plugins in Cytoscape [14–16]. Results with $P < 0.05$ and validated experimentally could be visualized. Protein–protein interaction (PPI) networks were constructed using the Search Tool for the Retrieval of Interacting Genes (STRING) database [17]. The annotation and biological process categories were downloaded in Metascape database [18]. Survival analysis was performed using the Gene Expression Profiling Interactive Analysis (GEPIA) database [19]. The Cancer Genome Atlas (TCGA) data of CRC was downloaded using cBioPortal database [20,21]. Immune cell infiltration was estimated using the RNA-seq data using the TIMER database [22].

Generation of adenoma organoids

For human adenoma organoids, fresh adenoma tissue was cut into small pieces, washed with PBS and dissociated in TrypLE (Gibco, 12,604–013) to achieve cell suspension. The culture medium was prepared using Sato's method [23]. The organoids were passaged (1:3) every week. Appropriate Cell Recovery Solution (Corning, REF354253) was added to the Matrigel and dissociated for 1 h at 4°C to obtain the organoid suspension. Then the supernatant was discarded after centrifugation, the organoids were embedded in the Matrigel.

For mouse adenoma organoids, the intestinal fragments containing adenomas from C57BL/6 J-APC^{Min} mice (male, 16 weeks old, Model Animal Resource Information Platform) were treated as described for the human adenoma tissue. The culture medium was prepared as follows: penicillin 1 U/mL, streptomycin 1 µg/mL, HEPES 10 mmol/L, Glutamax, 1x N2, 1x B27 and N-acetylcysteine 1 mmol/L in Advanced Dulbecco's Modified Eagle's Medium and F12.

RNA extraction and qRT-PCR analysis

Total RNA was extracted from HCT116 cells and adenoma organoids using TRIzol reagent (Thermo Scientific, 155,996,018) according to manufacturer's instructions. We used HCT116 cells to detect the knockdown efficiency of PLOD3. RNA concentrations of cells and adenoma organoids were measured by NanoDrop spectrophotometer (Thermo Scientific). Real-time quantitative PCR was performed using the SYBR-Green (Takara, RR096A), and the relative PLOD3 mRNA expression levels were normalized to that of GAPDH. Primers used in this experiment were as follows:

GAPDH (forward: 5'-GTCTCCTCTGACTTCAACAGCG-3',

reverse: 5'-ACCACCCTG-TTGCTGTAGCCAA-3');

PLOD3 (forward: 5'-CTGAAGAAGTTCGTCCAGAGTG-3',

reverse: 5'-ACCGATGAATCCACCAGAA TTG-3').

Lentiviral transduction of adenoma organoids

The organoids were harvested and dissociated in TrypLE for 5 min in a 37°C incubator. After terminating digestion, the organoids were centrifuged to pellet for 5 min at 500 × g. Then the organoids were resuspended in 500 µL lentiviral transduction medium and transferred into a 12-well culture plate. The plate was centrifuged at 600 × g at 32°C for 1 h and placed in the 37°C incubator for another 4 h. Transferred the organoid-virus mixture into a tube and centrifuged to discard the supernatant. The pellet was resuspended in Matrigel and cultured for 48 h at 37°C.

Immunofluorescence

Organoids were isolated from the Matrigel and then fixed with 4% paraformaldehyde for 1 h at room temperature. After three washes with PBS, the organoids were blocked and permeabilized with PBS containing 5% goat serum, 0.5% Triton X-100 and 0.5% NaN₃ for 2 h at room temperature. And then, the organoids were incubated with primary antibodies Ki67 (abcam, ab15580, 1:50) and PCNA (Origene, TA800875, 1:50) separately. After that, the organoids were incubated with Alexa Fluor 488 (Thermo, A-11034, 1:50) and Alexa Fluor 594 (Thermo, A-11032, 1:50) for 2 h at room temperature. Next, the organoids were stained with 5 µg/mL DAPI (Solarbio, C0065) for 15 mins at room temperature. After washing with PBS, the images were captured with a confocal microscope (Zeiss LSM 780).

3D cell viability assay

The adenoma organoids transfected with PLOD3-siRNA were used to detect cell viability. The CellTiter-Glo® 3D Reagent (Promega, G9681) was used to dissolve the matrigel. The adenoma organoid suspension was equally distributed to opaque-walled multiwell plates (Corning, REF3603). The volume of CellTiter-Glo® 3D Reagent equal to the volume of cell culture medium was added to each well and then the content was incubated for 30 min at room temperature. Cell viability was calculated based on ATP content of each well.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 8.0.1 (GraphPad Software, Inc., La Jolla, CA, USA). Quantitative data were analyzed using Student's *t*-test or ANOVA. Pearson's correlation analysis was performed to evaluate associations between sets of data. The Kaplan–Meier method was used for survival analysis. *P* < 0.05 was considered to indicate statistical significance.

Results

In the present study, we aimed to explore the typical protein features of TVA by proteomic

strategy and bioinformatic analysis. We found that the typical protein features of TVA are mainly enriched in metabolism related biological process and pathway. Then we identified PLOD3 as the potential malignant predictive protein of TVA and validated the function of PLOD3 on adenoma organoids. Furthermore, we find that PLOD3 might be related to the immune desert phenotype of CRC and be a risk factor of CRC.

Screening for the protein clusters representative of TVA

The clinical characteristics of the NM, TA, TVA and AC tissue samples obtained from patients are listed in Table S1. The workflow of this study is shown in Figure 1. A total of 5,665 proteins were identified according to the criteria described in the Methods section. The proteins were classified into seven clusters as shown in Figure 2a. This heatmap revealed that the proteins in the cluster 6 were highly expressed in TVA (950 proteins) and the proteins in cluster 2 were highly expressed in both TVA and AC (734 proteins). Based on the theory that TVA is more liable to evolve into cancer than

TA, we speculated that the proteins in cluster 2 might contribute to the carcinogenesis because of the similarity in the expression pattern between TVA and AC. The proteins in clusters 2 and 6 were regarded as the representative clusters of TVA and selected for further analysis.

Identification of the typical features of TVA

To identify the biological processes and pathways typical of TVA, we selected the highly expressed proteins related to TVA in cluster 6 with a fold-change (FC) in expression >1.5 in 127/128, 127/129 and 127/130 (265 proteins, Supplementary file 1). We selected the highly expressed proteins related to both TVA and AC in cluster 2 following the FC >1.5 in 127/128, 127/129, 130/129 and 130/128 (33 proteins, Supplementary file 1). The expression profile of the highly expressed proteins in clusters 6 and 2 are shown in Figure 2b-c. We performed Gene Ontology (GO) analysis of the highly expressed proteins in three aspects: GO-cellular component (CC), GO-molecular function (MF) and GO-biological process (BP). The aim of GO analysis was to investigate the main location of

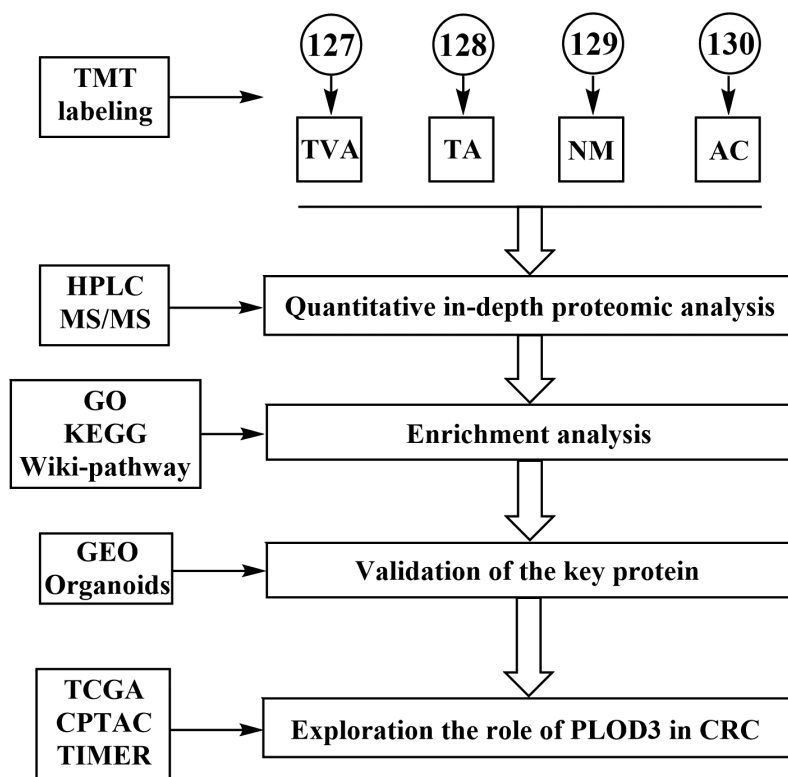


Figure 1. Workflow of the study.

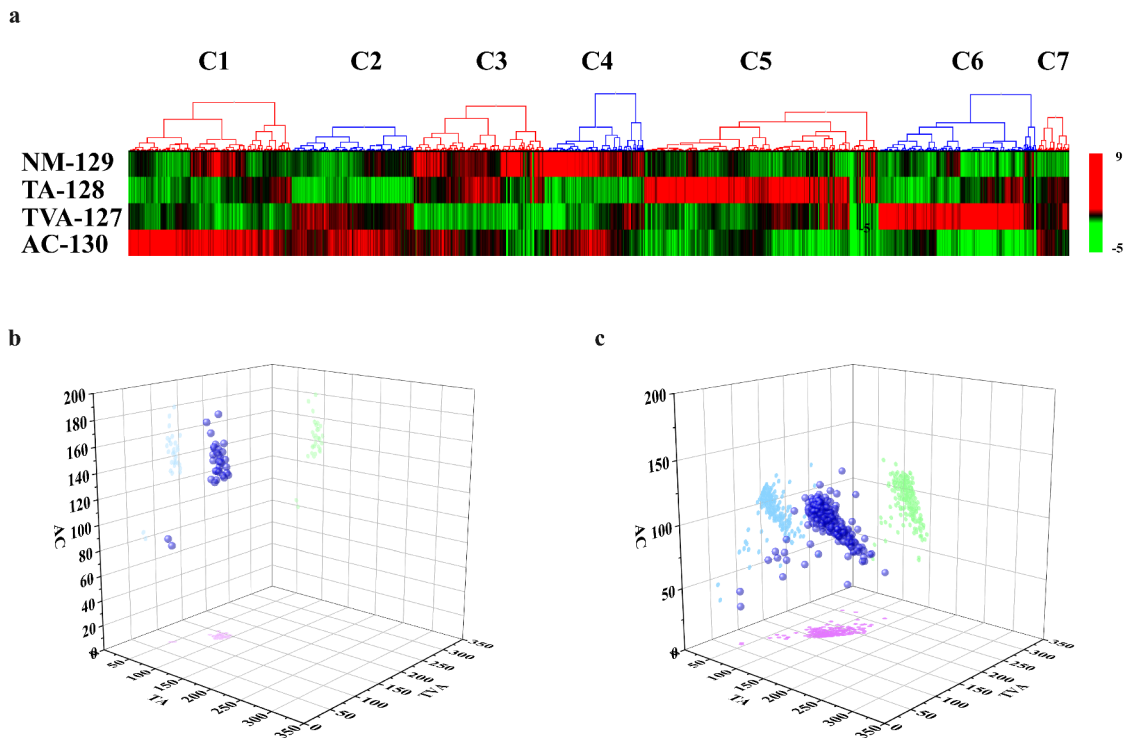


Figure 2. Identification of the highly expressed proteins in tubulovillous adenoma.

A. Heat map showing the relative abundance of 5,665 proteins in the normal mucosa (NM), tubular adenoma (TA), tubulovillous adenoma (TVA) and Adenocarcinoma (AC) group. The proteins were clustered hierarchically by Pearson correlation analysis. Green indicates downregulated proteins, whereas red indicates upregulated proteins. B-C. Figures 2b and 2c are three-dimensional diagrams of the expression patterns of the filtered proteins in clusters 2 and cluster 6. The filtered proteins of cluster 2 are the highly expressed proteins related to both TVA and AC. The filtered proteins of cluster 6 are the highly expressed proteins related to TVA only. Three-dimensional space was constructed with protein expression value of TVA, AC and TA tissues as the coordinate axis. The expression level of the filtered cluster 2 and cluster 6 proteins in TVA, AC and TA tissues are indicated by blue balls. The light blue dots are the projection of the protein on the AC and TVA coordinate axes, showing the expression level of the proteins in AC and TVA tissues. Similarly, the pink dots indicate the expression level of the protein in TA and TVA tissues, and the green dots indicate the expression level of the protein in TA and AC tissues. All the protein names in Figures 2b and 2c were listed in Supplementary file 1.

the cell, activities at the molecular level and biological phenomenon of the highly expressed proteins in cluster 6. GO-CC enrichment analysis showed that most of the proteins were located in the mitochondria (Figure 3a). Mitochondrial gene expression, acyl-CoA metabolic process and carboxylic acid metabolic process were mainly enriched in GO-BP analysis (Figure 3b). Oxidoreductase activity and CoA-ligase activity were mainly enriched in the GO-MF analysis (Figure 3c). We performed a KEGG pathway analysis to investigate which pathway the proteins were mainly involved in. The result showed that valine, leucine and isoleucine degradation, citrate cycle, oxidative phosphorylation and mitochondrial ribosome-related pathway were mainly enriched in the KEGG pathway analysis

(Figure 3d). The results of GO and KEGG analysis indicated that the formation of the villous component is related to enhanced metabolic activities. To further validate and supplement the results of GO and KEGG analysis, we performed Wiki-pathway analysis. The advantage of Wiki-pathway analysis is that Wiki-database supports crowdsourcing and is updated in a timely manner. Furthermore, Wiki-pathway analysis shows favorable interactivity in metabolic pathways [24]. The enrichment result of Wiki-pathway is similar to the KEGG analysis. The proteins were mainly enriched in metabolism-related pathway. Moreover, Wiki-pathway analysis showed that the proteins were related to metabolic reprogramming in colon cancer (Figure 3e). Metabolic reprogramming is one of the hallmarks of cancer [25]. This result indicates that the

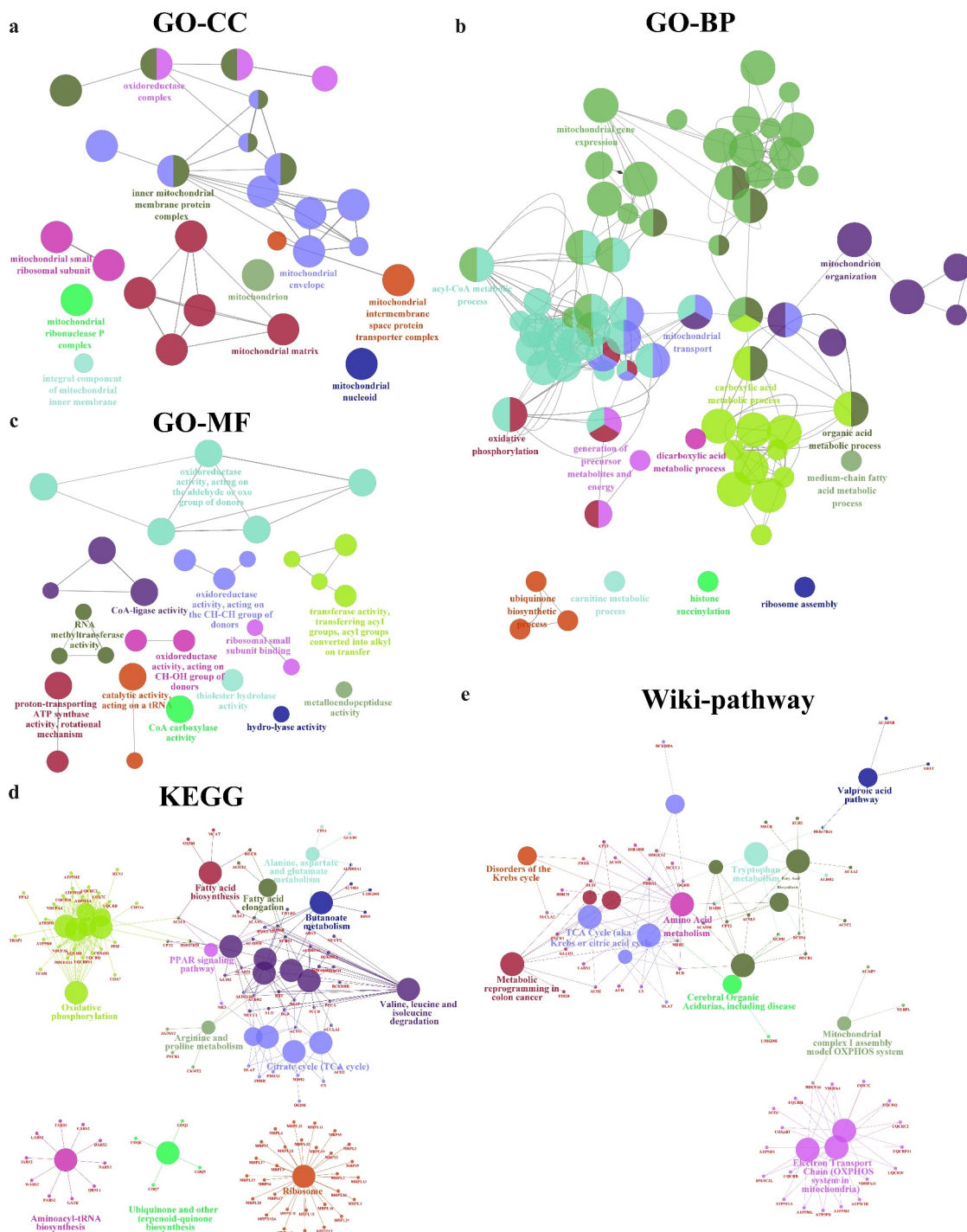


Figure 3. Enrichment analysis of the highly expressed proteins in cluster 6.

Gene ontology (GO), Kyoto Encyclopedia of Gene (KEGG) and Wiki-pathway analyses of the highly expressed proteins in cluster 6 were performed using the ClueGO and CluePedia plugins in Cytoscape. Circles shown in the same color represent similar enrichment results. (A) GO-cell component (CC) analysis of the highly expressed proteins in cluster 6. (B) GO-molecular function (MF) analysis of the highly expressed proteins in cluster 6. (C) GO-biological process (BP) analysis of the highly expressed proteins in cluster 6. (D) The KEGG analysis of cluster 6 (E) Wiki-pathway analysis of cluster 6.

metabolic reprogramming events in CRC might initiate in the TVA phase. The construction of the PPI network of the highly expressed proteins in cluster 2 showed that mitochondrial-, ribosome- and matrisome-related biological processes and pathways were enriched (Figure 4a-b). Matrisome-related pathway was the representative in cluster 2 and PLOD3, TGFBI and COL12A1 were the annotation proteins in this pathway. Moreover, PLOD3 and COL12A1 were related to the epithelial-mesenchymal transition (EMT) pathway, which is regarded as one of the core events that contribute to cancer metastasis. Therefore, we speculated that the matrisome-related pathway is associated with a higher potential for carcinogenesis of TVA compared to TA and we focused on PLOD3 and COL12A1 for further analyses.

PLOD3 promotes the viability and proliferation of adenoma organoids

To further explore the key proteins that may contribute to the carcinogenesis of TVA, we focused on PLOD3 and COL12A1, which are proteins

related to the EMT pathway. PLOD3 is a multifunctional protein that promotes collagen maturation by catalyzing hydroxylation and glycosylation of lysine residues in procollagen, while COL12A1 is a collagen protein that is a component of the extracellular matrix. Comparison of the expression of PLOD3 and COL12A1 in TVA and TA using the Gene Expression Omnibus database (GSE 117607) showed higher specificity of PLOD3 in TVA (Figure S1). Therefore, based on its multifunctional characteristics and higher specificity in TVA, we selected PLOD3 for further analysis. After confirming the knockdown ability of si-RNA-mediated knockdown of PLOD3 in adenoma organoids generated from fresh surgical specimens, we investigated its effects on viability. The viability of si-PLOD3 organoids was reduced compared with that of control organoids (Figure 5a-c). Then we explored the overexpression effect of PLOD3. Compared with the control group, adenoma organoids showed increased cell viability and more Ki67- and PCNA-positive cells (Figure S3). The functional experiments indicated that

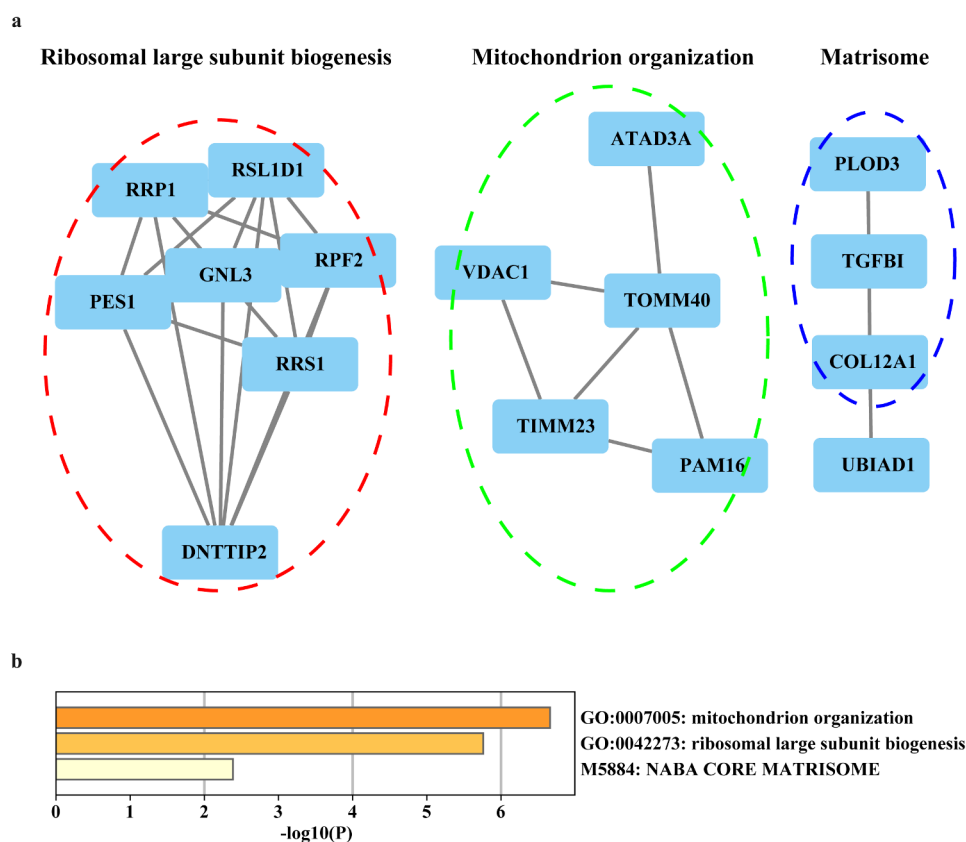


Figure 4. Enrichment analysis of the highly expressed proteins in cluster 2. (A) Protein-protein interaction (PPI) network of the highly expressed proteins in cluster 2. (B) GO-BP and pathway analysis of the highly expressed proteins in cluster 2.

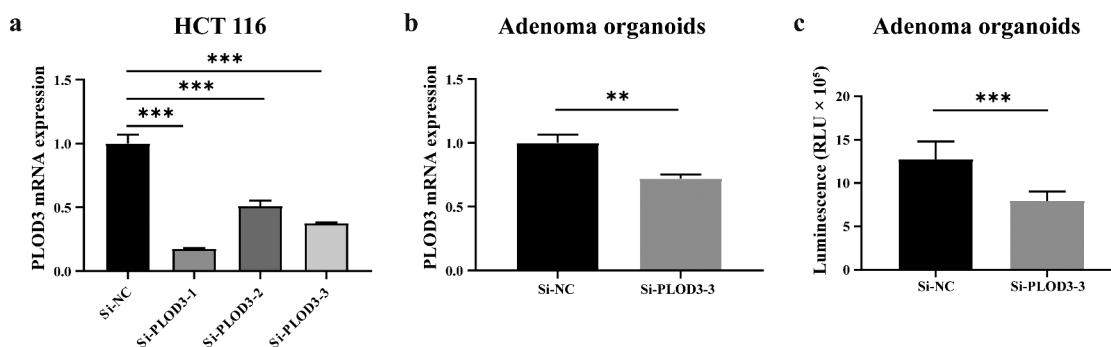


Figure 5. PLOD3 knockdown reduces viability of adenoma organoids. (A) Expression level of PLOD3 mRNA in HCT116 cells was evaluated by qPCR after knockdown mediated by three si-RNA sequences. (B) Expression level of PLOD3 mRNA in adenoma organoids was significantly decreased after knockdown mediated by the si-3 RNA sequence. (C) The viability of adenoma organoids was significantly decreased after PLOD3 knockdown.

PLOD3 promotes the viability and proliferation of adenoma. Since the acquisition of increased viability and proliferation of premalignant lesions is a pivotal event in carcinogenesis, we speculated that PLOD3 plays a key role in the carcinogenesis of CRC.

Exploration the role of PLOD3 in CRC

Analysis of the proteins in the cluster 2 indicated that PLOD3 has the potential to promote TVA carcinogenesis. We next evaluated the correlation between PLOD3 and survival, clinical stage and

the tumor microenvironment (TME) of CRC based on public databases. GEPIA is an online database that contains the survival data of the CRC patients in TCGA database. The high PLOD3 expression was associated with poor overall survival (OS) and disease-free survival (DFS) in the GEPIA database (Figure 6a-b). Furthermore, PLOD3 expression was significantly higher in the patients with distant metastasis (M1) than in the M0 patients based on TCGA database (Figure S2A). PLOD3 expression was higher in the patients with local lymph node metastasis (N+) than in the N0 patients, although this difference

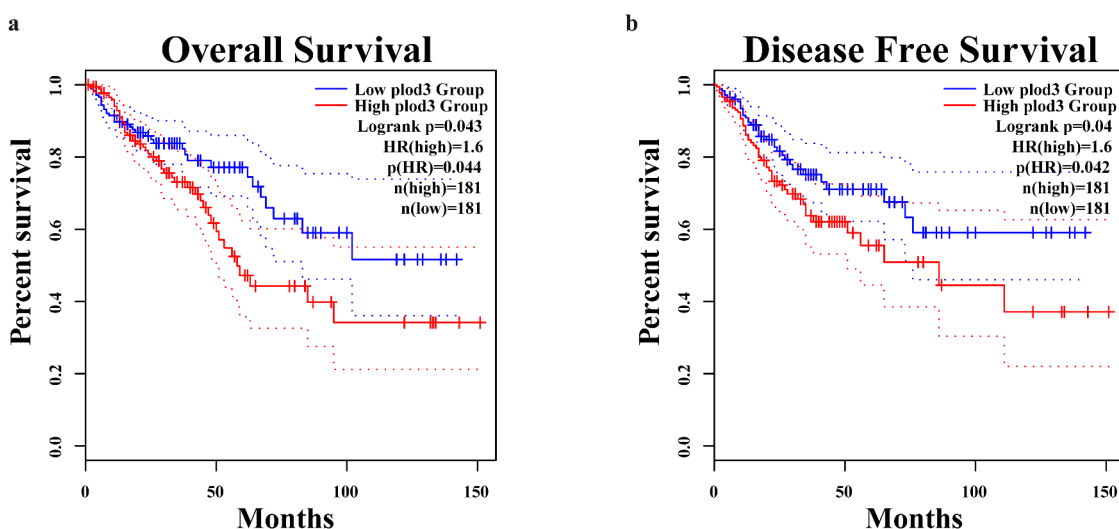


Figure 6. Survival analysis of PLOD3 in CRC patients. (A) The disease-free survival (DFS) of the high PLOD3 expression group is significantly poorer than that of the low PLOD3 expression group in the GEPIA database ($P = 0.04$, hazard ratio (HR) = 1.6, 362 patients). (B) The overall survival (OS) of the high PLOD3 expression group is significantly poorer than that of the low PLOD3 expression group in GEPIA database ($P = 0.043$, HR = 1.6, 362 patients).

was not statistically significant (Figure S2B). In the CPTAC database, there was no difference in PLOD3 expression between the M0 and M1 patients, although this result may be influenced by the imbalanced sample size (Figure S2C). The expression of PLOD3 was significantly higher in N + patients than in N0 patients (Figure S2D). We further explored the correlation between PLOD3 and tumor infiltrating immune cells. Based on the analysis of TCGA database using the TIMER method, the PLOD3 expression showed a significantly negative correlation with the abundance of B cells, CD8⁺ T cells, CD4⁺ T cells, neutrophils, macrophages and myeloid dendritic cells (Figure 7).

Discussion

In this study, we performed quantitative proteomics on NM, TA, TVA and AC tissues and conducted a systematic analysis on the proteomic profile of TVA. We found that the typical proteins of TVA were located mainly in the mitochondria and enriched in metabolic events. We then

compared the similar proteomic clusters identified in the analysis of TVA and AC samples. We found that matrisome-related proteins were highly expressed in both TVA and AC and these proteins might be relevant to potential malignant features of TVA. PLOD3, an EMT-related protein, was identified as the key protein in the TVA-AC process among the matrisome-related proteins. The function of PLOD3 was validated using adenoma organoids. Moreover, our findings indicated that PLOD3 is associated with local or distant metastasis, poor prognosis and the ‘immune desert’ phenotype in CRC.

Enrichment analysis of the typical proteins of TVA indicated that TVA possesses a high metabolic status compared with TA and AC. Moreover, we found that in the TVA stage metabolic reprogramming events could occur and the main metabolic mode of TVA was oxidative phosphorylation, branched chain amino acids (BCAA) degradation and mitochondrial ribosome-related pathway. TVA/VA has a higher proportion of hyperplastic epithelial cells, which reflects a greater proliferative ability and energy requirement compared with TA

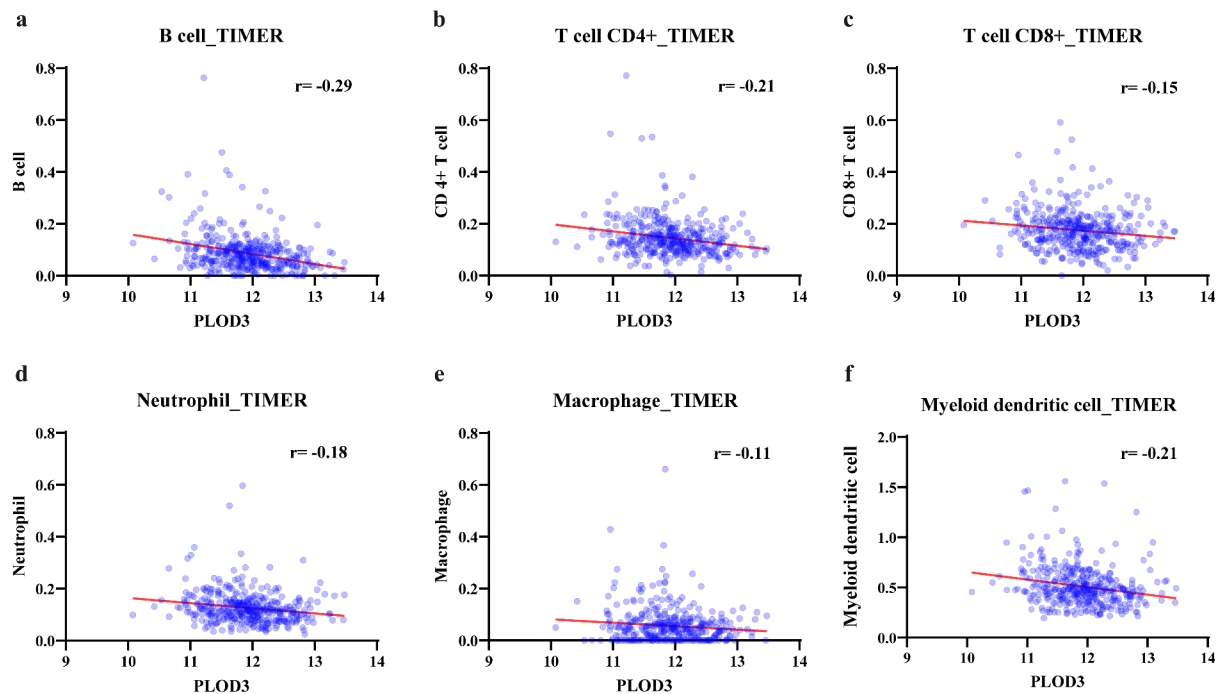


Figure 7. The correlation analysis between PLOD3 expression and immune cells. (A-F) PLOD3 negatively correlated with B cell, CD4 + T cell, CD8 + T cell, neutrophil, macrophage, myeloid dendritic cell infiltration in the TCGA database using TIMER algorithm ($P < 0.05$).

[26]. The histological features could partially explain the high metabolic status of TVA. Some studies showed analogous results that enhanced metabolic related pathway and metabolic reprogramming are initiated in the pre-malignant stage of CRC. In one study of the metabolic and mitochondrial reprogramming events in the pre-malignant lesions of the rectal mucosa, Cruz et al. found that the Warburg effect and metabolic reprogramming existed prior to the pre-malignant status and that mitochondrial numbers were increased in both biopsy tissues and animal models [27]. Similarly, Bensard et al. demonstrated that metabolic reprogramming of aerobic glycolysis, a pattern that is known as the Warburg effect and regarded as the hallmark of most cancer cells, promoted tumor initiation in CRC [25,28,29]. Satoh et al. found that metabolic reprogramming occurred at the adenoma stage and metabolism-related genes were regulated by MYC [30]. However, it is worth noting that in our study the main mechanism of energy generation in TVA is the mitochondrial oxidative phosphorylation rather than aerobic glycolysis. We therefore consider that the evolution of CRC is a complex process and the metabolic and mitochondrial reprogramming associated with the NM-TVA-AC sequence represents a fluctuation rather than a monotonic change. It possibly means that oxidative phosphorylation is the main metabolic pathway in the NM-TVA sequence and aerobic glycolysis are enhanced in the TVA-AC sequence. We have reviewed the relevant studies and found that the tumorigenesis process of some types of CRC showed similar metabolism fluctuating mode. A bimodal change in the mitochondrial DNA copy number has been identified in the evolution of ulcerative colitis-related cancer [31]. In one study of familial adenomatous polyposis tumorigenesis, the genes related to the TCA cycle also triggered bimodal changes in expression based on a pseudotime trajectory [32]. In addition to the oxidative phosphorylation pathway, we found that the BCAA degradation and mitochondrial ribosome-related pathways were significantly enriched in TVA. Also, plasma concentrations of BCAAs are negatively correlated to the risk of colorectal adenoma [33]. In the obesity colon cancer model, BCAA supplementation prevented the initiation of

colonic pre-neoplastic lesions [34]. Therefore, BCAA intake represents a potential chemoprevention strategy for CRC. Abnormal expression of mitochondrial ribosome proteins (MRPs) was related to mitochondrial metabolism disorder and cell dysfunction and was observed in a variety of cancer types [35]. MRPL35 was found to be upregulated in CRC and associated with poor survival in a Chinese cohort [36]. The specific role of MRPs in the initiation and development of CRC and the underlying mechanism remains to be fully elucidated.

In addition to exploring the clusters and proteins characteristic of TVA, we identified a cluster of proteins that showed similar expression modes across TVA and AC. These proteins might be associated with the malignant potential of TVA. In constructing the PPI network, we focused on the matrisome-related proteins and selected PLOD3, an EMT-related protein, as the key protein that represents the higher malignant potential of TVA compared with TA. EMT-related proteins were regarded as risk factors for CRC in the previous studies. Huang et al. found that an EMT-related gene expression signature combining infiltrating T-cell abundance has showed promising prognostic value in colon cancer [37]. Further, EMT-related gene expression signature can also be applied as a risk stratified tool in hepatocellular cancer [38]. In our study, PLOD3 plays an important role not only in CRC, but also in the TVA-AC sequence. At the adenoma organoid level, we found that PLOD3 can promote the viability and proliferation of adenoma. Thus, the functional experiments validated the key role of PLOD3 in the tumorigenesis of CRC. Besides the function of PLOD3 at the adenoma level, we also found that PLOD3 is upregulated in tumor tissue. PLOD3 is a multifunctional enzyme with lysyl hydroxylase, collagen galactosyltransferase as well as glucosyltransferase ability and is highly expressed and associated with the progression of lung cancer, glioma, ovarian cancer, gastric cancer, hepatocellular cancer and pancreatic cancer [39–45]. In line with our findings, Nicastri et al. found that PLOD3 was upregulated in CRC tissue using the quantitative proteomic technique focusing on N-linked glycoproteins [46]. In the hepatocellular cancer, PLOD3 was identified as one of the early-

stage decision markers that were upregulated in the premalignant lesions [44]. This result further has indicated the potential key role of PLOD3 in other cancer types. In our analysis of the CRC cohort in TCGA, we found that high PLOD3 expression was significantly associated with poor survival of patients, and PLOD3 was negatively correlated with the immune cells infiltrating in TME. This phenomenon conforms to the immune desert type of TME, in which the typical cancer category is Microsatellite stable-CRC (MSS-CRC) initiated from the adenoma-carcinoma sequence, and confirmed the reliability of our data analysis [47]. Therefore, we speculated that the immune cell-excluded events can occur in the pre-cancer stage.

Two limitations of this study are specified as follows. First, since the results are based mainly on proteomic analysis, the potential mechanisms underlying the role of PLOD3 in the tumorigenesis and development of CRC remain to be elucidated. Second, our analysis showed that metabolic and mitochondrial reprogramming of oxidative phosphorylation and the BCAA degradation and mitochondrial ribosome-related pathways occurred in the TVA stage. This phenomenon needs to be further validated in mechanistic and metabolomic studies.

Conclusion

In conclusion, we comprehensively analyzed the proteomic profiles of NM, TA, TVA and AC. Enhanced metabolic and mitochondrial reprogramming, such as oxidative phosphorylation, BCAA degradation and mitochondrial ribosome-related pathway were identified as typical features of TVA. Furthermore, we showed that PLOD3, which was correlated with an ‘immune desert’ type TME, might contribute to TVA tumorigenesis and CRC development.

Authors' Contributions

Y.X. and L.X. conceived the study; Y.Z. performed data pre-processing; Y.Z. performed primary analysis and wrote the article. C.L., M.P. and J.L. performed and supervised the experiments. All authors read and approved the final manuscript.

Acknowledgements

The authors thank Wen-yun Hou, Yan-pan Gao and Dan Luo for their guidance in organizing the manuscript and assisting in the experiments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by the National Natural Science Foundation of China (No. 81702933) and CAMS Innovation Fund for Medical Sciences(No. 2017-I2M-1-001);National Natural Science Foundation of China;CAMS Innovation Fund for Medical Sciences.

Data Availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the iProX partner repository with the dataset identifier: PXD023899.

References

- [1] Sung H, Ferlay J, Siegel RL, 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):209–249.
- [2] Strum WB. Colorectal Adenomas. *N Engl J Med.* 2016;374(11):1065–1075.
- [3] Hermsen M, Postma C, Baak J, et al. Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability. *Gastroenterology.* 2002;123(4):1109–1119.
- [4] Click B, Pinsky PF, Hickey T, et al. Association of Colonoscopy Adenoma Findings With Long-term Colorectal Cancer Incidence. *JAMA.* 2018;319(19):2021–2031.
- [5] Hornick JL, Odze RD. CHAPTER 19 - Polyps of the Large Intestine. In: Odze RD, Goldblum JR, editors. *Surgical Pathology of the GI Tract, Liver, Biliary Tract, and Pancreas. Second Edition* ed. Philadelphia: W.B. Saunders; 2009. p. 481–533.
- [6] Chen CD, Yen MF, Wang WM, et al. A case-cohort study for the disease natural history of adenoma-carcinoma and de novo carcinoma and surveillance of colon and rectum after polypectomy: implication for efficacy of colonoscopy. *Br J Cancer.* 2003;88(12):1866–1873.
- [7] Sekine S, Ogawa R, Oshiro T, et al. Frequent lack of GNAS mutations in colorectal adenocarcinoma

- associated with GNAS-mutated villous adenoma. *Genes Chromosomes Cancer*. 2014;53(4):366–372.
- [8] Yamada M, Sekine S, Ogawa R, et al. Frequent activating GNAS mutations in villous adenoma of the colorectum. *J Pathol*. 2012;228(1):113–118.
- [9] Kakar S, Deng G, Cun L, et al. CpG island methylation is frequently present in tubulovillous and villous adenomas and correlates with size, site, and villous component. *Hum Pathol*. 2008;39(1):30–36.
- [10] Galamb O, Sipos F, Spisák S, et al. Potential biomarkers of colorectal adenoma-dysplasia-carcinoma progression: mRNA expression profiling and in situ protein detection on TMAs reveal 15 sequentially upregulated and 2 downregulated genes. *Cell Oncol*. 2009;31:19–29.
- [11] Zhan S, Wang T, Wang M, et al. In-Depth Proteomics Analysis to Identify Biomarkers of Papillary Thyroid Cancer Patients Older Than 45 Years with Different Degrees of Lymph Node Metastases. *Proteomics Clin Appl*. 2019;13(5):e1900030.
- [12] Ma J, Chen T, Wu S, et al. iProX: an integrated proteome resource. *Nucleic Acids Res*. 2019;47(D1):D1211–D1217.
- [13] Seo J, Gordish-Dressman H and Hoffman EP. An interactive power analysis tool for microarray hypothesis testing and generation. *Bioinformatics*. 2006;22(7):808–814.
- [14] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504.
- [15] Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*. 2009;25(8):1091–1093.
- [16] Bindea G, Galon J and Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics*. 2013;29(5):661–663.
- [17] Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607–D613.
- [18] Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
- [19] Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45(W1):W98–W102.
- [20] Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal*. 2013;6(269):11.
- [21] Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multi-dimensional cancer genomics data. *Cancer Discov*. 2012;2(5):401.
- [22] Li T, Fu J, Zeng Z, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res*. 2020;48(W1):W509–W514.
- [23] Sato T, Stange DE, Ferrante M, et al. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology*. 2011;141(5):1762–1772.
- [24] Slenter DN, Kutmon M, Hanspers K, et al. WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res*. 2018;46(D1):D661–d667.
- [25] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646–674.
- [26] Mahajan D, Downs-Kelly E, Liu X, et al. Reproducibility of the villous component and high-grade dysplasia in colorectal adenomas <1 cm: implications for endoscopic surveillance. *Am J Surg Pathol*. 2013;37:427–433.
- [27] Cruz MD, Ledbetter S, Chowdhury S, et al. Metabolic reprogramming of the premalignant colonic mucosa is an early event in carcinogenesis. *Oncotarget*. 2017;8(13):20543–20557.
- [28] Bensard CL, Wisidagama DR, Olson KA, et al. Regulation of Tumor Initiation by the Mitochondrial Pyruvate Carrier. *Cell Metab*. 2020;31. DOI:10.1016/j.cmet.2019.11.002.
- [29] Warburg O. On the origin of cancer cells. *Science*. 1956;123(3191):309–314.
- [30] Satoh K, Yachida S, Sugimoto M, et al. Global metabolic reprogramming of colorectal cancer occurs at adenoma stage and is induced by MYC. *Proc Natl Acad Sci U S A*. 2017;114(37):E7697–E7706.
- [31] Ussakli CH, Ebae A, Binkley J, et al. Mitochondria and tumor progression in ulcerative colitis. *J Natl Cancer Inst*. 2013;105(16):1239–1248.
- [32] Li J, Wang R, Zhou X, et al. Genomic and transcriptomic profiling of carcinogenesis in patients with familial adenomatous polyposis. *Gut*. 2020;69(7):1283–1293.
- [33] Budhathoki S, Iwasaki M, Yamaji T, et al. Association of plasma concentrations of branched-chain amino acids with risk of colorectal adenoma in a large Japanese population. *Ann Oncol*. 2017;28(4):818–823.
- [34] Shimizu M, Shirakami Y, Iwasa J, et al. Supplementation with branched-chain amino acids inhibits azoxymethane-induced colonic preneoplastic lesions in male C57BL/KsJ-db/db mice. *Clin Cancer Res*. 2009;15:3068–3075.
- [35] Huang G, Li H and Zhang H. Abnormal Expression of Mitochondrial Ribosomal Proteins and Their Encoding Genes with Cell Apoptosis and Diseases. *Int J Mol Sci*. 2020;21. DOI:10.3390/ijms21228879
- [36] Zhang L, Lu P, Yan L, et al. MRPL35 Is Up-Regulated in Colorectal Cancer and Regulates Colorectal Cancer Cell Growth and Apoptosis. *Am J Pathol*. 2019;189(5):1105–1120.
- [37] Huang X, Chen C, Xu Y, et al. Infiltrating T-cell abundance combined with EMT-related gene expression as a prognostic factor of colon cancer. *Bioengineered*. 2021;12(1):2688–2701.

- [38] Xiong C, Wang G and Bai D. A novel prognostic models for identifying the risk of hepatocellular carcinoma based on epithelial-mesenchymal transition-associated genes. *Bioengineered*. 2020;11(1):1034–1046.
- [39] Baek JH, Yun HS, Kwon GT, et al. PLOD3 promotes lung metastasis via regulation of STAT3. *Cell Death Dis*. 2018;9:1138.
- [40] Baek JH, Yun HS, Kwon GT, et al. PLOD3 suppression exerts an anti-tumor effect on human lung cancer cells by modulating the PKC-delta signaling pathway. *Cell Death Dis*. 2019;10(3):156.
- [41] Tsai CK, Huang LC, Tsai WC, et al. Overexpression of PLOD3 promotes tumor progression and poor prognosis in gliomas. *Oncotarget*. 2018;9(21):15705–15720.
- [42] Guo T, Gu C, Li B, et al. PLODs are overexpressed in ovarian cancer and are associated with gap junctions via connexin 43. *Lab Invest*. 2021;101(5):564–569.
- [43] Wang B, Xu L, Ge Y, et al. PLOD3 is Upregulated in Gastric Cancer and Correlated with Clinicopathologic Characteristics. *Clin Lab*. 2019;65. DOI:10.7754/Clin.Lab.2018.180541.
- [44] Shen Q, Eun JW, Lee K, et al. Barrier to autointegration factor 1, procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3, and splicing factor 3b subunit 4 as early-stage cancer decision markers and drivers of hepatocellular carcinoma. *Hepatology*. 2018;67(4):1360–1377.
- [45] Schiarea S, Solinas G, Allavena P, et al. Secretome analysis of multiple pancreatic cancer cell lines reveals perturbations of key functional networks. *J Proteome Res*. 2010;9(9):4376–4392.
- [46] Nicastrì A, Gaspari M, Sacco R, et al. N-glycoprotein analysis discovers new up-regulated glycoproteins in colorectal cancer tissue. *J Proteome Res*. 2014;13(11):4932–4941.
- [47] O'Donnell JS, Teng MWL and Smyth MJ. Cancer immunoediting and resistance to T cell-based immunotherapy. *Nat Rev Clin Oncol*. 2019;16:151–167.