Identification of MTHFD1L as a potential biomarker for multiple cancers by pan-cancer analysis

Lin Yan
The Second Affiliated Hospital of Anhui Medical University

Xin Wang
The First Affiliated Hospital of Anhui Medical University

Jian-ming Yang (✉ Jmingyang88@163.com)
The Second Affiliated Hospital of Anhui Medical University

Research Article

Keywords: MTHFD1L, pan-cancer, prognosis, mitotic cell cycle

Posted Date: August 5th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1918153/v1

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

MTHFD1L is a monofunctional enzyme, which plays a vital role in the generation of tetrahydrofolate and maintains the balance of folate cycle. In the past ten years, it was reported that MTHFD1L may participate the growth and development of cancers. HPA (Human Protein Atlas) database was used to explored the consensus MTHFD1L tissues expression and MTHFD1L gene conservation analysis. The expression of MTHFD1L in different cancer types and the relationship between the level of expression of MTHFD1L and the cancer-associated fibroblast immune infiltration were showed in the TIMER2 database. Kaplan–Meier (K-M) analysis was performed to explore prognostic value of MTHFD1L in different cancers. The cBioPortal was used to investigate the MTHFD1L genetic mutation in various tumor types of TCGA. Finally, MTHFD1L-related genes enrichment analysis was performed to study the functional mechanism of MTHFD1L in carcinogenesis. In most cancers, the mRNA expression of MTHFD1L is higher in the tumor tissues compared to the normal tissues. Besides, higher expressions of MTHFD1L were significantly associated with shorter OS in ACC, BLCA, BRCA, CESC, HNSC, LGG, LIHC, LUAD, SKCM and shorter DFS in ACC, BLCA, CESC, LGG, PRAD and SKCM. The high expression of MTHDF1L was related to the advanced stage of BLCA, LIHC, LUAD, OV, SKCM, UCEC and UCS significantly. What’s more, MTHDF1L expression was positively linked with cancer-associated fibroblast infiltration in HNSC, KIRC, KIRP, LUAD and PAAD. The GO biological process (BP) enrichment includes mitotic cell cycle, cell cycle, mitotic cell cycle process and so on. MTHDF1L physically interacts with CLPP, CS, LRPPRC and MTIF2. This pan-cancer investigation suggested the prognostic value and oncogenic role of MTHFD1L for multiple tumor types.

Introduction

In the world, cancer became an increasing death cause for patients and increased the cost of social economy. According to estimation from Siegel et.al, more than 600,000 American will die of various of cancers in 2022[1]. What’s more, the increasing number of newly diagnosed cases of cancers will over 5000 per day[1], which obviously injury the physical and mental health of those people. Therefore, it is necessary to identify the new biomarkers for cancer in large-scale and muti-omics cancer pattern.

Methylenetetrahydrofolate dehydrogenase 1-like (MTHFD1L) is a monofunctional enzyme to produce formate and promote folic acid cycle[2, 3], which plays an important role in the synthesis of tetrahydrofolate (THF) in mitochondria[4, 5]. As known to all, when the folic acid cycle broken, the related diseases would turn up including immune deficiency diseases, acromegaly and cancers[6, 7]. Besides, MTHFD1L is an exosomal protein, increasing with activation of brown adipose tissues to predict the obesity and metabolic disease[8]. In the heart disease, polymorphism in MTHFD1L is associated with high risk of hypertension and cardiovascular disease[9]. And MTHFD1L modulate the process of pathologic cardiac growth[10]. In colorectal cancer, MTHFD1L promotes the cancer cell growth and high expression of this gene plays a vital role in the process of colorectal cancer[3]. Similar to osteosarcoma, MTHFD1L knockdown is not helpful for tumor cell proliferation and up-regulated expression level of MTHFD1L was
identified in this disease[11]. However, as far as we all know, there are not comprehensive and systematically analysis for the clinical value and function of MTHFD1L in pan-cancer to be reported.

In our research, we performed pan-cancer analysis to explore the mRNA expression level, genetic mutations, molecular function and prognostic validity of MTHFD1L and its association with the cancer-related fibroblast infiltration in a variety of cancers.

**Results**

**Different gene expression levels of MTHFD1L**

According to the information in the database of HAP, GTEx and FANTOM5, low tissues specificity of MTHFD1L was detected. However, the levels of MTHFD1L in a majority of normal tissues are pretty high ([Fig.1A and Supplementary Fig.S1](#)). The low tissue specificity of MTHFD1L hint the conservation of this gene in vertebrates ([Fig.1B](#)). In addition, on the basis of various of single cell types, high degrees of MTHFD1L expression are observed in extravillous trophoblasts, microglial cells, granulosa cells and excitatory neurons ([Supplementary Fig.S1](#)).

The MTHFD1L expression status were investigated in various cancer types. As shown in the [Fig.1C](#), the levels of MTHFD1L mRNA expression were higher in most tumor tissues than the corresponding normal tissues. Compared to the non-tumor tissues, the mRNA expression of MTHFD1L is higher in the tumor tissues of BLCA (bladder urothelial carcinoma), CHOL (cholangiocarcinoma), COAD (colon adenocarcinoma), ESCA (esophageal carcinoma), HNSC (head and neck squamous cell carcinoma), KIRC (kidney renal clear cell carcinoma), KIRP (kidney renal papillary cell carcinoma), LIHC (liver hepatocellular carcinoma), LUAD (lung adenocarcinoma), LUSC (lung squamous cell carcinoma) (all p<0.001), and BRCA (breast invasive carcinoma), CESC (cervical squamous cell carcinoma and endocervical adenocarcinoma) (p<0.05). what's more, the MTHFD1L mRNA expression is lower in the tumor tissues with HNSC-HPV + than the normal tissues with HNSC-HPV – (p<0.001) ([Fig.1C](#)).

**Prognostic value of MTHFD1L**

According to the TCGA datasets, we used GEPIA2 tool to investigate the relationship between MTHFD1L mRNA expression and prognostic survival time of suffers with different cancers. We detected that higher mRNA expression of MTHFD1L is related with shorter survival time in the patients with ACC (adrenocortical carcinoma) (p=2.8×10^{-2}), BLCA (p=8.7×10^{-3}), BRCA (p=4.8×10^{-2}), CESC (p=1.4×10^{-2}), HNSC (p=3.6×10^{-2}), LGG (brain lower grade glioma) (p=8.8×10^{-3}), LIHC (p=1.1×10^{-2}), LUAD (p=9.6×10^{-3}), and SKCM (skin cutaneous melanoma) (p=1.1×10^{-3}) ([Fig.2A-J](#)). Then, we analyzed the association between the mRNA expression of MTHFD1L and DFS. The outcomes suggested that higher MTHFD1L expression is associated with the poorer result for suffers with ACC (p=2.7×10^{-2}), BLCA (p=2.3×10^{-2}), CESC (p=8×10^{-3}), LGG (p=4.7×10^{-2}), PRAD (prostate adenocarcinoma) (p=2.2×10^{-2}) and SKCM (p=3×10^{-2}) ([Fig.3A-G](#)). Furthermore, the correlation between the mRNA expression of MTHFD1L and pathological
stages of cancers was explored in GEPIA2 websites. As you see, the high expression levels of MTHFD1L are related to the advanced stage of BLCA, LIHC, LUAD, OV (ovarian serous cystadenocarcinoma), SKCM, UCEC (uterine corpus endometrial carcinoma) and UCS (uterine carcinosarcoma) (Fig.3H-N). In summary, high mRNA expression of MTHFD1L was involved in poor prognostic value among patients with a majority of cancer types.

The genetic mutation profiles of MTHFD1L in various cancers

The investigation about the genetic mutation profiles of MTHFD1L in different cancers was performed by the cBioPortal tool based on TCGA datasets. The highest MTHFD1L genetic mutation frequency is 7.5% belonging to the cases with UVM (uveal melanoma) (Fig.4A; Supplementary Table S1). And in UVM tumor cases, all the types of genetic mutations are deep deletion. Besides, more than 6% of SARC (sarcoma) and UCEC cases showed genetic mutation of MTHFD1L (Fig.4A; Supplementary Table S1). A summary about MTHFD1L genetic mutations suggested that a total of 165 MTHFD1L mutations, including 131 missense mutations, 17 truncating mutations, 10 SV/Fusion mutations and 7 splice mutations (Fig.4B; Supplementary Table S2). As you see in the Fig.4B, the protein region of R326Q encoded by MTHFD1L had 5 mutations, which is the most frequently altered region in the MTHFD1L protein.

In the next step, we tried to find out the relationship between the genetic mutations of MTHFD1L and clinical prognosis of cancer sufferers. MTHFD1L deep deletion was correlated with poor prognosis of SARC suffers in OS (p=4.48×10^{-2}) and Disease-specific (p=4×10^{-2}) (Fig.4C/D; Supplementary Table S3). Besides, the group of MTHFD1L mutations in the patients with UCEC had a better clinical outcome in terms of PFS (p=2.68×10^{-2}) and OS (p=3.46×10^{-2}) (Fig.4E/F; Supplementary Table S4). In addition, BRCA suffers with MTHFD1L alterations had a pretty prognosis in PFS (p=3.7×10^{-4}) and DFS (p=1.74×10^{-2}) (Fig.4G/H; Supplementary Table S5).

Association between MTHFD1L expression and immune infiltration

We used the tool of GEPIA2 to analyze cancer-related fibroblast infiltration associated with MTHFD1L expression in a majority of tumor types. We detected that MTHFD1L expression was related to cancer-related fibroblast infiltration in HNSC, KIRC, KIRP, LUAD and PAAD (pancreatic adenocarcinoma) (Fig.5).

Enrichment of MTHFD1L-associated genes

Total 100 similar genes with expression patterns analogous to MTHFD1L from pan-cancer in the TCGA database was collected by GEPIA2 to pursue the function mechanism of MTHFD1L on tumorigenesis (Supplementary Table S6). Then, GO and KEGG enrichment analysis revealed that the similar genes were closely associated with metabolic process and cell cycle (Fig.6A). Following step, a PPI network was constructed based on the similar genes by STRIING database, which showed the co-expression between 50 genes and MTHFD1L expression (Fig.6B, Supplementary Table S7). In the database of BioGRID, we detected that MTHFD1L significantly interacts with CLPP, CS, LRPPRC and MTIF2 (Fig.6C), which play important roles in metabolic process and cell cycle (Supplementary Fig.S2). In addition, the level of
mRNA expression of MTHFD1L are positively involved in CLPP, CS, LRPPRC and MTIF2 (Fig.6D-G). According to the investigations, we guessed that MTHFD1L may promote tumor development by regulating cell cycle and influencing metabolic processes.

Discussion

More than 30 types of cancers with the date of gene expression and clinical outcome in the TCGA database, which provided very detailed information for pan-cancer analysis\(^1\). In the past five years, numerous studies have been reported to ascertain the biological and molecular markers and explore their functional roles by pan-cancer analysis\(^1\). In this study, we identified MTHFD1L as a potential biomarker for multiple cancers by pan-cancer analysis and explored the prognostic value and association with immune infiltration of MTHFD1L in numerous tumor types.

As a core enzyme in folate pathway, which participates in numerous cellular pathological and physiological process, MTHFD1L plays an important role in tumorigenesis and development of a majority tumor types\(^2\). In the osteosarcoma tissues, the increased expression level of MTHFD1L results in poor prognosis. And the AKT/mTOR pathway and cell cycle were regulated by MTHFD1L, which significantly promoted the tumor cell proliferation invasion and migration\(^2\). What’s more, MTHFD1L participate in vital molecular axis to impact the progression of disease. For example, MTHFD1L is an important partner of the axis of melatonin-p-CREB1-MTHFD1L-formate, which facilitated to the procession of HNSC\(^2\). And the accelerated level of MTHFD1L expression affected by the mTORC1-4EBP1-eIF4E axis significantly increases the tumor cells in TSCC (tongue squamous cell carcinoma)\(^2\). In bladder cancer, overexpression of MTHFD1L inhibits colony formation and invasion of bladder tumor cells, which hints a potential therapeutic target for this type tumor\(^2\). Inversely, MTHFD1L knockdown significantly decreased the tumor cell growth in papillary thyroid cancer (PTC)\(^2\). Similarly, decreased expression of MTHFD1L suppressed the tumor cell proliferation and promoted apoptosis in OSCC (oral squamous cell carcinoma)\(^2\). However, there are not systematic analysis about the functional roles of MTHFD1L in all kinds of cancer types. Hence, we comprehensively characterized MTHFD1L in all cancer types in TCGA database by investigating the association with gene expression, immune infiltration and genetic mutation.

In this article, we detected that MTHFD1L extensively expressed in a mass of tumor tissues and the levels of MTHFD1L expression are increased in various tumors. Furthermore, we explored the association between the MTHFD1L expression and the clinical data to investigate the prognostic value of MTHFD1L. Results revealed that upregulated expression of MTHFD1L is linked with worse OS, PFS and DFS. Higher mRNA expression of MTHFD1L is involved in poorer prognostic outcome in suffers with ACC, BLCA, BRCA, CESC, HNSC, LGG, LIHC, LUAD, and SKCM. And high MTHFD1L expression is related to the advanced tumor stage exhibiting malignant progression significantly. In the last decade, the roles of genetic mutation in carcinogenesis had been investigated in increasing reports\(^2\). In the present research, we found that genetic alterations of MTHFD1L were most common in UVM (> 6%), then came
back to SARC, UCEC and ACC. Association between mutations of MTHFD1L and clinical results revealed that MTHFD1L deep deletion was a high-risk factor for UCEC and BRCA patients. To sum, these results suggested that MTHFD1L may be an oncogene in the carcinogenesis of different cancers and be an accurate predictor for estimating survival time.

During the tumorigenesis and development of cancers, recruitment of immune cells is universal\cite{27}. In recent years, some reports revealed that the expression of immune microenvironment-associated genes has a positive impact on prognostic outcomes of cancer patients\cite{28,29}. Here, MTHFD1L expression was related to cancer-related fibroblast infiltration in patients with HNSC, KIRC, KIRP, LUAD and PAAD. In summary, MTHFD1L expression have significant effects in tumor immunity, which maybe play a crucial role in poor prognosis of various cancers.

Combined with the co-expression between with the similar genes and MTHFD1L, gene enrichment analysis hinted that these genes were obviously enriched in metabolic process and cell cycle, similar to that MTHFD1L with metabolic superiority promote HCC by the process of folate metabolism\cite{2}. In addition, our outcomes suggested that MTHFD1L strongly interacts with CLPP, CS, LRPPRC and MTIF2. And the mRNA expression of CLPP, CS, LRPPRC and MTIF2 were significantly involved in MTHFD1L expression. Previous study showed that overexpression of CLPP could induce apoptosis in human epithelial ovarian cancer (EOC)\cite{30}. Knockdown LRPPRC is strongly associated with the promotion of apoptosis in prostate cancer\cite{31}. Xu et.al found that upregulated MTIF2 have the ability of suppression in apoptosis and immune cell activity in HCC\cite{32}. The results of enrichment analysis provided a true direction for further study about MTHFD1L biologic process and molecular function.

At last, we concluded that overexpression of MTHFD1L extensively exist in a majority of cancers and its expression and genetic mutation are significantly relevant to clinical data in some definite cancers. Moreover, MTHFD1L-associated gene enrichment analysis and immune analysis showed that MTHFD1L may promote carcinogenesis by regulating cell cycle, metabolic process and tumor immunity. MTHFD1L may act as an oncogenic role in multiple cancers.

**Methods**

**Exploration for MTHFD1L expression levels**

The website, https://www.proteinatlas.org/, was used to analyze the gene expression of MTHFD1L in different tissues based on the function annotation of the mammalian genome (FANTOM5), GTEx and the Human Protein Atlas (HPA). The gene conservation analysis of MTHFD1L was performed in the UCSC genome browser, whose website is https://genome.ucsc.edu/. MTHFD1L expression levels in tumor and normal tissues of various cancers was detected in the module of “Gene DE” belonging to the website (http://timer.cistrome.org/), namely Tumor Immune Estimation Resource version 2 (TIMER2).
Exploration for the prognostic value of MTHFD1L

The module of “Similar Map” in the GEPIA2 database, scilicet Gene Expression Profiling Interactive Analysis version 2, was used to produce the information about overall survival (OS), Kaplan-Maier (K-M) and disease-free survival (DFS) of MTHFD1L. In the module of “Expression DIY” in the GEPIA2 database, we analyzed the association between the gene expression and clinical pathologic stage in multiple cancers. The progression-free survival (PFS) analysis of MTHFD1L was performed in the website (https://xenabrowser.net/) named UCSC Xena Browser. According to MTHFD1L expression level, we set the threshold as 50%.

Identification for genetic mutation of MTHFD1L

The cBioPortal website (https://www.cbioportal.org/) is a network tool for analysis of MTHFD1L genetic mutations. The modules of “Cancer types summary”, “Mutation” and “Comparison/Survival” are important modules in the procession of genetic mutation analysis, where the copy number mutation and frequency of MTHFD1L gene alteration were obtained. All analysis were performed on the bases of TCGA Pan-Cancer Atlas datasets.

Analysis for relationship between MTHFD1L expression and immune infiltration

TIMER2 database has a high degree of authority in the analysis of immune cells infiltration. And the “Immune Association” module where we set the research object as “Cancer associated fibroblast” is a good tool to explore the correlation between MTHFD1L expression and cancer-related fibroblast infiltration.

Enrichment analysis for MTHFD1L-associated genes

Total 100 similar genes of MTHFD1L from the TCGA datasets were collected in the module of “Similar Gene Detection” in the GEPIA2 website. Then Gene Ontology (GO) enrichment and the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment were performed to gain the significant signal pathway and biologic process. Next, we constructed a protein-protein interaction (PPI) network about MTHFD1L-related proteins in the STRING website (https://string-db.org/). Furthermore, the PPI network was embellished in the module of “Concentric Circles” in the website of BioGRID (https://thebiogrid.org/). At last, the function of “Correlation Analysis” in GEPIA2 was used to analyze the relationship of pairwise gene association.

Declarations

Ethics statement: Ethical review and approval were not needed for the investigation based on the local legislation.

Consent for publication: All authors consent for publication.
Data availability statement: Supplementary Tables and Supplementary Figures provided the detailed information about the online repositories which offered the datasets in this investigation.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding: This study was supported by National Natural Science Foundation of China (No. 82071055).

Author contributions: Jian-ming Yang contributed to concept and design of this article. Lin Yan participated in data collection, analysis and manuscript writing. Xin Wang analyzed and interpreted data. All authors contributed to the article and approved the submitted version.

References


Figures
Figure 1

MTHFD1L expression levels in different tumor and non-tumor tissues and gene conversation of MTHFD1L. (A) RNA tissue specificity of MTHFD1L based on datasets of Human Protein Atlas (HPA), GTEx, and function annotation of the mammalian genome (FANTOM5). (B) Visualization about MTHFD1L gene conversation in the UCSC genome browser. (C) MTHFD1L expression levels in different tumor tissues by using TIMER2 database. *p < 0.05; **p < 0.01; ***p < 0.001.
Figure 2

Relationship between MTHFD1L mRNA expression and prognostic survival time of suffers with different TCGA cancers. (A) OS map was built by GEPIA2. (B-J) Kaplan–Meier plots with significant outcomes are showed. The 95% (CI) confidence intervals of OS are displayed by red and blue dotted lines for high and low MTHFD1L groups, severally.
Figure 3

Relationship between MTHFD1L mRNA expression and DFS of suffers with different TCGA cancers. (A) DFS map was built by GEPIA2. (B-G) Kaplan–Meier plots with significant outcomes are showed. The 95% (CI) of DFS are displayed by red and blue dotted lines for high and low MTHFD1L groups, severally. (H-N) Correlation between MTHFD1L expression and pathological stages of BLCA, LIHC, LUAD, OV, SKCM, UCEC and UCS from TCGA datasets. Log2 (TPM + 1) was applied for log-scale.
Figure 4

The genetic mutation profiles of MTHFD1L in various TCGA cancers. The mutation frequency with MTHFD1L genetic mutation type (A) and MTHFD1L alteration site (B) were collected by cBioPortal. The relationship between MTHFD1L mutation status and OS and DFS of Sarcoma (C–D). The association between MTHFD1L mutation status and OS and PFS of UCEC (E–F). The link between MTHFD1L mutation status and DFS and PFS of Breast Invasive Carcinoma (G–H).
Figure 5

Relationship between MTHFD1L mRNA expression and DFS of suffers with different TCGA cancers. (A) DFS map was built by GEPIA2. (B-G) Kaplan–Meier plots with significant outcomes are showed. The 95% (CI) of DFS are displayed by red and blue dotted lines for high and low MTHFD1L groups, severally. (H-N) Correlation between MTHFD1L expression and pathological stages of BLCA, LIHC, LUAD, OV, SKCM, UCEC and UCS from TCGA datasets. Log2 (TPM + 1) was applied for log-scale.
Figure 6

The genetic mutation profiles of MTHFD1L in various TCGA cancers. The mutation frequency with MTHFD1L genetic mutation type (A) and MTHFD1L alteration site (B) were collected by cBioPortal. The relationship between MTHFD1L mutation status and OS and DFS of Sarcoma (C-D). The association between MTHFD1L mutation status and OS and PFS of UCEC (E–F). The link between MTHFD1L mutation status and DFS and PFS of Breast Invasive Carcinoma (G–H).
Supplementary Files

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

- SupplementFigures.docx
- SupplementTables.xlsx