Research Article

Xin Zhang*, Peiying Yang, Guangxian Meng

High expression of folate metabolic pathway gene MTHFD2 is related to the poor prognosis of patients and may apply as a potential new target for therapy of NSCLC

https://doi.org/10.1515/pteridines-2025-0052
received February 02, 2024; accepted June 06, 2025

Abstract

Objective – Data mining was applied to explore the expression, functional enrichment, and signal pathway of methylenetetrahydrofolate dehydrogenase 2 (MTHFD2) gene in non-small cell lung cancer (NSCLC) and its relationship with patient prognosis.

Methods – The expression of MTHFD2 gene in NSCLC and its adjacent tissues was studied by bioinformatics data analysis. The biological function and signal pathway of MTHFD2 gene were enriched, and the MTHFD2 protein–protein interaction (PPI) network was constructed in the string database. The relationship between MTHFD2 expression and overall survival (OS), disease free survival (DFS) of NSCLC patients was analyzed in GEPIA database.

Results – In NSCLC patients, the expression level of MTHFD2 in cancer tissues was significantly higher than that in adjacent normal tissues ($P < 0.05$). There were 21 related proteins in the PPI network, and the interaction relationship between proteins was 173 with the average local clustering coefficient of 0.881. The biological process of MTHFD2 and relevant genes was mainly enriched in tetrahydrofolate metabolic process, one-carbon metabolic process, and folic acid metabolic process. The KEGG pathway of MTHFD2 and relevant genes was mainly enriched in one carbon pool by folate, metabolic pathways, and antifolate resistance pathway. The OS of patients with high expression of MTHFD2 gene in lung adenocarcinoma (LUAD) was significantly lower than that of patients with low expression (hazard ratio [HR] = 1.6, $P = 0.0041$), while the expression level of MTHFD2 was not related to DFS of LUAD (HR = 1.4, $P = 0.05$), lung squamous cell carcinoma (LUSC; HR = 1.3, $P = 0.16$), and OS of LUSC (HR = 0.82, $P = 0.15$). MTHFD2 expression level was correlated with B cells, CD8+ T ($r = 0.143$, $P < 0.05$) and CD4+ T lymphocyte infiltration.

Conclusion – MTHFD2 gene is highly expressed in NSCLC and participates in the metabolism of folic acid and one carbon unit. Its high expression is related to the poor prognosis of patients and may apply as a potential new target for therapy of NSCLC. However, the primary findings are derived from bioinformatic analyses, which, while valuable, necessitate further validation through further empirical methods.

Keywords: methylene tetrahydrofolate dehydrogenase 2, non-small cell lung cancer, bioinformatics analysis

1 Introduction

Lung cancer, including non-small cell lung cancer (NSCLC) and small cell lung cancer, is the leading reason of malignant carcinoma associated mortality globally [1,2]. In year 2020, it was reported that 2.2 million new lung cancer cases were diagnosed and about 1.8 million were dead in the same year [3]. In China, 733.3 thousand new lung cancer cases and 610.2 death were identified in the year 2015 [4]. For US, the incidence and mortality of lung cancer were estimated as 236,740 and 130,180, respectively, in year 2022 [5]. Therefore, lung cancer is a heavy burden for the global medical system. However, the molecular mechanism of lung cancer is not completely clear yet [6].

* **Corresponding author: Xin Zhang,** Department of Oncology, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, No. 88 Changling Road, Xiqing District, Tianjin, 300381, China; National Clinical Research Center for Chinese Medicine Acupuncture and Moxibustion, Tianjin, 300381, China; Tianjin Cancer Institute of Traditional Chinese Medicine, Tianjin, 300381, China, e-mail: zhangxin10242022@163.com, tel: +86-13821183558
Peiying Yang: Department of Oncology, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, No. 88 Changling Road, Xiqing District, Tianjin, 300381, China; National Clinical Research Center for Chinese Medicine Acupuncture and Moxibustion, Tianjin, 300381, China; Tianjin Cancer Institute of Traditional Chinese Medicine, Tianjin, 300381, China
Guangxian Meng: Department of Central Sterile Supply, Tianjin Hospital, Tianjin, 300211, China

Folic acid is an essential B vitamin for human body [7]. It is transformed into tetrahydrofolate in human body, and tetrahydrofolate can carry a carbon unit to participate in the synthesis of nucleotide and deoxynucleoside nucleic acid [8]. Folic acid plays an important role in maintaining genome stability [9,10]. The folate metabolism pathway includes methionine synthase [11], methylenetetrahydrofolate reductase, methylenetetrahydrofolate dehydrogenase 1, methylenetetrahydrofolate dehydrogenase 2 (MTHFD2), and mitochondrial aminomethyltransferase 1 and 2, which play an important role in folate metabolism [12]. MTHFD2 is a nicotinamide adenine dinucleotide phosphate dependent bifunctional enzyme located in mitochondria, which is highly expressed in mitochondrial folate metabolism [13,14]. In recent years, many studies have shown that the overexpression of MTHFD2 can promote the occurrence and development of breast cancer [15], colon cancer [16], ovarian cancer [17], glioma [18], and other types of tumors. Yu and his colleges found the inhibitory effect of MTHFD2 knockdown on NSCLC may be mediated via suppressing cell cycle-related genes [19]. These findings delineate the role of MTHFD2 in the development of NSCLC and may have potential applications in the treatment of NSCLC. However, there are few reports on the expression of MTHFD2 in NSCLC and its relationship with the patient's prognosis. The investigation of MTHFD2 in NSCLC research is supported by its roles in tumor metabolism and immune modulation, which are distinct from other metabolic enzymes and align with emerging therapeutic vulnerabilities in NSCLC. MTHFD2's integration of metabolic reprogramming and immune evasion positions it as a priority target in NSCLC. This study proposes that MTHFD2 serves as a critical metabolic driver in NSCLC, with its overexpression in tumor tissues not only promoting tumor progression through dysregulation of folate-mediated one-carbon metabolism but also shaping the tumor immune microenvironment.

2 Data collection and analysis

2.1 MTHFD2 mRNA expression analysis

MTHFD2 relative expression of a variety of solid tumors, including NSCLC, were investigated in GEPIA (GEPIA2021/2021-7-2) (http://gepia.cancer-pku.cn/detail.php) and TCGA databases (data freeze: 2016-01-15) (https://portal.gdc.cancer.gov/) with the sample size of LUAD: n = 483 tumors, 347 normals; LUSC: n = 486 tumors, 338 normals. MTHFD2 relative expression in human cancer cell lines was also explored in the database of "THE HUMAN PROTEIN ATLAS" (release: 2023-08) (https://www.proteinatlas.org/). The keywords of "lung cancer/non-small cell lung cancer/carcinoma of the lung" were applied for data searching. MTHFD-encoded protein expression analysis was investigated in protein atlas database (https://www.proteinatlas.org/database).

2.2 Protein–protein interaction (PPI) network construction of MTHFD2 and relevant genes

In the STRING (v.11.5, accessed on 2023-10-01) (http://string-db.org/cgi/input.pl) database, the network of PPI encoded by MTHFD2 gene was constructed under the condition that the confidence degree was greater than 0.7 and the number of interacting proteins was more than 20. The over-representation analysis of MTHFD2 and relevant genes was performed using the built-in functionality of STRING.

2.3 MTHFD2 gene ontology enrichment

MTHFD2 and relevant genes were enriched in the aspect of biological process, cellular component, and molecular function for gene ontology in the database of DAVID (v.2021, accessed on 2023-10-01) (https://david.ncifcrf.gov/). The number of genes, genetic background, and the related intensity and P values were calculated in the enrichment.

2.4 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of MTHFD2 and relevant genes

MTHFD2 and relevant genes were enriched for KEGG pathway in the KEGG database through calculating the number of genes, genetic background, and the related intensity and P values by an online tool DAVID (https://david.ncifcrf.gov/). Significant KEGG pathways were identified using a dual-threshold approach: (1) p-value < 0.05 (hypergeometric test with Benjamini-Hochberg correction for multiple testing) and (2) enrichment score (ES) > 0 (permutation-based GSEA), with false discovery rate (FDR) < 0.25 for ES. Pathways meeting either criterion were considered significant unless overlapping gene sets required stricter FDR < 0.05 for functional specificity.

2.5 Correlation between MTHFD2 expression and prognosis

According to MTHFD2 median expression in cancer tissue of NSCLC, the NSCLC cases were divided into high

expression group (MTHFD2 ≥ median expression) and low expression group (MTHFD2 < median expression). The overall survival (OS) and disease-free survival (DFS) between MTHFD2 high and low expression groups were analyzed in the GEPIA databases.

2.6 Relationship between MTHFD2 expression and lymphocyte infiltration

The relationship between MTHFD2 expression and lymphocyte infiltration was analyzed by Spearman's Rank Correlation test by using the data from TIMER: Tumor IMmune Estimation Resource (v.2.0, accessed on 2023-10-01) (https://cistrome.shinyapps.io/timer/) database. The relationship analysis of MTHFD2 expression and lymphocyte infiltration was independent in each cell type therefore with small gene sets. The correlation analysis module visualizes pairwise gene expression scatterplots in NSCLC samples, annotated with Spearman's rank correlation coefficients and corresponding statistical significance estimates (*p*-values). Correlation module draws the expression scatterplots between MTHFD2 and lymphocyte infiltration in NSCLC, together with the Spearman's rho value and estimated statistical significance. Partial correlation conditioned on tumor purity or age are also provided. Partial correlation (partial.cor) measures the correlation between two variables while controlling for other factors.

2.7 Batch-effect handling, transformation, and outlier filtering

Batch effects were addressed using ComBat for known technical batches or Surrogate Variable Analysis for unknown confounders, with validation via PCA to confirm reduced batch variance. Data transformation involved converting transcript counts to transcripts per million (TPM) followed by log2 (TPM+1) to stabilize variance and approximate normality. Outlier filtering removed low-abundance genes (TPM < 1 in > 90% samples) and extreme sample outliers detected by *Z*-score ($|Z| > 3$) or interquartile range (IQR) (≥1.5 × IQR). Missing data imputation was performed using k-nearest neighbors for genes with >30% missingness, validated by Spearman correlation, while genes with >50% missingness were excluded.

2.8 Statistical analysis

R4.1.0 software (https://www.r-project.org/) was applied for statistical analysis. MTHFD2 relative expression between cancer and normal tissue was performed by *t*-test. The comparison of MTHFD2 expression between cancer and normal tissue was independent in each cancer type with small gene sets; therefore, the *p*-value was not adjusted for multiple testing. PPI network analysis was constructed in the online statistical tool (http://string-db.org/cgi/input.pl). The relationship between MTHFD2 and the clinical stage of NSCLC patients was analyzed by Spearman's Rank Correlation test. Survival analysis between MTHFD2 high and low expression groups was performed by log-rank test and demonstrated by hazard ratio (HR). $P < 0.05$ was considered statistically significant.

3 Results

3.1 MTHFD2 expression in human solid tumors

The expression level of MTHFD2 in solid tumor tissues such as bladder cancer, esophageal cancer, head and neck squamous cell carcinoma, gastric cancer, and colorectal cancer was significantly higher than that in corresponding normal tissues ($P < 0.05$), Figure 1a. In NSCLC patients, the expression level of MTHFD2 in cancer tissues was significantly higher than that in adjacent tissues ($P < 0.05$), Figure 1b. The expression level of MTHFD2 was related to clinical stage ($P < 0.05$), Figure 1c. MTHFD-encoded protein was mainly expressed in the cytoplasm with positive expression of brownish yellow particle staining, Figure 1d and e (https://www.proteinatlas.org/).

3.2 PPI network construction of MTHFD2 and relevant genes

In the protein interaction database (STRING), the PPI network of MTHFD2 was constructed. There were 21 proteins, which interacted with MTHFD2 in the network, and the interaction relationship between proteins was 173 with the average local clustering coefficient of 0.881, Figure 2. The 21 genes interacting with MTHFD2 are listed in Table 1.

3.3 GO enrichment of MTHFD2 gene

The biological process of MTHFD2 and relevant genes was mainly enriched in tetrahydrofolate metabolic process, one-carbon metabolic process, and folic acid metabolic process (Table 2). Cellular component of MTHFD2 and relevant genes was mainly enriched in mitochondrial matrix

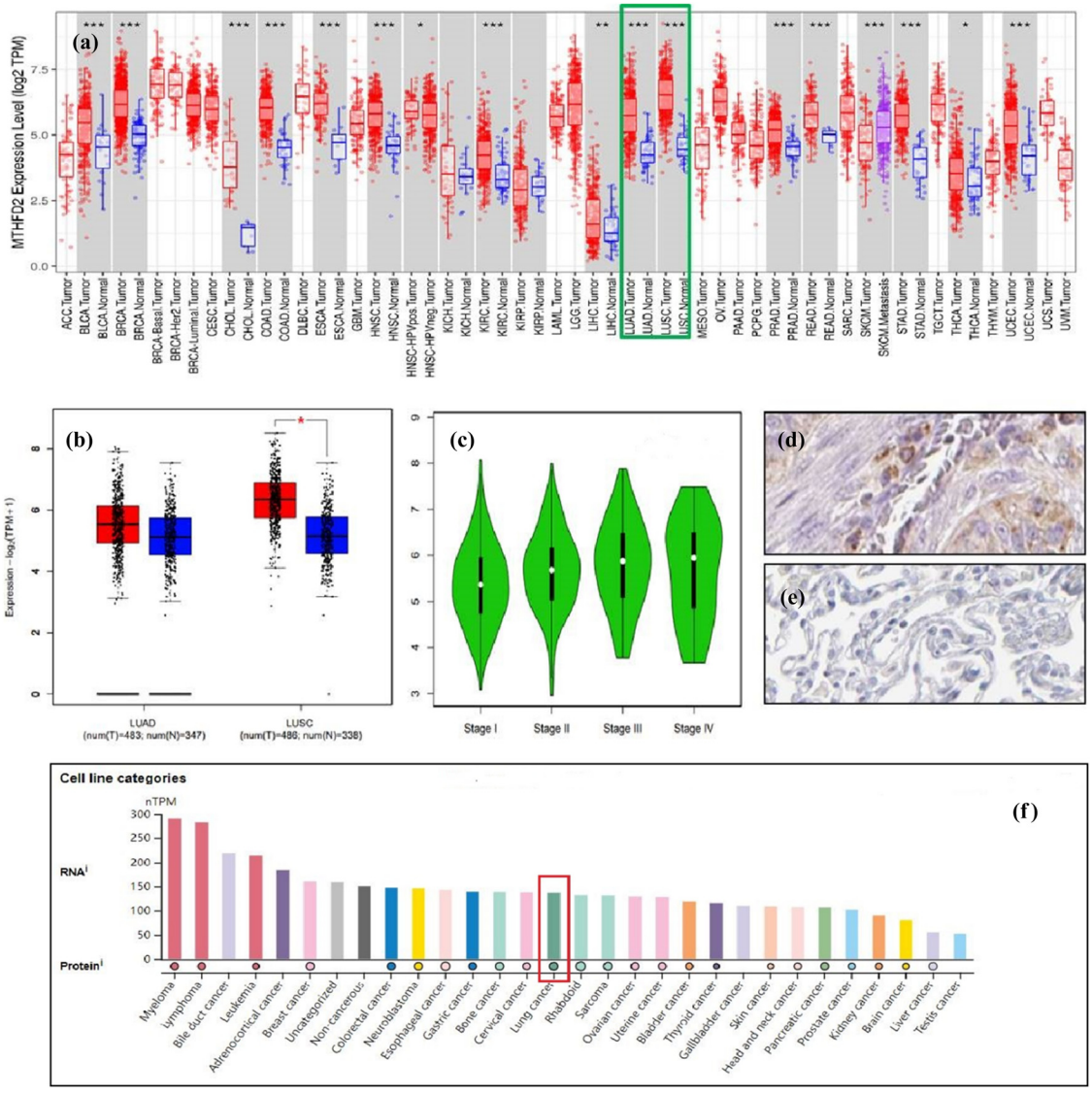

Figure 1: Relative expression of MTHFD2 gene in human solid tumors and NSCLC. (a) Expression of MTHFD2 in a variety of human tumors; (b) relative expression of MTHFD2 in cancer tissues and normal bronchial epithelial tissues of NSCLC patients; (c) relationship between MTHFD2 expression level and clinical stages of NSCLC patients; (d) positive expression of MTHFD2 in the carcinoma tissues of NSCLC patients with brown-yellow staining; (e) MTHFD2 gene has negative expression in normal lung tissue; (f) expression of MTHFD2 in a variety of human cancer cell lines. *P < 0.05; **P < 0.01, ***P < 0.001.

and mitochondrion. Molecular function of MTHFD2 and relevant genes was mainly enriched in hydroxymethyl-, formyl-, and related transferase activity, oxidoreductase activity, acting on the CH-NH group of donors, NAD, or NADP as acceptor and methylenetetrahydrofolate dehydrogenase (NADP+) activity (Table 3).

3.4 KEGG pathway enrichment of MTHFD2 gene

The KEGG pathway of MTHFD2 and relevant genes was mainly enriched in one carbon pool by folate, metabolic pathways, antifolate resistance pathway, etc. Figure 3 and Table 4.

3.5 MTHFD2 expression and prognosis

The OS of patients with high expression of MTHFD2 gene in lung adenocarcinoma (LUAD) was significantly lower than that of patients with low expression (HR = 1.6, P = 0.0041), while the expression level of MTHFD2 was not related to the DFS of LUAD (HR = 1.4, P = 0.05), lung squamous cell carcinoma (LUSC; HR = 1.3, P = 0.16), and OS of LUSC (HR = 0.82, P = 0.15), Figure 4.

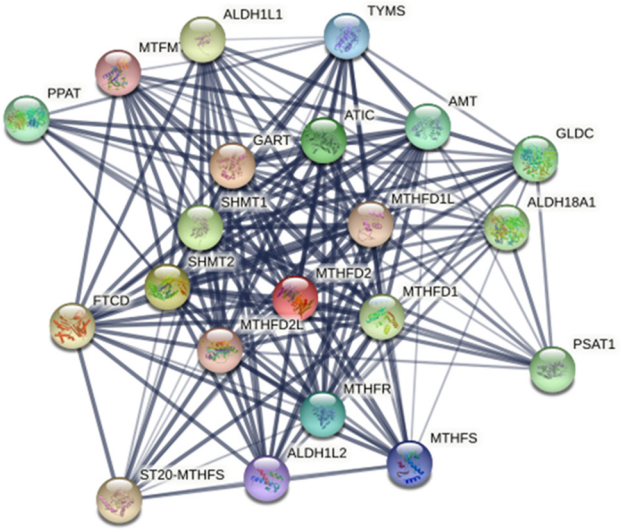

Figure 2: PPI network of MTHFD2 and relevant genes. 21 proteins interacting with MTHFD2 were constructed in the network, and the interaction relationship between proteins was 173 with the average local clustering coefficient of 0.881.The thickness of the PPI connections (edges) means the strength of data support (confidence score: 0.400). The nodes means the corresponding genes.

3.6 Relationship between MTHFD2 expression and lymphocyte infiltration

MTHFD2 expression level was correlated with B cells (r = −0.174, P < 0.05), CD8+ T cells (r = 0.143, P < 0.05), and CD4+ T cells (r = −0.236, P < 0.05) lymphocyte infiltration in LUAD. For LUSC, MTHFD2 expression level was correlated with CD4+ T cells (r = −0.195, P < 0.05) lymphocyte infiltration, Figure 5.

4 Discussion

MTHFD2 is an enzyme involved in folate metabolism and one-carbon metabolism, pathways critical for nucleotide synthesis and cellular proliferation. Its biological significance is particularly pronounced in the context of cancer, as aberrant MTHFD2 expression has been linked to various malignancies. Elevated levels of MTHFD2 are often associated with enhanced tumor growth and metastasis, suggesting its role as a facilitator of cancer cell bioenergetics and survival. This enzyme not only supports the biosynthesis of nucleotides necessary for DNA replication and repair but also influences the epigenetic landscape of cancer cells through its involvement in the generation of S-adenosylmethionine. Thus, targeting MTHFD2 presents a promising therapeutic strategy, highlighting the need for further research to fully elucidate its mechanisms and potential as a biomarker for treatment response in oncology. MTHFD2 plays a pivotal role in the mitochondrial folic acid cycle, which is crucial for cellular

Table 1: Description of MTHFD2 and relevant genes

Gene name	Gene description or protein
MTHFD2	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase
MTHFD1L	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1 like
MTHFD1	Methylenetetrahydrofolate dehydrogenase, cyclohydrolase, and formyltetrahydrofolate synthetase 1
MTHFR	Methylenetetrahydrofolate reductase
MTHFD2L	Methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2 like
SHMT2	Serine hydroxymethyltransferase 2
SHMT1	Serine hydroxymethyltransferase 1
GART	Phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase
AMT	Aminomethyltransferase
GLDC	Glycine decarboxylase
ALDH18A1	Aldehyde dehydrogenase 18 family member A1
PSAT1	Phosphoserine aminotransferase 1
MTHFS	Methenyltetrahydrofolate synthetase
ALDH1L2	Aldehyde dehydrogenase 1 family member L2
ST20-MTHFS	ST20-MTHFS readthrough
FTCD	Formimidoyltransferase cyclodeaminase
PPAT	Phosphoribosyl pyrophosphate amidotransferase
MTFMT	Mitochondrial methionyl-tRNA formyltransferase
ALDH1L1	Aldehyde dehydrogenase 1 family member L1
TYMS	Thymidylate synthetase

Table 2: Biological process enrichment for MTHFD2 gene

Biological process description	Observed gene count	Background gene count	Strength	*P*-value
Tetrahydrofolate metabolic process	14	18	2.86	5.01×10^{-32}
One-carbon metabolic process	12	39	2.46	2.22×10^{-23}
Tetrahydrofolate interconversion	10	10	2.97	7.29×10^{-23}
Dicarboxylic acid metabolic process	13	93	2.11	6.70×10^{-22}
Carboxylic acid metabolic process	19	853	1.32	5.62×10^{-21}
Folic acid metabolic process	10	19	2.69	5.62×10^{-21}
Alpha-amino acid metabolic process	13	191	1.8	2.34×10^{-18}
Water-soluble vitamin metabolic process	11	83	2.09	5.69×10^{-18}
Cellular amide metabolic process	15	773	1.26	3.37×10^{-14}
Carboxylic acid biosynthetic process	11	290	1.55	2.07×10^{-12}
Serine family amino acid metabolic process	7	37	2.25	1.28×10^{-11}
10-formyltetrahydrofolate metabolic process	5	5	2.97	1.30×10^{-10}

metabolism and proliferation. Recent studies have highlighted its significant impact on various cancer types, including colon cancer and glioma, where it influences critical processes such as cell proliferation, apoptosis, and drug resistance [19–21]. The enzyme's involvement in folate metabolism not only underscores its metabolic importance but also positions it as a potential therapeutic target in cancer treatment. Understanding the mechanisms by which MTHFD2 contributes to tumor biology may lead to the development of more effective strategies for combating resistance to conventional therapies and improving patient outcomes in malignancies linked to folate metabolism dysregulation. However, the role of MTHFD2 NSCLC remains largely underexplored, as there is a paucity of research concerning its expression, biological function, and prognostic implications. Despite its significant involvement in folate metabolism and cellular processes that influence tumor progression, studies focusing specifically on MTHFD2 in the context of NSCLC have been limited. This gap in knowledge underscores the need for further investigation to elucidate the potential contributions of MTHFD2 to tumor biology and patient outcomes, which could ultimately inform therapeutic strategies and improve clinical management of NSCLC.

In the present work we performed an integrated bioinformatics analysis of expression and significance of folate metabolic pathway gene MTHFD2 in NSCLC. In this present work, it reveals a significant correlation between the expression levels of MTHFD2 and the prognosis of NSCLC patients, indicating that elevated MTHFD2 expression is associated with poorer outcomes. Interestingly, contrasting findings have emerged from research focusing on liver cancer, where high levels of MTHFD2 expression correlate with improved disease-free survival. These divergent results underscore the complexity of MTHFD2's role in cancer biology, suggesting that its prognostic significance

Table 3: Molecular function enrichment for MTHFD2 gene

Molecular function description	Observed gene count	Background gene count	Strength	*P*-value
Hydroxymethyl-, formyl-, and related transferase activity	8	8	2.97	1.57×10^{-17}
Oxidoreductase activity, acting on the CH-NH group of donors, NAD, or NADP as acceptor	7	18	2.56	4.08×10^{-13}
Transferase activity, transferring one-carbon groups	10	212	1.64	7.31×10^{-12}
Methylenetetrahydrofolate dehydrogenase (NADP+) activity	4	4	2.97	4.41×10^{-8}
Catalytic activity	20	5,486	0.53	5.67×10^{-8}
Cyclohydrolase activity	4	5	2.87	5.67×10^{-8}
Vitamin binding	7	142	1.66	5.67×10^{-8}
Methenyltetrahydrofolate cyclohydrolase activity	3	3	2.97	6.86×10^{-6}
Methylenetetrahydrofolate dehydrogenase (NAD+) activity	3	3	2.97	6.86×10^{-6}
Oxidoreductase activity	9	726	1.06	7.43×10^{-6}
Transferase activity	13	2,170	0.75	8.44×10^{-6}

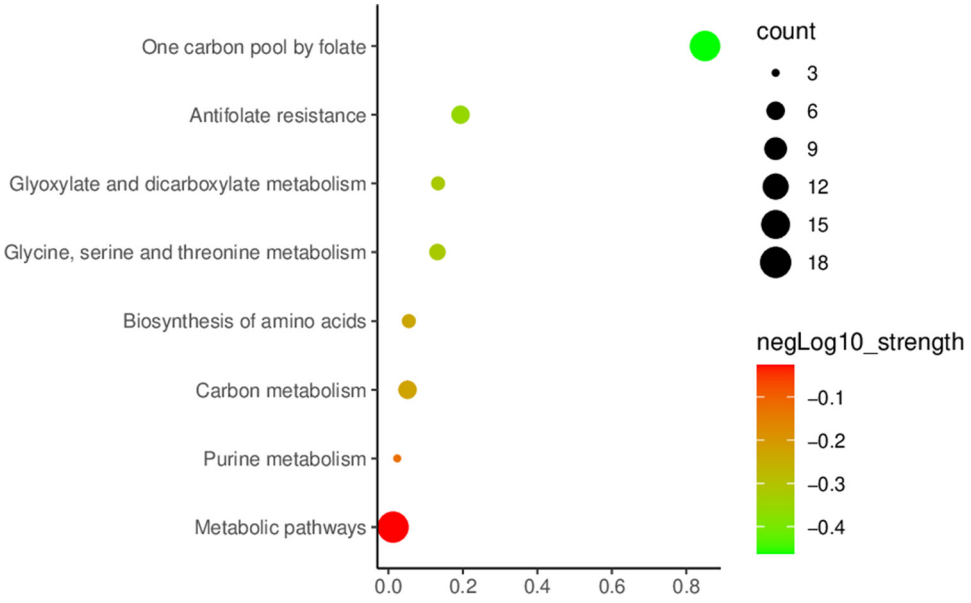

Figure 3: The KEGG pathway of MTHFD2 and relevant genes was mainly enriched in one carbon pool by folate, metabolic pathways and antifolate resistance pathway. The bubble (count = observed gene number) means observed genes and the negLog10_strength means the strength of data support [negLog10_strength = −log10 (*p*-value)].

may vary significantly across different tumor types. The differential prognostic impact of MTHFD2 in LUAD vs squamous cell carcinoma (LUSC) may stem from intrinsic biological and metabolic disparities between these NSCLC subtypes. LUAD, often driven by mutations in EGFR or KRAS, exhibits heightened reliance on one-carbon metabolism for nucleotide synthesis and redox balance, processes where MTHFD2 is pivotal [22]. This metabolic addiction may render LUAD more vulnerable to MTHFD2 overexpression, accelerating proliferation and therapy resistance, thereby worsening survival [22]. In contrast, LUSC, typically linked to PI3K/AKT or SOX2 alterations, may depend less on folate-driven pathways, instead prioritizing glycolytic or glutaminolytic routes, diminishing MTHFD2's prognostic relevance [23]. Additionally, the tumor microenvironment (TME) may differ between LUAD and LUSC. LUAD shows stronger correlations between MTHFD2 and immune infiltration (e.g., CD8+ T cells), suggesting its role in modulating immunosuppression. High MTHFD2 might enhance tumor immune evasion in LUAD through metabolic competition or checkpoint interactions, indirectly affecting survival. LUSC's TME, enriched with neutrophils and distinct stromal interactions, may buffer such effects. Molecularly, LUAD's higher anabolic demands and oncogene-metabolism crosstalk could amplify MTHFD2's oncogenic role, while LUSC's squamous differentiation pathways (e.g., TP63) may override folate metabolism's influence. These subtype-specific dependencies and TME dynamics likely explain the survival association divergence [24]. Consequently, to better understand these discrepancies and validate the implications of MTHFD2 expression on patient outcomes, further investigation through large-scale clinical trials is warranted.

The growth, invasion, metastasis, and recurrence of tumors are a complex pathological process. Tumor

Table 4: KEGG pathway enrichment of MTHFD2 gene

KEGG pathway description	Observed gene count	Background gene count	Strength	FDR
One carbon pool by folate	17	20	2.9	1.24×10^{-42}
Metabolic pathways	18	1,447	1.06	8.75×10^{-16}
Antifolate resistance	6	31	2.26	1.77×10^{-10}
Glycine, serine, and threonine metabolism	5	38	2.09	6.68×10^{-8}
Carbon metabolism	6	117	1.68	1.83×10^{-7}
Glyoxylate and dicarboxylate metabolism	4	30	2.09	2.48×10^{-6}
Biosynthesis of amino acids	4	73	1.71	6.03×10^{-5}
Purine metabolism	3	127	1.34	0.0146

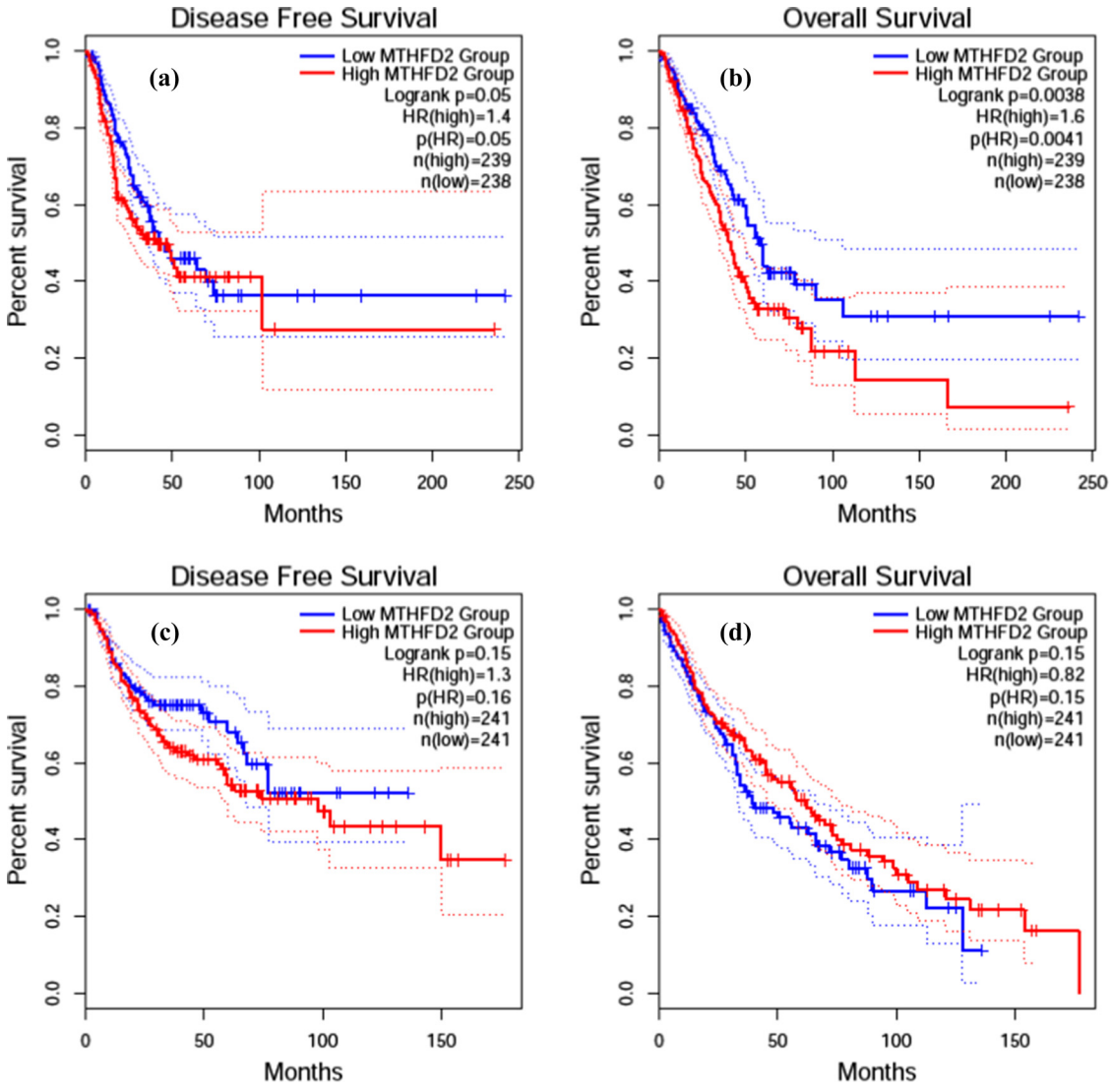

Figure 4: Survival curve of MTHFD2 expression and NSCLC patient's prognosis. (a) DFS of LUAD between MTHFD2 high and low expression groups; (b) OS of LUAD between MTHFD2 high and low expression groups; (c) DFS of LUSC between MTHFD2 high and low expression groups; and (d) OS of LUSC between MTHFD2 high and low expression groups.

lymphocyte infiltration may refer to the infiltrating lymphocytes in tumors and their surrounding tissues, which are the products of the specific immune response of the body to tumors and reflect the body's anti-tumor immune response. In the correlation analysis of the present work between MTHFD2 and immune infiltrating cells, it was found that MTHFD2 was associated with a variety of immune cell infiltrates, suggesting that folic acid metabolic pathway is related to body immunity. MTHFD2 has emerged as a significant player in the field of tumor immunity due to its important role in cellular metabolism and the regulation of folate pathways. Recent studies suggest that MTHFD2 not only contributes to the metabolic reprogramming of cancer cells but also influences cancer immune evasion through PD-L1 up-regulation [25]. Elevated expression of MTHFD2 in various malignancies has been associated with immune evasion tactics employed by tumors, including the suppression of T-cell activation and the promotion of an immunosuppressive milieu [26]. Consequently, targeting MTHFD2 presents a promising therapeutic strategy, potentially enhancing anti-tumor immunity and improving patient outcomes through the restoration of effective immune responses against cancer cells. Therefore, immunotherapy for NSCLC combined with drugs targeting folic acid metabolic pathway may achieve better therapeutic effect compared to immunotherapy alone. This work also identified that the biological process of MTHFD2 and relevant genes was mainly enriched in

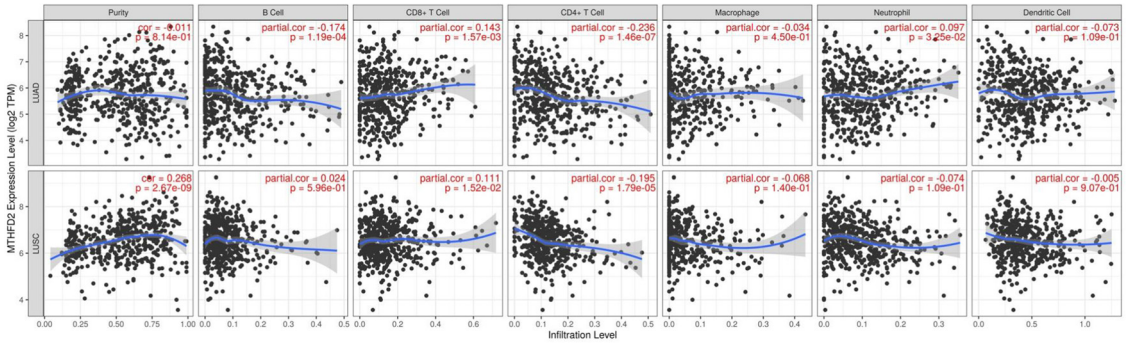

Figure 5: Scatter plot of relationship between MTHFD2 expression and lymphocyte infiltration. Correlation module draws the expression scatterplots between MTHFD2 and lymphocyte infiltration in NSCLC, together with the Spearman's rho value and estimated statistical significance. Partial correlation conditioned on tumor purity or age are also provided. Partial correlation (partial.cor) measures the correlation between two variables while controlling for other factors. log2 TPM refers to the logarithm (base 2) of TPM, which is a normalized measure of gene expression. LUSC: Lung Squamous Cell Carcinoma; LUAD: Lung Adenocarcinoma. MTHFD2 expression level was correlated with B cells ($r = -0.174$, $P < 0.05$), CD8+ T cells ($r = 0.143$, $P < 0.05$), and CD4+ T cells ($r = -0.236$, $P < 0.05$) lymphocyte infiltration in LUAD. For LUSC, MTHFD2 expression level was correlated with CD4+ T cells ($r = -0.195$, $P < 0.05$) lymphocyte infiltration.

tetrahydrofolate metabolic process, one-carbon metabolic process, and folic acid metabolic process. The KEGG pathway of MTHFD2 and relevant genes was mainly enriched in one carbon pool by folate, metabolic pathways and antifolate resistance pathway. MTHFD2 protein, also known as the four-hydrogen methylene folic acid hydrolase, is one of the mitochondrial enzymes encoded by MTHFD2 gene located in the second chromosome in human being. This enzyme was first identified in Ehrlich ascites tumor cells in 1960 by Scrimgeour and Huennekens [27]. MTHFD2 has the dual activities of methylene tetrahydrofolate dehydrogenase and cyclohydrolase, and is a key enzyme in the reciprocal conversion reaction of folic acid metabolites *in vivo* [28]. Therefore, MTHFD2 metabolic pathway may provide a potential cancer including NSCLC treatment target.

Tumor cells have rapid glycolysis and synthesis of amino acids, nucleotides, and lipids to support the rapid proliferation of tumor cells. In particular, the synthesis of one carbon unit carried by the tetrahydrofolate cofactor is important for cells proliferating, which are required for nucleotide synthesis and methylation reactions [29]. Previous publications have shown that MTHFD2 is expressed in fetal cells and transformed cell lines, but has low or absent expression in most adult tissues and cell types [30,31]. In recent years, MTHFD2 has been shown to be important for rapidly growing cells, such as embryonic or tumor cells, mainly by supporting the high levels of purine synthesis required [32]. Nisson and his colleges performed a study including 19 cancer types and found that MTHFD2 mRNA and protein expressions were elevated in the mitochondrial folate pathway in highly proliferative cancers [33]. Furthermore, a number of studies have shown that MTHFD2 is highly expressed in breast cancer [15], renal carcinoma [34], and is associated with poor prognosis. The increased expression of MTHFD2 has been proved to be closely related to the occurrence and development of tumors and patient's prognosis by clinical study.

The current study presents significant insights; however, it is important to acknowledge inherent limitations. The primary findings are derived from bioinformatic analyses, which, while valuable, necessitate further validation through empirical methods. To strengthen the conclusions drawn from this research, it is essential that subsequent investigations incorporate laboratory-based cell experiments and robust clinical data. Such comprehensive validation will not only enhance the reliability of the findings but also facilitate a more thorough understanding of the underlying biological mechanisms.

5 Conclusion

In conclusion, the key gene of folate metabolism pathway MTHFD2 may play an important role in the occurrence and development of lung cancer. In addition to affecting the regulation of folate metabolic pathway, the changes in MTHFD2 in patients with lung cancer can also affect the immune microenvironment of the human body, which may involve in the occurrence and development of lung cancer by affecting the immune microenvironment and metabolic pathways.

Funding information: This work was supported by Tianjin Municipal Education Commission Scientific Research Project (Project Number: 2024KJ023).

Author contributions: Xin Zhang designed the research framework, performed bioinformatics data mining, conducted functional enrichment and survival analysis, and drafted the manuscript with figures and tables. Peiying Yang validated the statistical significance of results, explored the correlation between MTHFD2 expression and immune infiltration, and implemented software-based data preprocessing. Guangxian Meng supervised the study, secured computational resources, revised the manuscript, and cross-validated key findings. All authors contributed to data interpretation, manuscript revisions, and final approval.

Conflict of interest: The authors state no conflict of interest.

Foundation: This work was supported by scientific research project of Tianjin education commission (2019KJ051).

Data availability statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

[1] Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. Int J Cancer. 2021. doi: 10.1002/ijc.33588.

[2] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.

[3] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.

[4] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66(2):115–32.

[5] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. CA Cancer J Clin. 2022;72(1):7–33.

[6] Chang L, Fang S, Gu W. The molecular mechanism of metabolic remodeling in lung cancer. J Cancer. 2020;11(6):1403–11.

[7] Goldsmith GA. Vitamin B complex. Thiamine, riboflavin, niacin, folic acid (folacin), vitamin B12, biotin. Prog Food Nutr Sci. 1975;1(9):559–609.

[8] Menezo Y, Elder K, Clement A, Clement P. Folic acid, folinic acid, 5 methyl tetrahydrofolate supplementation for mutations that affect epigenesis through the folate and one-carbon cycles. Biomolecules. 2022;12(2):197.

[9] Teo T, Fenech M. The interactive effect of alcohol and folic acid on genome stability in human WIL2-NS cells measured using the cytokinesis-block micronucleus cytome assay. Mutat Res. 2008;657(1):32–8.

[10] Kimura M, Umegaki K, Higuchi M, Thomas P, Fenech M. Methylenetetrahydrofolate reductase C677T polymorphism, folic acid and riboflavin are important determinants of genome stability in cultured human lymphocytes. J Nutr. 2004;134(1):48–56.

[11] Shi Q, Zhang Z, Li G, Pillow PC, Hernandez LM, Spitz MR, et al. Polymorphisms of methionine synthase and methionine synthase reductase and risk of lung cancer: a case-control analysis. Pharmacogenet Genomics. 2005;15(8):547–55.

[12] Binia A, Contreras AV, Canizales-Quinteros S, Alonzo VA, Tejero ME, Silva-Zolezzi I. Geographical and ethnic distribution of single nucleotide polymorphisms within genes of the folate/homocysteine pathway metabolism. Genes Nutr. 2014;9(5):421.

[13] Zhang F, Wang D, Li J, Su Y, Liu S, Lei QY, et al. Deacetylation of MTHFD2 by SIRT4 senses stress signal to inhibit cancer cell growth by remodeling folate metabolism. J Mol Cell Biol. 2022;14(4):mjac020.

[14] Yang C, Zhang J, Liao M, Yang Y, Wang Y, Yuan Y, et al. Folate-mediated one-carbon metabolism: a targeting strategy in cancer therapy. Drug Discov Today. 2021;26(3):817–25.

[15] Huang J, Qin Y, Lin C, Huang X, Zhang F. MTHFD2 facilitates breast cancer cell proliferation via the AKT signaling pathway. Exp Ther Med. 2021;22(1):703.

[16] Ju HQ, Lu YX, Chen DL, Zuo ZX, Liu ZX, Wu QN, et al. Modulation of redox homeostasis by inhibition of MTHFD2 in colorectal cancer: Mechanisms and therapeutic implications. J Natl Cancer Inst. 2019;111(6):584–96.

[17] Cui X, Su H, Yang J, Wu X, Huo K, Jing X, et al. Up-regulation of MTHFD2 is associated with clinicopathological characteristics and poor survival in ovarian cancer, possibly by regulating MOB1A signaling. J Ovarian Res. 2022;15(1):23.

[18] Shi LF, Zhang Q, Shou XY, Niu HJ. Expression and prognostic value identification of methylenetetrahydrofolate dehydrogenase 2 (MTHFD2) in brain low-grade glioma. Int J Gen Med. 2021;14:4517–27.

[19] Li Q, Yang F, Shi X, Bian S, Shen F, Wu Y, et al. MTHFD2 promotes ovarian cancer growth and metastasis via activation of the STAT3 signaling pathway. FEBS Open Bio. 2021;11(10):2845–57.

[20] Wei Y, Liu P, Li Q, Du J, Chen Y, Wang Y, et al. The effect of MTHFD2 on the proliferation and migration of colorectal cancer cell lines. Onco Targets Ther. 2019;12:6361–70.

[21] Liu X, Liu S, Piao C, Zhang Z, Zhang X, Jiang Y, et al. Non-metabolic function of MTHFD2 activates CDK2 in bladder cancer. Cancer Sci. 2021;112(12):4909–19.

[22] Yao S, Peng L, Elakad O, Küffer S, Hinterthaner M, Danner BC, et al. One carbon metabolism in human lung cancer. Transl Lung Cancer Res. 2021;10(6):2523–38.

[23] Na J, Zhou W, Yin M, Hu Y, Ma X. GNA13 promotes the proliferation and migration of lung squamous cell carcinoma cells through regulating the PI3K/AKT signaling pathway. Tissue Cell. 2022;76:101795.

[24] Shi X, Peng X, Chen Y, Shi Z, Yue C, Zuo L, et al. Overexpression of MTHFD2 represents an inflamed tumor microenvironment and precisely predicts the molecular subtype and immunotherapy response of bladder cancer. Front Immunol. 2023;14:1326509.

[25] Shang M, Yang H, Yang R, Chen T, Fu Y, Li Y, et al. The folate cycle enzyme MTHFD2 induces cancer immune evasion through PD-L1 up-regulation. Nat Commun. 2021;12(1):1940.

[26] Sugiura A, Andrejeva G, Voss K, Heintzman DR, Xu X, Madden MZ, et al. MTHFD2 is a metabolic checkpoint controlling effector and regulatory T cell fate and function. Immunity. 2022;55(1):65–81.e9.

[27] Scrimgeour KG, Huennekens FM. Occurrence of a DPN-linked, N 5,N 10-methylene tetrahydrofolic dehydrogenase in Ehrlich ascites tumor cells. Biochem Biophys Res Commun. 1960;2(3):230–3.
[28] Christensen KE, Mackenzie RE. Mitochondrial methylenetetrahydrofolate dehydrogenase, methenyltetrahydrofolate cyclohydrolase, and formyltetrahydrofolate synthetases. Vitam Horm. 2008;79:393–410.
[29] Li AM, Ducker GS, Li Y, Seoane JA, Xiao Y, Melemenidis S, et al. Metabolic profiling reveals a dependency of human metastatic breast cancer on mitochondrial serine and one-carbon unit metabolism. Mol Cancer Res. 2020;18(4):599–611.
[30] Mejia NR, MacKenzie RE. NAD-dependent methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate cyclohydrolase in transformed cells is a mitochondrial enzyme. Biochem Biophys Res Commun. 1988;155(1):1–6.
[31] Rozen R, Hum DW, Mackenzie RE. In vitro synthesis of the trifunctional protein, methylenetetrahydrofolate dehydrogenase–methenyltetrahydrofolate cyclohydrolase–formyltetrahydrofolate synthetase, in normal and transformed cells. Can J Biochem Cell Biol. 1985;63(11):1189–93.
[32] Di Pietro E, Wang XL, MacKenzie RE. The expression of mitochondrial methylenetetrahydrofolate dehydrogenase-cyclohydrolase supports a role in rapid cell growth. Biochim Biophys Acta. 2004;1674(1):78–84.
[33] Nilsson R, Jain M, Madhusudhan N, Sheppard NG, Strittmatter L, Kampf C, et al. Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. Nat Commun. 2014;5:3128.
[34] Lin H, Huang B, Wang H, Liu X, Hong Y, Qiu S, et al. MTHFD2 overexpression predicts poor prognosis in renal cell carcinoma and is associated with cell proliferation and vimentin-modulated migration and invasion. Cell Physiol Biochem. 2018;51(2):991–1000.