### Original Article Abnormal 5-methylcytosine IncRNA methylome is involved in human high-grade serous ovarian cancer

Li Meng\*, Qianqian Zhang\*, Xianghua Huang

Department of Gynecology, The Second Hospital of Hebei Medical University, 215 Heping West Road, Shijiazhuang 050011, Hebei, China. \*Equal contributors.

Received June 24, 2021; Accepted November 15, 2021; Epub December 15, 2021; Published December 30, 2021

Abstract: Methylcytosine (m5C) is an important posttranscriptional RNA methylation modification. Studies have reported that aberrant RNA methylation can regulate tumorigenesis and development, indicating the importance of exploring the distribution and biological functions of m5C modification in human high-grade serous ovarian cancer (HGSOC) IncRNAs. In the current study, we identified 2,050 dysregulated m5C peaks, 1,767 of which were significantly upregulated, while 283 were significantly downregulated by performing methylated RNA immunoprecipitation sequencing on 3 pairs of human HGSOC tissues and paired normal tissues. GO enrichment analysis showed that genes altered by the m5C peak played a key role in phylogeny, protein metabolism, and gene mismatch repair. KEGG pathway analysis revealed that these genes were enriched in some important pathways in cancer regulation, such as the PI3K-Akt signalling pathway, transcriptional dysregulation in cancer, and mismatch repair pathways. In addition, through joint analysis of MeRIP-seg and RNA-seg data, we identified 1671 differentially methylated m5C peaks and synchronous differentially expressed genes. These genes play a key role in cell growth or maintenance, RNA metabolism and material transport. We analyzed expression of the m5C modification regulatory gene collagen type IV alpha 3 chain (COL4A3) in 80 HGSOC tissue samples by immunohistochemistry and found that high expression of COL4A3 was significantly correlated with CA125 level (P=0.016), lymph node metastasis (P<0.001), degree of interstitial invasion (P<0.001) and FIGO staging (P<0.001) and indicated a poorer prognosis. Our results revealed the critical role of m5C methylation of IncRNAs in HGSOC, and provided a reference for the prognostic stratification and treatment strategy of HGSOC.

Keywords: IncRNA, 5-methylcytosine, high-grade serous cancer, MeRIP sequencing, prognosis

#### Introduction

RNA modifications, serving as novel epigenetic markers, have been demonstrated to be critically implicated in diverse biological regulatory processes and involved in the occurrence of various diseases, including cancer [1]. Decades of research have revealed more than 150 different types of posttranscriptional RNA modifications, which are widely distributed among rRNAs, tRNAs, mRNAs, and long noncoding RNAs (IncRNAs) [2-4]. N6-methyladenosine (m6A), the most abundant type of modification in eukaryotic RNAs, has been shown to be involved in most types of RNA metabolism, including RNA translation, degradation, splicing, and export [5, 6]. In addition, growing evidence has demonstrated that dysregulated m6A modifications are implicated in cell initiation, progression and even drug responses in multiple cancers. Decreased m6A methylation caused by the METTL14 mutation promotes activation of the AKT pathway, leading to enhanced endometrial cell proliferation, migration and invasion [7]. Elevated m6A levels in epithelial ovarian cancer contribute to poly (ADP-ribose) polymerase inhibitor (PARPi) resistance by upregulating the Wnt/ $\beta$ -catenin pathway [8]. These lines of emerging evidence all suggest that RNA methylation modification may play a central role in gynaecological tumours.

5-Methylcytosine (m5C), another important RNA modification that was first identified in abundant tRNAs and rRNAs, was shown to be involved in multiple processes of tumorigenesis, development, and metastasis, attracting increasing attention in biomedical research [9,

10]. In particular, a comprehensive view of m5C modification has been identified in mRNAs and noncoding RNAs (ncRNAs) based on highthroughput sequencing approaches [11-13]. Studies have shown that the m5C site in mammalian mRNAs is primarily formed by NSUN2 and is preferentially enriched in the region immediately downstream of the translation initiation site [14]. More recently, research has confirmed that m5C methylation disorders are related to the occurrence of hepatocellular carcinoma [15], and it is possible that NSUN2mediated m5C modification of H19 IncRNA promotes tumour occurrence and development by recruiting the G3BP1 oncoprotein [16]. Some evidence that dysregulated m5C methylation in IncRNAs promotes cancer progression has been demonstrated in previous reports [17-19]. However, studies concerning dysregulated m5C modification of IncRNAs and the occurrence of high-grade serous ovarian cancer (HGSOC) are still limited.

As the deadliest ovarian tumour among female malignancies, HGSOC accounts for 70% of ovarian cancer-related deaths [20]. Although great breakthroughs have been made in first-line treatments, such as surgery and paclitaxel/ platinum-based chemotherapy, the overall survival of HGSOC patients has not been improved for several decades [21-23]. In addition, recent emerging evidence has confirmed that reversible RNA modifications are involved in HGSOC tumorigenesis and progression, indicating that RNA methylation may represent a novel therapeutic target for HGSOC [24-26].

In the present study, we investigated the transcriptional role of m5C in HGSOC IncRNAs by MeRIP-seq. We found that the number of m5C methylation peaks in HGSOC was significantly less than that in paired non-tumor tissues, and these differentially methylated genes were widely derived from most chromosomes. Bioinformatics analysis showed that these differentially expressed methylation-targeted genes are involved in important biological processes and cancer pathways. Finally, we used immunohistochemical methods to confirm that the m5C IncRNA-regulated gene collagen type IV alpha 3 chain (COL4A3) was significantly associated with poor prognosis in HGSOC. Our findings provide a comprehensive description of m5C methylation modification in HGSOC, expanding ideas for the development of new therapies at the epitranscriptional level.

### Materials and methods

### Sample collection and RNA isolation

Three pairs of HGSOC tissues and paired normal tissues were obtained from the Second Hospital of Hebei Medical University. This study was approved by the Ethics Committee of the Second Hospital of Hebei Medical University (No. 2021-R255). Total RNA was extracted from tissues using TRIzol reagent (Invitrogen Corporation, CA, USA) as recommended by the manufacturer. The Ribo-Zero rRNA Removal Kit (Illumina, Inc., CA, USA) was used to reduce ribosomal RNA content. Total RNA was fragmented into approximately 100 bp fragments using the GenSeq m5C RNA IP Kit (GenSeq Inc., China). The quality of RNA was measured by a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) with OD260/OD280 ratios between 1.8 and 2.1.

RNA MeRIP-seq library construction & sequencing

Briefly, m5C RNA MeRIP-seq was performed using a GenSeqTM m5C RNA IP Kit (GenSeq Inc., China). Both the input samples without immunoprecipitation and the m5C IP samples were used for RNA-seq library generation with the NEBNext<sup>®</sup> Ultra II Directional RNA Library Prep Kit (New England Biolabs, Inc., USA). The library quality was evaluated with a BioAnalyzer 2100 system (Agilent Technologies, Inc., USA). cDNA library sequencing was performed on an Illumina HiSeq 4000 instrument with 150 bp paired-end reads.

### Sequencing data analysis

Briefly, after sequencing with the Illumina NovaSeq 6000 sequencer, image analysis, base identification and quality control, raw reads (raw data) were generated. Q30 was used to identify the mapping quality of the paired-end reads. Then, Cutadapt software (v1.9.3) [27] was used to remove the connectors and low-quality reads to obtain high-quality clean reads. STAR software [28] was used to match the clean reads of the input library to the reference genome (HG38), and then DCC soft-



Figure 1. Flowchart of the study.

ware [29] was used to identify circRNA. Next, clean reads of all libraries were aligned to the reference genome using Hisat2 software (v2.0.4) [30], followed by identifying methylated peaks on RNAs with Model-based Analysis of ChIP-Seq (MACS) software [31]. Differentially methylated sites were identified by diffReps with a fold change cut-off of  $\geq$ 2 and false discovery rate cut-off of  $\leq$ 0.0001 [31]. Peaks identified by both MACS and diffReps overlapping with exons of IncRNA were chosen by inhouse constructed scripts, which were further annotated accordingly.

### Bioinformatics analysis

DAVID is an online database with gene annotation, visualization and integrated discovery function, which was performed to explore potential functions and pathways of differential m5C IncRNA targeted genes [32, 33]. Gene ontology (GO) analysis, covering biological process (BP), cellular component (CC) and molecular function (MF), was used to provide potential functional annotation of target genes. Kyoto Encyclopedia of Genes and Genomes pathway (KEGG) enrichment analysis was performed to investigate biological pathways and functions. *P*-values <0.05 were considered the cut-off criteria. All processes are described in **Figure 1**.

### Immunofluorescence staining

For immunohistochemical staining, a total of 80 paraffin-embedded HGSOC tissue samples were collected from patients at the Second Hospital of Hebei Medical University (Shijiazhuang, China). Anti-rabbit COL4A3 (1:500, Bioss Antibodies, USA) was used as the primary antibody. IHC staining was performed using a previously described protocol [34, 35]. Image-Pro Plus (version 6.0) was used to quantify expression level of COL4A3.

### Statistical analysis

Spearman's rho test was performed to analyse the relationship between the expression of COL4A3 and various clinicopathological parameters. Univariate and multivariate Cox regression proportional hazards models were used to identify potential prognostic factors. The survival curve was generated by the Kaplan-Meier method. All statistical analyses were performed using SPSS 25.0 software, and significance was set to *P*-values <0.05. LncRNA m5C methylome is involved in HGSOC



**Figure 2.** Characteristics of m5C methylation in HGSOC. A. Venn diagram of m5C sites identified in IncRNA and m5C targeted genes from human HGSOC and adjacent tissues. B. Cluster analysis of methylation in normal and HGSOC tissues. C. Percentage of IncRNAs harbouring different numbers of m5C peaks in the two tissue types. D. The most reliable m5C motif sequence in human HGSOC and adjacent tissues.

### Results

## Overview of the m5C methylation map in HGSOC and adjacent normal tissues

To investigate the m5C methylation of the whole transcriptome, we performed IncRNA MeRIP-seq on tumour tissues of 3 HGSOC patients and normal tissues adjacent to the tumours. In normal tissues, we detected

15,091 m5C peaks and mapped 13,421 annotated genes. In the HGSOC organization, we detected 11,968 m5C peaks and mapped 10,570 annotated genes. Among them, 5,547 m5C peaks were observed in both HGSOC and normal tissues, mapping 5,639 annotated genes (**Figure 2A**). In addition, we performed a methylation heatmap analysis and cluster analysis of the total data. The results of the cluster analysis showed the apparent expression dif-

Am J Transl Res 2021;13(12):13625-13639

	· -	_			
Regulation	Chrom	Peak ID	Gene name	Fold change	P-value
up	chr15	diffreps_peak_675966	GOLGA8N	9.503566334	7.43521E-12
up	chr17	diffreps_peak_887213	LINC01974	5.059602649	8.95821E-12
up	chr5	diffreps_peak_1740107	PPIAP77	8.445714286	9.52351E-12
up	chr1	diffreps_peak_130590	AL365181.2	4.609708738	1.17803E-11
up	chr1	diffreps_peak_111064	AC241952.1	5.644416719	1.20355E-11
up	chr10	diffreps_peak_299550	AL158835.1	6.876344086	1.217E-11
up	chr8	diffreps_peak_2153791	AC009686.2	4.272667971	1.38983E-11
up	chr1	diffreps_peak_57652	TIE1	5.704784131	1.96945E-11
up	chr1	diffreps_peak_160318	LGR6	4.435920578	2.20924E-11
up	chr18	diffreps_peak_993921	AC023090.1	7.26042841	2.45183E-11
down	chr11	diffreps_peak_352729	OR5BN2P	460.1666667	1.61304E-08
down	chr12	diffreps_peak_467723	AC020611.1	62.49484536	2.322E-08
down	chr19	diffreps_peak_1072532	SIGLEC17P	200.7216495	2.43066E-08
down	chr16	diffreps_peak_782627	AC145350.2	15.02836879	3.29821E-08
down	chr16	diffreps_peak_746358	E4F1	4.430357143	3.80473E-08
down	chr8	diffreps_peak_2103137	LINC00208	95.46341463	3.9969E-08
down	chr11	diffreps_peak_332527	CBX3P1	83.58823529	4.80162E-08
down	chr1	diffreps_peak_102048	LAMTOR5	37.9471831	5.96517E-08
down	chr1	diffreps_peak_117459	AC239859.5	9.831683168	6.14172E-08
down	chr4	diffreps_peak_1586700	GAK	2.084809076	6.39882E-08

Table 1. Top ten upregulated and downregulated m5C methylation peaks

ference between the groups and relatively consistent within the group. And these differences may be closely related to the occurrence and development of HGSOC (Figure 2B). Through further statistical analysis of methylation peaks and corresponding IncRNAs (Figure 2C), it was found that greater than 80% of m5C-modified RNAs have only one methylation site, and there was no significant difference in peaks per IncRNA between HGSOC and paired normal tissues. Furthermore, we observed that the number of IncRNAs with two or more methylation peaks on one IncRNA in HGSOC was slightly greater than that of adjacent tissues (P<0.01). To determine the presence of the m5C motif, we used Dreme software to scan the methylation peak sequence of each group of samples (50 bp on both sides of the apex). Results showed that CCAGSCUG (S=C/G) was the most conserved sequence motif and the most likely retained methylation site motif (P-value =1.2e-057). The conserved motif in adjacent normal tissues was CAGSCUGG, with a P-value of 2.9e-075 (S=G/C) (Figure 2D).

Gene expression regulated by m5C modification is enriched in important biological functions and signalling pathways

Using DiffReps software to analyze the difference between the methylation peaks of the two

sets of samples, we identified 1,767 upregulated methylation peaks and 283 downregulated methylation peaks. The top 10 methylation peaks with the most significant changes are listed in **Table 1**.

To explore the potential function associated with these significantly different m5C methylation regulatory genes, we performed gene ontology analysis on the data obtained by sequencing. In terms of the biological processes (BP) category, we found that genes with hypermethylated m5C sites were significantly richer in system development, regulation of plasma membrane bounded cell projection organization, and regulation of cell projection organization (Figure 3A), while genes with hypomethylated m5C sites were significantly richer in system development, plasma membrane bounded cell projection organization, and neuron projection development (Figure 3D). For the cellular components (CC) category, analysis revealed a notable enrichment of synapses, postsynaptic specialization, and postsynaptic density for genes with up-methylated m5C sites (Figure 3B), while synapses, supramolecular polymers, and supramolecular fibres were the most unique significant enrichment for downmethylated m5C sites (Figure 3E). For the

### LncRNA m5C methylome is involved in HGSOC



**Figure 3.** GO and KEGG enrichment analyses of HGSOC and adjacent tissues. (A) Biological processes, (B) cellular components, and (C) molecular functions of genes annotated by upregulated m5C peaks. (D) Biological processes, (E) cellular components, and (F) molecular functions of genes annotated by down-methylated m5C peaks. KEGG pathways of genes annotated by (G) up-methylated and (H) down-methylated m5C peaks. The top ten most significant terms are shown for each analysis.

Regulation	Transcript ID	P-value	logFC	Gene name
up	ENST00000441160	1.83196E-11	12.17832326	LINC02474
up	ENST00000618966	1.09892E-10	9.264405099	AL161431.1
up	ENST00000478686	2.29249E-09	5.8495534	NXNL2
up	ENST00000433576	1.38558E-08	9.589770001	LINC02257
up	ENST00000488210	2.43414E-08	7.961399865	LINC02006
up	ENST00000557736	2.47801E-08	8.789494787	AL136018.1
up	ENST00000485020	5.08094E-08	5.798610675	AC068647.2
up	ENST00000412445	1.24252E-07	9.035352394	LINC02257
up	ENST00000450480	2.11907E-07	4.996777919	AC004888.1
up	ENST00000646231	5.35194E-07	7.741784527	AL591485.1
down	ENST00000611525	3.60677E-09	-10.27999915	AC026336.3
down	ENST00000462739	3.91191E-09	-10.04963893	GAD1
down	ENST00000478562	1.71001E-08	-9.725211221	GAD1
down	ENST00000436895	6.09494E-08	-9.268401261	TUBB8P6
down	ENST00000484239	1.01079E-07	-9.111916481	SLC30A10
down	ENST00000414049	1.17983E-07	-9.066561807	SP3P
down	ENST00000567788	1.55385E-07	-8.985364382	LINC02418
down	ENST00000539163	1.5909E-07	-8.978389034	HNF1A-AS1
down	ENST00000435236	2.81654E-07	-8.80798087	XRCC6P5
down	ENST00000508955	3.48559E-07	-8.743706272	AC084866.1

Table 2. Top ten upregulated and downregulated differentially expressed IncRNAs

molecular functions, genes with upregulated m5C sites in HGSOC were primarily related to tubulin binding, steroid dehydrogenase activity, and omega peptidase activity (**Figure 3C**), while genes with downregulated m5C sites were primarily related to transmembrane receptor protein tyrosine phosphatase activity, transmembrane receptor protein phosphatase activity, and transmembrane-ephrin receptor activity (**Figure 3F**).

Furthermore, KEGG analysis results revealed significant enrichment for the hypermethylated genes of steroid hormone biosynthesis, ABC transporters, the cAMP signalling pathway, protein digestion and absorption, and the PI3K-Akt signalling pathway (**Figure 3G**), while sulphur relay system, ABC transporters, folate biosynthesis, transcriptional dysregulation in cancer, and mismatch repair were the most unique significant pathways for the downregulated genes (**Figure 3H**).

### Joint analysis of MeRIP-seq and RNA-seq data

The transcriptome profiles of tumour tissues from three HGSOC patients and normal tissues adjacent to the tumours were analysed using RNA-Seq. edgeR software was used to detect differentially expressed IncRNAs. Compared to adjacent tissues, tumour tissues exhibited 1,388 significantly upregulated IncRNAs and 1040 significantly downregulated IncRNAs (P<0.05). The top 10 up- and downregulated differentially expressed IncRNAs are listed in 
 Table 2. Next, all differentially methylated m5C
peaks with differential IncRNA levels (1,671) were divided into four groups by joint analysis. We detected 1,458 hypermethylated m5C peaks in IncRNAs that were significantly upregulated (797; hyper-up) or downregulated (661; hyper-down), while 312 hypomethylated m5C peaks in IncRNAs were obviously upregulated (84; hypo-up) or downregulated (129; hypodown) (Figure 4A). In addition, we conducted a



**Figure 4.** Joint analysis of MeRIP-seq and RNA-seq data. (A) Four-quadrant plots showing differentially expressed IncRNAs with significant changes in both the m5C modification and IncRNA levels. The top 10 GO terms of (B) biological processes, (C) cellular components, and (D) molecular functions were significantly enriched for IncRNA-annotated genes with significant changes in both the m5C modification and IncRNA levels. (E) The top ten KEGG path-

ways of IncRNA-annotated genes with significant changes in both the m5C modification and IncRNA levels.

joint analysis of differentially expressed IncRNAs with annotated genes and identified 83 upregulated annotated genes and 10 downregulated annotated genes (**Figure 4B**).

We further performed ontology and pathway analysis on these 93 annotated genes with differentially expressed IncRNAs to investigate their biological significance. GO analysis indicated that these genes were primarily enriched in signal transduction (26.8%), cell communication (24.4%), and cell growth and/or maintenance (12.2%) with respect to BP. Membrane fraction (11.8%), perinuclear region (8.8%), and perinuclear region of cytoplasm (5.9%) were the most unique significant enrichment for the cellular components. Cytoskeletal protein binding (7.3%), auxiliary transport protein activity (7.3%), and RNA binding (4.9%) were the significantly enriched molecular functions. GO analysis of CC found that the genes were primarily related to membrane fraction (11.8%), perinuclear region (8.8%), and perinuclear region of cytoplasm (5.9%). GO analysis of MF found that the genes were primarily enriched in cytoskeletal protein binding (7.3%), auxiliary transport protein activity (7.3%), and RNA binding (4.9%) (Figure 4B-D). Pathway analysis revealed that SLC-mediated transmembrane transport (18.8%), transport of glucose and other sugars, bile salts and organic acids, metals and amine compounds (12.5%), and AKT phosphorylation targets in the cytosol (6.3%) were the most significant enrichment for these genes (**Figure 4E**).

# Clinical association between gene expression regulated by m5C modification and HGSOC prognosis

Finally, we found that there were 84 hypomethylated m5C peaks with upregulated in IncRNAs, of which COL4A3 is an oncogene. We observed that the m5C regulatory gene COL4A3 was enriched in several important pathways, such as the 'PI3K-Akt signalling pathway', 'ECMreceptor interaction', and 'focal adhesion'. Therefore, we obtained 80 clinical samples to investigate the potential relationship between the abnormal m5C regulatory gene COL4A3 and the prognosis of HGSOC patients. Figure 5A shows representative images of COL4A3 staining in HGSOC tissue. Table 3 summarizes the association between the expression of COL4A3 and clinicopathological characteristics of HGSOC patients. There were 32/80 cases with low COL4A3 expression and 48/80 cases with high COL4A3 expression. Expression of COL4A3 was correlated with CA125 level (P=0.016), lymph node metastasis (P<0.001), depth of myometrial invasion (P<0.001) and FIGO stage (P<0.001). Cox regression analysis revealed that the overall survival of HGSOC patients was significantly correlated with lymph node metastasis (P=0.039), depth of myometrial invasion (P=0.008), and FIGO stage (P=0.010). High expression of COL4A3 also conveyed a poor prognosis (P=0.013).

### Discussion

Increasing evidence suggests that m5C RNA modification is involved in diverse fundamental biological functions and in the tumorigenesis and development of multiple cancers [36]. However, previous studies on m5C and ovarian cancer were at the level of DNA methylation, and the focus was more on the production of 5-hydroxymethylcytosine (5-hmC) after m5C was oxidized by TET family enzymes [37-39]. In this study, we first demonstrated that m5C methylation modification occurred in IncRNAs of HGSOC using MeRIP-seq. By comparing the full transcripts of 3 pairs of HGSOC and adjacent pairs of normal tissues, we identified 2,050 dysregulated m5C peaks in IncRNA, of

which 1,767 m5C peaks were noticeably upregulated, and 283 m5C peaks were clearly downregulated. In addition, through GO and KEGG pathways, we identified the potential biological functions of these differentially methylated transcripts and cancer-related pathways. Furthermore, by combining MeRIP-seq and RNAseq data, we discovered differentially methylated m5C peaks and synchronously differentially expressed genes, as well as their potential functions. Finally, through immunohistochemical staining, we confirmed the expression of the m5C modification regulatory gene COL4A3 in human HGSOC tissues and its correlation with HGSOC clinicopathological parameters and prognosis. Taken together, these findings expand our knowledge of the epitranscriptional role of 5mC in regulating the pathogenesis of HGSOC.

Numerous studies have proved that m5C methylation modification plays an important role in the pathogenesis of many diseases, including tumors [40]. In addition, the distribution of m5C sites is critical to the stability of RNA and the regulation of translation [41]. We identified thousands of methylation peaks in IncRNAs, and found significant differences in the number and distribution of methylation peaks between the two samples. The results indicated that the methylation frequency and methylation genes of HGSOC are far less than the frequency of adjacent tissues, proving that there is a clear connection between m5C and HGSOC. A recent study in a BRCA2-deficient mouse embryonic stem cell model showed that the deletion of the m5C demethylase TET2 leads to a decrease in the abundance of m5c oxidation to 5-hydroxymethycytosine (5hmC), and the negative feedback of 5hmc and apurinic/apyrimidinic endonuclease 1 (APE1) is recruited to stalled replication forks (RFS), which leads to the instability of the genome and promotes the progression and drug resistance of cancer cells [37]. In addition, the m5C demethylase TET1 is highly expressed in ovarian cancer and is associated with a poor prognosis [42]. However, due to the limitations of m5C in HGSOC research, we can only make some suggestions: normal ovarian tissue undergoes abnormal m5C demethylation modification at the post-RNA transcription level, which in turn leads to the occurrence of HGSOC. This mechanism needs more sophisticated experiments to further verify.



**Figure 5.** Relationship between COL4A3 and HGSOC clinical prognosis. (A) Low and high expression of COL4A3 was detected by IHC in 80 clinical HGSOC samples (×200). Association of overall survival with (B) lymph node metastasis, (C) depth of myometrial invasion, (D) FIGO stage, and (E) COL4A3 expression.

Mounting evidence indicates that m5C plays crucial roles in mediating cellular proliferation [43] and differentiation [44], cellular migration [45], maintaining RNA stability [46], and miRNA processing and cleavage [47] through the NSUN family of methyltransferases. Using GO functional enrichment analysis, we determined that the differentially expressed m5C lncRNAannotated genes were involved in cell development, cell adhesion, and cell metabolism. In a mouse model, mutations in the writer protein NSUN7 of m5C cause defects in sperm motility, leading to infertility or complete infertility in male mice [48]. In addition, overexpression and copy number increases of NSUN2 have been detected in human cancers [49], and mutations within the NSUN2 gene are associated with autosomal recessive mental retardation [50]. It is worth noting that we found that m5C IncRNA-regulated genes were enriched in some important cancer pathways, such as ABC transporters, the cAMP signalling pathway, the PI3K-Akt signalling pathway, transcriptional dysregulation in cancer, and the mismatch repair path-

Characteristics			COL4A3		
Characteristics	Low (%)	High (%)	P		
Age	<50 years	31	12 (38.7)	19 (61.3)	0.851
	>50 years	49	20 (40.8)	29 (59.2)	
CA125*	<150 U/L	23	14 (60.9)	9 (9.1)	0.016
	>150 U/L	57	18 (31.6)	39 (68.4)	
Tumor size	<3 cm	52	19 (36.5)	33 (36.5)	0.389
	>3 cm	28	13 (46.4)	15 (53.6)	
Lymph node metastasis	no	30	27 (90.0)	3 (10.0)	< 0.001
	yes	50	5 (90.0)	45 (10.0)	
Depth of myometrial invasion*	<1/2	21	18 (85.7)	3 (14.3)	< 0.001
	>1/2	59	14 (23.7)	45 (76.3)	
FIGO stage*	l or ll	24	12 (50.0)	12 (50.0)	<0.001
	111	34	20 (58.8)	14 (41.2)	
	IV	22	0 (0.0)	22 (100.0)	

Table 3. Clinicopathological variables and the expression status of COL4A3 in HGSOC patients

Spearman-rho test was used. \*P<0.05.

way. These results indicated that the cancer pathways involved in abnormal m5C transcripts are extensive. However, we do not know enough about the exact mechanisms at present, so more experiments were needed to determine the role and mechanism of m5C, and ultimately provide new therapeutic targets for HGSOC. Previous studies have shown that the ABC transporter pathway makes an important contribution to the progression, platinum resistance and poor prognosis of ovarian cancer [51, 52]. It has been reported that FTO-dependent m6A modification inhibits the selfrenewal of ovarian cancer stem cells by preventing cAMP signalling [53]. In summary, these results indicate that differentially expressed m5C IncRNA modification exerts precise functions in regulating the occurrence of HGSOC.

Combining analysis of MeRIP sequence and RNA sequence data, we identified 93 genes with different m5C methylation peaks and synchronous expression in HGSOC. Similarly, these genes were enriched in some important functional pathways, such as signal transduction, cell growth, RNA binding, and AKT phosphorylation targets in the cytoplasm. These results indicate that differentially expressed IncRNAs also make important contributions to HGSOC. LncRNAs act as key regulators in various biological processes, such as cell fate determination, transcription splicing, translation efficiency, T-cell homeostasis, and sex determination, and abnormal IncRNA expression may be critical for the occurrence and development of cancer [54]. Marianna Buttarelli et al. showed that the IncRNA MEG3 exerts a tumour suppressor effect by regulating the phosphatase and tensin homologue network, and low expression of MEG3 was correlated with better progressionfree and overall survival of HGSOC [55]. LncRNA NEAT1, regulated by LIN28B, promotes cell proliferation and migration by sponging miR-506 in HGSOC [56]. LncRNA SOCAR upregulates MMP-9 through the Wnt/ $\beta$ -catenin pathway to partially promote the proliferation and invasion of HGSOC cells, and high expression of SOCAR is positively correlated with HGSOC progression [57]. Consistent with these findings, we analyzed expression of the differentially expressed m5C IncRNA regulatory gene COL4A3 in 80 human HGSOC tissues and found that COL4A3 is highly expressed in HGSOC tissues and is correlated with poor prognosis in HGSOC patients. A previous study showed that COLA43 is the angiogenic-related gene of HGSOC patients and indicates a poor prognosis [58]. In addition, A recent study in an ovarian cancer cell model showed that COL4A3 is a candidate gene that promotes the occurrence of ovarian cancer and is enriched in the extracellular matrix-receptor interaction pathway [59]. However, the specific mechanism by which the m5C modified gene COL4A3 leads to the poor prognosis of HGSOC needs further study. Taken together, our research shows that the regulation of m5C IncRNA modification is a promising therapeutic strategy for HGSOC in the future.

### Conclusions

These findings are of great significance for understanding the important molecular mechanism of RNA m5C modification in the pathogenesis of HGSOC at the level of epitranscription. Targeting the regulation of m5C IncRNA represents a new therapeutic approach for HGSOC, but the specific mechanisms involved need to be further clarified.

### Acknowledgements

This work was supported by the Key R&D Projects in Hebei Province (21377706D). In addition, we thank Professor Helin Feng (Department of Orthopaedics, The Fourth Hospital of Hebei Medical University) for his experimental support. We thank Cloud-Seq Biotech Ltd. Co. (Shanghai, China) for the MeRIP-Seq service and the subsequent bioinformatics analysis.

### Disclosure of conflict of interest

None.

Address correspondence to: Xianghua Huang, Department of Gynecology, The Second Hospital of Hebei Medical University, 215 Heping West Road, Shijiazhuang 050011, Hebei, China. Tel: +86-15803210557; E-mail: xianghuahuang0311@163. com

### References

- [1] Shi H, Wei J and He C. Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers. Mol Cell 2019; 74: 640-650.
- [2] Keffer-Wilkes LC, Veerareddygari GR and Kothe U. RNA modification enzyme TruB is a tRNA chaperone. Proc Natl Acad Sci U S A 2016; 113: 14306-14311.
- [3] Agris PF, Narendran A, Sarachan K, Väre V and Eruysal E. The importance of being modified: the role of RNA modifications in translational fidelity. Enzymes 2017; 41: 1-50.
- [4] Roundtree IA, Evans ME, Pan T and He C. Dynamic RNA modifications in gene expression regulation. Cell 2017; 169: 1187-1200.
- [5] Chen J, Fang X, Zhong P, Song Z and Hu X. N6methyladenosine modifications: interactions with novel RNA-binding proteins and roles in

signal transduction. RNA Biol 2019; 16: 991-1000.

- [6] He L, Li H, Wu A, Peng Y, Shu G and Yin G. Functions of N6-methyladenosine and its role in cancer. Mol Cancer 2019; 18: 176.
- [7] Liu J, Eckert MA, Harada BT, Liu SM, Lu Z, Yu K, Tienda SM, Chryplewicz A, Zhu AC, Yang Y, Huang JT, Chen SM, Xu ZG, Leng XH, Yu XC, Cao J, Zhang Z, Liu J, Lengyel E and He C. m(6) A mRNA methylation regulates AKT activity to promote the proliferation and tumorigenicity of endometrial cancer. Nat Cell Biol 2018; 20: 1074-1083.
- [8] Fukumoto T, Zhu H, Nacarelli T, Karakashev S, Fatkhutdinov N, Wu S, Liu P, Kossenkov AV, Showe LC, Jean S, Zhang L and Zhang R. N(6)methylation of adenosine of FZD10 mRNA contributes to PARP inhibitor resistance. Cancer Res 2019; 79: 2812-2820.
- [9] Bohnsack KE, Höbartner C and Bohnsack MT. Eukaryotic 5-methylcytosine (m<sup>5</sup>C) RNA methyltransferases: mechanisms, cellular functions, and links to disease. Genes (Basel) 2019; 10: 102.
- [10] Gilbert WV, Bell TA and Schaening C. Messenger RNA modifications: form, distribution, and function. Science 2016; 352: 1408-1412.
- [11] Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, Suter CM and Preiss T. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 2012; 40: 5023-5033.
- [12] Schumann U, Zhang HN, Sibbritt T, Pan A, Horvath A, Gross S, Clark SJ, Yang L and Preiss T. Multiple links between 5-methylcytosine content of mRNA and translation. BMC Biol 2020; 18: 40.
- [13] David R, Burgess A, Parker B, Li J, Pulsford K, Sibbritt T, Preiss T and Searle IR. Transcriptome-wide mapping of RNA 5-methylcytosine in Arabidopsis mRNAs and noncoding RNAs. Plant Cell 2017; 29: 445-460.
- [14] Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, Li A, Wang X, Bhattarai DP, Xiao W, Sun HY, Zhu Q, Ma HL, Adhikari S, Sun M, Hao YJ, Zhang B, Huang CM, Huang N, Jiang GB, Zhao YL, Wang HL, Sun YP and Yang YG. 5-methylcytosine promotes mRNA export-NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res 2017; 27: 606-625.
- [15] He Y, Shi Q, Zhang Y, Yuan X and Yu Z. Transcriptome-wide 5-methylcytosine functional profiling of long non-coding RNA in hepatocellular carcinoma. Cancer Manag Res 2020; 12: 6877-6885.
- [16] Sun Z, Xue S, Zhang M, Xu H, Hu X, Chen S, Liu Y, Guo M and Cui H. Aberrant NSUN2-mediated m(5)C modification of H19 IncRNA is associat-

ed with poor differentiation of hepatocellular carcinoma. Oncogene 2020; 39: 6906-6919.

- [17] Deng W, Wang J, Zhang J, Cai J, Bai Z and Zhang Z. TET2 regulates LncRNA-ANRIL expression and inhibits the growth of human gastric cancer cells. IUBMB Life 2016; 68: 355-364.
- [18] Silva JP and van Booven D. Analysis of diet-induced differential methylation, expression, and interactions of IncRNA and protein-coding genes in mouse liver. Sci Rep 2018; 8: 11537.
- [19] Hu H, Shu M, He L, Yu X, Liu X, Lu Y, Chen Y, Miao X and Chen X. Epigenomic landscape of 5-hydroxymethylcytosine reveals its transcriptional regulation of IncRNAs in colorectal cancer. Br J Cancer 2017; 116: 658-668.
- [20] Lisio MA, Fu L, Goyeneche A, Gao ZH and Telleria C. High-grade serous ovarian cancer: basic sciences, clinical and therapeutic standpoints. Int J Mol Sci 2019; 20: 952.
- [21] Bowtell DD, Böhm S, Ahmed AA, Aspuria PJ, Bast RC Jr, Beral V, Berek JS, Birrer MJ, Blagden S, Bookman MA, Brenton JD, Chiappinelli KB, Martins FC, Coukos G, Drapkin R, Edmondson R, Fotopoulou C, Gabra H, Galon J, Gourley C, Heong V, Huntsman DG, Iwanicki M, Karlan BY, Kaye A, Lengyel E, Levine DA, Lu KH, Mc-Neish IA, Menon U, Narod SA, Nelson BH, Nephew KP, Pharoah P, Powell DJ Jr, Ramos P, Romero IL, Scott CL, Sood AK, Stronach EA and Balkwill FR. Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. Nat Rev Cancer 2015; 15: 668-679.
- [22] Freimund AE, Beach JA, Christie EL and Bowtell D. Mechanisms of drug resistance in highgrade serous ovarian cancer. Hematol Oncol Clin North Am 2018; 32: 983-996.
- [23] Hoppenot C, Eckert MA, Tienda SM and Lengyel E. Who are the long-term survivors of high grade serous ovarian cancer. Gynecol Oncol 2018; 148: 204-212.
- [24] Papakonstantinou E, Androutsopoulos G, Logotheti S, Adonakis G, Maroulis I and Tzelepi V. DNA methylation in epithelial ovarian cancer: current data and future perspectives. Curr Mol Pharmacol 2020; [Epub ahead of print].
- [25] Wada M, Kukita A, Sone K, Hamamoto R, Kaneko S, Komatsu M, Takahashi Y, Inoue F, Kojima M, Honjoh H, Taguchi A, Kashiyama T, Miyamoto Y, Tanikawa M, Tsuruga T, Mori-Uchino M, Wada-Hiraike O, Osuga Y and Fujii T. Epigenetic modifier SETD8 as a therapeutic target for high-grade serous ovarian cancer. Biomolecules 2020; 10: 1686.
- [26] Kukita A, Sone K, Oda K, Hamamoto R, Kaneko S, Komatsu M, Wada M, Honjoh H, Kawata Y, Kojima M, Oki S, Sato M, Asada K, Taguchi A, Miyasaka A, Tanikawa M, Nagasaka K, Matsu-

moto Y, Wada-Hiraike O, Osuga Y and Fujii T. Histone methyltransferase SMYD2 selective inhibitor LLY-507 in combination with poly ADP ribose polymerase inhibitor has therapeutic potential against high-grade serous ovarian carcinomas. Biochem Biophys Res Commun 2019; 513: 340-346.

- [27] Kechin A, Boyarskikh U, Kel A and Filipenko M. cutPrimers: a new tool for accurate cutting of primers from reads of targeted next generation sequencing. J Comput Biol 2017; 24: 1138-1143.
- [28] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M and Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013; 29: 15-21.
- [29] Cheng J, Metge F and Dieterich C. Specific identification and quantification of circular RNAs from sequencing data. Bioinformatics 2016; 32: 1094-1096.
- [30] Kim D, Langmead B and Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015; 12: 357-360.
- [31] Shen L, Shao NY, Liu X, Maze I, Feng J and Nestler EJ. diffReps: detecting differential chromatin modification sites from ChIP-seq data with biological replicates. PLoS One 2013; 8: e65598.
- [32] Huang da W, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009; 4: 44-57.
- [33] Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC and Lempicki RA. DAVID bioinformatics resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res 2007; 35: W169-175.
- [34] Liu X, Zhang Q, Zhao Y, Xun J, Wu H and Feng H. Association of JMJD2B and hypoxia-inducible factor 1 expressions with poor prognosis in osteosarcoma. Anal Cell Pathol (Amst) 2020; 2020: 2563208.
- [35] Diao Y, Zhang P, Dai R, Xu J and Feng H. H3K-27me3 and VEGF is associated with poor prognosis in patients with synovial sarcoma. Pathol Res Pract 2018; 214: 974-977.
- [36] Chen YT, Shen JY, Chen DP, Wu CF, Guo R, Zhang PP, Lv JW, Li WF, Wang ZX and Chen YP. Identification of cross-talk between m(6)A and 5mC regulators associated with onco-immunogenic features and prognosis across 33 cancer types. J Hematol Oncol 2020; 13: 22.
- [37] Kharat SS, Ding X, Swaminathan D, Suresh A, Singh M, Sengodan SK, Burkett S, Marks H, Pamala C, He Y, Fox SD, Buehler EC, Muegge K, Martin SE and Sharan SK. Degradation of 5hmC-marked stalled replication forks by

APE1 causes genomic instability. Sci Signal 2020; 13: eaba8091.

- [38] Tucker DW, Getchell CR, McCarthy ET, Ohman AW, Sasamoto N, Xu S, Ko JY, Gupta M, Shafrir A, Medina JE, Lee JJ, MacDonald LA, Malik A, Hasselblatt KT, Li W, Zhang H, Kaplan SJ, Murphy GF, Hirsch MS, Liu JF, Matulonis UA, Terry KL, Lian CG and Dinulescu DM. Epigenetic reprogramming strategies to reverse global loss of 5-hydroxymethylcytosine, a prognostic factor for poor survival in high-grade serous ovarian cancer. Clin Cancer Res 2018; 24: 1389-1401.
- [39] Liu Y, Hu Z, Cheng J, Siejka-Zielińska P, Chen J, Inoue M, Ahmed AA and Song CX. Subtractionfree and bisulfite-free specific sequencing of 5-methylcytosine and its oxidized derivatives at base resolution. Nat Commun 2021; 12: 618.
- [40] Nombela P, Miguel-López B and Blanco S. The role of m(6)A, m(5)C and  $\Psi$  RNA modifications in cancer: novel therapeutic opportunities. Mol Cancer 2021; 20: 18.
- [41] Chen X, Li A, Sun BF, Yang Y, Han YN, Yuan X, Chen RX, Wei WS, Liu Y, Gao CC, Chen YS, Zhang M, Ma XD, Liu ZW, Luo JH, Lyu C, Wang HL, Ma J, Zhao YL, Zhou FJ, Huang Y, Xie D and Yang YG. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol 2019; 21: 978-990.
- [42] Han X, Zhou Y, You Y, Lu J, Wang L, Hou H, Li J, Chen W, Zhao L and Li X. TET1 promotes cisplatin-resistance via demethylating the vimentin promoter in ovarian cancer. Cell Biol Int 2017; 41: 405-414.
- [43] Tang H, Fan X, Xing J, Liu Z, Jiang B, Dou Y, Gorospe M and Wang W. NSun2 delays replicative senescence by repressing p27 (KIP1) translation and elevating CDK1 translation. Aging (Albany NY) 2015; 7: 1143-1158.
- [44] Sajini AA, Choudhury NR, Wagner RE, Bornelöv S, Selmi T, Spanos C, Dietmann S, Rappsilber J, Michlewski G and Frye M. Loss of 5-methylcytosine alters the biogenesis of vault-derived small RNAs to coordinate epidermal differentiation. Nat Commun 2019; 10: 2550.
- [45] Atala A. Re: 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. J Urol 2020; 203: 884-885.
- [46] Zhang X, Liu Z, Yi J, Tang H, Xing J, Yu M, Tong T, Shang Y, Gorospe M and Wang W. The tRNA methyltransferase NSun2 stabilizes p16INK<sup>4</sup> mRNA by methylating the 3'-untranslated region of p16. Nat Commun 2012; 3: 712.
- [47] Yuan S, Tang H, Xing J, Fan X, Cai X, Li Q, Han P, Luo Y, Zhang Z, Jiang B, Dou Y, Gorospe M and Wang W. Methylation by NSun2 represses the levels and function of microRNA 125b. Mol Cell Biol 2014; 34: 3630-3641.

- [48] Harris T, Marquez B, Suarez S and Schimenti J. Sperm motility defects and infertility in male mice with a mutation in Nsun7, a member of the Sun domain-containing family of putative RNA methyltransferases. Biol Reprod 2007; 77: 376-382.
- [49] Okamoto M, Hirata S, Sato S, Koga S, Fujii M, Qi G, Ogawa I, Takata T, Shimamoto F and Tatsuka M. Frequent increased gene copy number and high protein expression of tRNA (cytosine-5-)-methyltransferase (NSUN2) in human cancers. DNA Cell Biol 2012; 31: 660-671.
- [50] Khan MA, Rafiq MA, Noor A, Hussain S, Flores JV, Rupp V, Vincent AK, Malli R, Ali G, Khan FS, Ishak GE, Doherty D, Weksberg R, Ayub M, Windpassinger C, Ibrahim S, Frye M, Ansar M and Vincent JB. Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. Am J Hum Genet 2012; 90: 856-863.
- [51] Christie EL, Pattnaik S, Beach J, Copeland A, Rashoo N, Fereday S, Hendley J, Alsop K, Brady SL, Lamb G, Pandey A, deFazio A, Thorne H, Bild A and Bowtell D. Multiple ABCB1 transcriptional fusions in drug resistant high-grade serous ovarian and breast cancer. Nat Commun 2019; 10: 1295.
- [52] Hedditch EL, Gao B, Russell AJ, Lu Y, Emmanuel C, Beesley J, Johnatty SE, Chen X, Harnett P, George J, Williams RT, Flemming C, Lambrechts D, Despierre E, Lambrechts S, Vergote I, Karlan B, Lester J, Orsulic S, Walsh C, Fasching P, Beckmann MW, Ekici AB, Hein A, Matsuo K, Hosono S, Nakanishi T, Yatabe Y, Pejovic T, Bean Y, Heitz F, Harter P, du Bois A, Schwaab I, Hogdall E, Kjaer SK, Jensen A, Hogdall C, Lundvall L, Engelholm SA, Brown B, Flanagan J, Metcalf MD, Siddiqui N, Sellers T, Fridley B, Cunningham J, Schildkraut J, Iversen E, Weber RP, Berchuck A, Goode E, Bowtell DD, Chenevix-Trench G, deFazio A, Norris MD, MacGregor S, Haber M and Henderson MJ. ABCA transporter gene expression and poor outcome in epithelial ovarian cancer. J Natl Cancer Inst 2014; 106: dju149.
- [53] Huang H, Wang Y, Kandpal M, Zhao G, Cardenas H, Ji Y, Chaparala A, Tanner EJ, Chen J, Davuluri RV and Matei D. FTO-dependent N (6)-methyladenosine modifications inhibit ovarian cancer stem cell self-renewal by blocking cAMP signaling. Cancer Res 2020; 80: 3200-3214.
- [54] Aravindhan S, Younus LA, Hadi Lafta M, Alexander M, Enina YI, Yushchenko NA, Thangavelu L, Mostafavi SM, Pokrovskii MV and Ahmadi M. P53-long noncoding RNA regulatory network in cancer development. Cell Biol Int 2021; 45: 1583-1598.

- [55] Buttarelli M, De Donato M, Raspaglio G, Babini G, Ciucci A, Martinelli E, Baccaro P, Pasciuto T, Fagotti A, Scambia G and Gallo D. Clinical Value of IncRNA MEG3 in high-grade serous ovarian cancer. Cancers (Basel) 2020; 12: 966.
- [56] Yong W, Yu D, Jun Z, Yachen D, Weiwei W, Midie X, Xingzhu J and Xiaohua W. Long noncoding RNA NEAT1, regulated by LIN28B, promotes cell proliferation and migration through sponging miR-506 in high-grade serous ovarian cancer. Cell Death Dis 2018; 9: 861.
- [57] Guo Q, Wang L, Zhu L, Lu X, Song Y, Sun J, Wu Z, Shi J, Wang Z and Zhou X. The clinical significance and biological function of IncRNA SO-CAR in serous ovarian carcinoma. Gene 2019; 713: 143969.
- [58] Siamakpour-Reihani S, Owzar K, Jiang C, Turner T, Deng Y, Bean SM, Horton JK, Berchuck A, Marks JR, Dewhirst MW and Alvarez Secord A. Prognostic significance of differential expression of angiogenic genes in women with high-grade serous ovarian carcinoma. Gynecol Oncol 2015; 139: 23-29.
- [59] Peng X, Yu M and Chen J. Transcriptome sequencing identifies genes associated with invasion of ovarian cancer. J Int Med Res 2020; 48: 300060520950912.