



OPEN Genetic variants reshape the m⁶A epitranscriptome and drive transcriptomic reprogramming in colorectal cancer

Seung Hun Han¹, Seongmin Jang¹, Yeongwon Kim¹, Kun Tan², Miles F. Wilkinson^{2,3}✉, Hyobin Jeong⁴✉ & Junho Choe^{1,5,6,7}✉

Genome-wide association studies (GWAS) have identified numerous single-nucleotide polymorphisms (SNPs) associated with various diseases, including cancer. However, the mechanisms by which these SNPs contribute to disease susceptibility remain largely unclear. While recent studies have explored the transcriptional impact of disease-associated SNPs, their role in post-transcriptional regulation has been less extensively investigated. In this study, we investigated whether cancer-associated SNPs influence gene expression by altering N⁶-methyladenosine (m⁶A) RNA methylation. We collected GWAS-identified SNPs across nine cancer types and integrated these with matched tumor and normal m⁶A RNA immunoprecipitation sequencing (m⁶A-seq) and RNA sequencing (RNA-seq) datasets. We first identified differentially methylated m⁶A sites and assessed whether cancer-associated SNPs were enriched within these regions. These analyses revealed that cancer-associated SNPs were significantly enriched within hypermethylated m⁶A regions in colon cancer. Integrative analysis revealed that SNPs enriched in m⁶A-modified regions are associated with altered gene expression and RNA splicing, suggesting that m⁶A methylation mediates the post-transcriptional impact of genetic variants. Experimental validation further confirmed altered gene expression following ALKBH5 knockdown, consistent with patient-derived data. Collectively, our findings support a novel mechanistic connection between genetic variants and RNA methylation-driven transcriptomic regulation in colorectal cancer, underscoring the epitranscriptome as a potential axis of oncogenic control.

Keywords N⁶-methyladenosine, Genome-wide association study, Colorectal cancer, Alternative splicing, Epitranscriptomic regulation, Single-nucleotide polymorphisms

Abbreviations

A3SS	Alternative 3' splice site
A5SS	Alternative 5' splice site
CDS	Coding sequence
DEG/DEGs	Differentially expressed gene(s)
DMP/DMPs	Differential m ⁶ A peak(s)
DRACH	D = A/G/U, R = A/G, H = A/C/U
eQTL	Expression quantitative trait loci
FPKM	Fragments per kilobase of transcript per million mapped reads
GWAS	Genome-wide association study
LD	Linkage disequilibrium
lncRNA	Long non-coding RNA

¹Department of Life Science, College of Natural Sciences, Hanyang University, Seoul 04763, Republic of Korea.

²Department of Obstetrics, Gynecology, and Reproductive Sciences, School of Medicine, University of California San Diego, La Jolla, San Diego, CA 92093, USA. ³Institute of Genomic Medicine, University of California San Diego, La Jolla, San Diego, CA 92093, USA. ⁴Department of Systems Biology, College of Life Science and Biotechnology, Yonsei University, Seoul, Republic of Korea. ⁵Hanyang Institute of Bioscience and Biotechnology, Hanyang University, Seoul 04763, Republic of Korea. ⁶Research Institute for Natural Sciences, Hanyang University, Seoul 04763, Republic of Korea. ⁷Research Institute for Convergence of Basic Sciences, Hanyang University, Seoul 04763, Republic of Korea. ✉email: mfwilkinson@health.ucsd.edu; hyobinjeong@yonsei.ac.kr; jcho2711@hanyang.ac.kr

m ⁶ A	N6-methyladenosine
m ⁶ A-seq	m ⁶ A RNA immunoprecipitation sequencing
MXE	Mutually exclusive exon
RI	Retained intron
RNA-seq	RNA sequencing
SE	Skipped exon
SNP	Single nucleotide polymorphism
sQTL	Splicing quantitative trait loci
SRA	Sequence read archive
TCGA	The cancer genome atlas
TTS	Transcription termination site
UTR	Untranslated region
VSEA	Variant Set Enrichment Analysis

N6-methyladenosine (m⁶A), the most abundant internal modification in eukaryotic messenger RNAs (mRNA)^{1,2}, plays a pivotal role in shaping the fate of transcripts by influencing RNA stability^{3,4}, splicing^{5,6}, localization^{5,7}, and translation^{4,8,9}. The deposition and removal of m⁶A are orchestrated by methyltransferase (“writer”) and demethylase (“eraser”) enzymes, whereas RNA-binding proteins (“readers”) specifically recognize the modification and mediate downstream regulatory effects on RNA metabolism, establishing a dynamic and reversible control layer^{10,11}. Accumulating evidence suggests that dysregulation of m⁶A contributes to a wide range of human diseases, including cancer^{12–14}, where it contributes to tumor initiation, progression, and immune evasion by reshaping transcriptomic outputs in both oncogenes and tumor suppressors^{11,14–16}.

Despite increasing interest in m⁶A-dependent gene regulation, the extent to which genetic variants contribute to shaping the m⁶A landscape remains poorly defined, although epigenetic and environmental factors have been shown to influence its dynamics^{17,18}. Given that m⁶A deposition depends on specific sequence and structural contexts such as the DRACH motif (D = A/G/U, R = A/G, H = A/C/U)¹⁹, it is plausible that single-nucleotide polymorphisms (SNPs) may disrupt or create m⁶A consensus sites²⁰, thus altering the modification status of target transcripts in a genotype-dependent manner²¹. Such changes could in turn affect RNA stability or processing, ultimately influencing post-transcriptional gene regulation^{22,23}.

Genome-wide association studies (GWAS) have uncovered numerous SNPs associated with cancer susceptibility^{24,25}. However, the majority of these SNPs are located in noncoding regions of the genome²⁶, making it difficult to assign their biological functions. Traditional functional interpretations have focused on the potential for noncoding SNPs to affect transcription factor binding, enhancer activity, promoter accessibility, or chromatin organization^{27,28}. Although these models are well established, the potential of noncoding SNPs to act through post-transcriptional mechanisms remains underexplored. Recent work has suggested that RNA modifications, particularly m⁶A, may constitute an emerging layer of gene regulation through which genetic variants influence the transcriptome^{29–31}. Although some studies have begun to explore these possibilities, the functional consequences of SNP-driven m⁶A modulation remain largely uncharacterized. Nevertheless, comprehensive investigations into the mechanistic interplay between cancer-associated SNPs and the m⁶A methylome are still lacking and are needed to define their functional impact on the transcriptome.

In this study, we hypothesize that cancer-associated SNPs may contribute to transcriptomic changes by altering m⁶A RNA modification sites. To test this hypothesis, we performed an integrative analysis of GWAS data from nine cancer types, alongside matched m⁶A RNA immunoprecipitation sequencing (m⁶A-seq) and RNA sequencing (RNA-seq) datasets derived from tumor and normal tissues. We examined whether cancer-associated SNPs are enriched in regions with differential m⁶A methylation, and whether these variants are associated with changes in transcript levels or alternative splicing patterns. Our results reveal that m⁶A may function as a post-transcriptional mediator of genetic variant effects in cancer.

Results

Transcriptome-wide m⁶A profiling reveals distinct methylation landscapes between tumor and normal tissues

To examine the relationship between cancer-associated SNPs and alterations in m⁶A methylation patterns, we first analyzed m⁶A-seq data from nine cancer types and matched normal tissues. The m⁶A-seq datasets were obtained from the Sequence Read Archive (SRA) database (Supplementary Fig. S1A)³². Raw sequencing reads were preprocessed using Trim-Galore to remove adapters and low-quality bases, and aligned to the human reference genome (GENCODE hg38) using STAR³³. The aligned data were subsequently used for gene expression quantification, alternative splicing detection, and m⁶A peak identification as summarized in the overall analysis pipeline (Fig. 1A). Gene expression levels in tumor and normal samples were quantified using HTSeq³⁴, followed by differential expression analysis with DESeq2 to identify differentially expressed genes (DEGs) in tumors³⁵. Alternative splicing differences were detected using rMATS-turbo to identify differentially spliced genes between tumor and normal tissues³⁶.

We next identified m⁶A-enriched peaks from the m⁶A-seq data using exomePeak and annotated their locations within transcript features³⁷. Motif enrichment analysis confirmed that these peaks were significantly enriched for the canonical DRACH sequence, a known motif for m⁶A deposition (Supplementary Fig. S1B)^{38,39}. We also annotated the genomic distribution of all detected m⁶A peaks, including intronic, exonic, intergenic, and regulatory regions, to provide a more complete reference for downstream analyses (Supplementary Fig. S1C). Differential m⁶A peaks (DMPs) were identified by comparing m⁶A enrichment levels between tumor and normal samples. These DMPs served as the basis for subsequent analyses, including the evaluation of their overlap with cancer-associated SNPs and transcriptomic features in multiple cancer types.

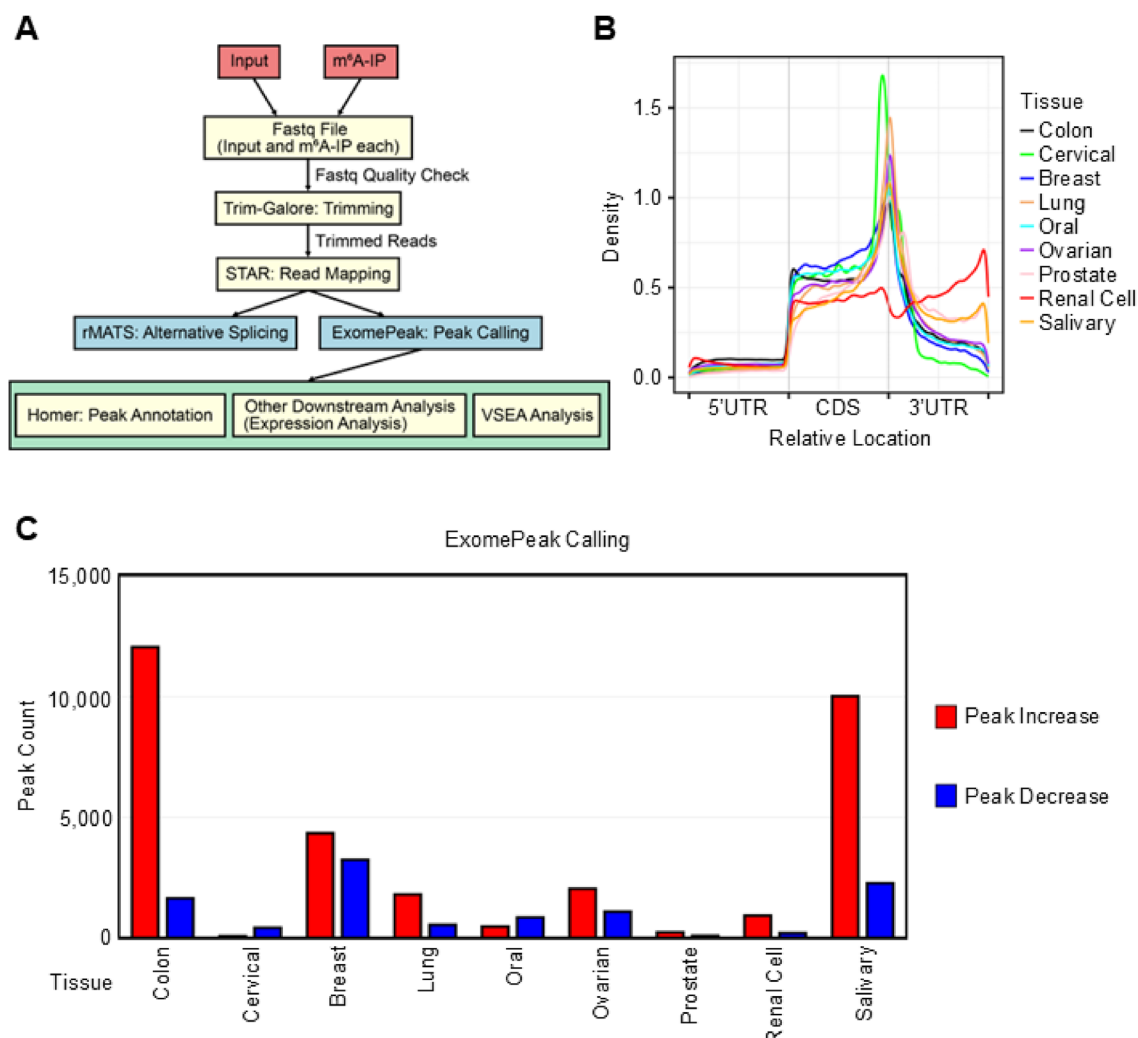


Fig. 1. Overview of m⁶A-seq analysis pipeline and summary of peak distribution across various cancer types. **(A)** Schematic workflow of the analysis pipeline used in this study. **(B)** Metagene plot showing the distribution of m⁶A peaks along transcripts from various tissue types. **(C)** Bar plot of differential m⁶A peaks identified by exomePeak, showing the number of peaks increased (red) or decreased (blue) in tumor tissues across different cancer types (adjusted $P < 0.05$, after Benjamini-Hochberg multiple correction).

To further characterize the biological relevance of m⁶A dynamics, we examined both the positional distribution and cancer-specific variation of m⁶A methylation. Using metaplotR, we visualized the density of m⁶A peaks across transcript regions and found that peaks were predominantly enriched in the 3' untranslated regions (3'UTRs) and coding sequences (CDSs) across the nine cancer types analyzed (Fig. 1B)⁴⁰. This distribution pattern is consistent with prior findings on the positional preference of m⁶A across transcripts⁴¹. In addition to these spatial patterns, previous studies have also reported that m⁶A levels vary across cancer types, with some cancers exhibiting increased m⁶A methylation while others display decreased levels compared to normal tissues^{11,42,43}. In line with these findings, m⁶A modification changes exhibited substantial heterogeneity across cancer types, with both up-regulated and down-regulated DMPs identified (Fig. 1C). In particular, cancers exhibited increased m⁶A methylation in colon, breast, lung, ovarian, prostate, renal cell, and salivary cancers, with colon and salivary cancers showing a greater number of up-regulated DMPs compared to down-regulated DMPs. These results suggest that cancer-type-specific reprogramming of m⁶A methylation may contribute to transcriptomic remodeling in the tumor microenvironment, a possibility we explore further in the following analyses.

Colon cancer-specific enrichment of genetic variants in differential m⁶A regions

To evaluate whether m⁶A changes were associated with genetic variation, Variant Set Enrichment Analysis (VSEA) was used to test for enrichment of cancer-associated SNPs within DMP regions (Figs. 1A and 2A)⁴⁴. In this analysis, we integrated GWAS data with m⁶A-seq profiles across nine cancer types. We retrieved lead SNPs and their linkage disequilibrium (LD) SNPs from the GWAS Catalog to construct a comprehensive set of cancer-associated variants⁴⁵. These SNPs were intersected with DMPs previously identified between tumor

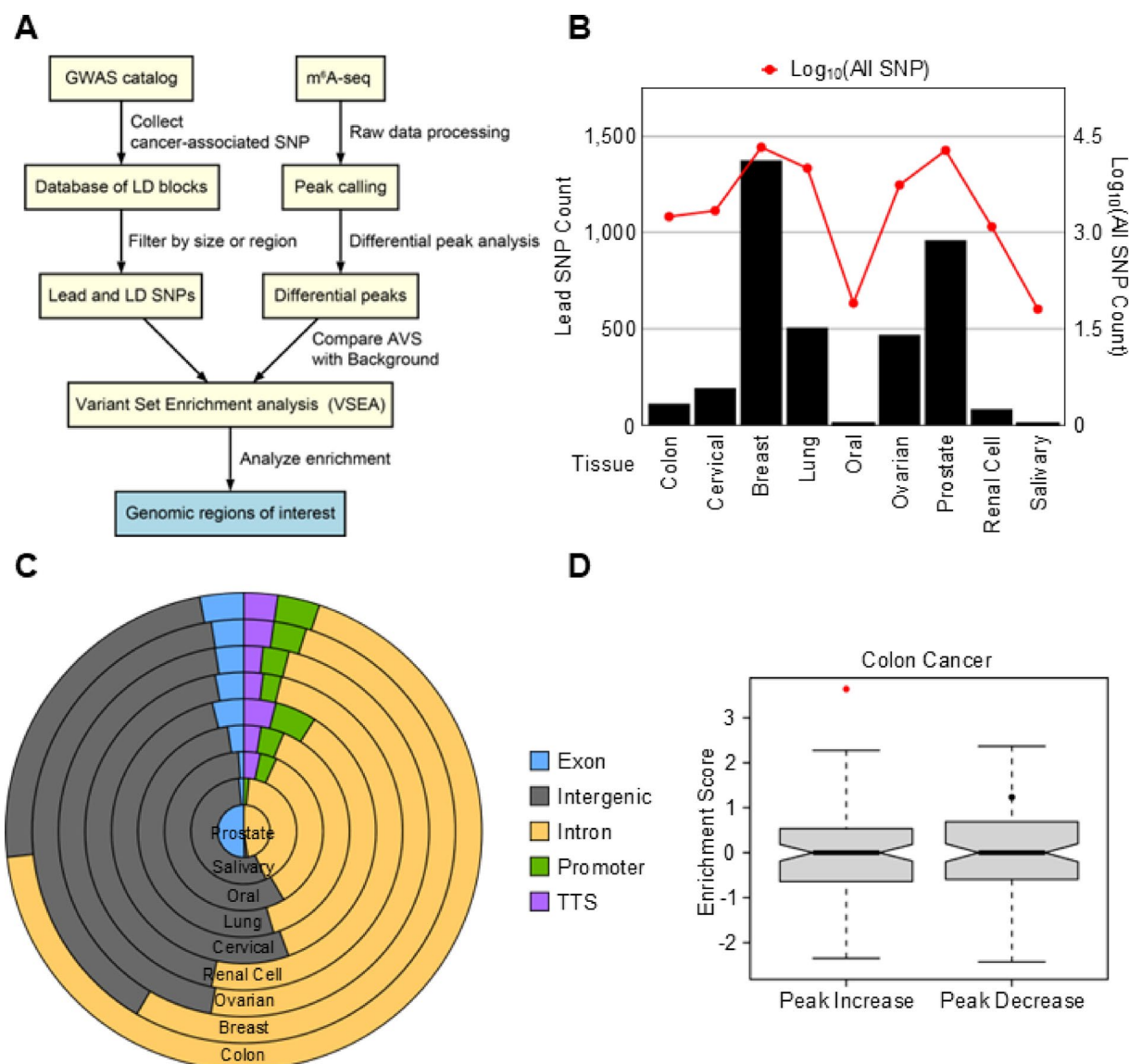


Fig. 2. Integration of differential m⁶A peaks in multiple tumor tissues with cancer-associated SNPs via Variant Set Enrichment Analysis (VSEA). **(A)** Workflow of VSEA combining cancer associated SNPs from GWAS and m⁶A-seq data from different cancer types. **(B)** Bar plot shows the counts of lead SNP and the line plot shows the number of total SNPs (lead SNPs and LD SNPs) across cancer types. **(C)** Stacked pie graph shows the genomic distribution of cancer-associated lead SNPs in different cancer types. **(D)** Enrichment scores of colon cancer-associated SNPs within the differential m⁶A peaks based on VSEA. Enrichment scores were calculated separately for increased and decreased peaks in tumors. A red dot indicates statistical significance ($P < 0.05$, permutation test, $n = 1,000$).

and normal samples (Fig. 1C). To assess whether this overlap was greater than expected by chance, we applied VSEA, comparing observed intersections against a randomized background (Fig. 2A). This analysis was designed to determine whether cancer-associated SNPs are non-randomly distributed within regions of altered m⁶A methylation, providing a potential mechanistic link between genetic variation and m⁶A epitranscriptomic regulation.

We first quantified the number of lead SNPs and total SNPs (lead + LD) associated with each cancer type to estimate the differences in genetic variant burden in different cancer types. As shown in Fig. 2B, breast and prostate cancers exhibited the highest number of lead SNPs, while colon, cervical, oral, renal cell, and salivary cancers displayed substantially fewer. This disparity implies that the contribution of inherited genetic variants to m⁶A regulatory dynamics may differ across tissue types.

To understand the genomic distribution of these SNPs, we annotated them by region, including exonic, intronic, intergenic, promoter, and transcription termination site (TTS) regions (Fig. 2C). As expected, most SNPs were located in intronic and intergenic regions, followed by exonic and regulatory elements, consistent with prior reports that disease-associated variants often reside in noncoding regions^{46,47}.

We next examined whether DMPs, classified as increased peaks or decreased peaks in tumors compared to normal, were significantly enriched for cancer-associated SNPs in each corresponding cancer type. Strikingly, among all nine cancer types analyzed, only colon cancer showed a significant enrichment of cancer-associated SNPs within increased peaks (Fig. 2D; Supplementary Table S1; adjusted $P=0.0023$ based on VSEA after Benjamini-Hochberg multiple correction). This enrichment was assessed using permutation-based sampling, comparing the observed overlap between SNPs and DMPs to a null distribution generated from randomly selected genomic regions. This finding indicates that genetic variation may play a unique role in shaping m⁶A remodeling in colon cancer, highlighting a genetic-epitranscriptomic axis that is specifically observed across certain tumor types.

Colon cancer-associated SNPs overlapping with m⁶A methylation may contribute to transcriptomic changes

Given that colon cancer exhibited significant enrichment of cancer-associated SNPs in their differentially m⁶A-modified regions, we next investigated whether these SNPs are directly positioned within altered m⁶A peaks. We performed an overlap analysis between colon cancer-associated SNPs and hypermethylated m⁶A peaks, followed by annotation of their genomic positions (Fig. 3A). Genomic annotation revealed that a majority of SNPs overlapping with hypermethylated m⁶A peaks in colon cancer were located in intronic regions (60%). This annotation summarizes the positional distribution of m⁶A-overlapping colon cancer-associated SNPs. A permutation test confirmed that the observed proportion of intronic SNPs was within the expected range given the genomic distribution of colon cancer-associated SNPs (adjusted $P=0.99986$), suggesting no statistical enrichment. By contrast, SNPs located in exonic and promoter regions were statistically enriched compared to the null distribution (adjusted $P=0.000025$ and 0.001625 , respectively), indicating potential regulatory relevance. To facilitate further investigation, we compiled a comprehensive list of colon cancer-associated SNPs overlapping with hypermethylated m⁶A peaks, including rsIDs, genomic locations, eQTL/sQTL annotations, and m⁶A peak information (Supplementary Table S2). These overlapping SNPs were unevenly distributed across chromosomes, with higher numbers observed on chromosomes 11, 15, and 21 (Fig. 3B). To determine whether this chromosomal distribution reflected meaningful enrichment rather than random variation, we conducted a permutation test. In each of 100,000 iterations, we randomly sampled the same number of SNPs from the full set of colon cancer-associated SNPs and recorded their chromosomal locations. This generated a background distribution for SNP counts per chromosome. Compared to this background, chromosome 21 was significantly enriched (adjusted $P<0.05$), while chromosomes 11 and 15 did not show enrichment beyond expectation. Previous cancer genomics studies have reported the recurrent chromosome arm level genetic rearrangement in chromosome 21q in microsatellite instability-positive (MSI) primary colon cancers, which aligns with our enrichment findings⁴⁸. These findings suggest that chromosome 21 may harbor a higher density of functionally relevant cancer-associated SNPs within hypermethylated m⁶A regions.

To gain functional insight into the 52 colon cancer-associated SNPs located within DMPs and narrow down to the functionally associated SNPs, we annotated them as expression quantitative loci (eQTLs) or splicing quantitative trait loci (sQTLs) using the snpXplorer platform⁴⁹. This analysis identified 25 SNPs located within hypermethylated m⁶A peaks in tumors, which were predicted to influence transcript expression and/or splicing (Supplementary Fig. S3A and Table S2). We subsequently extracted the target genes of these SNPs and categorized them according to QTL type to identify candidates potentially regulated at the level of expression, splicing, or both (Fig. 3C).

To assess whether these QTL-associated genes display differential transcript levels in tumors, we analyzed input RNA-seq data matched to the m⁶A-seq samples from colon cancer and corresponding normal tissues, as illustrated in the workflow in Fig. 1A. First, we identified DEGs in colon cancer compared to normal tissues and generated a heatmap representation of global expression changes (Supplementary Fig. S3B). By integrating these expression data with our eQTL results, we identified *RIBC2* and *MCM3AP-AS1* as genes that not only carried m⁶A-associated eQTL signals but were also significantly upregulated in colon tumors compared to normal tissues (Fig. 3D). These observations were further supported by TCGA pan-cancer data showing elevated expression of *RIBC2* mRNA and *MCM3AP-AS1* lncRNA in colon adenocarcinoma compared to normal tissues (Supplementary Fig. S3C and S3D). These findings support a possible mechanistic link between m⁶A-associated SNPs and altered gene expression in colon cancer. According to the TCGA pan-cancer data, both genes are also upregulated in multiple other cancer types compared to normal. This broader expression pattern is consistent with previous studies reporting that *RIBC2* and *MCM3AP-AS1* are differentially expressed and clinically relevant in various human cancers beyond colon cancer, likely through mechanisms independent of m⁶A regulation^{50–52}. While these findings highlight the potential importance of these genes across multiple cancer types, the specific contribution of m⁶A-associated SNPs to their regulation remains to be elucidated. We compiled an extended list of colon cancer-associated m⁶A SNPs located within the *RIBC2* and *MCM3AP-AS1* loci, identified through integrative annotation including HaploReg (Supplementary Table S3). As a next step, we examined the potential involvement of m⁶A-associated SNPs in alternative splicing regulation in colon cancer. Using the same RNA-seq dataset, we performed a genome-wide splicing analysis to identify differentially spliced genes (DSGs) in colon cancer relative to normal tissue. This analysis revealed widespread splicing alterations in skipped exons (SE), alternative 5' or 3' splice sites (A5SS, A3SS), mutually exclusive exons (MXE), and retained introns (RI) (Supplementary Fig. S3E). Among the genes containing m⁶A-SNP overlaps, *FAM118A* and *TRIM4* showed significant A3SS alterations (Fig. 3E). These findings suggest that m⁶A-associated SNPs may contribute to alternative splicing regulation in a transcript-specific manner, potentially influencing cancer-associated splicing patterns.

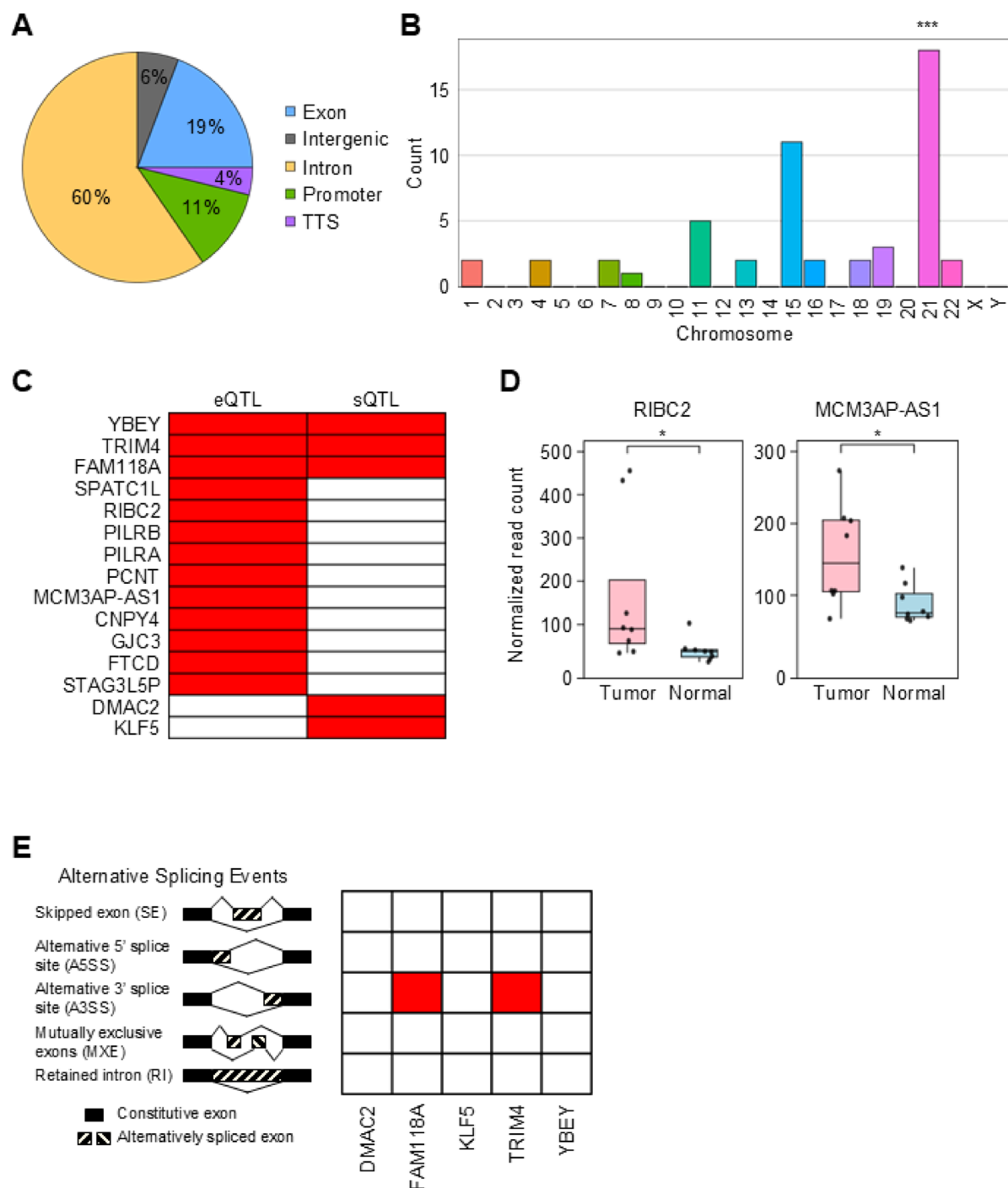


Fig. 3. Functional impact of m⁶A-associated SNPs in colon cancer. (A) Genomic distribution of colon cancer-associated SNPs overlapping with increased m⁶A peaks. We define these SNPs as m⁶A-associated SNPs in colon cancer. (B) Chromosomal distribution of m⁶A-associated SNPs. (C) Annotation of m⁶A-associated SNPs using eQTLs and sQTL information based on snpXplorer⁴⁸. Red box indicates the presence of eQTLs and/or sQTLs of corresponding SNPs. (D) Expression of *RIBC2* and *MCM3AP-AS1* in tumor vs. normal tissues based on RNA-seq data. Statistical significance was determined based on adjusted *P*-values (Wald test in DESeq2 followed by Benjamini-Hochberg multiple correction). (E) Evaluation of alternative splicing events for five sQTL-related genes detected using rMATS-turbo. Red indicates the presence of alternative splicing events of corresponding genes in colon cancers compared to normal. The asterisks indicate statistical significance: * *P* < 0.05, ** *P* < 0.01, *** *P* < 0.001.

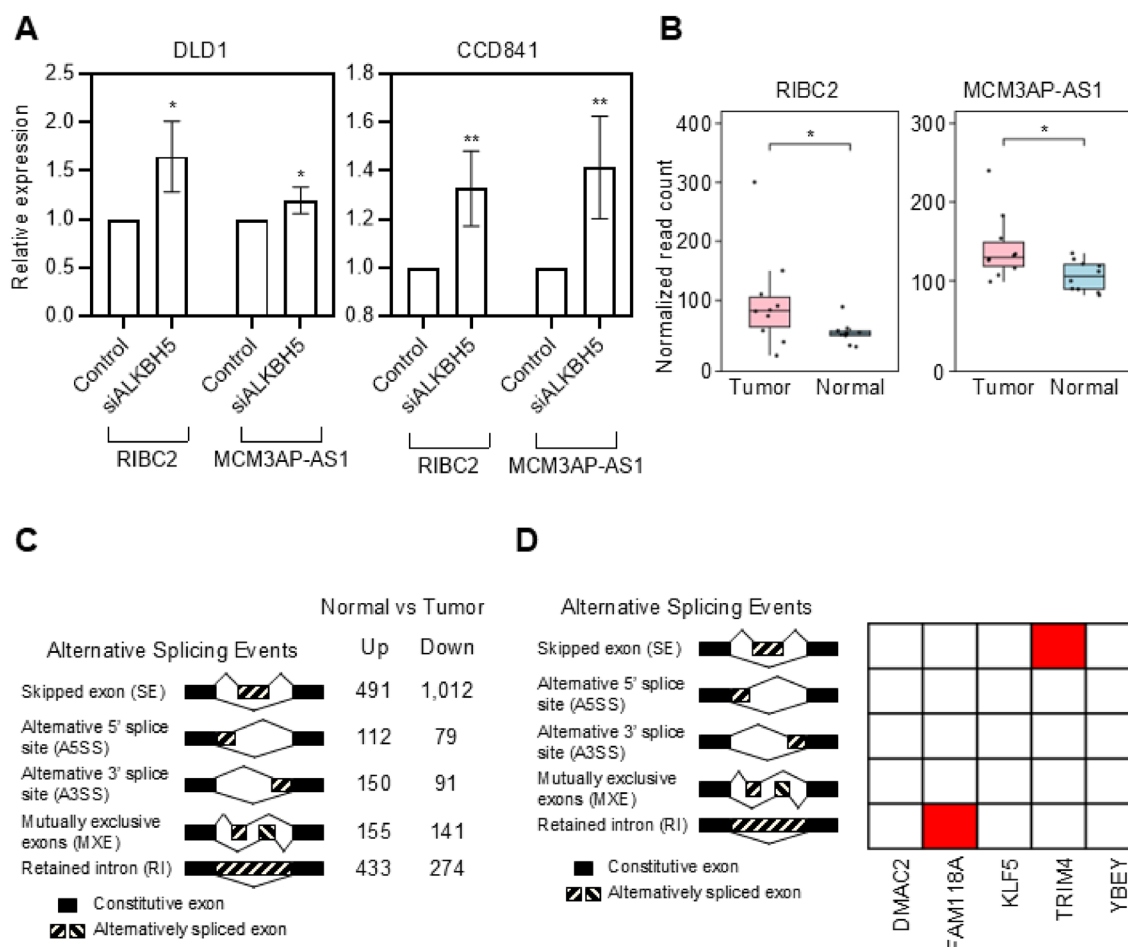


Fig. 4. Orthogonal validation of differentially expressed genes and differential splicing events in colon cancer potentially linked with m⁶A-associated SNPs. **(A)** RT-qPCR showing increased expression of *RIBC2* and *MCM3AP-AS1* after ALKBH5 knockdown in DLD1 and CCD841 cells. Statistical significance was determined using two-tailed unpaired t-test. Error bars represent standard deviation (SD). **(B)** Upregulation of *RIBC2* and *MCM3AP-AS1* in colon tumor tissues compared to normal tissues based on orthogonal RNA-seq dataset (PRJNA742008). Statistical significance was determined based on adjusted *P*-values (Wald test in DESeq2 followed by Benjamini-Hochberg correction). **(C)** Global identification of differential splicing events between tumor and normal tissues categorized by event type. **(D)** Alternative splicing events of selected genes identified from the PRJNA742008 dataset using rMATS-turbo. Red boxes indicate the presence of alternative splicing events of corresponding genes in colon cancers compared to normal. The asterisks indicate statistical significance: * *P* < 0.05, ** *P* < 0.01.

Genetic variants modulate m⁶A modification and drive post-transcriptional rewiring in colorectal cancer

The observed association between colon cancer-associated SNPs and altered m⁶A methylation, along with accompanying changes in gene expression and splicing, led us to investigate whether these transcriptomic alterations are mediated by m⁶A modification. We performed functional assays using the colon cancer cell line DLD1 and the normal colon epithelial cell line CCD841. In both cell lines, we inhibited m⁶A demethylation by knocking down ALKBH5, an m⁶A demethylase known to exert tumor-suppressive functions in colon cancer. In parallel, to assess concordance with patient-derived data and clinical relevance, we analyzed a publicly available dataset (PRJNA742008) containing matched m⁶A-seq and RNA-seq data from ten colon cancer patients. This complementary analysis allowed us to assess the regulatory effects of m⁶A-SNPs in both experimental systems and patient-derived tissues. Upon ALKBH5 depletion, both *RIBC2* and *MCM3AP-AS1* transcripts were upregulated, supporting the hypothesis that m⁶A modification promotes the expression of these genes. (Fig. 4A). Consistent with this observation, analysis of dataset from an orthogonal cohort of colon cancer patients (PRJNA742008) revealed that both genes were significantly upregulated in tumor tissues, supporting the reproducibility of our findings in an independent dataset (Fig. 4B).

Using transcriptome data from PRJNA742008, we additionally investigated differential splicing events potentially regulated by m⁶A-SNPs in colon cancer. Analysis with rMATS identified differential splicing events between tumor and normal tissues (Fig. 4C). Notably, *FAM118A* showed RI events, and *TRIM4* showed SE events in tumor tissues (Fig. 4D). These splicing alterations were distinct from the A3SS events observed in

our earlier analysis, suggesting that m⁶A-associated SNPs may regulate splicing in a context- and transcript-dependent manner. Moreover, while earlier analyses emphasized increased A3SS in tumor tissues, our current findings reveal a broader spectrum of splicing changes, underscoring the complex regulatory roles of m⁶A-SNP interactions in colorectal cancer.

Discussion

Our findings uncover a previously underappreciated regulatory axis suggesting that single-nucleotide polymorphisms (SNPs) may modulate the transcriptome through m⁶A RNA methylation. By integrating m⁶A-seq, RNA-seq, and genome-wide association studies (GWAS) across nine cancer types, we identified significant enrichment of cancer-associated SNPs within dynamically remodeled m⁶A regions.

Notably, among the nine cancer types analyzed, only colon cancer exhibited significant enrichment of SNPs within m⁶A peaks that gained methylation in tumor tissues, passing multiple testing correction. While we cannot definitively resolve the reason for this tissue-specific enrichment without deeper understanding of tissue-specific epitranscriptomic regulation, it is possible that colorectal tissues harbor biological features that increase the likelihood of interaction between germline variation and m⁶A modification. In addition, although variant burden alone does not fully explain the observed enrichment, differences in the number and genomic distribution of GWAS-identified SNPs across cancer types (Fig. 2B and C) may have contributed in part to detection sensitivity or statistical power. For example, although colon cancer harbored fewer associated SNPs than breast or prostate cancers, it showed significant enrichment, whereas other low-SNP cancers such as oral or renal did not. This suggests that variant burden may interact with tissue-specific regulatory environments to shape the observed enrichment pattern. Consistent with this interpretation, Liu et al. (2020) reported that m⁶A-related SNPs are strongly associated with colorectal cancer based on independent Disease Ontology enrichment analysis, reinforcing the biological plausibility of our findings⁵³.

To explore potential functional consequences of m⁶A-associated SNPs, we focused on cancer-associated variants that overlapped with hypermethylated m⁶A peaks and were annotated as expression or splicing quantitative trait loci (eQTLs or sQTLs). These regions exhibited the strongest enrichment of disease-linked variants and showed frequent associations with changes in transcript abundance. In contrast, SNPs located in regions with decreased m⁶A methylation showed no comparable enrichment and were excluded from downstream analyses. Among the QTL-linked transcripts identified, *MCM3AP-AS1* and *RIBC2* were consistently upregulated in colon cancer.

The observed upregulation of *MCM3AP-AS1* and *RIBC2* was supported by both experimental and patient-derived data. In DLD1 and CCD841 cell lines, ALKBH5 depletion increased the expression of these transcripts, indicating that they are responsive to elevated m⁶A levels. Consistent expression changes were observed in an independent cohort of colon cancer patients (PRJNA742008), reinforcing the reproducibility of our findings across both experimental and clinical contexts. These results support the functional relevance of m⁶A in regulating these genes, in agreement with predictions from our integrative m⁶A-QTL analysis.

MCM3AP-AS1 is a long noncoding RNA that has been implicated in the proliferation and invasion of colon cancer cells, potentially *via* Wnt/β-catenin⁵⁴. *RIBC2*, although not previously characterized as a cancer-associated gene, also exhibited consistently elevated expression in colon cancer. While its role in oncogenesis remains to be experimentally validated, its regulation by m⁶A-sensitive loci highlights it as a gene of interest for future investigation.

We also observed that m⁶A-associated SNPs are linked to alternative splicing events in a manner that appears transcript- and context-dependent. In our discovery dataset, *FAM118A* and *TRIM4* exhibited A3SS usage in tumors. In an independent patient cohort, however, splicing patterns differed. In particular, *FAM118A* showed retained introns, while *TRIM4* exhibited exon skipping, indicating transcript-specific differences in splicing outcomes across datasets. These results suggest that m⁶A-SNP interactions may modulate splicing in variable ways depending on cellular context, RNA structure, or the availability of splicing regulators. Given the frequent positioning of m⁶A peaks within exons and promoter-proximal regions, these modifications may facilitate the recruitment of splicing-regulatory RNA-binding proteins, thereby influencing isoform selection in tumors.

Our findings extend current models of how noncoding genetic variants contribute to cancer biology. While many GWAS-identified SNPs lie outside protein-coding regions, their mechanistic interpretation has remained a major challenge. The present study provides evidence that m⁶A methylation may act as a regulatory intermediary, linking static DNA variation to dynamic RNA-level outcomes in tumor transcriptomes. Although tumor and normal tissues are nearly identical at the DNA sequence level, SNPs may shape m⁶A profiles that influence RNA stability, splicing, and abundance.

We define a class of functional variants, which we refer to as m⁶A-QTLs, that operate through epitranscriptomic mechanisms rather than classical transcriptional regulation. These variants reveal a genetically encoded layer of post-transcriptional control that is dynamic and cell-type specific. Although our findings were most prominent in colon cancer, this framework may also inform studies in other tumor types or molecular subgroups where m⁶A dysregulation and regulatory variation co-occur. Future work should further characterize m⁶A-QTLs across diverse biological contexts and investigate whether targeting m⁶A enzymes can mitigate the phenotypic effects of pathogenic variants. As small molecule inhibitors of m⁶A regulators move toward clinical development, our results underscore the importance of considering inherited variation in RNA regulatory mechanisms when designing targeted cancer therapies.

Materials and methods

Preprocessing of RNA-seq and m⁶A-seq data

Gene expression profiles (RNA-seq and m⁶A-seq) from tumor tissues and matched normal tissues for nine cancer types (including colorectal, breast, lung, ovarian, renal cell, and salivary gland cancers) were retrieved from the SRA (<https://www.ncbi.nlm.nih.gov/sra>) database as Fastq format. Detailed dataset information, including accession numbers and sample sizes is summarized in Supplementary Fig. S1A³². Fastq files underwent preprocessing with Trim-Galore (v.0.6.6; <https://github.com/FelixKrueger/TrimGalore>) to remove adapter sequence and low-quality reads, ensuring data quality for subsequent analyses. For RNA-seq data, processed reads were aligned to the human reference genome (GRCh38) with the STAR aligner (v.2.7.10a) using the GTF file from GENCODE (GRCh38.gencode.v43.basic.annotation.gtf) with default parameters³³. Read counts of each gene was calculated using HTSeq package (v.2.0.2)³⁴.

Identification of differential m⁶A peaks (DMPs)

For m⁶A-seq data, peak calling was conducted using R package exomePeak (v.2.16.0)³⁷, followed by annotation of peaks according to genomic features (UTR, exon, intron, promoter, and intergenic regions) using custom R scripts and the GENCODE hg38 annotation. Differential m⁶A peaks (DMPs) were identified using exomePeak with the following criteria: $P < 0.05$.

Identification of differential expressed genes (DEGs)

For RNA-seq data, differentially expressed genes (DEGs) between tumor and normal tissues were identified using the DESeq2 R package (v.1.42.1)³⁵. Genes with adjusted P value < 0.05 and absolute log₂ fold-change > 0.58 were considered as DEG. Heatmaps were generated to visualize DEGs using the R packages ggplot2 (v.3.5.1) and pheatmap (v.1.0.12), respectively^{55,56}.

Identification of differentially spliced genes (DSGs)

To identify differentially spliced genes (DSGs) that contains alternative splicing events in tumor samples, RNA-seq reads were first aligned to the human reference genome (GENCODE hg38) using STAR (v.2.7.10a) with the following options (--twopassMode Basic --alignSJOverhangMin 1 --alignSJoverhangMin 8 --alignEndsType EndToEnd --outSAMattributes NH HI AS NM MD XS --outSAMstrandField intronMotif), which ensured high-confidence spliced alignment and preservation of strand-specific attributes required for downstream splicing analysis. The aligned RNA-seq data were analyzed using rMATS-turbo (v.4.1.2)³⁶. Alternative splicing events including skipped exons (SE), retained introns (RI), alternative 5' splice sites (A5SS), alternative 3' splice sites (A3SS), and mutually exclusive exons (MXE) were quantified. Events with FDR < 0.05 were considered as statistically significant differential splicing events in tumors compared to normal tissues.

Variant set enrichment analysis (VSEA)

Cancer-associated SNPs were obtained from the GWAS catalog⁴⁵. Lead SNPs and their LD-associated SNPs ($r^2 \geq 0.8$) were extracted to establish a comprehensive set of cancer-specific genetic variants using R package haploR (v4.0.7) bioinformatics package³⁷. To assess the functional significance of SNP enrichment within DMPs, Variant Set Enrichment Analysis (VSEA) was performed using the method previously described by Ahmed et al. (2017)⁴⁴. Briefly, VSEA calculates enrichment scores by comparing the observed frequency of cancer-associated SNPs in DMPs to the estimated frequency of cancer-specific SNPs in the same number of randomly selected genomic regions. Statistical significance was determined by generation of null distributions based on 1,000 permutations and random selection of genomic regions each time, with $P < 0.05$ considered significant enrichment.

eQTL and sQTL analysis

Expression quantitative trait loci (eQTL) and splicing quantitative trait loci (sQTL) analyses were conducted using snpXplorer⁴⁹, an integrative database containing functional annotations of SNPs. Candidate cancer-associated SNPs overlapping with differential m⁶A peaks were queried against the snpXplorer database to narrow down SNPs that are potentially associated with the gene expression or alternative splicing events. Statistical significance of eQTL and sQTL associations was defined as $P < 0.05$.

Functional validation via ALKBH5 knockdown

To validate the functional role of m⁶A methylation, ALKBH5 knockdown was performed using siRNA transfection (target-specific siRNA purchased from Genolution) in colon cancer (DLD1) and normal colon epithelial (CCD841) cell lines. Transfections were performed using Lipofectamine RNAiMAX reagent (Invitrogen) according to the manufacturer's protocol. The following ALKBH5 siRNA sequence was used in this study: 5'-r(GCGCCGUCAUCAACGACUA)d(TT)-3'. Gene expression changes of selected targets were quantified via quantitative real-time PCR (qRT-PCR). Gene-specific oligonucleotides used in the study are as follows: 5'- TGGCAAATTCATGGCACC-3' (sense) and 5'- AGAGATGATGACCCTTTTG-3' (antisense) for GAPDH mRNA; 5'- CAGAGAATGGTCTTTGCAGCAGC-3' (sense) and 5'- TTCTGGAGGTGCTTGGCTGTCT-3' (antisense) for RIBC2 mRNA; and 5'- CTGGAAGCAGAAAGAGGCTGG-3' (sense) and 5'- ACTGGAAGGAGCACAGAGTG-3' (antisense) for MCM3AP-AS1 lncRNA. All experiments were performed in triplicate.

Statistical analysis

All statistical analyses were conducted using R software (v.4.2.1)⁵⁶. Data are expressed as mean \pm standard deviation. Comparisons between two groups were performed using Student's t-tests. Spearman's correlation

analysis was performed using R packages ggpubr (v.0.6.0) and stats (base package)⁵⁸. $P < 0.05$ was considered statistically significant.

Data availability

The RNA-seq and m⁶A-seq datasets analyzed in this study are publicly available from the NCBI Sequence Read Archive (SRA) under accession number PRJNA488293, PRJNA1011293, PRJNA659478, PRJNA679771, PRJNA901504, PRJNA814496, PRJNA719065, PRJNA1092401, PRJNA1039844, PRJNA786917, and PRJNA742008. Additional processed data supporting the findings of this study are available from the corresponding authors upon reasonable request.

Received: 18 June 2025; Accepted: 6 October 2025

Published online: 11 November 2025

References

- Helm, M. & Motorin, Y. Detecting RNA modifications in the epitranscriptome: predict and validate. *Nat. Rev. Genet.* **18**, 275–291 (2017).
- Nachtergaele, S. & He, C. Chemical modifications in the life of an mRNA transcript. *Annu. Rev. Genet.* **52**, 349–372 (2018).
- Wang, Y. et al. N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat. Cell Biol.* **16**, 191–198 (2014).
- Huang, H. et al. Publisher Correction: Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* **22**, 1288 (2020).
- Zheng, G. Q. et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell.* **49**, 18–29 (2013).
- Liu, N. A. et al. N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* **45**, 6051–6063 (2017).
- Roundtree, I. A. et al. YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *Elife* **6** <https://doi.org/10.7554/eLife.31311> (2017).
- Wang, X. et al. N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388–1399 (2015).
- Shi, H. et al. YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. *Cell Res.* **27**, 315–328 (2017).
- Jia, G., Fu, Y. & He, C. Reversible RNA adenosine methylation in biological regulation. *Trends Genet.* **29**, 108–115 (2013).
- Han, S. H. & Choe, J. Diverse molecular functions of m(6)A mRNA modification in cancer. *Exp. Mol. Med.* **52**, 738–749 (2020).
- Han, S. H. & Choe, J. Deciphering the molecular mechanisms of epitranscriptome regulation in cancer. *BMB Rep.* **54**, 89–97 (2021).
- An, Y. & Duan, H. The role of m6A RNA methylation in cancer metabolism. *Mol. Cancer.* **21**, 14 (2022).
- Jiang, X. et al. The role of m6A modification in the biological functions and diseases. *Signal. Transduct. Target. Ther.* **6**, 74 (2021).
- Sang, H., Liu, J., Chen, X. & Zeng, Y. METTL16-dependent miR-146b-5p m6A modification remodeling sensitizes NSCLC to osimertinib via activating PI3K/AKT signaling. *BMC Cancer.* **25**, 641. <https://doi.org/10.1186/s12885-025-14041-y> (2025).
- Lu, Z. M., Lyu, Z., Dong, P. X., Liu, Y. M. & Huang, L. N6-methyladenosine RNA modification in stomach carcinoma: novel insights into mechanisms and implications for diagnosis and treatment. *Biochimica Et Biophys. Acta Mol. Basis Disease* **1871** (2025).
- Zhao, Q. et al. m(6)A RNA modification modulates PI3K/Akt/mTOR signal pathway in Gastrointestinal cancer. *Theranostics* **10**, 9528–9543. <https://doi.org/10.7150/thno.42971> (2020).
- Wang, Y. et al. Epigenetic modification of m(6)A regulator proteins in cancer. *Mol. Cancer.* **22**, 102. <https://doi.org/10.1186/s12943-023-01810-1> (2023).
- Meyer, K. D. et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646 (2012).
- Ly, J. et al. N6-methyladenosine-related single-nucleotide polymorphism analyses identify oncogene RNFT2 in bladder cancer. *Cancer Cell Int.* **22**, 301 (2022).
- Tan, L. et al. Positive natural selection of N6-methyladenosine on the RNAs of processed pseudogenes. *Genome Biol.* **22**, 180 (2021).
- Shastry, B. S. SNPs: impact on gene function and phenotype. *Methods Mol. Biol.* **578**, 3–22 (2009).
- Komar, A. A. Silent snps: impact on gene function and phenotype. *Pharmacogenomics* **8**, 1075–1080 (2007).
- Stadler, Z. K. et al. Genome-Wide association studies of cancer. *J. Clin. Oncol.* **28**, 4255–4267 (2010).
- Stadler, Z. K. et al. Genome-wide association studies of cancer predisposition. *Hematol. Oncol. Clin. N. Am.* **24**, 973–996 (2010).
- Schipper, M. & Posthuma, D. Demystifying non-coding GWAS variants: an overview of computational tools and methods. *Hum. Mol. Genet.* **31**, R73–R83 (2022).
- Degtyareva, A. O., Antontseva, E. V. & Merkulova, T. I. Regulatory snps: altered transcription factor binding sites implicated in complex traits and diseases. *Int J. Mol. Sci.* **22** (2021).
- Pang, B. X., van Weerd, J. H., Hamoen, F. L. & Snyder, M. P. Identification of non-coding silencer elements and their regulation of gene expression. *Nat. Rev. Mol. Cell. Biol.* **24**, 383–395 (2023).
- Sun, X. R., Dai, Y. S., Tan, G. L., Liu, Y. Q. & Li, N. Integration analysis of m6A-SNPs and eQTLs associated with sepsis reveals platelet degranulation and infection are mediated by m6A mRNA methylation. *Front. Genet.* **11** <https://doi.org/10.3389/fgene.2020.00007> (2020).
- Li, T. et al. N6-methyladenosine-associated genetic variants in NECTIN2 and HPCAL1 are risk factors for abdominal aortic aneurysm. *Iscience* **27** <https://doi.org/10.1016/j.isci.2024.109419> (2024).
- Chen, M., Lin, W. M., Yi, J. R. & Zhao, Z. H. Exploring the epigenetic regulatory role of m6A-Associated SNPs in type 2 diabetes pathogenesis. *Pharmacogenomics Pers. Med.* **14**, 1369–1378 (2021).
- Leinonen, R., Sugawara, H. & Shumway, M. International nucleotide sequence Database, C. The sequence read archive. *Nucleic Acids Res.* **39**, D19–21 (2011).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Anders, S., Pyl, P. T. & Huber, W. HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
- Love, M. I., Huber, W. & Anders, S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15** (2014).
- Wang, Y. Y. et al. rMATS-turbo: an efficient and flexible computational tool for alternative splicing analysis of large-scale RNA-seq data. *Nat Protoc.* **19** (2024).
- Meng, J. et al. A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomepeak R/Bioconductor package. *Methods* **69**, 274–281 (2014).

38. Linder, B. et al. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods*. **12**, 767–772 (2015).
39. Huang, H., Weng, H. & Chen, J. m(6)A modification in coding and Non-coding rnas: roles and therapeutic implications in cancer. *Cancer Cell*. **37**, 270–288 (2020).
40. Olarerin-George, A. O. & Jaffrey, S. R. MetaPlotR: a Perl/R pipeline for plotting metagenes of nucleotide modifications and other transcriptomic sites. *Bioinformatics* **33**, 1563–1564 (2017).
41. Domissini, D. et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201–206 (2012).
42. Gao, R. et al. m6A modification: A Double-Edged sword in tumor development. *Front. Oncol.* **11**, 679367 (2021).
43. Sun, T., Wu, R. & Ming, L. The role of m6A RNA methylation in cancer. *Biomed. Pharmacother.* **112**, 108613 (2019).
44. Ahmed, M. et al. Variant set enrichment: an R package to identify disease-associated functional genomic regions. *Biodata Min* **10** (2017).
45. Sollis, E. et al. The NHGRI-EBI GWAS catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* **51**, D977–D985 (2023).
46. Bartonicek, N. et al. Intergenic disease-associated regions are abundant in novel transcripts. *Genome Biol.* **18**, 241 (2017).
47. Alsheikh, A. J. et al. The landscape of GWAS validation; systematic review identifying 309 validated non-coding variants across 130 human diseases. *BMC Med. Genomics*. **15**, 74 (2022).
48. Cornish, A. J. et al. The genomic landscape of 2,023 colorectal cancers. *Nature* **633**, 127–136 (2024).
49. Tesi, N., van der Lee, S., Hulsman, M., Holstege, H. & Reinders, M. J. T. SnpXplorer: a web application to explore human SNP-associations and annotate SNP-sets. *Nucleic Acids Res.* **49**, W603–W612 (2021).
50. Zhou, M. Y. et al. Long noncoding RNA MCM3AP-AS1 enhances cell proliferation and metastasis in colorectal cancer by regulating miR-193a-5p/SENPI. *Cancer Med-U.S.* **10**, 2470–2481. <https://doi.org/10.1002/cam4.3830> (2021).
51. Ghafouri-Fard, S., Khoshbakht, T., Hussien, B. M., Taheri, M. & Samadian, M. A review on the role of MCM3AP-AS1 in the carcinogenesis and tumor progression. *Cancer Cell Int.* **22** <https://doi.org/10.1186/s12935-022-02644-5> (2022).
52. Li, N., Yu, K., Lin, Z. & Zeng, D. Y. Identifying a cervical cancer survival signature based on mRNA expression and genome-wide copy number variations. *Exp. Biol. Med.* **247**, 207–220 (2022).
53. Liu, J. E. et al. Landscape and regulation of m6A and m6Am methylome across human and mouse tissues. *Mol. Cell.* **77**, 426–440 (2020).
54. Wang, Y. et al. A novel lncRNA MCM3AP-AS1 promotes the growth of hepatocellular carcinoma by targeting miR-194-5p/FOXAI axis. *Mol. Cancer*. **18**, 28. <https://doi.org/10.1186/s12943-019-0957-7> (2019).
55. Wang, J. et al. PROS1 shapes the immune-suppressive tumor microenvironment and predicts poor prognosis in glioma. *Front. Immunol.* **13**, 1052692. <https://doi.org/10.3389/fimmu.2022.1052692> (2022).
56. Zhang, X. et al. Single-cell RNA and transcriptome sequencing profiles identify immune-associated key genes in the development of diabetic kidney disease. *Front. Immunol.* **14**, 1030198. <https://doi.org/10.3389/fimmu.2023.1030198> (2023).
57. Zhabannikov, I. Y., Arbeev, K., Ukrantseva, S. & Yashin A. I. haploR: an R package for querying web-based annotation tools. *F1000Research* **6**, 97 (2017).
58. Zhu, N. et al. SIRT1 and ZNF350 as novel biomarkers for osteoporosis: a bioinformatics analysis and experimental validation. *Mol. Biol. Rep.* **51**, 530 (2024).

Author contributions

S.H.H. designed and conducted all experiments, performed data analysis, and drafted the manuscript. S.J., Y.K. and K.T. assisted with data preprocessing and figure preparation. M.F.W., H.J., and J.C. conceptualized and supervised the study, and revised the manuscript. All authors read and approved the final version of the manuscript.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korean government (Ministry of Science and ICT): RS-2020-NR050915, RS-2022-NR066714 and RS-2023-00207840 to J.C.; and RS-2024-00336518 to H.J. Additional support was provided to J.C. by the Korea Basic Science Institute (National Research Facilities and Equipment Center) grant funded by the Ministry of Education (2023R1A6C101A009), and by the Bio&Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. RS-2024-00411768), and to H.J. by the POSCO Science Fellowship from the POSCO TJ Park Foundation (2024-11-1819).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-23397-x>.

Correspondence and requests for materials should be addressed to M.F.W., H.J. or J.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025