Check for updates

Received: 5 May 2026 | Accepted: 16 September 2026

DOI: 10.1002/tpg2.70327

The Plant Genome

ORIGINAL ARTICLE

The novel regulatory influence of Argonaute 1 on messenger RNA 5-methylcytosine modification in *Arabidopsis thaliana*

Zedi Feng[1] | **Xiu Yin**[2] | **Zhifang Jiang**[1] | **Jia Yao**[1] | **Yiling Dong**[1] | **Xiaoxia Ma**[2] | **Yijun Meng**[1]

[1]College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou, China

[2]School of Pharmacy, Hangzhou Normal University, Hangzhou, China

Correspondence
Yijun Meng, College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou, 311121, China.
Email: mengyijun@zju.edu.cn

Xiaoxia Ma, School of Pharmacy, Hangzhou Normal University, Hangzhou, 311121, China.
Email: xiaoma_0910@163.com

Assigned to Associate Editor Meixia Zhao.

Funding information
National Natural Science Foundation of China, Grant/Award Number: 32370707

Abstract

5-Methylcytidine (m^5C) modifications play a critical role in RNA metabolism. TRM4B (tRNA-specific methytransferase 4B) is the only known component of the plant m^5C machinery that mediates mRNA (messenger RNA) methylation. To date, no regulators governing site-specific m^5C modifications have been identified, apart from several degenerated sequence motifs. Drawing on relevant findings in animals, we investigated the involvement of AGO1 (Argonaute 1) in m^5C regulation using an *ago1* mutant of *Arabidopsis*. m^5C-RIP-seq (m^5C RNA immunoprecipitation followed by deep sequencing) identified high-confidence 5981 and 4559 m^5C peaks in WT (wild type) and *ago1*, respectively. A substantial proportion of these peaks overlapped between WT and *ago1*, indicating their high gene specificity. Compared to WT, *ago1* exhibited 366 hyper-methylated and 168 hypo-methylated peaks. While both hyper- and hypo-methylation could correlate with either up- or down-regulated gene expression, actively transcribed genes with relatively complex structures were preferentially targeted by m^5C modification. A significant fraction of m^5C-associated genes are functionally involved in mRNA surveillance and ribosome biogenesis. Intriguingly, over 60% of m^5C peaks identified in WT were associated with the miRNA (microRNA) binding sites. Furthermore, AGO1 binding signals could be detected on the corresponding transcripts for >40% of the m^5C peaks based on the AGO1-RIP-seq experiment. Immunoprecipitation coupled to mass spectrometry validated the interaction between TRM4B and TRANSPORTIN1 (TRN1)—a known AGO1-interacting protein that is also implicated in the miRNA

Abbreviations: AGO1, Argonaute 1; CDS, coding sequence; FPKM, fragments per kilobase of exon per million fragments mapped; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; lncRNA, long non-coding RNA; m^5C, 5-methylcytidine; m^5C-RIP-seq, m^5C RNA immunoprecipitation followed by deep sequencing; m^6A, N^6-methyladenosine; MBS, microRNA binding site; miRNA, microRNA; mRNA, messenger RNA; PCR, polymerase chain reaction; pre-miRNA, precursor miRNA; RNAi, RNA interference; sRNA-seq, small RNA sequencing; TAIR, The *Arabidopsis* Information Resource; TRM4B, tRNA-specific methytransferase 4B; UTR, untranslated region; WT, wild type.

Zedi Feng, Xiu Yin, and Zhifang Jiang contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2026 The Author(s). *The Plant Genome* published by Wiley Periodicals LLC on behalf of Crop Science Society of America.

pathway. These findings support a model in which the miRNA-guided AGO1 complex plays a critical role in mediating site-specific m^5C modifications.

Plain Language Summary

The high conservation of the AGO proteins, the miRNA pathway, and the m^5C modifications implies their biological importance in the post-transcriptional pathways. In the present study, transcriptome-wide m^5C-RIP-seq uncovered the prevalent m^5C methylation in mRNAs, and a portion of these modifications are influenced by the Argonaute 1 (AGO1) activity. Interestingly, both the microRNA binding sites and the AGO1 binding signals were found to be associated with many m^5C peaks. These discoveries drove us to propose a novel model that the miRNA-guided AGO1 complex played a critical regulatory role in m^5C modifications in plants, which could serve as an important supplement to the molecular mechanisms underlying the site-specific methylation in addition to the degenerated motifs identified previously.

1 | INTRODUCTION

In plants, gene expression is governed by multi-layered regulatory mechanisms that are essential for normal growth, development, reproduction, and environmental adaptability. Epigenetic modifications contribute a critical regulatory layer in plant gene expression networks. It is well established that these modifications occur not only at the DNA level but also at the RNA level (Liang et al., 2020). Accordingly, the terms "epigenomics" and "epitranscriptomics" have been formally established to describe the mechanistic and functional studies of genome-wide DNA modifications and transcriptome-wide RNA modifications, respectively. Although the regulatory roles of RNA modifications in gene expression have emerged as a prominent research focus only in the past decade (A. Burgess et al., 2016; Saletore et al., 2012), chemical modifications of RNA molecules were first identified in calf liver over 70 years ago (Cohn, 1951). Subsequent studies have confirmed that RNA modifications are widespread across eukaryotes, including the 5′ cap structure, N^6-methyladenosine (m^6A), and 5-methylcytidine (m^5C) (Desrosiers et al., 1974; Dubin & Taylor, 1975; Shatkin, 1976). To date, >170 distinct types of post-transcriptional covalent RNA modifications have been documented across all kingdoms of life (Cappannini et al., 2024; Wiener & Schwartz, 2021). Notably, most RNA modifications, along with their core regulatory machineries (including "writers" for signal deposition, "erasers" for signal removal and, "readers" for signal interpretation), are highly conserved in plants, animals, and microorganisms, highlighting their indispensable roles in organismal survival (Shen et al., 2023; Sun et al., 2023). Beyond m^6A, m^5C represents one of the most prevalent RNA methylation marks in both plants and animals. In the model plant *Arabidopsis* (*Arabidopsis thaliana*), m^5C modifications were initially discovered on tRNAs, with this methylation event catalyzed by tRNA-specific methyltransferases encoded by *DNMT2* (*At5g25480*, *DNA methyltransferase 2*, also designated *TRDMT1*) and *TRM4B* (*At2g22400*, *tRNA-specific methytransferase 4B*) (A. L. Burgess et al., 2015; Goll et al., 2006). Importantly, *TRM4B* also mediates m^5C methylation on messenger RNAs (mRNAs) and long non-coding RNAs (lncRNAs). Two independent studies have demonstrated that transcriptome-wide m^5C levels are significantly reduced in the *trm4b* mutant, a phenotype specifically restricted to roots rather than other tested organs, indicating that TRM4B functions as a tissue-specific RNA m^5C methyltransferase in *Arabidopsis* (Cui et al., 2017; David et al., 2017). Additionally, m^5C modifications have also been identified on rRNAs, with *NOP2A* (*At5g55920*, *nucleolar protein 2A*), *NOP2B* (*At4g26600*), *NOP2C* (*At1g06560*), *NSUN5* (*At5g26180*, *NOP2/Sun domain protein 5*), and *RNMT* (*At3g13180*, *RNA methyltransferase*) identified as the corresponding rRNA-specific methyltransferase (A. L. Burgess et al., 2015). To date, m^5C modifications have been shown to be extensively involved in multiple RNA regulatory processes, including RNA processing, alternative polyadenylation, structural remodeling, stability control, translational regulation, subcellular localization, and intracellular trafficking (Shen et al., 2023; Sun et al., 2023; Wiener & Schwartz, 2021). In plants, m^5C also plays critical roles in organ development and stress response (Zhang et al., 2020). However, fundamental knowledge of RNA m^5C regulatory mechanisms remains largely limited, particularly in plant systems. For instance, TRM4B is the only well-characterized m^5C writer for mRNAs and lncRNAs in *Arabidopsis*, while no corresponding m^5C erasers or readers have been reported to date.

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Future studies should prioritize the systematic identification of plant epitranscriptomic regulatory codes and the full characterization of the complete m^5C regulatory machinery.

Recent technological advances in high-throughput sequencing platforms have greatly promoted the m^5C research. In 2017, two independent research groups successively reported transcriptome-wide profiling of RNA m^5C modification sites in *Arabidopsis*. David and colleagues (2017) combined RNA bisulfite conversion with high-throughput sequencing (bsRNA-seq) to map m^5C sites at single-nucleotide resolution. This work identified over 1000 m^5C sites on mRNAs and lncRNAs across *Arabidopsis* siliques, shoots, and roots, and these sites exhibited striking tissue-specific distribution patterns (David et al., 2017). To address the inherent sequencing artifacts of bsRNA-seq introduced during the C-to-U conversion step, Cui and coworkers (2017) developed the m^5C-RIP-seq (m^5C RNA immunoprecipitation followed by deep sequencing) approach, which achieved significantly higher specificity and sensitivity by employing an m^5C-specific antibody. Using this optimized method, a total of 6045 high-confidence m^5C peaks were detected across >4400 genes in *Arabidopsis* seedlings (Cui et al., 2017). Unlike plant m^6A sites, which are consistently enriched at the 3′ ends of the transcripts, the distribution landscapes of m^5C sites are highly divergent across different plant species (Shen et al., 2023). For example, m^5C sites are predominantly localized to the 3′ ends of *Arabidopsis* transcripts, whereas in rice they are preferentially enriched at the 5′ ends (Shen et al., 2019; Zhang et al., 2020). Collectively, these findings demonstrate that RNA m^5C modifications are governed by a multi-layered regulatory logic that integrates site specificity, tissue specificity, and species specificity.

In Cui et al. (2017) study, several conserved sequence motifs, including HACCR (H = A/C/U; R = A/G) and CUYCUYC (Y = C/U), were predicted to govern site-specific m^5C modifications in *Arabidopsis*. However, these degenerate cis-elements are clearly insufficient to confer the high sequence specificity required for precise RNA methylation, as it has been estimated that only 0.02%–0.09% of total cellular cytosines undergo m^5C modification (Sun et al., 2023). A critical mechanistic clue emerged from a pioneering investigation of m^5C sites in human mRNAs. In that work, Squires and colleagues (2012) uncovered a striking transcriptome-wide correlation between mRNA m^5C sites and the binding regions of Argonaute (AGO) proteins (Squires et al., 2012). Guided by microRNAs (miRNAs) or small interfering RNAs, AGO proteins serve as core effectors in RNA interference (RNAi) pathways via a sequence-dependent targeting mechanism (Fang & Qi, 2016; Tolia & Joshua-Tor, 2007). Remarkably, both RNAi and m^5C modifications are deeply conserved across eukaryotic lineages, including animals and plants. Thus, it will be interesting to investigate the potential functional crosstalk between these two evolutionarily conserved pathways. For instance, the recruitment of AGO protein complexes to target transcripts could sterically block or facilitate the access of m^5C "writer" or "eraser" enzymes, thereby directly shaping the landscape of site-specific RNA methylation. The *Arabidopsis* genome encodes 10 members of the AGO protein family (Vaucheret, 2008). Among them, Argonaute 1 (AGO1) is the most well-characterized paralog, which predominantly associates with miRNAs to mediate sequence-specific post-transcriptional gene regulation. In the present study, we performed m^5C-RIP-seq on *ago1-25*, a hypomorphic *ago1* mutant of *Arabidopsis*, to systematically dissect the potential regulatory role of AGO1 in RNA m^5C deposition. Integrative analysis of these transcriptome-wide datasets led us to propose a working model that the miRNA-guided AGO1 complex played an important regulatory role in the site-specific m^5C modifications. This work provides fundamental new insights into the elusive molecular mechanisms governing site-specific m^5C deposition in plant transcriptomes.

Core Ideas

- Thousands of m^5C peaks enriched with CU/AC motifs were identified from *Arabidopsis*.
- Genes with higher expression and more exons are more susceptible to m^5C modifications.
- The influence of the *ago1* mutation on mRNA m^5C modifications was reported.
- MiRNAs play a potential regulatory role in the site-specific m^5C modifications.

2 | MATERIALS AND METHODS

2.1 | Plant materials and growth condition

All seeds of *Arabidopsis* were surface-sterilized in 70% ethanol for 3 min and then in 10% NaClO for 10 min. It was followed by four washes with sterile distilled water. The seeds were sown on the Gamborg B5 basal medium (PhytoTechnology Laboratories, product No. G398) at pH 5.8 under sterile condition. The formula of the B5 basal medium was referenced from a previous study (Gamborg et al., 1968), containing 134 mg/L ammonium sulfate, 3 mg/L boric acid, 113.24 mg/L anhydrous calcium chloride, 0.025 mg/L cobalt chloride·$6H_2O$, 0.025 mg/L cupric sulfate·$5H_2O$, 37.26 mg/L $Na_2EDTA{\cdot}2H_2O$, 27.8 mg/L ferrous sulfate·$7H_2O$, 122.09 mg/L anhydrous magnesium sulfate, 10 mg/L manganese sulfate·H_2O, 0.25 mg/L molybdic acid (sodium salt). Note that·$2H_2O$, 0.75 mg/L potassium

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

iodide, 2500 mg/L potassium nitrate, 150 mg/L sodium phosphate monobasic, 2 mg/L zinc sulfate·7H_2O, 100 mg/L myo-inositol, 1 mg/L nicotinic acid (free acid), 1 mg/L pyridoxine·HCl, and 10 mg/L thiamine·HCl. The seeds were incubated at 4°C in darkness for 2 days before being moved to a growth chamber at 24°C for 16 h in the light (150 μmol/m^2/s). The mutant *ago1-25* was kindly provided by Prof. Yijun Qi (School of Life Sciences, Tsinghua University), and the mutant was described by a previous study (Morel et al., 2002).

2.2 | Generation and screening of the transgenic plants

The transgenic plants over-expressed with *AGO1* and *TRANSPORTIN1* (*TRN1*) fused with the 3×FLAG-tag were generated by using the floral dip method (Clough & Bent, 1998). The TAIR10 (Reiser et al., 2024)-registered full-length coding sequences (CDSs) of *AGO1* and *TRN1* were engineered into the pEGOEP35S vector separately for transformation. Besides, the empty vector was also transformed into the wild-type (WT) plant of Col-0, treated as the control. The T_0 transgenic seeds were sterilized, germinated, and selected on the agar medium containing hygromycin B (50 mg/mL). For each T_1 line, two-step screening was performed: (1) pooled-sampling followed by the polymerase chain reaction (PCR)-based detection of the hygromycin B resistance gene, and (2) pooled-sampling followed by reverse transcription-PCR-based examination of the gene expression levels. Finally, the screened transgenic plants were sent for IP-MS (immunoprecipitation coupled to mass spectrometry).

2.3 | m^5C-RIP-seq and data analysis

Total RNAs were isolated from the 16-day-old seedlings by using the TRIzol reagent (Thermo Scientific, cat. no. 15596026) based on the user manual. The total RNAs were treated with DNase I (NEB, cat. no. M0303L) in order to eliminate potential DNA contamination. The purity of the RNA samples was assessed by measuring the A260/A280 ratio by using Nanodrop OneC (Thermo Scientific). The integrity of the RNA samples was verified by using LabChip GX Touch (Revvity). The concentrations of the qualified RNA samples were then determined by using a Qubit 3.0 fluorometer with the QubitTM RNA Broad Range Assay kit (Thermo Scientific, cat. no. Q10210). The poly(A)-tailed RNAs were enriched from the total RNAs by using VAHTS mRNA Capture Beads (Vazyme, cat. no. N401), following the user manual.

To perform the IP experiment for the m^5C methylated RNAs, the poly(A)-tailed RNAs were chemically fragmented into approximately 100-nt fragments by treating with 20 mM $ZnCl_2$ at 95°C for 5 to 10 min. A portion (10%) of the fragments was reserved as "input." The remaining was subjected to IP by using the m^5C monoclonal antibody (Diagenode, cat. no. C15200081) at a final concentration of 40 U/μL and RNasin (Promega, cat. no. N2615). The IP was performed by incubation at 4°C for 2 h with rotation. The immunoprecipitated complexes were then incubated with the protein-A magnetic beads (Thermo Scientific, cat. no. 88848) under 4°C for 1 h with rotation. The protein-A magnetic beads were thoroughly washed to remove non-specifically bound RNAs, and the specifically bound RNAs were extracted using the TRIzol reagent.

Both the "input" and the immunoprecipitated RNAs ("IP") were subject to sequencing library preparation by using the KC Digital mRNA Library Prep Kit (Seqhealth Tech. Co., Ltd.), according to the manufacturer's instructions. The kit labeled the cDNAs with 12-nt unique molecular identifier (UMI) in order to eliminate duplication bias and errors possibly introduced by the PCR and sequencing steps. PCR products ranging from 200 to 500 bp were enriched for the sequencing libraries. Sequencing was performed on the NovaSeq X Plus (Illumina) platform using the PE150 mode. For m^5C-RIP-seq, there are three biological replicates for WT and *ago1-25*, respectively.

Raw sequencing data were pre-treated (removal of the low-quality reads and those contaminated with adaptors) by using fastp (Chen et al., 2018). Clean reads were further treated with in-house scripts to eliminate duplication bias introduced during PCR and sequencing. Briefly, the clean reads were first clustered based on the UMI tags. The reads containing the identical UMI tag were grouped into the same cluster. Reads assigned to the same cluster were then compared to each other by pairwise alignment. Based on the alignment, the reads sharing > 95% sequence identity were further assigned to a sub-cluster. Subsequently, multiple sequence alignment was performed to achieve one consensus sequence for each sub-cluster. After these steps, the possible errors and biases introduced by PCR or sequencing were eliminated. The deduplicated consensus sequences were used for m^5C peak analysis by mapping them onto the *Arabidopsis* genome retrieved from TAIR10 (The *Arabidopsis* Information Resource, release 10; https://www.arabidopsis.org/) (Reiser et al., 2024) using STAR (Dobin et al., 2013). For the protein-coding genes, the software exomePeak (Meng et al., 2014) was used for peak calling. For the non-coding genes, exomePeak2 (https://rdrr.io/github/ZhenWei10/exomePeak2/) (Meng et al., 2013, 2014) was used for peak calling by using the "whole-genome" mode. The differential m^5C peaks were identified by using QNB (Liu et al., 2017). Based on the "input" data, featureCounts (Liao et al., 2014) was used to calculate the gene expression levels in FPKM (fragments per kilobase of exon per million fragments mapped), and edgeR (M. D. Robinson et al., 2010) was used to discover the differentially expressed genes. To assign

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

m^5C peaks to distinct transcriptional regions (5′UTRs, CDSs, 3′UTRs [where UTRs are untranslated regions], and intronic and intergenic regions), the peaks were first filtered to retain only those exhibiting > 50% sequence overlap with a corresponding gene. For peaks successfully assigned to a host gene, the genomic feature harboring the longest overlapping segment of the peak was identified as its final annotation. For example, a peak was classified as CDS if its maximum overlap length resided within the CDS region. Peaks that could not be mapped to any annotated protein-coding gene were designated as intergenic regions.

Staem5 (Chai et al., 2021) was used to evaluate the m^5C modification potential of the peak regions identified from the m^5C-RIP-seq data. All peak regions subject to the analysis were segmented into 41-nt fragments, and the 41-nt sequences of the control sets were randomly selected from the cDNAs retrieved from TAIR10. The metagene profiles showing the distribution patterns of the m^5C peaks in mRNAs were drawn by using Guitar (Cui, Wei, et al., 2016). To do this analysis, the mRNAs were aligned and divided into three regions (5′UTRs, CDSs, and 3′UTRs).

2.4 | RIP-seq of AGO1 and data analysis

To perform the RIP-seq experiment of AGO1 of *Arabidopsis* (TAIR10 gene ID: *AT1G48410*), the 2-week-old WT seedlings of Col-0 were collected and frozen in liquid nitrogen, and ground by Tissue Lyser II (Qiagen, cat. no. 69982). The ground powder was resuspended in cell lysis buffer. A total of 10% of the cell lysate was saved as the input control. The remaining 80% of the lysate was incubated with anti-AGO1 antibody (Agrisera, cat. no. AS09 527-100) for IP. The immunoprecipitated complexes (referred to as the IP fraction) were then collected and analyzed. And the remaining 10% of the lysate was incubated with rabbit IgG (Cell Signaling Technology, cat. no. 2729S) and served as a negative control. Then, fresh protein G magnetic beads were prepared for capture. For each sample, aliquot 40 μL of protein G beads into a new tube. After washing the beads three times with 1 mL wash buffer, the beads were resuspended by an equal volume of wash buffer. Then, the washed bead suspension was added to the antibody-lysate mixture, followed by 4°C incubation for 1 h with rotation to allow the beads to capture the antibody-protein-RNA complexes. After incubation, the tubes were placed on a magnetic stand for 1 min, then carefully remove and discard the supernatant. To wash the bead-bound complexes, first take the tube off the magnetic stand, add 1 mL of ice-cold wash buffer, and gently resuspend the beads by slow pipetting, followed by a 10-min incubation at 4°C on a rotating mixer. After incubation, place the tube back on the magnetic stand for 5 min to fully settle the beads, then aspirate and discard the supernatant. Repeat this entire wash procedure three times in total, and it is strongly recommended to transfer the bead suspension to a fresh new microcentrifuge tube after each wash cycle to effectively minimize non-specific background contamination. After completing the final wash, gently resuspend the pelleted beads in 1 mL of ice-cold wash buffer by careful pipetting, and the immunoprecipitated RNA-protein complexes are now fully prepared for the subsequent RNA extraction procedure. RNA from the "input" and "IP" samples was extracted using TRIzol (Invitrogen, cat. no. 15596018CN). Strand-specific RNA sequencing libraries were constructed using KC-Digital Stranded mRNA Library Prep Kit for Illumina (Seqhealth, cat. no. DR08502) according to the user manual. By labeling the pre-amplified cDNAs with 8-nt UMI, duplication bias introduced during PCR and sequencing steps could be eliminated. The libraries enriched with 200- to 500-bp products were quantified. Finally, paired-end sequencing (PE150) was performed by SeqHealth on Illumina Novaseq 6000. A total of two biological replicates were included for the RIP-seq experiment.

Raw sequencing data were pre-treated (removal of the low-quality reads and those contaminated with adaptors) by using Trimmomatic (Bolger et al., 2014). Clean reads were further treated with in-house scripts to eliminate duplication bias introduced during PCR and sequencing based on the UMI tags. The deduplicated sequences were used for protein binding site analysis by mapping them onto the *Arabidopsis* genome retrieved from TAIR10 by using STAR (Dobin et al., 2013) with default parameters. Peak calling was performed by using exomePeak (Meng et al., 2014), and peak annotation was performed by using BEDTools (Quinlan & Hall, 2010). The RIP peaks ($p < 0.01$) were annotated with genes if they were entirely resided within the gene bodies.

2.5 | Small RNA sequencing and data processing

TruSeq Small RNA Sample Prep Kit (Illumina, cat. no. RS200-0012) was used for sRNA-seq (small RNA sequencing) library construction according to the user manual. Briefly, the RNAs <200 nt were separated from 1 μg total RNAs by polyacrylamide gel electrophoresis and were ligated to the adapters by T4 RNA ligase at both ends. By using SuperScript II Reverse Transcriptase (Invitrogen, cat. no. 18064071), the adapter-containing sRNAs were converted to the single-stranded cDNAs. These cDNAs were treated as the PCR templates in order to generate the double-stranded cDNAs, which were then purified by 16% TBE gel electrophoresis. Sequencing was performed by LC-Bio with the Illumina HiSeq 2500 platform following the vendor's recommended protocol. For sRNA-seq, there are three biological replicates for WT and *ago1-25*, respectively.

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) was used for adapter trimming and low-quality read removal. Then, the pre-treated sRNA sequences ranging from 15 to 45 nt were retained. For a dataset, the expression level of one sRNA was normalized by dividing the raw read count of this sRNA by the total raw read counts of this dataset and then multiplied by 10^6. The expression levels normalized in reads per million enabled cross-library comparisons.

"Omicstudio" (https://www.omicstudio.cn/) (Lyu et al., 2023) was used to draw the heatmaps. TargetFinder (Allen et al., 2005) was used to predict the microRNA binding sites (MBSs) under the cutoff "penalty score ≤ 6." The metagene profiles showing the distribution patterns of the MBSs in mRNAs were drawn by using Guitar (Cui, Wei, et al., 2016). To do this analysis, the mRNAs were aligned and divided into three regions (5′UTRs, CDSs, and 3′UTRs).

2.6 | IP-MS

The experiment was performed according to the previously reported procedure (Zhu et al., 2025). Briefly, the Co-IP assays were performed using 16-day-old transgenic seedlings over-expressed with *AGO1* and *TRN1*-3×FLAG, respectively. The seedlings were ground into a fine powder using liquid nitrogen and homogenized in lysis buffer (25 mM Tris–HCl pH 7.4, 25 mM KCl, 10 mM $MgCl_2$, 0.1% NP-40, and 2× complete protease inhibitor cocktail). The lysates were sonicated in an ice-water bath, followed by centrifugation (21,000 × *g*, 30 min, 4°C). The supernatants were collected and incubated with pre-equilibrated anti-Flag magnetic beads (Smart-Lifescience, B25081109) overnight at 4°C with constant rotation. The beads were washed three times with 500 μL of wash buffer (20 mM Tris–HCl pH 7.9, 0.5 M KCl, 10% glycerol, 1 mM EDTA, 5 mM DTT, 0.4% Nonidet P-40, and 0.2 mM phenylmethylsulfonyl fluoride) and eluted with 0.1 M glycine (pH = 2).

For protein digestion, the eluted samples were neutralized and precipitated with four volumes of ice-cold acetone at −20°C for 1 h, followed by centrifugation (12,000 × *g*, 10 min, 4°C). The protein pellets were air-dried and resuspended in a buffer containing 50 mM NH_4HCO_3, digested with trypsin (at a 50:1 protein-to-enzyme ratio) overnight at 37°C. Following reduction with 5 mM tris-carboxyethylphosphine and alkylation with 11 mM iodoacetamide for 1 h in the dark at room temperature, and digestion was terminated by adding 5% formic acid (FA). Finally, the peptides were desalted using StageTips, resuspended in 0.1% FA, and subjected to HPLC-MS/MS analysis.

Peptides were analyzed on a Q Exactive HF-X instrument (ThermoFisher Scientific) coupled to an Easy-nLC 1200 system (Zhu et al., 2025). Separation was performed with a reversed-phase C18 column (75 μm × 15 cm, 1.9 μm, 1 μm tip) using a 60 min gradient at 500 nL min^{-1} with 0.1% FA in water (solvent A) and 0.1% FA in 80% acetonitrile (solvent B): 6%–8% solvent B for 2 min, 8%–30% B for 37 min, 30%–40% B in 14 min, 40%–90% B for 3 min, and 90% B for 4 min. Full MS scans (400–1400 m/z) were acquired at 60,000 resolution with a 3×10^6 ion count target and a maximum injection time of 45 ms. Data-dependent acquisition was operated in top-speed mode (1 s cycle time). Precursors of +1, +6, +8, or above, or unassigned charge states, were rejected. The top precursors (intensity > 4×10^5) were isolated (1.6 m/z window) and fragmented by higher-energy collisional dissociation (27% NCE). MS/MS spectra were detected at 15,000 resolution (AGC 1×10^5, max injection time 22 ms), with dynamic exclusion set to 30 s (10 ppm tolerance). Peptides were identified using pFindStudio software. The protein sequences were downloaded from TAIR10.

3 | RESULTS

3.1 | The influence of the *ago1* mutation on transcriptome-wide m^5C modifications

To systematically dissect the potential regulatory role of AGO1 in m^5C modification, the hypomorphic *ago1-25* mutant (hereafter referred to as *ago1*) of *Arabidopsis* (Morel et al., 2002) was selected for m^5C-RIP-seq analysis. The homozygous mutation of *ago1-25* was validated by PCR amplification followed by Sanger sequencing. Besides, the WT plants of Col-0 were included for comparison. To ensure the statistical robustness of downstream analyses, three independent biological replicates were included for m^5C-RIP-seq of the 16-day-old seedlings of *ago1* and WT, respectively.

Based on the sequencing data, transcriptome-wide m^5C peak calling identified a total of 5981 and 4559 high-confidence m^5C peaks ($p < 0.01$) were identified in WT and *ago1*, respectively (Tables S1 and S2). The R package QNB (Liu et al., 2017) was applied to identify statistically significant differentially methylated peaks in the *ago1* mutant relative to WT. This workflow yielded 366 hyper-methylated peaks ($\log_2FC > 0.5$, $p < 0.01$) and 168 hypo-methylated peaks ($\log_2FC < -0.5$, $p < 0.01$) in *ago1* (Table S3). All identified m^5C peaks were categorized into four distinct groups for downstream analyses: (1) all peaks identified in WT; (2) all peaks identified in *ago1*; (3) the hyper-methylated peaks identified in *ago1*; and (4) the hypo-methylated peaks identified in *ago1*. Genome-wide distribution profiling revealed that m^5C peaks from all four groups were generally evenly distributed across the five *Arabidopsis* chromosomes, with notable depletion in centromeric and telomeric regions (Figure 1A–D). To further validate the reliability of these m^5C peaks, the computational tool Staem5 (Chai et al., 2021) was used to evaluate their association with the predicted single-nucleotide

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

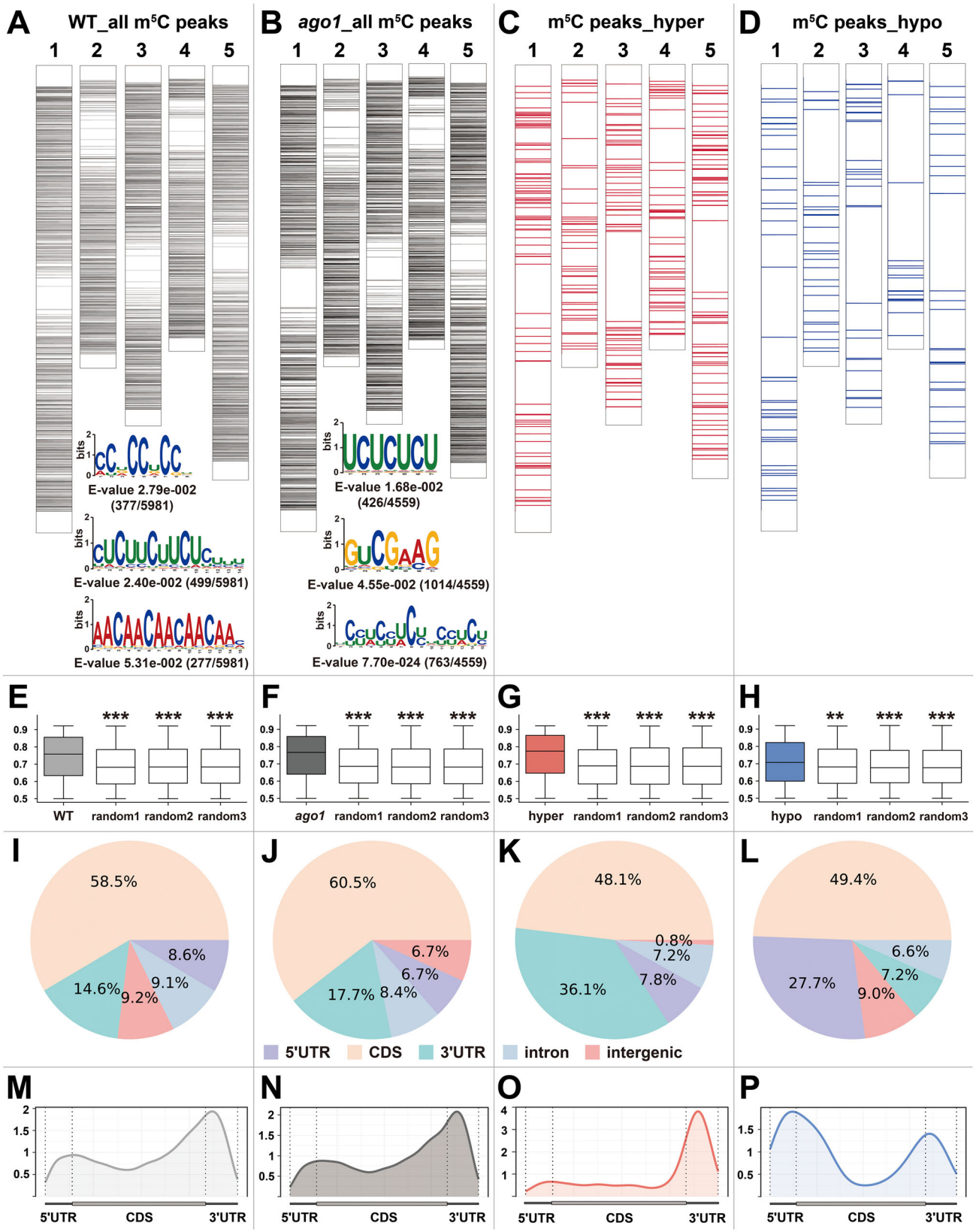


FIGURE 1 Transcriptome-wide identification of the m^{5}C peaks affected by the *ago1* mutation. (**A**) Genome-wide distribution pattern of all m^{5}C peaks identified from the wild-type (WT) seedlings of *Arabidopsis* ($p < 0.01$). The three illustrations present the enriched motifs identified from these m^{5}C peaks by using SEA (simple enrichment analysis), a functional module of the MEME Suite (5.5.7) (Bailey et al., 2015). (B) Genome-wide distribution pattern of all m^{5}C peaks identified from the seedlings of the *ago1* mutant ($p < 0.01$). The three illustrations present the enriched motifs

(Continues)

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 1 (Continued)
identified from these m^5C peaks by using SEA of the MEME Suite (5.5.7). (C) Genome-wide distribution pattern of the hyper-methylated peaks identified from *ago1* ($log_2FC > 0.5$, $p < 0.01$). (D) Genome-wide distribution pattern of the hypo-methylated peaks identified from *ago1* ($log_2FC < -0.5$, $p < 0.01$). Referring to the m^5C peaks from (A) to (D) and (E) to (H), they show the higher potential of these peaks for harboring the m^5C sites, as compared to the random sets (two-tailed *t*-test: $^{***}p < 0.001$ and $^{**}p < 0.01$). The *y*-axes indicate the m^5C prediction scores reported by Staem5 (Chai et al., 2021). Three random sets are included for each box plot. Referring to the m^5C peaks from (A) to (D), (I) to (L) show the statistics of their genomic locations, and (M) to (P) show the density of these m^5C peaks in mRNAs, which were aligned and divided into three regions (CDS, coding sequence; UTR, untranslated region) for drawing the metagene profiles by using Guitar (Cui, Wei, et al., 2016).

m^5C sites. Consistent with high peak quality, all four groups of m^5C peaks exhibited significantly higher predicted probabilities of containing the predicted m^5C sites compared to randomly shuffled genomic control regions (two-tailed *t*-test, $p < 0.01$) (Figure 1E–H).

Previous work identified several degenerated sequence motifs, including HACCR (H = A/C/U; R = A/G) and CUYCUYC (Y = C/U), as candidate *cis*-regulatory elements governing site-specific m^5C deposition in *Arabidopsis* (Cui et al., 2017). In this regard, all of the m^5C peaks identified in WT ($n = 5981$) and *ago1* ($n = 4559$) were independently subjected to the *de novo* motif enrichment analysis using the simple enrichment analysis module within the MEME Suite (5.5.7) (Bailey et al., 2015). This analysis recovered three significantly enriched, prominent sequence motifs in WT and *ago1*, respectively (Figure 1A,B). Notably, among the six total identified motifs, two from WT and two from *ago1* were CU-rich sequences that closely resembled the previously reported "CUYCUYC" consensus. This observation strongly suggests that CU-rich genomic regions represent preferential hotspots that either overlap with or lie in close proximity to bona fide m^5C modification sites. Additionally, the remaining WT-specific motif was highly enriched in AC dinucleotides, matching the previously characterized "HACCR" consensus sequence. However, this AC-rich motif was completely undetectable in the *ago1* mutant background, indicating that its enrichment was likely to be dependent on functional AGO1.

Next, we performed a comprehensive genomic feature annotation to map all m^5C peaks to distinct transcriptional regions, including UTRs, CDSs (also referred to as exonic regions), introns, and intergenic regions. Across all four peak categories, 48.1%–60.5% of the m^5C peaks were consistently mapped to CDS regions, while <10% of the peaks were annotated to introns and intergenic regions, respectively (Figure 1I–L). A strikingly distinct distribution pattern was observed for the peaks annotated to UTR regions. Specifically, 36.1% of the hyper-methylated peaks in the *ago1* mutant were mapped to 3′UTRs, whereas only 7.2%–17.7% of the peaks from the other three categories were localized to this region. Besides, 27.7% of the hypo-methylated peaks in *ago1* were annotated to 5′UTRs, whereas only 6.7%–8.6% of the peaks from the other three categories were localized to this region. Notably, the above annotation did not consider the substantial differences in the total sequence length of UTRs versus CDSs across the *Arabidopsis* transcriptome. To eliminate this length bias and further resolve the precise distribution landscape of the m^5C peaks, all mRNAs associated with the m^5C peaks were aligned to a normalized metagene structure consisting of three canonical regions: 5′UTR, CDS, and 3′UTR. Metagene profiles generated using the R package Guitar (Cui, Wei, et al., 2016) revealed highly similar global m^5C distribution patterns between WT and *ago1*, with a prominent enrichment peak localized to the 3′UTR region (Figure 1 M,N). However, the distribution of differentially methylated m^5C peaks in *ago1* showed dramatic deviations from this global pattern. Notably, the hyper-methylated peaks exhibited a far more pronounced, concentrated enrichment specifically within the 3′UTR region (Figure 1O). While the hypo-methylated peaks also showed moderate enrichment in the 3′UTR, they displayed a distinct, dominant enrichment signal that was largely restricted to the 5′UTR (Figure 1P). Collectively, these results demonstrate that among the m^5C modifications affected by the *ago1* mutation, the hyper-methylation events are predominantly localized to 3′UTRs, whereas the hypo-methylation events are strongly enriched in 5′UTRs.

3.2 | Structure and expression of the m^5C peak-associated genes

To systematically characterize the genomic context of m^5C deposition, we first annotated the 5981 m^5C peaks from WT and 4559 peaks from *ago1* to protein-coding genes in the TAIR10 (The *Arabidopsis* Information Resource, release 10) reference genome annotation (Reiser et al., 2024). As a result, 5559 peaks from WT were mapped to 4696 distinct genes, and 4325 peaks from *ago1* were mapped to 3716 distinct genes. Notably, 3040 genes are shared between the two datasets, the vast majority of which are protein-coding genes (Figure 2A). The substantial overlap strongly indicates that m^5C modifications are preferentially targeted to a conserved subset of protein-coding genes in the 16-day-old seedlings of *Arabidopsis*. Comparative analysis revealed that m^5C-associated genes in both WT and *ago1* contained significantly more exons than non-methylated genes (two-tailed *t*-test, $p < 0.001$) (Figure 2B), indicating that genes with more complex exon-intron architectures are preferentially targeted by the m^5C methylation machinery. However, no significant difference was observed between WT and *ago1*.

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

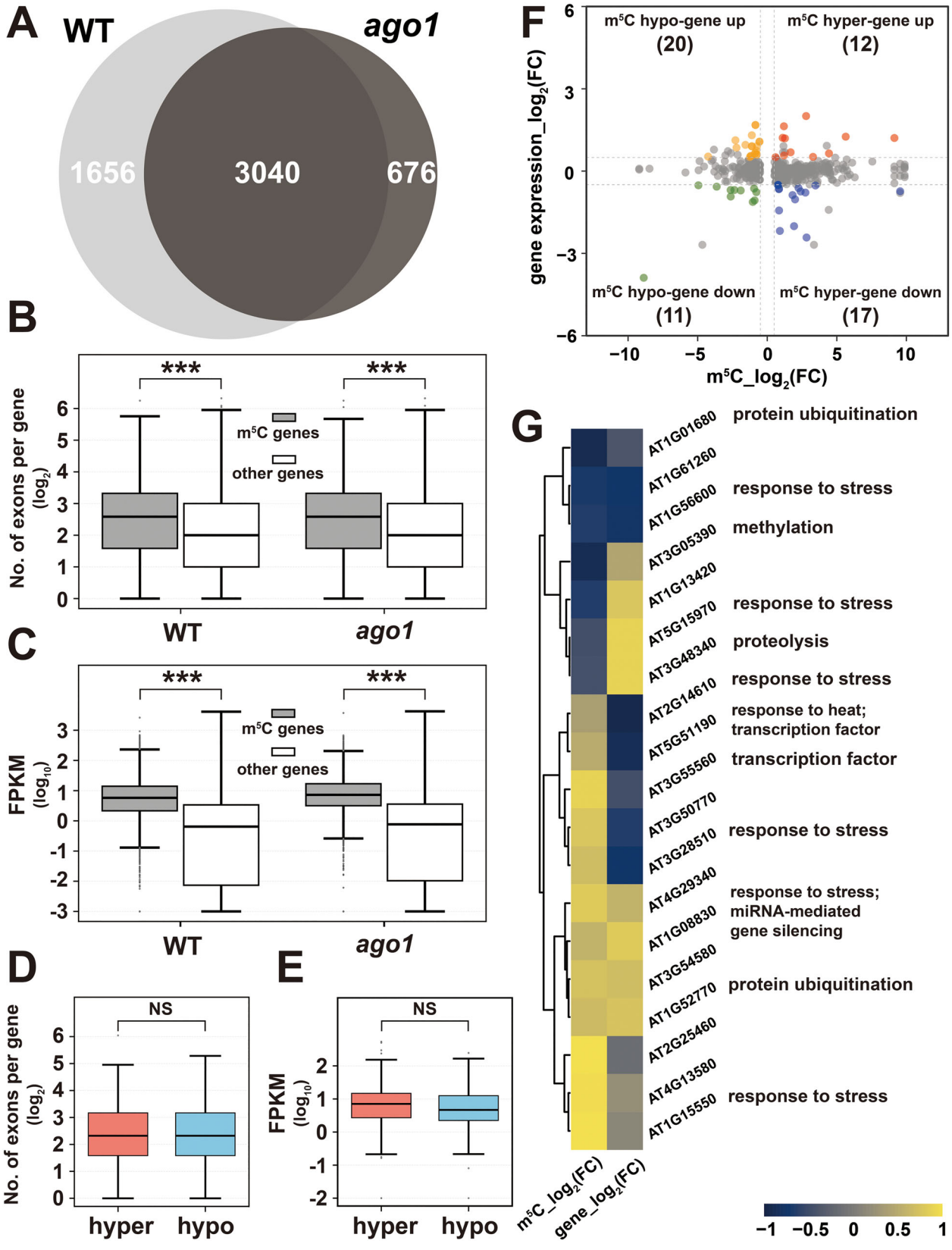


FIGURE 2 Structure and expression of the m^{5}C peak-associated genes identified from the wild-type (WT) plants and the *ago1* mutant of *Arabidopsis*. (A) Venn diagram showing the numbers of the m^{5}C peak-associated genes identified from WT and *ago1*, respectively. (B) Comparison of the gene structure complexity (number of exons per gene, measured by the *y*-axis, $\log_2$-scaled) between the m^{5}C peak-associated genes and the other non-m^{5}C genes in both WT and *ago1*. (C) Comparison of the gene expression levels (measured by the *y*-axis, in fragments per kilobase of exon

(Continues)

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 2 (Continued)
per million fragments mapped [FPKM], $\log_{10}$-scaled) between the m^{5}C peak-associated genes and the other non-m^{5}C genes in both WT and *ago1*. (D) Comparison of the gene structure complexity between the hyper-methylated and the hypo-methylated genes identified from *ago1*. (E) Comparison of the gene expression levels between the hyper-methylated and the hypo-methylated genes identified from *ago1*. (F) Four-quadrant chart showing the relationship between m^{5}C regulation and gene expression in *ago1*. The genes with altered m^{5}C modification levels ($|\log_2 FC| > 0.5$, $p < 0.01$) and altered expression levels ($|\log_2 FC| > 0.5$, FDR < 0.05) were marked by color (red, orange, blue, and green) dots. (G) Heatmap showing the differentially expressed genes ($|\log_2 FC| > 1$, FDR < 0.05) associated with differential m^{5}C peaks ($|\log_2 FC| > 0.5$, $p < 0.01$) identified from *ago1*. The heatmap was drawn based on the $\log_2$-scaled fold changes of m^{5}C modification and gene expression levels as compared to WT. According to the annotations, many genes are involved in transcriptional or post-transcriptional regulation, protein metabolism, or stress response. ***Two-tailed *t*-test, $p < 0.001$.

Using FPKM values quantified from the "input" data of m^{5}C-RIP-seq, we found that m^{5}C-associated genes exhibited significantly higher expression levels than non-methylated genes in both WT and *ago1* (two-tailed *t*-test, $p < 0.001$) (Figure 2C). Notably, a distinct discrepancy was observed between WT and the mutant. Specifically, the mean FPKM value of m^{5}C-modified genes in WT was 17.25, while the mean FPKM of its corresponding random control was 5.98. By comparison, the mean FPKM of m^{5}C-modified genes in *ago1* reached 20.77, with its matched random control exhibiting a mean FPKM of 5.92. These observations suggest that the m^{5}C-modified genes in *ago1* display either more active transcriptional status or higher transcript stability. We next performed the same comparison on genes linked to hyper-methylated and hypo-methylated m^{5}C peaks in the *ago1* mutant. However, no statistically significant differences in exon number or steady-state expression level were detected between hyper-methylated and hypo-methylated gene groups (Figure 2D,E).

To further dissect the functional impact of m^{5}C deposition on transcript abundance, we systematically integrated our m^{5}C peak datasets with transcriptomic profiles (the "input" datasets of m^{5}C-RIP-seq) from WT and *ago1*. We focused on the 534 differentially methylated peaks ($|\log_2 FC| > 0.5$, $p < 0.01$) identified in *ago1*, comprising 366 hyper-methylated and 168 hypo-methylated peaks, along with their corresponding host genes. We specifically intersected these host genes with the set of differentially expressed genes in *ago1* relative to WT ($|\log_2 FC| > 0.5$, FDR < 0.05). This overlap identified 12 hyper-methylated genes with up-regulated expression, 17 hyper-methylated genes with down-regulated expression, 20 hypo-methylated genes with up-regulated expression, and 11 hypo-methylated genes with down-regulated expression (Figure 2F). No significantly biased relationship was observed between m^{5}C methylation and gene expression, indicating that both hyper- and hypo-methylation could correlate with either up- or down-regulation of transcript levels. When applying a more stringent cutoff "$|\log_2 FC| > 1$, FDR < 0.05" for differentially expressed genes, a total of 19 high-confidence genes associated with differential m^{5}C peaks were recovered from *ago1*. Based on TAIR10 functional annotations, these genes are enriched in pathways related to transcriptional or post-transcriptional regulation, protein metabolism, or stress response (Figure 2G). Several representative loci were selected for graphic illustrations using the integrative genomics viewer software (J. T. Robinson et al., 2011). For example, a specific hyper-methylated m^{5}C peak was localized to the 5′UTR of transcript AT4G29340.1 in *ago1*. Based on the "input" data of m^{5}C-RIP-seq, the expression of the gene *AT4G29340* that was annotated to regulate cytoskeleton organization was up-regulated in the mutant (Figure 3A). A second example shows a broad hyper-methylated m^{5}C peak spanning the entire gene body of *AT3G54580*, a cell wall organization-related gene, whose transcript level was correspondingly up-regulated in *ago1* (Figure 3B). We also identified a hypo-methylated m^{5}C peak overlapped with the 5′UTR and the first exon of the transcript AT5G15970.1 in *ago1*. This cold acclimation and salt tolerance-related gene was concurrently up-regulated in the mutant (Figure 3C). At the 5′ end of transcript AT4G13580.1, a hyper-methylated m^{5}C peak was detected in *ago1*. This dirigent protein-encoding gene, a key regulator of lignin biosynthesis, was significantly up-regulated in the mutant (Figure 3D). Finally, a hyper-methylated m^{5}C peak was mapped to the first exon of transcript AT1G72210.1 in *ago1*. This gene, which encodes a basic helix-loop-helix transcription factor, showed an elevated transcript level in the mutant (Figure 3E). Taken together, these lines of evidence demonstrate that m^{5}C deposition can functionally modulate gene expression in a subset of target transcripts, and many of these m^{5}C-associated genes fulfill critical biological roles in *Arabidopsis* development and stress responses.

3.3 | Functional analysis of the m^{5}C peak-associated genes

The analyses above revealed that a subset of the m^{5}C-associated genes fulfilled essential biological functions in *Arabidopsis*. To systematically dissect their functional properties, we set out to perform a comprehensive survey on the biological roles of these m^{5}C-associated genes based on gene ontology (GO) terms (Ashburner et al., 2000) and Kyoto Encyclopedia of Genes and Genomes (KEGG)

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

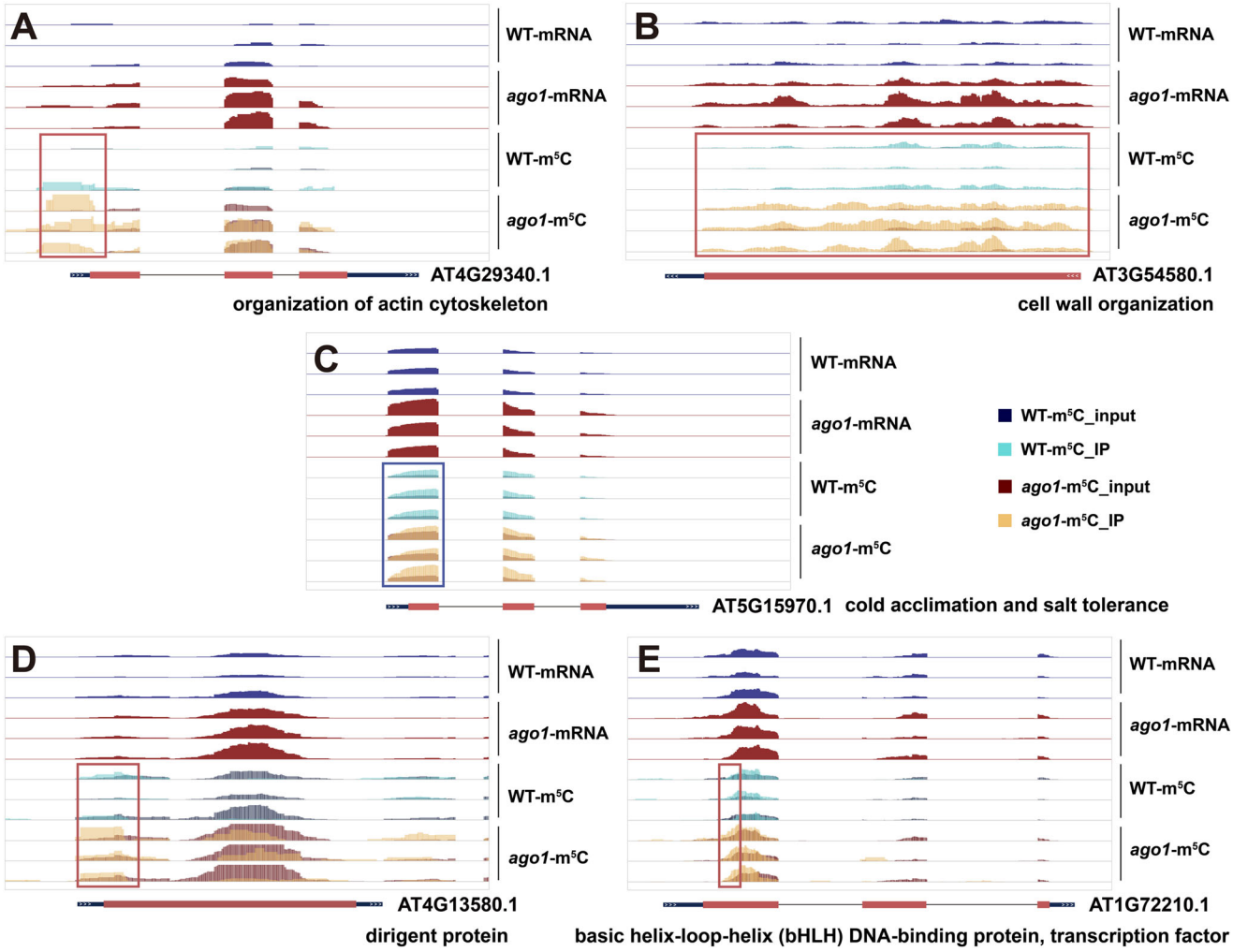


FIGURE 3 Graphic illustrations showing the effect of differential m^5C modifications on gene expression. From (A) to (E), five genes associated with the differential m^5C peaks identified from the *ago1* mutant were taken, for example. Based on the m^5C-RIP-seq data, the hyper- and hypo-methylated peaks were marked by the red and blue boxes, respectively. The gene expression comparison between the wild-type (WT) plants and *ago1* was performed based on the "input" data of m^5C-RIP-seq. The IGV (integrative genomics viewer) software (http://www.igv.org) (J. T. Robinson et al., 2011) was used for visualization.

pathways (Ogata et al., 1999). Consistent with the classification of the m^5C peaks, we stratified all m^5C-associated genes into four independent groups for downstream analyses. These four groups are defined as (1) genes associated with m^5C peaks identified in WT; (2) genes associated with m^5C peaks identified in *ago1*; (3) genes associated with the hyper-methylated peaks in *ago1*; and (4) genes associated with the hypo-methylated peaks in *ago1*.

First, we systematically annotated all m^5C peaks to four distinct gene categories: protein-coding genes, pseudogenes, transposable element genes, and non-coding genes (including tRNAs, miRNA precursors, and other non-coding RNAs). As a result, 4626 protein-coding genes, seven tRNAs, six pre-miRNAs (precursor miRNAs), and 1028 other ncRNAs (non-coding RNAs) were associated with the m^5C peaks identified from WT. A total of 3662 protein-coding genes, five tRNAs, four pre-miRNAs, and 862 other ncRNAs were associated with the m^5C peaks identified from *ago1*. Notably, no m^5C peaks were annotated to pseudogenes or transposable element genes in either genotype (Figure 4A,B). For both protein-coding and non-coding gene groups, the vast majority of loci harbored only a single m^5C peak, while a small subset of genes contained multiple distinct m^5C peaks. Separately, we found that in the *ago1* mutant, 317 protein-coding genes and nine ncRNAs were associated with hyper-methylated m^5C peaks (Figure 4C), and 141 protein-coding genes and six ncRNAs were linked to hypo-methylated peaks (Figure 4D). Importantly, the small set of m^5C peaks located on tRNAs and pre-miRNAs showed no significant changes in methylation level in the *ago1* mutant compared to WT.

Next, functional analysis was performed for the m^5C peak-associated genes based on GO and KEGG annotations.

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

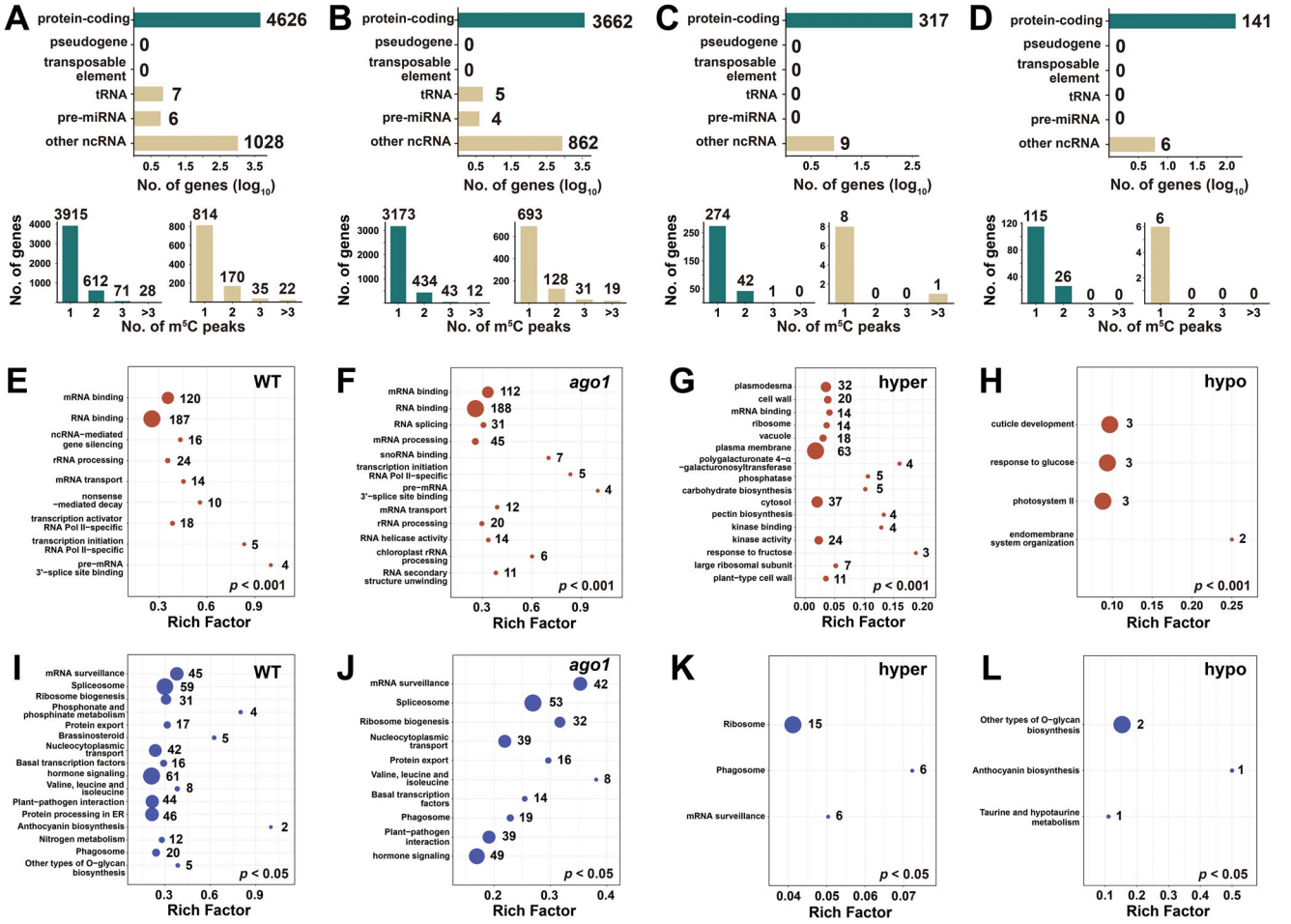


FIGURE 4 Functional annotations of the m^{5}C peak-associated genes identified from the wild-type (WT) plants and the *ago1* mutant of *Arabidopsis*. (A) Numbers of the m^{5}C peak-associated genes assigned to different categories in WT. The two histograms below show the numbers of the protein-coding and the non-coding genes associated with different numbers of the m^{5}C peaks, respectively. (B) Numbers of the m^{5}C peak-associated genes assigned to different categories in *ago1*. The two histograms below show the numbers of the protein-coding and the non-coding genes associated with different numbers of the m^{5}C peaks, respectively. (C) Numbers of the hyper-methylated genes assigned to different categories in *ago1*. The two histograms below show the numbers of the protein-coding and the non-coding genes associated with different numbers of the hyper-methylated peaks, respectively. (D) Numbers of the hypo-methylated genes assigned to different categories in *ago1*. The two histograms below show the numbers of the protein-coding and the non-coding genes associated with different numbers of the hypo-methylated peaks, respectively. (E) GO (gene ontology) term enrichment analysis of the m^{5}C peak-associated genes in WT. (F) GO term enrichment analysis of the m^{5}C peak-associated genes in *ago1*. (G) GO term enrichment analysis of the hyper-methylated genes in *ago1*. (H) GO term enrichment analysis of the hypo-methylated genes in *ago1*. (I) KEGG (Kyoto Encyclopedia of Genes and Genomes) term enrichment analysis of the m^{5}C peak-associated genes in WT. (J) KEGG term enrichment analysis of the m^{5}C peak-associated genes in *ago1*. (K) KEGG term enrichment analysis of the hyper-methylated genes in *ago1*. (L) KEGG term enrichment analysis of the hypo-methylated genes in *ago1*

GO enrichment results showed that in both WT and *ago1*, a large fraction of the m^{5}C peak-associated genes were significantly enriched in core RNA-related functional categories, including RNA binding, processing, splicing, transportation, and metabolism (Figure 4E,F). However, further comparative analysis revealed distinctively enriched GO terms between the m^{5}C-modified genes identified in WT and *ago1*. For example, the enriched GO terms "RNA splicing" and "RNA secondary structure unwinding" were specifically detected in *ago1*. Notably, 14 genes linked to the hyper-methylated m^{5}C peaks in *ago1* were functionally related to mRNA binding or ribosome (Figure 4G). Consistent with the GO results, KEGG pathway analysis revealed that the m^{5}C-associated genes from both WT and *ago1* were significantly overrepresented in two core post-transcriptional regulatory pathways: the mRNA surveillance pathway and ribosome biogenesis (Figure 4I,J). For the hyper-methylated gene subset in *ago1*, 15 genes were functionally related to ribosome, and six genes were assigned to the mRNA surveillance pathway, respectively (Figure 4K). The 15 hyper-methylated ribosome-related genes could be further divided into two distinct subgroups: nine genes (*AT1G74050*, *AT1G33120*,

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

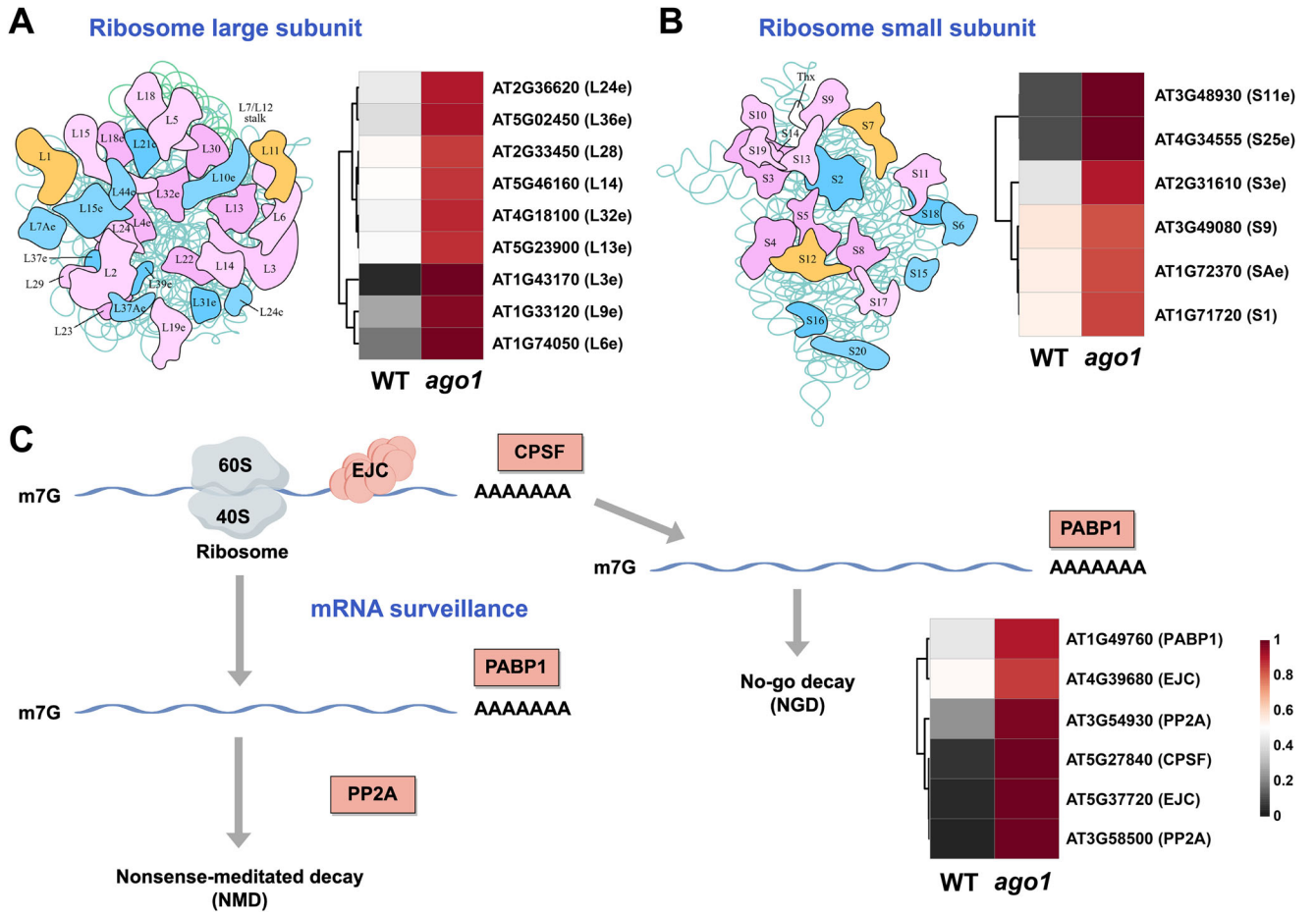


FIGURE 5 The genes involved in translation and mRNA surveillance are hyper-methylated in *ago1*. (A) Nine genes encoding different ribosome large subunits are hyper-methylated in *ago1*. (B) Six genes encoding different ribosome small subunits are hyper-methylated in *ago1*. For both (A) and (B), the model diagrams of the ribosome large and small subunits were retrieved from KEGG (Kyoto Encyclopedia of Genes and Genomes) at https://www.kegg.jp/. (C) Six genes implicated in the mRNA surveillance pathway are hyper-methylated in *ago1*. The pathway was referenced from KEGG and drawn by using FigDraw (https://www.figdraw.com/static/index.html#/). CPSF, cleavage and polyadenylation specificity factor; EJC, exon junction complex; PABP, polyadenylate-binding protein; PP2A, protein phosphatase 2A. Based on the raw data of m^5C-RIP-seq, the heatmaps were drawn for the methylation level comparison between the wild-type (WT) plants and the *ago1* mutant. The hyper-methylated peaks were identified under the cutoff "$\log_2FC > 0.5$, $p < 0.01$."

AT1G43170, *AT5G23900*, *AT4G18100*, *AT5G46160*, *AT2G33450*, *AT5G02450*, and *AT2G36620*) encoding different large ribosomal subunit proteins (Figure 5A), and six genes (*AT1G71720*, *AT1G72370*, *AT3G49080*, *AT2G31610*, *AT4G34555*, and *AT3G48930*) encoding different small ribosomal subunit proteins (Figure 5B). Notably, all genes in both subgroups showed consistent hyper-methylation at their m^5C peaks in *ago1* compared to WT. However, transcriptome quantification showed that none of these ribosomal genes exhibited statistically significant changes in steady-state mRNA levels in *ago1* relative to WT (FDR > 0.1) (Table S3). On the other hand, another six hyper-methylated genes in *ago1* were mapped to the mRNA surveillance pathway, including *AT5G37720* and *AT4G39680* encoding the components of exon junction complex, *AT3G54930* and *AT3G58500* encoding protein phosphatase 2A, *AT1G49760* encoding a polyadenylate-binding protein, and *AT5G27840* encoding a component of cleavage and polyadenylation specificity factor (Figure 5C). Similarly, none of these mRNA surveillance pathway genes showed significant differential expression in *ago1* (Table S3). Although the differential m^5C modifications on the above two gene sets did not lead to detectable changes in their steady-state transcript levels, we hypothesized that these hyper-methylation marks might be specifically recognized by unknown m^5C "readers" to exert downstream post-transcriptional regulatory effects independent of mRNA abundance. For example, the recognition of the m^5C signals by specific "readers" may modulate the translational efficiency of the target transcripts.

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.4 | Potential involvement of miRNAs in the AGO1-mediated site-specific m^{5}C regulation

Our results above demonstrate that the *ago1* loss-of-function mutation induces widespread, quantitative alterations to the m^{5}C methylome in *Arabidopsis*. Given the well-established, central role of AGO1 in the canonical plant miRNA effector pathway (Jones-Rhoades et al., 2006; Vaucheret, 2008), we hypothesized that miRNAs themselves may directly contribute to the site-specific deposition of m^{5}C. To test this, we used the TargetFinder algorithm (Allen et al., 2005) to predict high-confidence miRNA binding sites (MBSs) that either overlapped with, or were located within 500 nucleotides of, the differential m^{5}C peaks in *ago1*, using a strict penalty score cutoff ≤ 6. This analysis revealed that 217 out of 366 hyper-methylated peaks (59.29%) and 93 out of 168 hypo-methylated peaks (55.36%) in *ago1* were directly associated with at least one predicted MBSs (Table S4). Specifically, a total of 873 MBSs were predicted to be associated with the differential m^{5}C peaks. Among these, 424 MBSs associated with the hyper-methylated peaks were mapped to 3′UTRs, and 155 MBSs associated with the hypo-methylated peaks were mapped to 5′UTRs. The observation that much more MBSs are enriched in 3′UTRs as compared to 5′UTRs may contribute to the distinct influence of the *ago1* mutation on m^{5}C modifications in the two UTR regions, which needs further investigations. Several examples of the MBSs associated with the differential m^{5}C peaks were provided here. The MBS of ath-miR773a fully overlaps with a hyper-methylated m^{5}C peak located on its target transcript AT1G19190.1 (Figure 6A). A second example shows that the predicted MBS for ath-miR5654-3p lies immediately adjacent to a hyper-methylated m^{5}C peak on the target transcript AT2G47170.1 (Figure 6B). Conversely, the MBS of ath-miR167a-5p is positioned near a hypo-methylated m^{5}C peak on its target transcript AT2G17250.1 (Figure 6C). We also identified an MBS for ath-miR393a-5p that overlaps with a hypo-methylated m^{5}C peak on transcript AT2G05940.1 (Figure 6D). Finally, the predicted MBS of ath-miR403-3p is located directly adjacent to a hyper-methylated m^{5}C peak on transcript AT1G52770.1 (Figure 6E). Consistent with this spatial association, our sRNA-seq data of the 16-day-old seedlings of *ago1-25* and WT shows that ath-miR5654-3p, ath-miR167a-5p, ath-miR393a-5p, and ath-miR403-3p are all significantly down-regulated in the *ago1-25* mutant. Based on our sRNA-seq data of *ago1-27*, another hypomorphic mutant, these miRNAs except for ath-miR167a-5p are also down-regulated. Besides, the sRNA-seq datasets reported by a previous study (Montgomery et al., 2008) were also used to validate the expression patterns of the above four miRNAs. Although the public sequencing libraries were prepared from flower tissue, ath-miR167a-5p and ath-miR403-3p were observed to be significantly repressed in *ago1-25*. Collectively, the results of miRNA expression analysis provide independent experimental evidence linking miRNA abundance to the observed changes in m^{5}C deposition.

To further validate the spatial association between m^{5}C peaks and predicted MBSs, we analyzed the full set of 5981 m^{5}C peaks identified in WT based on our own dataset. Filtering for peaks mapped to annotated protein-coding genes yielded 5559 non-redundant m^{5}C peaks with defined host genes based on the TAIR10 annotation, and these peaks were carried forward for downstream analysis. We ran TargetFinder on these peak-hosting genes using the cutoff "penalty score ≤ 6", and retained the MBSs that either overlapped directly with the m^{5}C peak or resided within a 500-nt flanking window. This analysis revealed that 3351 out of the 5559 m^{5}C peaks (60.28%) in WT were linked to at least one predicted MBS. Among these MBS-associated peaks, 35.5% had the MBS fully overlapping the m^{5}C peak, 33.5% had the MBS within 200 nt of the peak, and the remaining 31.0% had the MBS located 200–500 nt away from the peak (Figure 6F). To confirm this pattern in an independent published dataset, we re-analyzed the public m^{5}C methylome from 9-day-old *Arabidopsis* seedlings generated by a previous study (Cui et al., 2017). From this dataset, 5989 m^{5}C peaks mapped to TAIR10-annotated genes were included in the identical MBS prediction pipeline. Consistently, this re-analysis showed that 3696 out of the 5989 m^{5}C peaks (61.71%) were also associated with predicted MBSs. Among these MBS-associated peaks, 31.0% had the MBS fully overlapping the m^{5}C peak, 34.1% had the MBS within 200 nt of the peak, and the remaining 35.0% had the MBS located 200 to 500 nt away from the peak (Figure 6F). To assess the statistical significance of the overlap between two finite gene sets, we applied the hypergeometric distribution test, a standard probabilistic framework widely utilized in omics-based investigations. The total number of protein-coding genes annotated in the TAIR10 *Arabidopsis* reference genome is 28775. Under the null hypothesis of random sampling, the expected overlap between the two gene sets—2917 genes harboring m^{5}C-associated MBSs identified in 16-day-old WT seedlings and 2960 genes with m^{5}C-associated MBSs derived from 9-day-old WT seedlings—is calculated to be approximately 300. Despite the difference in developmental stages between the two samples, the observed intersection of these two sets reaches 1182 genes, which represents a statistically highly significant enrichment far exceeding the random expectation. Next, the distribution patterns of the MBSs associated with the m^{5}C peaks were analyzed. In both our dataset and the published dataset, MBSs were significantly enriched in the 3′UTRs (Figure 6G), a pattern that closely mirrors the 3′UTR-preferential distribution of m^{5}C peaks we observed in WT (Figure 1M). Collectively, these independent lines of evidence from our own dataset and a public methylome strongly support a conserved, genome-wide spatial

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

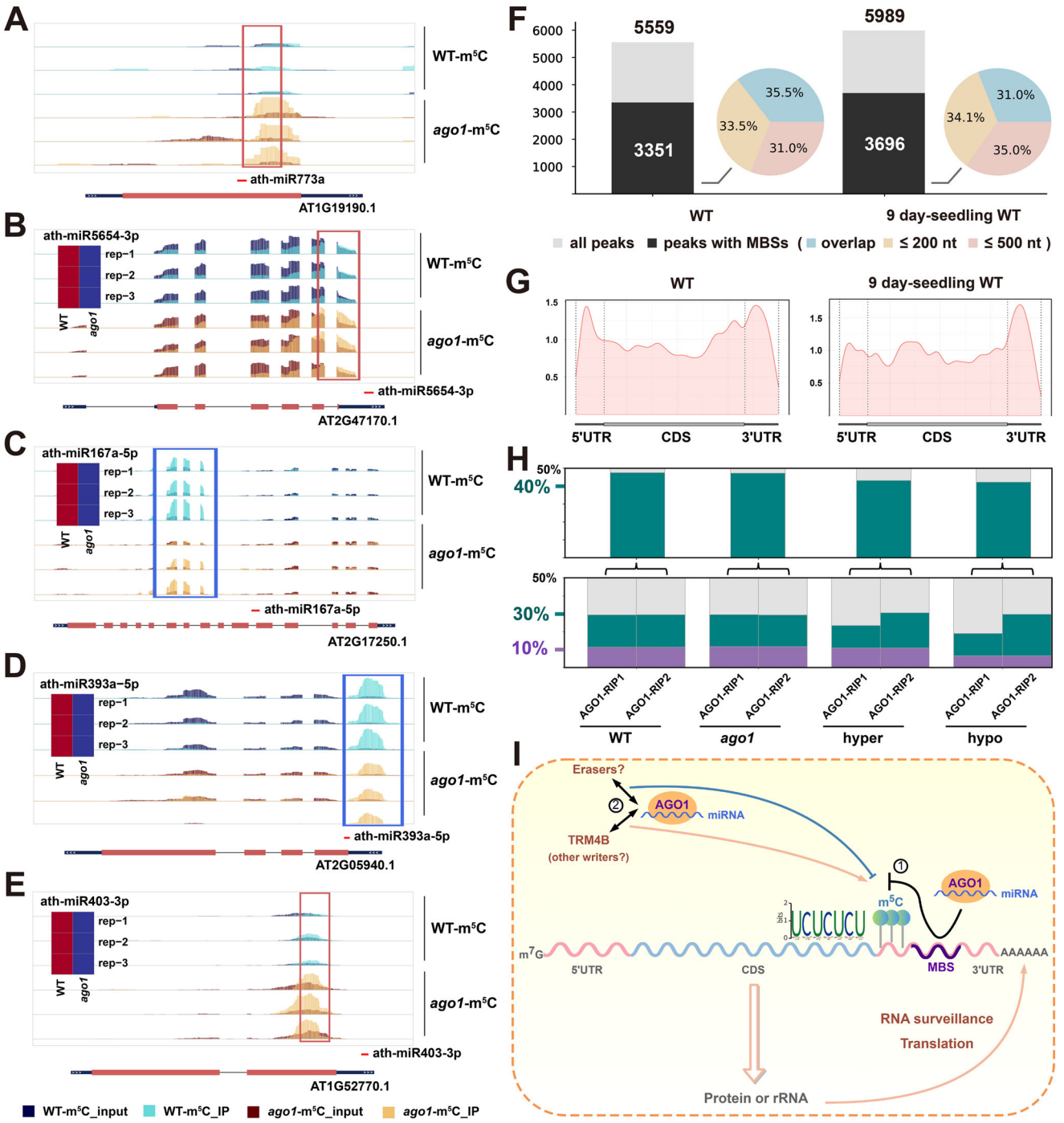


FIGURE 6 Model proposed for the site-specific m^{5}C regulation mediated by the microRNA (miRNA)-guided AGO1 protein complex. (A) to (E) Examples of the differential m^{5}C peaks associated with microRNA binding sites (MBSs). The IGV (integrative genomics viewer) software (http://www.igv.org) (J. T. Robinson et al., 2011) was used for visualization. Based on the m^{5}C-RIP-seq data, the hyper- and hypo-methylated peaks identified from *ago1* were marked by the red and blue boxes, respectively. The MBSs overlapping with or neighboring to the differential peaks were marked by short red lines. From (B) to (E), the differential expression patterns of the four miRNAs are shown by the heatmaps. For m^{5}C-RIP-seq, there are three biological replicates for the wild-type plants (WT) and the *ago1-25* mutant, respectively. (F) Prevalent association of the m^{5}C peaks with the MBSs. Based on the present study (WT) and the previous study (9-day-seedling WT) (Cui et al., 2017), a total of 5559 and 5989 m^{5}C peaks were analyzed, respectively. Among these, 3351 and 3696 peaks are associated with the MBSs. The pie charts show the percentages of the MBS-associated m^{5}C peaks assigned to different categories: peaks containing the MBSs (overlap); peaks with the neighboring MBSs ≤ 200 nt; and peaks with the neighboring MBSs ≤ 500 nt but > 200 nt. The MBSs were predicted by using TargetFinder (Allen et al., 2005) under the cutoff "penalty score ≤ 6." (G) Distribution patterns of the MBSs associated with the m^{5}C peaks in mRNAs. The metagene profiles were drawn by using Guitar (Cui, Wei, et al., 2016). The mRNAs were aligned and divided into three regions (CDS, coding sequence; UTR, untranslated region). (H)

(Continues)

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

FIGURE 6 (Continued)
RIP-seq data-based evidence showing the direct binding of the AGO1 protein onto the transcripts associated with the m^5C peaks. Four categories of the m^5C peaks were analyzed separately: 5981 peaks identified from WT; 4559 peaks identified from *ago1*; 366 hyper-methylated peaks identified from *ago1*; and 168 hypo-methylated peaks identified from *ago1*. Two biological replicates were included for the RIP-seq experiment of AGO1. The *y* axes measure the percentages of the m^5C peaks overlapping with the AGO1-bound transcripts. The lower histogram shows that around 30% of the m^5C peaks are associated with the AGO1-bound transcripts based on each replicate of the RIP-seq experiment. Among these peaks, around 10% are supported by both replicates. The upper histogram shows that >40% of the m^5C peaks are associated with the AGO1-bound transcripts, as supported by either of the two replicates. (I) The miRNA-guided AGO1 complex is potentially involved in the site-specific regulation of m^5C modifications. In this model, two possible pathways are proposed. First, the binding of the AGO1 complex onto the target transcripts blocks the action of the m^5C machinery on the neighboring m^5C site. Second, by association with the m^5C machinery, the AGO1 complex may facilitate the action of the m^5C machinery on the specific m^5C site neighboring to its binding region. For both pathways, the m^5C machinery may involve "writers" such as tRNA-specific methytransferase 4B (TRM4B) or the unidentified "erasers." In this model, the miRNAs play a critical role in guiding the AGO1 complex to the specific region based on the complementary microRNA binding sites (MBSs), which are preferentially enriched in 3′UTRs. Functional analysis indicated that some of the m^5C modifications regulated by AGO1 were likely to be implicated in the post-transcriptional pathways such as RNA surveillance or translation. The model was drawn by using FigDraw (https://www.figdraw.com/static/index.html#/).

association between m^5C deposition and miRNA binding sites in *Arabidopsis*.

Our results demonstrate that the *ago1* mutation drives widespread differential m^5C methylation across the *Arabidopsis* transcriptome and strongly suggest a mechanistic link between the miRNA-guided AGO1 effector complex and site-specific m^5C deposition. To obtain direct, biochemical evidence that AGO1 physically associates with these m^5C-modified transcripts, we performed RIP-seq (RNA immunoprecipitation followed by deep sequencing) of the AGO1 protein using the AGO1-specific antibody. Two biological replicates of the RIP-seq experiment were performed for the 2-week-old WT seedlings of Col-0. A total of 2890 and 4370 genes were identified as harboring significant AGO1 binding signals in the two biological replicates, respectively. Notably, only 797 genes were shared by the two replicates. The relatively low reproducibility between the two biological replicates of AGO1-RIP-seq can be attributed to the highly dynamic nature of AGO1-binding transcripts and the inherent instability of these targets, since many of them are miRNA targets. Then, we systematically overlapped the full set of AGO1-bound transcripts identified from RIP-seq with our four pre-defined categories of m^5C peaks. These four groups included 5559 m^5C peaks mapped to TAIR10-annotated genes in WT; 4295 peaks mapped to TAIR10-annotated genes in *ago1*; 366 hyper-methylated peaks in *ago1*; and 168 hypo-methylated peaks in *ago1*. In the WT dataset, 29.43% and 29.54% of all m^5C peaks were located on AGO1-bound transcripts in the two RIP-seq replicates, respectively. On the other hand, 11.39% of all m^5C peaks in WT are supported by the two RIP-seq replicates, and the ratio is increased to 47.58% when considering the peaks supported by at least one RIP-seq replicate (Figure 6H; Table S5). In the *ago1* mutant dataset, 29.59% and 29.38% of all m^5C peaks were located on AGO1-bound transcripts in the two RIP-seq replicates, respectively. On the other hand, 11.69% of all m^5C peaks in *ago1* are supported by the two RIP-seq replicates, and the ratio is increased to 47.29% when considering the peaks supported by at least one RIP-seq replicate (Figure 6H; Table S5). It is not surprising that relatively consistent observations were obtained between WT and *ago1*. The *ago1-25* mutant is not a complete loss-of-function allele. Consequently, the AGO1 protein in this mutant does not fully lose its capacity to bind target transcripts; instead, the binding affinity between AGO1 and its target transcripts is attenuated in most scenarios. Actually, in the present study, we reported only 366 hyper-methylated peaks ($\log_2FC > 0.5$, $p < 0.01$) and 168 hypo-methylated peaks ($\log_2FC < -0.5$, $p < 0.01$) were discovered from *ago1-25* (Table S3). This suggests that the *ago1* mutation does not affect m^5C modifications across the entire transcriptome but exerts regulatory effects on the m^5C modification of specific gene transcripts, reminiscent of the well-documented specificity of miRNA-mediated target gene regulation. By focusing on the 366 hyper-methylated m^5C peaks in *ago1*, 23.50% and 30.60% were mapped to AGO1-bound transcripts in the two RIP-seq replicates, respectively. On the other hand, 10.9% of the 366 hyper-methylated peaks in *ago1* are supported by the two RIP-seq replicates, and the ratio is increased to 43.17% when considering the peaks supported by at least one RIP-seq replicate (Figure 6H; Table S5). For the 168 hypo-methylated peaks in *ago1*, 19.05% and 29.76% were mapped to AGO1-bound transcripts in the two RIP-seq replicates, respectively. On the other hand, 6.55% of the 168 hypo-methylated peaks in *ago1* are supported by the two RIP-seq replicates, and the ratio is increased to 42.26% when considering the peaks supported by at least one RIP-seq replicate (Figure 6H; Table S5). To assess the significance of the overlap between hyper-methylated genes and AGO1-binding genes, we performed a one-tailed hypergeometric test. Among the 321 hyper-methylated genes identified in *ago1*, 139 were annotated as AGO1-binding targets. Using the 28,775 genes annotated in TAIR10, where 6463 genes carry AGO1-RIP signals, our observed overlap is ~1.93-fold higher than the random expectation of 72 genes,

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

representing a statistically highly significant enrichment ($p = 1.1 \times 10^{-17}$). A one-tailed hypergeometric test also confirmed that the 63 AGO1-binding targets among 142 hypo-methylated genes represent a ~1.98-fold significant enrichment over the random expectation of 32 genes ($p = 1.2 \times 10^{-9}$). Collectively, these RIP-seq results provide robust biochemical evidence that many m^{5}C-modified transcripts are direct, in vivo targets of the AGO1 protein in *Arabidopsis*.

Integrating our multi-omic evidence from m^{5}C methylome profiling, MBS co-localization analysis, and AGO1 RIP-seq, we proposed a working model in which the miRNA-guided AGO1 effector complex acts as a critical upstream regulator of site-specific m^{5}C deposition in *Arabidopsis*. This model posits two distinct, non-mutually exclusive mechanisms by which AGO1 can shape the m^{5}C landscape on its target transcripts. First, occupancy of the AGO1 complex at its target MBS creates a local steric block that prevents the m^{5}C machinery from accessing and modifying the adjacent m^{5}C site. Second, AGO1 can act as a molecular scaffold: by interacting with the m^{5}C enzymatic machinery, it actively recruits m^{5}C writers or erasers to deposit or remove m^{5}C at specific sites proximal to its own binding region. To date, the m^{5}C methyltransferase TRM4B remains the only well-characterized component of the plant m^{5}C enzymatic machinery that specifically mediates m^{5}C deposition on mRNA transcripts. We therefore hypothesized that TRM4B could be a direct or indirect interacting partner of the AGO1 complex in vivo. To test this, we generated two independent transgenic plants of *Arabidopsis*: one over-expressing *AGO1* and the other over-expressing FLAG-tagged *TRN1*, a previously validated AGO1 interactor with established roles in the miRNA pathway (Cui, Fang, et al., 2016). Using these two transgenic plants, we performed IP-MS assays using AGO1-specific and FLAG-specific antibodies, respectively. While no direct physical interaction between AGO1 and TRM4B was detected in our IP-MS data, we identified robust, high-confidence interaction signals between TRN1 and TRM4B (Figure S1). On the other hand, based on the "input" data of m^{5}C-RIP-seq, no significant difference in transcript abundance of *TRM4B* was detected between WT and the *ago1* mutant (Table S6). This observation rules out the possibility that altered miRNA activity in the *ago1* mutant indirectly influences m^{5}C deposition by changing the steady-state expression level of *TRM4B*. These results provide experimental support for the second mechanism of our proposed working model, where AGO1 can bridge to the m^{5}C machinery via intermediate protein partner TRN1. Besides, it is notable that TRN1 has been reported to play a critical role in promoting the association of miRNA with AGO1 in *Arabidopsis* (Cui, Fang, et al., 2016). However, for both proposed mechanisms, the m^{5}C enzymatic machinery includes the known m^{5}C writer TRM4B, alongside additional currently uncharacterized m^{5}C writers and erasers that represent promising targets for future investigation. Besides, miRNAs serve as the primary sequence-specific guiders that direct the AGO1 complex to its target regions via sequence complementarity to MBSs, which we show are preferentially enriched in the 3′UTRs of *Arabidopsis* transcripts. Moreover, our functional enrichment analyses indicate that a subset of these AGO1-dependent m^{5}C marks are functionally linked to core post-transcriptional processes, including the mRNA surveillance pathway and translational regulation (Figure 6I).

4 | DISCUSSION

The discovery of m^{5}C modifications started from the tRNA-specific methyltransferases DNMT2 and TRM4B (A. L. Burgess et al., 2015; Goll et al., 2006). Here, we performed high-resolution whole-transcriptome m^{5}C-RIP-seq and identified thousands of high-confidence m^{5}C peaks in both WT and the *ago1* mutant (Tables S1 and S2). While a fraction of these peaks mapped to non-coding genomic regions (intergenic regions and introns), the vast majority were localized to annotated protein-coding genes (Figures 1I,J and 4A,B), demonstrating widespread mRNA m^{5}C modification across the *Arabidopsis* transcriptome. Motif analysis of these m^{5}C peaks identified two dominant sequence contexts: CU-rich and AC-rich motifs, which closely resemble the previously reported HACCR and CUYCUYC consensus sequences in *Arabidopsis* (Cui et al., 2017).

Comparative methylome analysis identified hundreds of statistically significant differential m^{5}C peaks between WT and the *ago1* mutant (Table S3), directly implicating AGO1 as a key regulator of transcriptome-wide m^{5}C deposition. Peak distribution pattern analysis uncovered a striking positional difference between hyper-methylated and hypo-methylated m^{5}C peaks. Hyper-methylated peaks in *ago1* were strongly and specifically enriched in 3′UTRs. In contrast, while hypo-methylated peaks also showed moderate enrichment in 3′UTRs, they exhibited a far more pronounced enrichment in 5′UTRs (Figure 1O,P). The molecular mechanism driving this distinct positional partitioning of differential m^{5}C peaks remains to be fully elucidated in future work. Notably, m^{5}C peaks mapped to tRNAs and pre-miRNAs showed no significant changes in the *ago1* mutant, indicating that AGO1 preferentially regulates m^{5}C deposition on mRNA transcripts rather than on non-coding RNAs.

While both up-regulation and down-regulation of gene expression can be correlated with hyper-methylation or hypo-methylation (Figure 2F), highly transcribed genes are significantly enriched for m^{5}C peaks across the transcriptome (Figure 2C). We propose two non-exclusive mechanistic explanations for this observed pattern. First, as higher substrate availability generally accelerates enzymatic reactions, the elevated steady-state levels of mRNAs could directly

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

promote more efficient targeting and catalysis by the m^{5}C writer complex. Second, m^{5}C deposition itself can act as a functional mark that promotes transcript stability, creating a positive correlation between m^{5}C occupancy and mRNA abundance. Besides, our functional analysis reveals that a large fraction of hyper-methylated genes in the *ago1* mutant are functionally annotated to the mRNA surveillance pathway or ribosome biogenesis (Figure 5C). This suggests that the modulatory effects of m^{5}C on transcript stability and translation can be functionally amplified through these core regulatory pathways.

We performed independent MBS prediction on m^{5}C peak-hosting transcripts derived from both our WT dataset and the previously published *Arabidopsis* m^{5}C methylome. Across both datasets, over 60% of all m^{5}C peaks are associated with the MBSs with distance $\leq$ 500 nt (Figure 6F). It is predictable that this association percentage will further increase if a more permissive distance cutoff is adopted. Indeed, some distal MBSs can remain functionally active via long-range RNA secondary structure interactions within the same transcript. Besides, both m^{5}C peaks and their linked MBSs exhibited a strong enrichment in 3′UTR. In the *ago1* mutant, 59.29% of the hyper-methylated peaks and 55.36% of the hypo-methylated peaks were also found to be associated with predicted MBSs (Table S4). Finally, our AGO-RIP-seq experiment provided direct biochemical evidence showing the direct binding signals of AGO1 on the m^{5}C-associated transcripts (Figure 6H; Table S5). It has been well-known that the *Arabidopsis* AGO1 protein requires small RNAs for its binding to target transcripts (Fang & Qi, 2016; Tolia & Joshua-Tor, 2007). These highly consistent lines of evidence led us to hypothesize that AGO1 is recruited to specific transcript regions via miRNA guidance to mediate site-specific m^{5}C methylation. However, it remains an open question whether AGO1 can modulate m^{5}C methylation via a miRNA-independent regulatory pathway.

Several key points from our results merit further clarification and discussion. First, the total number of the differential m^{5}C peaks identified in the *ago1-25* mutant is substantially lower than the number of the MBS-associated m^{5}C peaks or the number of peaks with AGO1 binding signals. We propose two non-mutually exclusive explanations for this observation: (1) *ago1-25* is one of the hypomorphic *ago1* mutants of *Arabidopsis*, in which AGO1 catalytic activity is not completely abolished. (2) The site-specific regulation of AGO1 guided by the miRNAs is also constrained by the spatiotemporal expression patterns of these miRNAs. Consistent with this, both the m^{5}C writer TRM4B and global m^{5}C deposition have been demonstrated to have tissue-specific distribution patterns in *Arabidopsis* (Cui et al., 2017; David et al., 2017). Given these constraints, it is predictable that a larger set of hyper-methylated and hypo-methylated m^{5}C peaks would be recovered when analyzing stronger AGO1 null alleles or profiling distinct tissues and developmental time points in the *ago1-25* background. Second, over 40% of the m^{5}C peak-hosting transcripts were recovered as AGO1-bound targets in only one of the two RIP-seq biological replicates, while only ~10% were detected as overlapping targets in both replicates (Figure 6H). This moderate overlap between replicates can be explained not only by inherent technical variability in RIP-seq assays but also by the highly transient and dynamic nature of AGO1—transcript interactions in vivo. Third, based on the canonical model, the miRNA-guided target regulation is mostly performed through AGO1-mediated target cleavages (Jones-Rhoades et al., 2006). However, based on the working model proposed by the present study (Figure 6I), the miRNA-guided target regulation may also involve m^{5}C modifications, which have a potential impact on transcript stability and translation efficiency. At the current stage, we cannot clearly distinguish the effects of miRNA-guided target cleavage and m^{5}C modification, as both transcript cleavage and m^{5}C modification can influence transcript stability. However, based on degradome sequencing data reported by our previous study (NCBI Sequence Read Archive accession ID: PRJNA1441820), target prediction and cleavage site maping were performed according to the previously reported workflow (Shao et al., 2013). As a result, no cleavage sites for potential miRNA regulators were identified on the transcripts exemplified in Figures 3 and 6A–E. This result further indicates that the differential m^{5}C modifications detected on these transcripts are unlikely to be artifacts caused by miRNA-guided target cleavages.

AUTHOR CONTRIBUTIONS

Zedi Feng: Data curation; formal analysis. **Xiu Yin**: Data curation; formal analysis. **Zhifang Jiang**: Resources. **Jia Yao**: Resources. **Yiling Dong**: Resources. **Xiaoxia Ma**: Conceptualization; data curation; resources; writing—original draft. **Yijun Meng**: Conceptualization; data curation; funding acquisition; resources; writing—original draft.

ACKNOWLEDGMENTS

We thank Prof. Yijun Qi (School of Life Sciences, Tsinghua University) for providing the mutant *ago1-25*. This study was supported by grants from the National Natural Science Foundation of China (32370707).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All of the sequencing data reported in this study have been deposited in the NCBI Sequence Read Archive (SRA) database under the accession IDs PRJNA1250198 (the m^{5}C-RIP-seq data of the WT plants and the *ago1-25* mutant), PRJNA1250199 (the sRNA-seq data of WT, *ago1-25*, and

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ago1-27), PRJNA1164813 (the AGO1 RIP-seq data of the 2-week-old WT seedlings of Col-0), and PRJNA1441820 (degradome-seq data of WT and *ago1-25*). The m^5C-RIP-seq data reported by Cui and his colleagues (2017) are publicly available in the NCBI Gene Expression Omnibus (GEO) under the accession ID GSE94065, and the information of the genomic loci and associated genes of the m^5C peaks discovered in the WT plants is available in the supplementary table of that study.

ORCID

Xiaoxia Ma https://orcid.org/0009-0007-0831-3086
Yijun Meng https://orcid.org/0000-0002-4654-0622

REFERENCES

Allen, E., Xie, Z., Gustafson, A. M., & Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, *121*, 207–221. https://doi.org/10.1016/j.cell.2005.04.004

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics*, *25*, 25–29. https://doi.org/10.1038/75556

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME suite. *Nucleic Acids Research*, *43*, W39–W49. https://doi.org/10.1093/nar/gkv416

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Burgess, A., David, R., & Searle, I. R. (2016). Deciphering the epitranscriptome: A green perspective. *Journal of Integrative Plant Biology*, *58*, 822–835. https://doi.org/10.1111/jipb.12483

Burgess, A. L., David, R., & Searle, I. R. (2015). Conservation of tRNA and rRNA 5-methylcytosine in the kingdom Plantae. *BMC Plant Biology*, *15*, Article 199. https://doi.org/10.1186/s12870-015-0580-8

Cappannini, A., Ray, A., Purta, E., Mukherjee, S., Boccaletto, P., Moafinejad, S. N., Lechner, A., Barchet, C., Klaholz, B. P., Stefaniak, F., & Bujnicki, J. M. (2024). MODOMICS: A database of RNA modifications and related information. 2023 update. *Nucleic Acids Research*, *52*, D239–D244. https://doi.org/10.1093/nar/gkad1083

Chai, D., Jia, C., Zheng, J., Zou, Q., & Li, F. (2021). Staem5: A novel computational approachfor accurate prediction of m5C site. *Molecular Therapy—Nucleic Acids*, *26*, 1027–1034. https://doi.org/10.1016/j.omtn.2021.10.012

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Clough, S. J., & Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal*, *16*(6), 735–743. https://doi.org/10.1046/j.1365-313x.1998.00343.x

Cohn, W. E. (1951). Some results of the applications of ion-exchange chromatography to nucleic acid chemistry. *Journal of Cellular and Comparative Physiology*, *38*, 21–40. https://doi.org/10.1002/jcp.1030380405

Cui, X., Liang, Z., Shen, L., Zhang, Q., Bao, S., Geng, Y., Zhang, B., Leo, V., Vardy, L. A., Lu, T., Gu, X., & Yu, H. (2017). 5-Methylcytosine RNA methylation in *Arabidopsis thaliana*. *Molecular Plant*, *10*, 1387–1399. https://doi.org/10.1016/j.molp.2017.09.013

Cui, X., Wei, Z., Zhang, L., Liu, H., Sun, L., Zhang, S.-W. U., Huang, Y., & Meng, J. (2016). Guitar: An R/bioconductor package for gene annotation guided transcriptomic analysis of RNA-related genomic features. *BioMed Research International*, *2016*, 8367534. https://doi.org/10.1155/2016/8367534

Cui, Y., Fang, X., & Qi, Y. (2016). TRANSPORTIN1 promotes the association of MicroRNA with ARGONAUTE1 in *Arabidopsis*. *The Plant Cell*, *28*, 2576–2585. https://doi.org/10.1105/tpc.16.00384

David, R., Burgess, A., Parker, B., Li, J., Pulsford, K., Sibbritt, T., Preiss, T., & Searle, I. R. (2017). Transcriptome-wide mapping of RNA 5-methylcytosine in Arabidopsis mRNAs and noncoding RNAs. *The Plant Cell*, *29*, 445–460. https://doi.org/10.1105/tpc.16.00751

Desrosiers, R., Friderici, K., & Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proceedings of the National Academy of Sciences*, *71*, 3971–3975. https://doi.org/10.1073/pnas.71.10.3971

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dubin, D. T., & Taylor, R. H. (1975). The methylation state of poly A-containing messenger RNA from cultured hamster cells. *Nucleic Acids Research*, *2*, 1653–1668. https://doi.org/10.1093/nar/2.10.1653

Fang, X., & Qi, Y. (2016). RNAi in plants: An Argonaute-centered view. *The Plant Cell*, *28*, 272–285. https://doi.org/10.1105/tpc.15.00920

Gamborg, O. L., Miller, R. A., & Ojima, K. (1968). Nutrient requirements of suspension cultures of soybean root cells. *Experimental Cell Research*, *50*, 151–158. https://doi.org/10.1016/0014-4827(68)90403-5

Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C.-L., Zhang, X., Golic, K. G., Jacobsen, S. E., & Bestor, T. H. (2006). Methylation of $tRNA^{Asp}$ by the DNA methyltransferase homolog Dnmt2. *Science*, *311*, 395–398. https://doi.org/10.1126/science.1120976

Jones-Rhoades, M. W., Bartel, D. P., & Bartel, B. (2006). MicroRNAS and their regulatory roles in plants. *Annual Review of Plant Biology*, *57*, 19–53. https://doi.org/10.1146/annurev.arplant.57.032905.105218

Liang, Z., Riaz, A., Chachar, S., Ding, Y., Du, H., & Gu, X. (2020). Epigenetic modifications of mRNA and DNA in plants. *Molecular Plant*, *13*, 14–30. https://doi.org/10.1016/j.molp.2019.12.007

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656

Liu, L., Zhang, S.-W., Huang, Y., & Meng, J. (2017). QNB: Differential RNA methylation analysis for count-based small-sample sequencing data with a quad-negative binomial model. *BMC Bioinformatics*, *18*, Article 387. https://doi.org/10.1186/s12859-017-1808-4

Lyu, F., Han, F., Ge, C., Mao, W., Chen, L., Hu, H., Chen, G., Lang, Q., & Fang, C. (2023). OmicStudio: A composable bioinformatics cloud platform with real-time feedback that can generate high-quality graphs for publication. *Imeta*, *2*, e85. https://doi.org/10.1002/imt2.85

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Meng, J., Cui, X., Rao, M. K., Chen, Y., & Huang, Y. (2013). Exome-based analysis for RNA epigenome sequencing data. *Bioinformatics*, *29*, 1565–1567. https://doi.org/10.1093/bioinformatics/btt171

Meng, J., Lu, Z., Liu, H., Zhang, L., Zhang, S., Chen, Y., Rao, M. K., & Huang, Y. (2014). A protocol for RNA methylation differential analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. *Methods*, *69*, 274–281. https://doi.org/10.1016/j.ymeth.2014.06.008

Montgomery, T. A., Yoo, S. J., Fahlgren, N., Gilbert, S. D., Howell, M. D., Sullivan, C. M., Alexander, A., Nguyen, G., Allen, E., Ahn, J. H., & Carrington, J. C. (2008). AGO1-miR173 complex initiates phased siRNA formation in plants. *Proceedings of the National Academy of Sciences*, *105*, 20055–20062. https://doi.org/10.1073/pnas.0810241105

Morel, J.-B., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., Proux, F., & Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *The Plant Cell*, *14*, 629–639. https://doi.org/10.1105/tpc.010358

Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., & Kanehisa, M. (1999). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, *27*, 29–34. https://doi.org/10.1093/nar/27.1.29

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Reiser, L., Bakker, E., Subramaniam, S., Chen, X., Sawant, S., Khosa, K., Prithvi, T., & Berardini, T. Z. (2024). The Arabidopsis information resource in 2024. *Genetics*, *227*, iyae027. https://doi.org/10.1093/genetics/iyae027

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*, 24–26. https://doi.org/10.1038/nbt.1754

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616

Saletore, Y., Meyer, K., Korlach, J., Vilfan, I. D., Jaffrey, S., & Mason, C. E. (2012). The birth of the Epitranscriptome: Deciphering the function of RNA modifications. *Genome Biology*, *13*, 175. https://doi.org/10.1186/gb-2012-13-10-175

Shao, C. G., Chen, M., & Meng, Y. J. (2013). A reversed framework for the identification of microRNA-target pairs in plants. *Briefings in Bioinformatics*, *14*(3), 293–301. https://doi.org/10.1093/bib/bbs040

Shatkin, A. J. (1976). Capping of eucaryotic mRNAs. *Cell*, *9*, 645–653. https://doi.org/10.1016/0092-8674(76)90128-8

Shen, L., Liang, Z., Wong, C. E., & Yu, H. (2019). Messenger RNA modifications in plants. *Trends in Plant Science*, *24*, 328–341. https://doi.org/10.1016/j.tplants.2019.01.005

Shen, L., Ma, J., Li, P., Wu, Y., & Yu, H. (2023). Recent advances in the plant epitranscriptome. *Genome Biology*, *24*, Article 43. https://doi.org/10.1186/s13059-023-02872-6

Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., Suter, C. M., & Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Research*, *40*, 5023–5033. https://doi.org/10.1093/nar/gks144

Sun, H., Li, K., Liu, C., & Yi, C. (2023). Regulation and functions of non-m^6A mRNA modifications. *Nature Reviews Molecular Cell Biology*, *24*, 714–731. https://doi.org/10.1038/s41580-023-00622-x

Tolia, N. H., & Joshua-Tor, L. (2007). Slicer and the argonautes. *Nature Chemical Biology*, *3*, 36–43. https://doi.org/10.1038/nchembio848

Vaucheret, H. (2008). Plant ARGONAUTES. *Trends in Plant Science*, *13*, 350–358. https://doi.org/10.1016/j.tplants.2008.04.007

Wiener, D., & Schwartz, S. (2021). The epitranscriptome beyond m^6A. *Nature Reviews Genetics*, *22*, 119–131. https://doi.org/10.1038/s41576-020-00295-8

Zhang, P., Wang, Y., & Gu, X. (2020). RNA 5-methylcytosine controls plant development and environmental adaptation. *Trends in Plant Science*, *25*, 954–958. https://doi.org/10.1016/j.tplants.2020.07.004

Zhu, T.-Y., Chen, S.-Y., Zhang, M., Li, H., Wu, T., Ajiboye, E., Wang, J. W., Jin, B.-K., Liu, D.-D., Zhou, X., Huang, H., Wan, X., Sun, K., Lu, P., Fu, Y., Yuan, Y., Song, H., Sablina, A. A., Tong, C., … Yang, B. (2025). Genetically encoding epsilon-N-methacryllysine into proteins in live cells. *Nature Communications*, *16*, 2623. https://doi.org/10.1038/s41467-025-57969-2

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Feng, Z., Yin, X., Jiang, Z., Yao, J., Dong, Y., Ma, X., & Meng, Y. (2026). The novel regulatory influence of Argonaute 1 on messenger RNA 5-methylcytosine modification in *Arabidopsis thaliana*. *The Plant Genome*, *19*, e70327. https://doi.org/10.1002/tpg2.70327

19403372, 2026, 4, Downloaded from https://acsess.onlinelibrary.wiley.com/doi/10.1002/tpg2.70327 by Hangzhou Normal University, Wiley Online Library on [12/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License