Computational and Structural Biotechnology Journal 21 (2023) 837–846

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

CUCUME: An RNA methylation database integrating systemic mRNAs signals, GWAS and QTL genetic regulation and epigenetics in different tissues of *Cucurbitaceae*

Xiaojun Li [a,1], Shujin Lin [b,1], Chenggang Xiang [c], Wenqian Liu [a], Xiaojing Zhang [a], Cuicui Wang [a], Xiaohong Lu [a], Mengshuang Liu [a], Tao Wang [a], Zixi Liu [a], Naonao Wang [a], Lihong Gao [a], Xiao Han [b,*], Wenna Zhang [a,*]

[a] *Beijing Key Laboratory of Growth and Developmental Regulation for Protected Vegetable Crops, China Agricultural University, Beijing, China*
[b] *College of Biology Science and Engineering, Fuzhou University, Fuzhou, China*
[c] *College of Life Science and Technology, HongHe University, Mengzi, Yunnan, 661100, China*

ARTICLE INFO

Article history:
Received 16 October 2022
Received in revised form 9 January 2023
Accepted 9 January 2023
Available online 10 January 2023

Keywords:
RNA methylation
Graft-transmissible mRNA
QTL
RNA motifs

ABSTRACT

As an internal modification of transcripts, RNA methylation determines RNA fate by changing RNA–protein binding affinity. In plants, RNA methylation is ubiquitous and is involved in all aspects of RNA post-transcriptional regulation. For instance, long-distance mobile RNAs, strongly influenced by their methylation status, play important roles in plant growth, development and environmental adaptation. Cucumber/pumpkin heterografts are widely used to improve stress tolerance of cucumber and to study mobile RNA signals due to their strong developed vasculature system. Here, we developed the Cucume (Cucurbit RNA methylation, http://cucume.cn/) database for these two important vegetables, cucumber (*Cucumis sativus* L.) and pumpkin (*Cucurbita moschata*) with high productivity worldwide. We identified mRNAs harboring 5-methylcytosine (m^5C) and N^6-methyladenosine (m^6A) sites in pumpkin and cucumber at the whole genome level via Methylated RNA Immunoprecipitation sequencing (MeRIP-seq) of different tissues and the vascular exudates. In addition to RNA methylation sites, the Cucume database includes graft-transmissible systemic mRNAs identified in previous studies using cucumber/pumpkin heterografts. The further integration of cucumber genome-wide association analysis (GWAS) and quantitative trait loci (QTL) allows the study of RNA methylation-related genetic and epigenetic regulation in cucurbits. Therefore, the here developed Cucume database will promote understanding the role of cucurbit RNA methylation in RNA mobility and QTL, ultimately benefitting future breeding of agronomic crop germplasms.

1. Background

In plants, N^6-methyladenosine (m^6A) and 5-methylcytosine (m^5C) are the most common RNA methylation markers. These markers play dynamic roles in many developmental processes, e.g. embryonic development, stem cell fate determination, trichome branching, leaf morphogenesis, floral transformation, stress response, fruit maturation and root development. m^6A is the most abundant internal mRNA modification and plays a vital role in almost every aspect of RNA fate through RNA metabolic pathways, including alternative precursor splicing [1,2], 3′-end processing [3], nucleocytoplasmic shuttling [4], polyadenylation site choice [5,6], translation [7–9], mRNA decay [10] and stability [11–13]. In-depth research on m^6A revealed that it plays an important role in plant growth, development and stress responses [14–16]. By contrast, studies on m^5C in eukaryotes only focused on mapping the modification transcriptome-wide in coding and non-coding RNAs [17–27]. Utilizing RNA Immunoprecipitation sequencing (RIP-seq), RNA bisulfite sequencing (RNA-BisSeq), and Nanopore-seq, recent studies explored the m^5C and m^6A landscapes and their functions in development, stress responses and RNA metabolism in *Arabidopsis thaliana* and *Oryza sativa* [28–30]. It has been reported that the

* Corresponding authors.
E-mail addresses: hanxiao@fzu.edu.cn (X. Han), zhangwenna@cau.edu.cn (W. Zhang).
[1] These authors contributed equally.

https://doi.org/10.1016/j.csbj.2023.01.012

writer, the eraser and the reader proteins responsible for m^6A methylation, demethylation and recognition are conserved in regulating stress responses of *Arabidopsis thaliana, Oryza sativa* and *Zea mays* [31–35]. An m^5C writer loss-of-function mutant exhibits weaker oxidative stress resistance in root development of Arabidopsis and rice [18].

Cucumber (*Cucumis sativus*) and pumpkin (*Cucurbita moschata* and *Cucurbita argyrosperma*) are two economically important members from the *Cucurbitaceae* family (cucurbits/gourds). Cucumber is one of the main vegetable crops and it often experiences stress conditions that limit its growth, productivity, and fruit quality. Notably, cucurbitacin-B, which is abundant in cucumbers, exhibits anti-cancer activity against several types of cancers and has great potential of application in tumor prevention and treatment [36]. Grafting using a vigorous pumpkin rootstock is widely used to improve the stress resistance of cucumber, suggesting that systemic communication signals play crucial roles between above and underground tissues/organs via well-developed vasculature. Well-known systemic signals delivered by the re-union vasculatures of graft junction include small molecules such as hormones, carbohydrates, non-coding RNAs, and macromolecules such as mRNAs and proteins [37–40]. Recent advances in molecular biotechnology and bioinformatics have stimulated research of long-distance mRNA signals in the interaction between rootstocks and scions [41]. Numerous endogenous mRNAs migrate between the rootstock and the scion to regulate plant height, apical meristem development, tuber growth, and root development [42–44]. These mRNAs include a member of the NAC domain gene family (*NACP*) [45] and *GA-INSENSITIVE* (*GAI*) [46] from pumpkin (*Cucurbita maxima*).

Systemic and intercellular mRNA transport is facilitated by the specific motifs or structures, including 3′-UTR [46,47], poly-CU [43], tRNA-related sequence (TLS) [44] etc. and some cytoplasmic ribosomes and chaperones so as to alter RNA self-conformation and the plasmodesma exclusion limit (SEL) in the phloem. In addition, the m^5C modification is required for the transport of *TRANSLATIONALLY CONTROLLED TUMOR PROTEIN 1* (*TCTP1*) mRNA to its target cells, ultimately influencing root growth in *Arabidopsis thaliana* [20]. However, due to inadequate mutant resources, our knowledge remains limited regarding the function of mobile mRNAs in controlling cucumber and pumpkin agronomic traits and edibility (i.e., reproductive development, fruit-related traits, seed-related traits, disease resistance, and abiotic stress tolerance). Thus, it is of great importance to characterize the mobile mRNAs related to agronomic traits and edibility concerning their potential non-cell-autonomous signaling, tissue-specific m^5C or m^6A markers, the presence of TLS motifs, linkage to Quantitative Trait Loci (QTL) influencing agronomic traits and associated Single Nucleotide Polymorphisms (SNPs) identified through Genome-Wide Association Studies (GWAS). To understand the relevance of RNA methylation in plant growth, development, stress adaptation and non-cell-autonomous mRNA signaling, it is necessary to curate the available information from the literature and published data sources.

Here, we developed the Cucume database (http://cucume.cn/) as a free source convenient for retrieving m^5C, m^6A RNA methylation sites, exploring their associations with biological functions and evaluating potential functions of these methylation sites involved in mobile RNA signaling in cucumber and pumpkin. In this database, we incorporated graft-transmissible mRNAs from hypocotyl-grafted cucumber and pumpkin seedlings [48], multiple GWAS-SNPs, QTLs mainly obtained from the literature [49], motifs including TLS [50], and vascular-delivered mRNAs identified by high-throughput sequencing of cucumber and pumpkin vascular exudate. Furthermore, the user-friendly interface allows mobile transcripts to be linked to the genome maps of cucumber (*Cucumis sativus* L., Gy14 genome v2 http://cucurbitgenomics.org/organism/16) and pumpkin (*Cucurbita moschata* Rifu http://cucurbitgenomics.org/organism/9) [51], presenting an overview of potential RNA methylation sites at the genome level. Therefore, Cucume will be a useful tool for elucidating the role of RNA methylation in cucurbit RNA mobility and QTLs, ultimately benefitting the future breeding of agronomic crop germplasms.

2. Utility and discussion

2.1. Overview of the Cucume database

We developed Cucume for a collection of RNA-seq, RNA methylation, RNA signaling, and QTL information from the genomes of cucumber (*Cucumis sativus* L.) and pumpkin (*Cucurbita moschata*) with a standardized pipeline. Firstly, 5-methylcytosine (m^5C) and N^6-methyladenosine (m^6A) methylation sites were obtained from our MeRIP-seq data using both the mixed tissues with equal mass of roots, stems, flowers and leaves of adult plants, and also the vascular exudates of the stem. Mixed tissues (designated as “Adult plant” in the database) and vascular exudates (designated as “Vascular” in the database) of cucumber and pumpkin. 5123 mRNAs harboring m^5C and 11042 mRNAs harboring m^6A sites were identified in pumpkin, 1127 mRNAs harboring m^5C and 8953 mRNAs harboring m^6A methylation sites were identified in cucumber (Fig. 1). Secondly, the graft-transmissible mobile RNA dataset, including 3923 cucumber and 1788 pumpkin endogenous mRNAs, was integrated into Cucume based on a previous study [48]. This database also included a total of 287 cucumber QTLs with 31 quantitative traits obtained from 91 published studies [49]. These QTLs, along with their associated phenotypes such as abiotic stress tolerance and disease resistance, involved 9260 genes, thereby providing a basis for in-depth research on *Cucurbitaceae* epigenetics.

2.2. Browsing the Cucume database

Cucume provides tabular-organized browse panels to retrieve graft-transmissible mRNAs, GWAS-SNPs, TLS, m^5C and m^6A methylation in different tissues. To conveniently access methylated mRNA with mobile mRNA and coding SNP genes from an external database, the “original ID” (gene ID with hyperlink) was obtained from a published cucurbit database (CuGenDB, http://cucurbitgenomics.org/) [52]. Each original ID entry lists plant species (cucumber/pumpkin), graft-transmissible mRNA (mobile direction of shoot-to-root and root-to-shoot), TLS motifs (the number of TLS in the CDS region determined using PlaMoM), biological functions, tissue-specific m^5C and m^6A methylation (Vascular/Adult plant) with methylation fold changes, QTL and GWAS-SNPs of cucumber (yes/no) (Fig. 2A). The original ID of each queried gene is linked to an annotation page, which provides details of the CuGenDB gene annotation, gene location, a genome browser link, the mobility, TLS, tissue-specific methylation, and biological functions (Fig. 2B). More comprehensive data is also included, such as: 1) Graft-transmissible mRNA: mobile direction, number of TLSs; 2) RNA methylation: methylation type, tissue, methylation location; 3) GWAS-SNP and QTL: ID, physical location, related agronomic traits. Within the information provided, “Genome browser” linking to the JBrowse panel displays m^5C and m^6A methylation sites and methylation peaks in the genome based on the input control and IP results from our MeRIP-seq analysis.

2.3. Searching and downloading at Cucume

Cucume provides a “search” interface to query detailed information on RNA methylation sites, mobility, and associated biological functions of each gene ID. Queries can be initiated by either “gene ID” or “condition”. By “gene ID”, users can submit multiple gene IDs (cucumber: *Cucumis sativus* Gy14 genome v2, pumpkin:

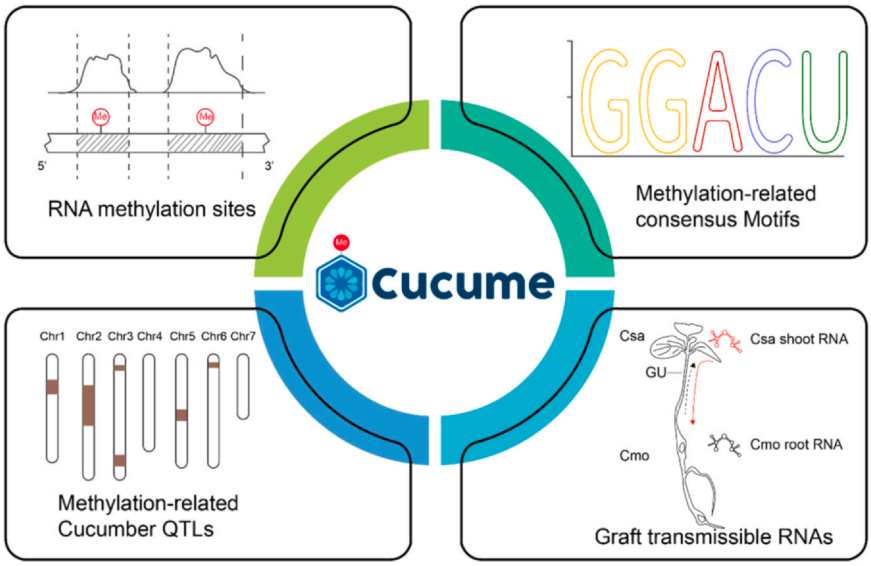


Fig. 1. Overview of Cucume. Diagram of the association network which includes the gene, methylation sites, QTLs, GWAS-SNPs RNA motifs, and mobile mRNAs. Cucume is a collection of RNA methylation sites and a visualization platform on which users can retrieve known or predicted methylated RNAs, potential methylation-related mRNA motifs and tRNA-like structure (TLS) motifs, as well as graft-transmissible mRNAs in cucumber and pumpkin.

Cucurbita moschata Rifu genome) from CuGenDB (http://cucurbitgenomics.org/) [52] (Fig. 3A). By "condition", users can select their search criteria from drop-down lists of species (pumpkin/cucumber), mobile RNA (Root-to-shoot/Shoot-to-root/Bidirectional/no), TLS ("no" means that this RNA is undetectable, "0" means that the TLS is non-detected in the RNA, other numbers indicate the number of TLS motifs detected), GWAS-SNP (yes/no), RNA methylation types (m^5C/m^6A) and methylation tissues (Vascular/Adult plant) (Fig. 3B). The search engine also allows for combined queries. For instance, to search for the number of shoot-to-root mobile mRNAs that harbor m^5C methylation and GWAS-SNPs in cucumber vascular exudates, the user should select "cucumber" for species, "shoot-to-root" for mobile RNA, "yes" for the GWAS-SNP option, and "Vascular" in the m^5C drop-down box. The result page displays four mRNAs meeting the specified criteria. The user can also choose to display their function, GWAS-SNP, and QTL annotation (Fig. 3C).

Furthermore, in the "Download" interface, we provide the mobile mRNA datasets of cucumber and pumpkin, all methylation sites from the whole plant samples and the vascular exudates, cucumber GWAS-SNPs, the genes contained in each QTL, and the TLS sequences of the CDS regions.

2.4. Blast and JBrowse features of Cucume

Cucume provides a user-friendly blast tool to query homologous gene IDs and their corresponding gene sequence in cucumber or pumpkin without knowing the exact name. By providing a FASTA sequence (select "Text" as type of query) or uploading a FASTA format file (select "File" as type of query), Cucume will display an auto-completed list of blast results including E-value, sequence length and max target seqs. It has been reported that some RNA methylation sites between species are conserved within a plant family [27]. Therefore, this module is also useful for predicting RNA methylation sites in other plant species, especially other members of the *Cucurbitaceae* family, such as *Cucumis melo, Citrullus lanatus,* and *Momordica charantia*.

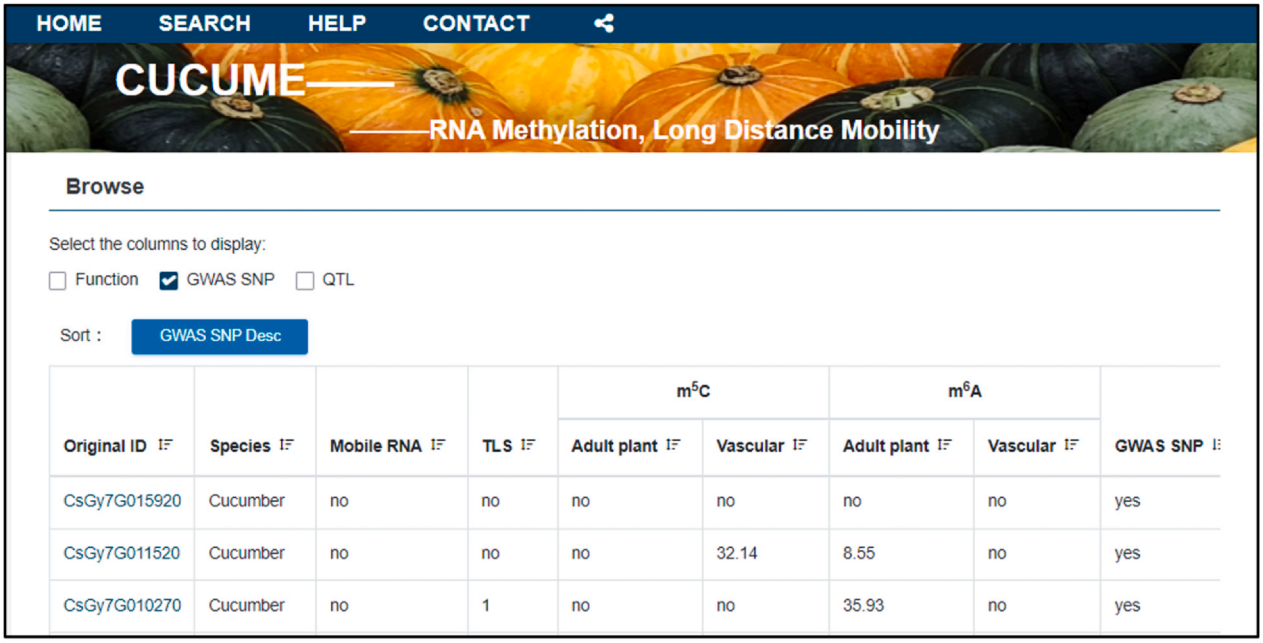
In addition, the built-in m^5C and m^6A methylation data from our MeRIP-seq experiments allowed us to create a JBrowse tool to visualize RNA methylation intensity and their corresponding locations at the genome level for candidate genes (Fig. 4A). For example, the m^5C methylation peak intensity at immunoprecipitated *CmoCh15G009950* RNA in replicate 1 of vascular exudates is higher than the RNA Input. This methylation peak is located at the junction between the 3′-UTR and the last intron of *CmoCh15G009950*. Together, the results will provide the users the possible methylation intensity and location at the genes of their intereste (Fig. 4B).

2.5. Cucumber QTL and GWAS-related SNPs in Cucume

The datasets of QTL regions in our database were obtained from published literature [49], which contains 287 cucumber QTL with 31 quantitative traits from 91 studies and published phenotypes. These traits include QTL for the whole plant vegetative growth and development, reproductive development, fruit-related traits, seed-related traits, disease resistance, and abiotic stress tolerance. By selecting the cucumber traits from the drop-down list, the corresponding QTL ID and the chromosome location are displayed in the cucumber genetic map (Fig. 5A,B). The hyperlink of QTL IDs will lead to detailed QTL characteristics, including QTL-related agronomic traits, location in the genome, length, and the QTL-related genes with methylated mRNAs in different tissues (from three independent experiments). While combining the m^5C- and m^6A-methylated genes and the QTL-related genes, we can further explore the relation between RNA methylation and QTL-related agronomic traits.

2.6. Statistics and motif analysis in Cucume

Cucume provides a preliminary "Statistics" module that helps users gain a comprehensive and intuitive understanding of the effects of RNA methylation on QTL-related agronomic traits and systemic RNA signaling in different tissues of cucumber and pumpkin. The "Statistics" module includes the distribution of RNA methylation peaks in transcripts, the enrichment of methylated mRNAs in QTL and the relationship between mRNA methylation and grafting-transmissible mRNAs. However, it still needs to be verified by additional experiments whether the differential enrichment of methylated mRNAs in QTL-associated genes is related to the specific agronomic phenotype and how it participates in genetic regulation. Statistical analysis indicated that the proportion of mobile RNAs harboring m^5C methylation was higher in the vascular exudate in cucumber and pumpkin compared to the whole plant tissue mixture (Fig. 6A). This result suggests that RNA methylation intensity is



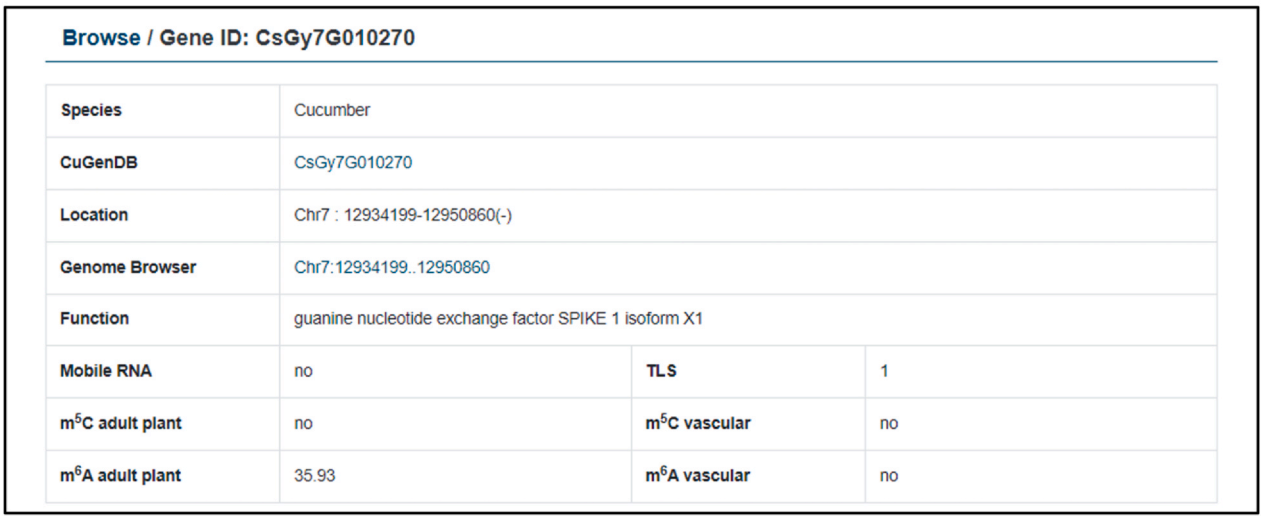


Fig. 2. The "Browse" module, detailed gene annotation page, and genome browser panel. (A) Screenshot showing the Cucume browsing module listing all or subsets of all genes. (B) Detailed gene ID linking to CuGenDB gene annotation, gene location, mobile mRNA, TLS, methylation tissue, genome browser, and functional annotations.

related to RNA mobility in the vasculature. Worth noting that the proportion of grafting-transmissible mRNAs harboring either m^5C or m^6A methylation in cucumber or pumpkin varies according to the position, number, and maximum value of the methylation peaks.

To help users understand the relationship between methylated RNA, potential methylation-related motifs, and RNA systemic signaling in different tissues, Cucume also provides *de novo* motif analysis using our MeRIP-seq data. Motif analysis was conducted using 50-bp sequences flanking methylation peaks from cucumber and pumpkin. The detailed "Motif" of each methylated RNA sample is available with hyperlinks to the DREME program (http://meme-suite.org/doc/dreme.html) (Fig. 6B).

3. Conclusions and outlook

Cucume is a comprehensive web-based resource that presents MeRIP-seq-derived m^5C and m^6A RNA methylation peaks from different tissues of cucumber and pumpkin. The RNA methylation data from the mixed tissues of adult plant tissues and vascular exudates will promote the rapid development of functional research on RNA methylation. In addition to exploring RNA methylation sites and methylation intensity at the genome level, users can also easily find additional relevant information in Cucume: systemic mobile mRNAs in cucumber or pumpkin, cucumber QTL and GWAS-SNPs, TLS and other RNA motifs conserved across species. These integrated datasets provide the basis for functional and epigenetic studies related to cucurbit RNA methylation. With the continuous development of technology, higher resolution RNA methylation profiles of *Cucurbitaceae* species will be obtained. Due to the great regulatory potential of RNA methylation, Cucume will reinforce the profiling of dynamic RNA methylation landscapes under diverse environmental conditions or at different developmental stages. We will continue to track developments in the field and frequently upload new data from cucumber and pumpkin, and also other vegetable crops into Cucume to keep up with the latest discoveries.

In our database, the RNA-seq library construction of both the input control sample and m^5C or m^6A immunoprecipitation samples was performed by using whole transcriptomic sequencing method

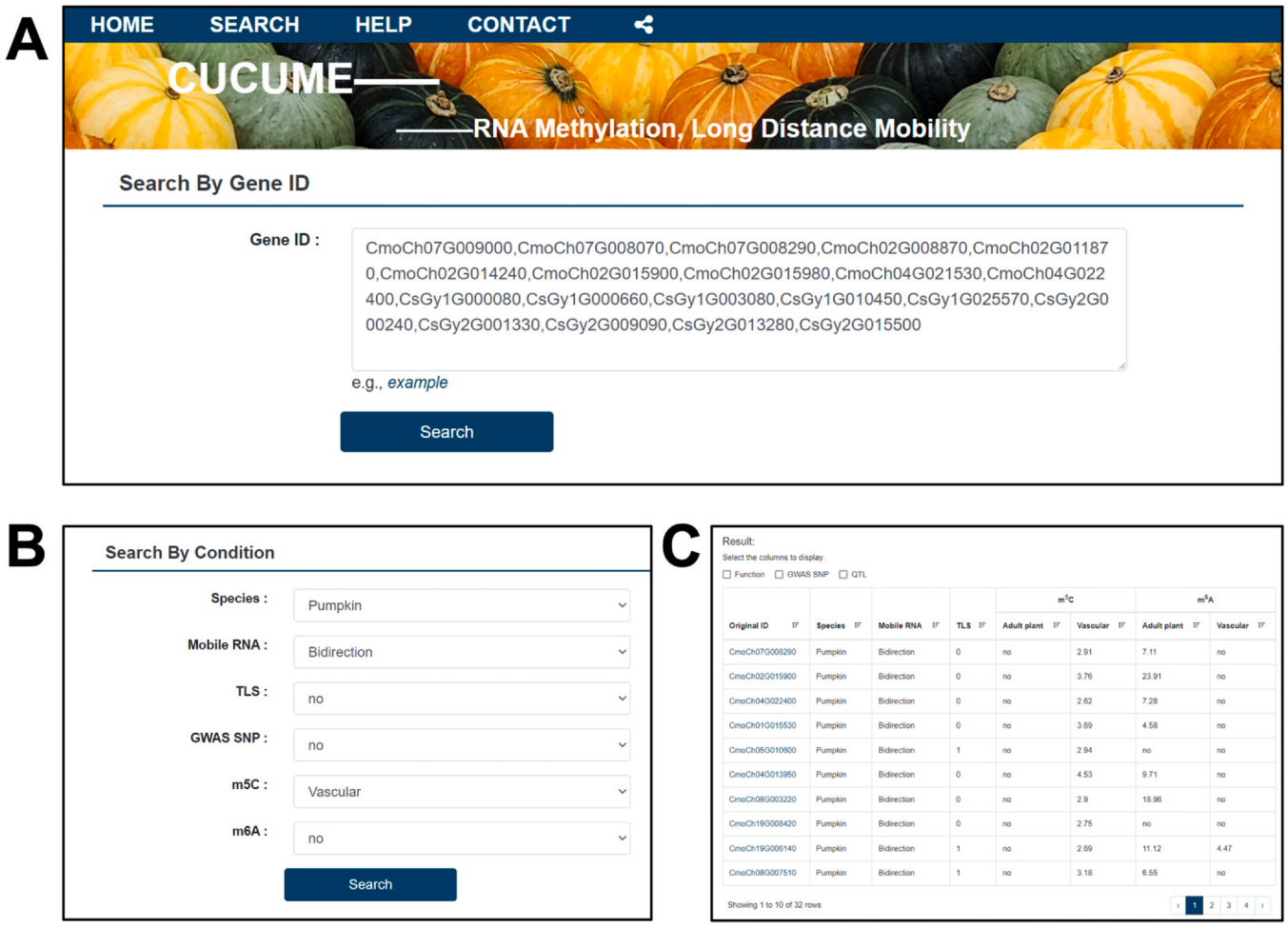


Fig. 3. The "Search" interface. (A-B) Querying detailed information on the methylation, mobility, and associated functions of each gene ID by either (A) inputting a gene ID or list of IDs ("Gene ID"), or (B) selecting specific search criteria ("Condition") from the drop-down lists. (C) The final result page displaying four mRNAs. The user may choose to display gene function, GWAS-SNP, and QTL annotation columns in combined queries.

with poly(A)-tails according to the protocol [55]. That's why we only can get the poly(A)-tailed methylated RNAs including mRNA, cirRNA, lncRNA etc. However, tRNA, rRNA and other small RNAs don't have poly(A) tails. If we want to get m^6A and m5C from them, we have to perform the tRNA input and immunoprecipitation library construction by adding poly(A) based on another protocol [27]. In the further updated version, we will add more other types of RNAs.

We believe that Cucume will be valuable for researchers interested in RNA methylation and its roles in RNA mobility and QTLs, ultimately promoting crop breeding for the future.

4. Materials and methods

4.1. Plant growth conditions, tissue pools, and harvesting vascular exudate

Cucumber (*Cucumis sativus* L. cv. Xintai Mici; "Csa") and pumpkin (*Cucurbita moschata*. cv. *Qianglishi*; "Cmo") seeds were sown in 50-hole seedling trays with mix matrix (peat: vermiculite: perlite, 2:1:1, by vol.) and the plants were grown in the greenhouse under conditions as described in Liu et al. [48]: light intensity between 1200 and 1500 mmol m^{-2} s^{-1} and day/night temperatures of approximately 28 ℃/20 ℃. Adult flowering cucumber and pumpkin plants with 20 leaves were used for sampling. For the mixture of different tissues, leaves, roots, and stems with same fresh weight were harvested from 3 to 6 individual cucumber and pumpkin plants and pooled as one biological replicate. At least three replicates were used to isolate RNA for Methylated RNA Immunoprecipitation sequencing (MeRIP-seq).

Vascular exudates from 18 adult cucumber and pumpkin plants with 20 leaves were harvested according to the protocol from Zhang et al. [53]. After cutting the stem/petiole, the surface liquid was removed 4–5 times by sterilized filter paper to avoid cytochylema contamination. 100 μL Vascular exudates was harvested with 10 μL sterile RNase-free pipette tips and transferred into 1 mL TRIzol® Reagent (Thermo Scientific, Beijing, China) on ice. The collected vascular exudate was stored at −80 °C prior to total RNA extraction.

4.2. RNA isolation and RNA-seq library construction

RNA isolation from vascular exudates and tissue pools was performed using TRIzol® Reagent according to the manufacturer's instructions [54]. The three replicates of mixed root, stem, and leaf samples were flash frozen in liquid nitrogen and then ground using a mortar and pestle. 1 mL of TRIzol® Reagent was added to about 50 mg of pulverized sample. Approximately 0.5–1 μg of total RNA was extracted per 100 μL of harvested vascular exudate. 1 μg of total RNA is required per replicate for Methylated RNA immunoprecipitation (MeRIP)-sequencing.

The RNA concentration of each sample was measured using a NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA, USA).

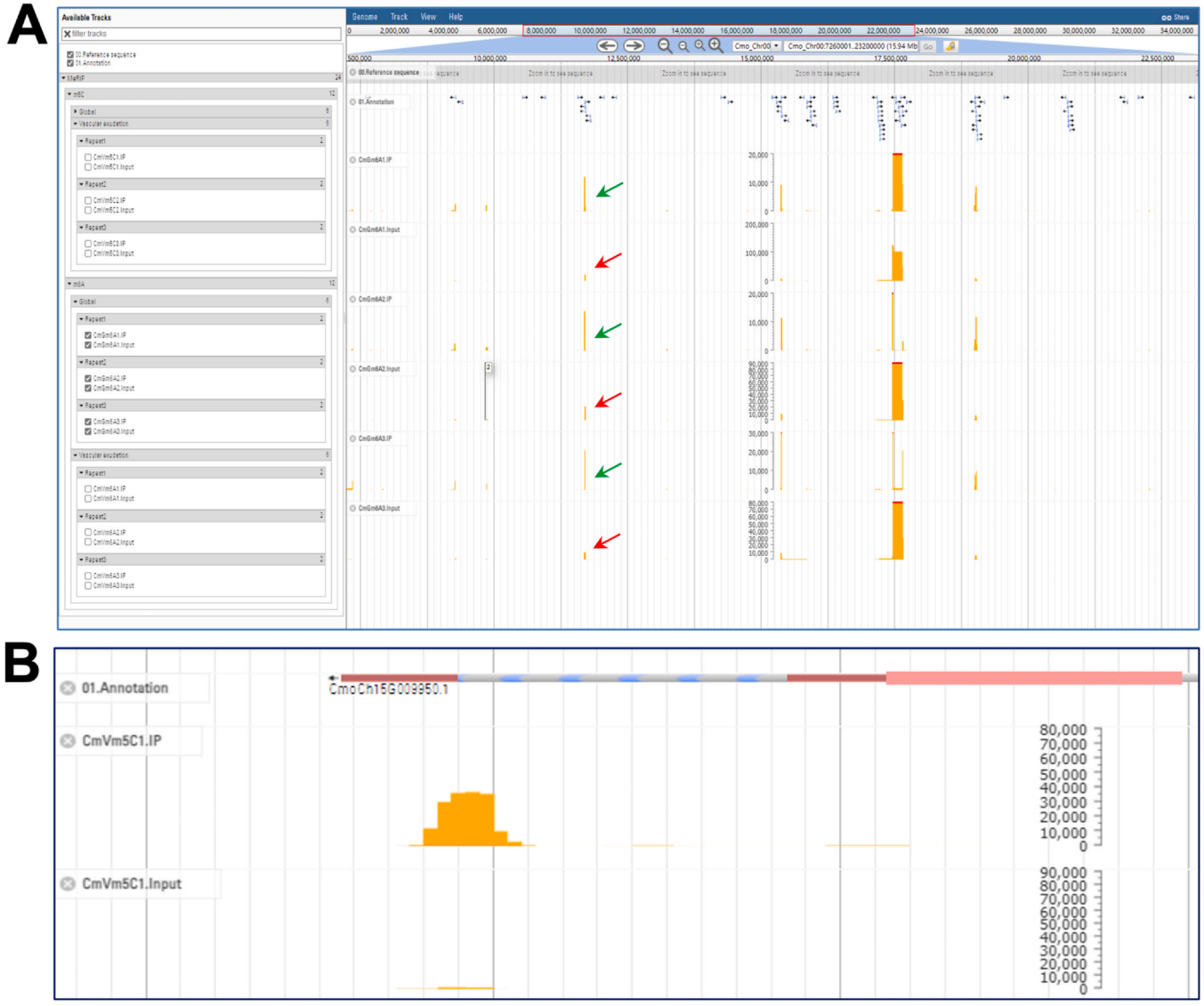


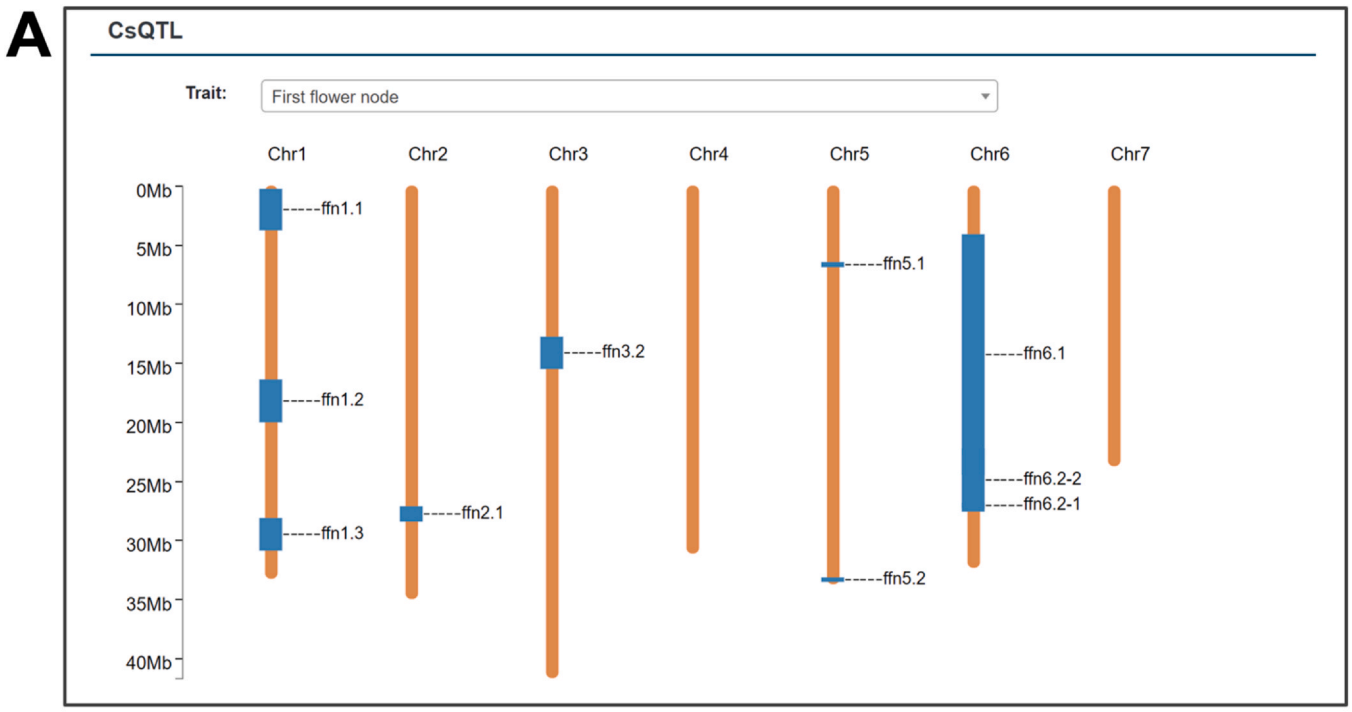
Fig. 4. The JBrowse module. (A) Visualization of the MeRIP-seq or Input results. Data sets selected on the left will be displayed in the right panel. All three replicates for each experimental group can be queried. Peak height indicates the degree of enrichment of the sequence at the peak location in the corresponding experiment. In the figure, m^6A MeRIP-seq data from the pumpkin mixed tissue sample shows that peaks of the three IP replicates (green arrows) at the same position in the genome are significantly higher than those of the Input (red arrows). This indicates that methylation has occurred at this position. (B) In the gene model, dark red represents the UTR, light red represents the CDS, stripes represent the intron, and the remaining regions are intergenic sequences. m^5C MeRIP-seq data from the sample of pumpkin vasculature saps shows that peaks of IP replicate 1# (orange peaks) at the same position in the genome are significantly higher than those of the Input. This indicates that methylation has occurred at this position.

The OD260/OD280 ratio was used as the index of RNA purity; an RNA sample with an OD260/OD280 ratio ranging from 1.8 to 2.1 was deemed pure. RNA integrity and genomic DNA contamination were assessed by denaturing agarose gel electrophoresis.

m^5C and m^6A RNA-seq library construction was performed by Cloudseq Biotech Inc. (Shanghai, China) based on Meyer et al. [55]. Briefly, m^5C RNA or m^6A RNA immunoprecipitation was performed with the GenSeqTM m^5C RNA IP Kit (GenSeq Inc., China) following the manufacturer's instructions. Both the input control sample and the m^5C or m^6A IP samples were used for RNA-seq library construction with NEBNext® Ultra II Directional RNA Library Prep Kit (New England Biolabs, Inc., USA). The library quality was evaluated using the BioAnalyzer 2100 system (Agilent Technologies, Inc., USA). Library sequencing was performed on an Illumina HiSeq instrument 3000 with 150 bp paired-end reads.

4.3. Methylated RNA immunoprecipitation (MeRIP)-sequencing and peak analysis

Briefly, paired-end reads were obtained from the Illumina HiSeq 4000 sequencer, and were quality controlled by Q30 (Q30 > 80 % indicates acceptable sequencing results). These reads were then subjected to 3′ adapter trimming and low-quality reads were removed using the cutadapt [56] software (v1.9.3). Next, clean reads from input libraries were aligned to the genome using the bowtie2 software [57]. Clean reads of input libraries were mapped to the genome by hisat2 [58] software (v2.04), and assembled using the stringtie [59] software (v1.3.4d). Methylated sites on RNAs (methylation peaks) were identified by the MACS software [60]. The detailed parameters of all software were described in Supplementary data. The MeRIP-sequencing raw data is hosted at Genome Sequence Archive (GSA) database (https://ngdc.cncb.ac.cn/gsa/) [61,62].



B

QTL ID	Trait	Chr	Start Position	End Position	Length	m5C methylated mRNA	m6A methylated mRNA	total mRNA
ffn1.1	First flower node	1	465777	3986626	3520849	22	221	536
ffn6.1	First flower node	6	4310799	24681050	20370251	106	817	2028

Fig. 5. The "QTL" interface. (A) By selecting cucumber traits, the corresponding QTL ID and chromosome location can be visualized in the cucumber genetic map. (B) The hyperlink of QTL IDs shows detailed QTL characteristics, including QTL-related agronomic traits, QTL location in the genome, QTL length, and QTL-related genes with methylated mRNAs in different tissues (from three independent experiments).

4.4. Data source and integration

The Cucume database mainly contains m^5C and m^6A RNA methylation region in vascular exudates and tissue pools of flowering cucumber and pumpkin. We also incorporated graft-transmissible mRNAs identified by RNA-seq and genome re-sequencing using SNPs from hypocotyl-grafted cucumber and pumpkin seedlings [48], multiple sources of QTLs and GWAS-SNPs. QTL mapping of 31 quantitative traits were obtained from the literature using multiple recombinant inbred lines such as the following crosses: 9110Gt × 9930, Gy14 × B10 and PI 183967 × 931. The original data was downloaded from Wang et al. [49]. The QTL interval information and RNA methylation sequencing results were integrated through the database, and the hypergeometric distribution made in excel (hypgemodist) was used to analyze the distribution of m^6A and m^5C methylation sites in QTLs.

4.5. De novo methylation-related motif prediction

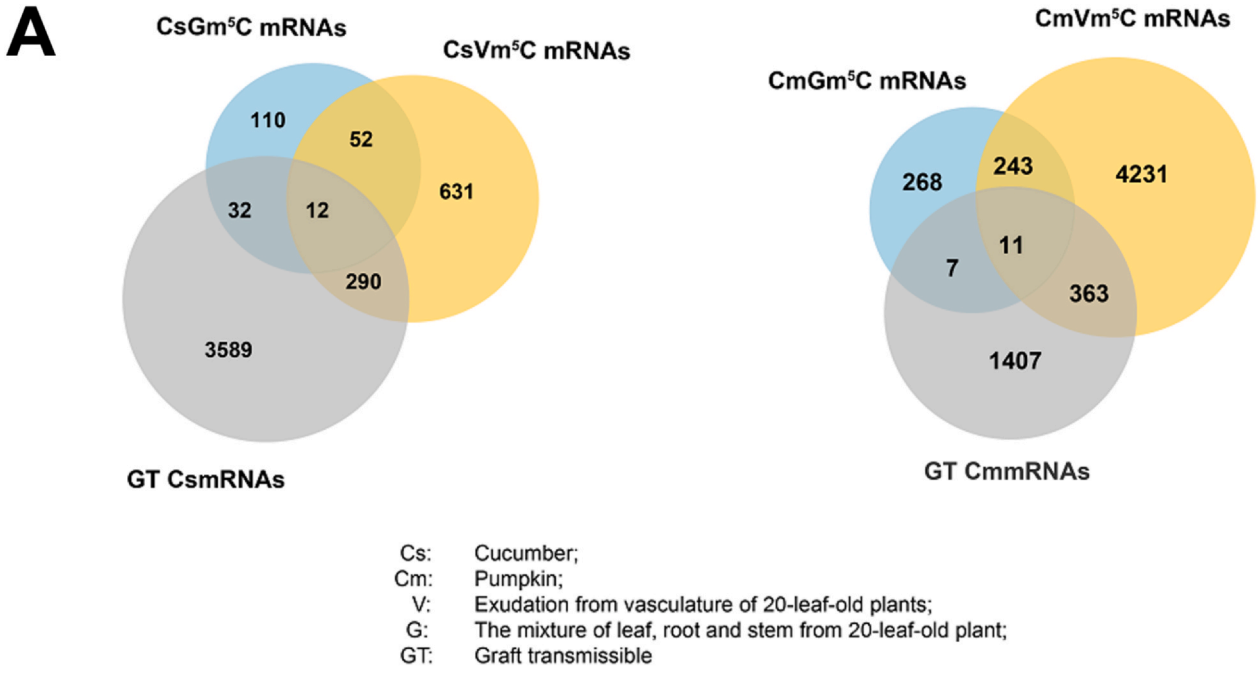
To identify conserved methylation-related motifs within the RNA, a 50 bp sequence from each side of the methylation peak vertex was used as the input sequence for conserved motif analysis using DREME [63] with default parameters. Moreover, TLS search was performed using PlaMoM with default parameters [50].

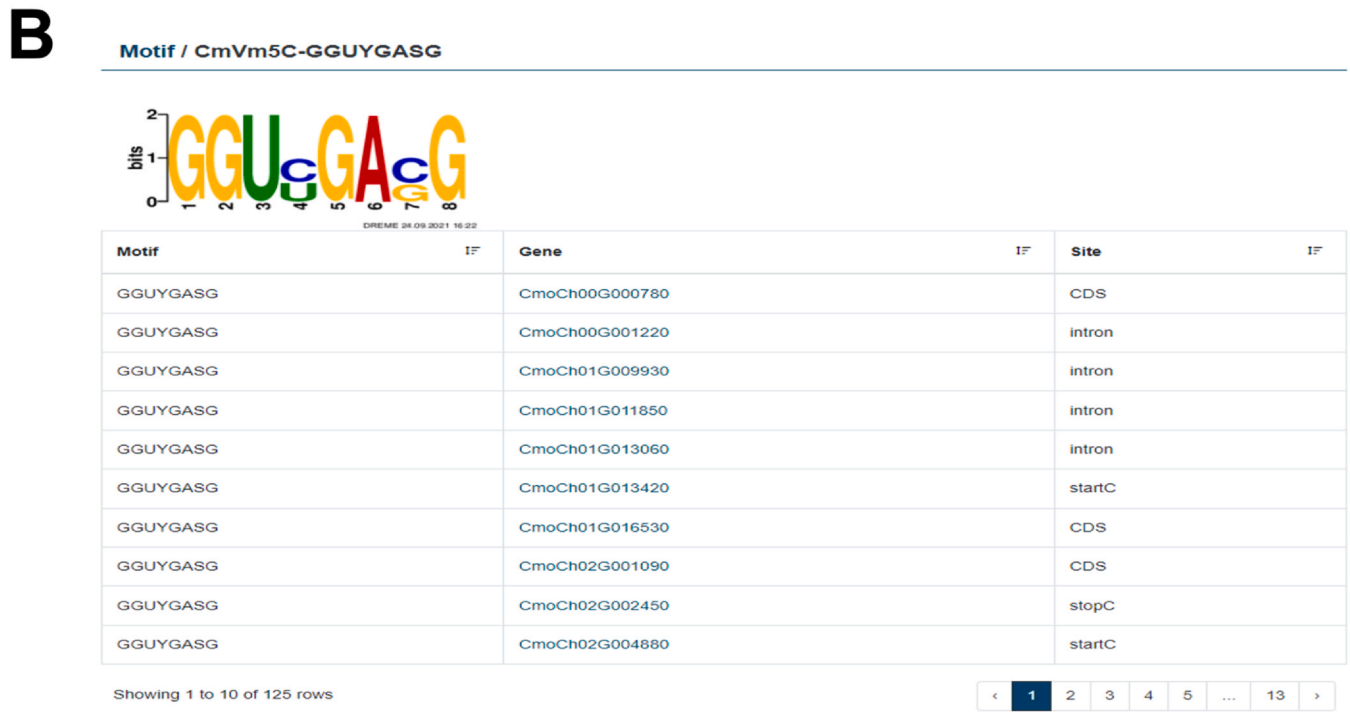
4.6. Implementation

The Cucume database was deployed in the Ubuntu 16.04 operating system and developed by AKKA 2.6.5 (web server), MySQL 5.7 (database server), Scala 2.13.2 and SBT 1.3.9. All data in the database were managed and stored using the MySQL Database Management System. The query function was enforced based on Slick3.3.2 middleware tier. To visualize the genome, we used JBrowse 1.16.6 [64]. The website interface components were designed and implemented by Bootstrap4.6.0 and Play Framework 2.8.7. The website has been tested in several popular web browsers, including Firefox, Google Chrome and Internet Explorer. Source code is available on Gitee (https://gitee.com/moxic/cap_db).

CRediT authorship contribution statement

Xiaojun Li, Shujin Lin, Xiao Han and Wenna Zhang wrote the original manuscript. Xiaojun Li, Shujin Lin, Xiao Han and Wenna Zhang analyzed data. Xiaojing Zhang, Tao Wang, Xiaohong Lu, Wenqian Liu, Naonao Wang, Cuicui Wang, Zixi Liu, Mengshuang Liu and Linhong Gao provided the resources and references. Xiao Han and Wenna Zhang revised the paper. All authors read the final version of this manuscript and approved it for publication.





Motif	Gene	Site
GGUYGASG	CmoCh00G000780	CDS
GGUYGASG	CmoCh00G001220	intron
GGUYGASG	CmoCh01G009930	intron
GGUYGASG	CmoCh01G011850	intron
GGUYGASG	CmoCh01G013060	intron
GGUYGASG	CmoCh01G013420	startC
GGUYGASG	CmoCh01G016530	CDS
GGUYGASG	CmoCh02G001090	CDS
GGUYGASG	CmoCh02G002450	stopC
GGUYGASG	CmoCh02G004880	startC

Showing 1 to 10 of 125 rows

‹ 1 2 3 4 5 ... 13 ›

Fig. 6. The "Statistics" interface. (A) The relationship between m^5C RNA methylation and mRNA mobility in different tissues. The proportion of grafting-transmissible RNAs harboring m^5C methylation in the vascular exudate was higher either in cucumber or pumpkin. Cm, Pumpkin. Cs, Cucumber. G, mixed tissues of adult plants. V, vascular exudates. (B) Pumpkin m^5C RNAs from vascular exudates harboring the motif "*GGUNGANG*". The detailed motif of each methylated RNA sample is available with hyperlinks to the DREME program (http://meme-suite.org/doc/dreme.html).

Conflict of interests

The authors declare no conflict of interests.

Acknowledgments

We thank ikann-editorial team for language editing and writing suggestions. This work was supported by grants from the National Key Research and Development Program of China, China, 2019YFD1000300; 2115 Talent Development Program of China Agricultural University, China; Fuzhou University Research Start-up Project, China, GXRC-19025R; Fundamental Research Funds of China Agricultural University, China, 2022TC171.

Ethics approval and consent to participate

This paper was completed within the laws of the People's Republic of China. No specific permits were required for our field

research. The study species is not included in the 'List of Protected Plants in China'.

Consent for publication

Not applicable.

Supplementary data

The parameters of the bioinformatics analysis used in Methylated RNA immunoprecipitation (MeRIP)-sequencing and peak analysis.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2023.01.012.

References

[1] Zhao X, Yang Y, Sun BF, Shi Y, Yang X, Xiao W, et al. FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res 2014;24:1403–19.
[2] Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, et al. Nuclear m(6)A reader YTHDC1 regulates mRNA splicing. Mol Cell 2016;61:507–19.
[3] Holmes MJ, Padgett LR, Bastos MS, Sullivan WJ. m6A RNA methylation facilitates pre-mRNA 3′-end formation and is essential for viability of Toxoplasma gondii. Plos Pathog 2021:**17**.
[4] Srinivas KP, Depledge DP, Abebe JS, Rice SA, Mohr I, Wilson AC. Widespread remodeling of the m(6)A RNA-modification landscape by a viral regulator of RNA processing and export. Proc Natl Acad Sci U S A 2021:**118**.
[5] Ke SD, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, et al. A majority of m (6)A residues are in the last exons, allowing the potential for 3 ' UTR regulation. Gene Dev 2015;29:2037–53.
[6] Yue Y, Liu J, Cui X, Cao J, Luo G, Zhang Z, et al. VIRMA mediates preferential m (6)A mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. Cell Discov 2018;4:10.
[7] Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, et al. 5 ' UTR m (6)A promotes cap-independent translation. Cell 2015;163:999–1010.
[8] Slobodin B, Han RQ, Calderone V, Vrielink JAFO, Loayza-Puch F, Elkon R, et al. Transcription impacts the efficiency of mRNA translation via Co-transcriptional N6-adenosine methylation. Cell 2017;169:326–37.
[9] Luo JH, Wang Y, Wang M, Zhang LY, Peng HR, Zhou YY, et al. Natural variation in RNA m(6)A methylation and its relationship with translational status. Plant Physiol 2020;182:332–44.
[10] Kim J, Lee G. Metabolic control of m(6)A RNA modification. Metabolites 2021:**11**.
[11] Wang X, Lu ZK, Gomez A, Hon GC, Yue YN, Han DL, et al. N-6-methyladenosine-dependent regulation of messenger RNA stability. Nature 2014;505. 117-+.
[12] Shi HL, Wang X, Lu ZK, Zhao BXS, Ma HH, Hsu PJ, et al. YTHDF3 facilitates translation and decay of N-6-methyladenosine-modified RNA. Cell Res 2017;27:315–28.
[13] Huang HL, Weng HY, Sun WJ, Qin X, Shi HL, Wu HZ, et al. Recognition of RNA N-6- methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol 2018;20. 285-+.
[14] Bodi Z, Zhong SL, Mehra S, Song J, Graham N, Li HY, et al. Adenosine methylation in Arabidopsis mRNA is associated with the 3 ' end and reduced levels cause developmental defects. Front Plant Sci 2012:**3**.
[15] Shen LS, Liang Z, Gu XF, Chen Y, Teo ZWN, Hou XL, et al. N-6-Methyladenosine RNA modification regulates shoot stem cell fate in arabidopsis. Dev Cell 2016;38:186–200.
[16] Yu Q, Liu S, Yu L, Xiao Y, Zhang SS, Wang XP, et al. RNA demethylation increases the yield and biomass of rice and potato plants in field trials. Nat Biotechnol 2021;39:1581.
[17] Jian H, Zhang C, Qi Z, Li X, Lou Y, Kang Y, et al. Alteration of mRNA 5-methylcytosine modification in neurons after OGD/R and potential roles in cell stress response and apoptosis. Front Genet 2021;12:633681.
[18] Cui X, Liang Z, Shen L, Zhang Q, Bao S, Geng Y, et al. 5-Methylcytosine RNA methylation in Arabidopsis thaliana. Mol Plant 2017;10:1387–99.
[19] David R, Burgess A, Parker B, Li J, Pulsford K, Sibbritt T, et al. Transcriptome-wide mapping of RNA 5-methylcytosine in arabidopsis mRNAs and noncoding RNAs. Plant Cell 2017;29:445–60.
[20] Yang L, Perrera V, Saplaoura E, Apelt F, Bahin M, Kramdi A, et al. m(5)C methylation guides systemic transport of messenger RNA over graft junctions in plants. Curr Biol 2019;29. 2465-2476 e2465.
[21] Edelheit S, Schwartz S, Mumbach MR, Wurtzel O, Sorek R. Transcriptome-wide mapping of 5-methylcytidine RNA Modifications in Bacteria, Archaea, and Yeast Reveals m(5)C within Archaeal mRNAs. Plos Genet 2013:**9**.
[22] Amort T, Rieder D, Wille A, Khokhlova-Cubberley D, Riml C, Trixl L, et al. Distinct 5-methylcytosine profiles in poly(A) RNA from mouse embryonic stem cells and brain. Genome Biol 2017:**18**.
[23] Hussain S, Sajini AA, Blanco S, Dietmann S, Lombard P, Sugimoto Y, et al. NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep 2013;4:255–61.
[24] Khoddami V, Cairns BR. Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat Biotechnol 2013;31:458.
[25] Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, et al. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 2012;40:5023–33.
[26] Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, et al. 5-methylcytosine promotes mRNA export-NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res 2017;27:606–25.
[27] Burgess AL, David R, Searle IR. Conservation of tRNA and rRNA 5-methylcytosine in the kingdom Plantae. Bmc Plant Biol 2015:**15**.
[28] Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, et al. MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res 2018;46:D303–7.
[29] Agris PF. Bringing order to translation: the contributions of transfer RNA anticodon-domain modifications. Embo Rep 2008;9:629–35.
[30] Chow CS, Larnichhane TN, Mahto SK. Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. Acs Chem Biol 2007;2:610–9.
[31] Zhong SL, Li HY, Bodi Z, Button J, Vespa L, Herzog M, et al. MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. Plant Cell 2008;20:1278–88.
[32] Arribas-Hernandez L, Bressendorff S, Hansen MH, Poulsen C, Erdmann S, Brodersen P. An m(6)A-YTH module controls developmental timing and morphogenesis in Arabidopsis([OPEN]). Plant Cell 2018;30:952–67.
[33] Duan HC, Wei LH, Zhang C, Wang Y, Chen L, Lu ZK, et al. ALKBH10B Is an RNA N-6-methyladenosine demethylase affecting arabidopsis floral transition. Plant Cell 2017;29:2995–3011.
[34] Hu J, Manduzio S, Kang H. Epitranscriptomic RNA methylation in plant development and abiotic stress responses. Front Plant Sci 2019;10:500.
[35] Miao Z, Zhang T, Qi Y, Song J, Han Z, Ma C. Evolution of the RNA N (6)-methyladenosine methylome mediated by genomic duplication. Plant Physiol 2020;182:345–60.
[36] Gao YJ, Islam MS, Tian J, Lui VWY, Xiao D. Inactivation of ATP citrate lyase by Cucurbitacin B: A bioactive compound from cucumber, inhibits prostate cancer growth. Cancer Lett 2014;349:15–25.
[37] Zhang B, Tolstikov V, Turnbull C, Hicks LM, Fiehn O. Divergent metabolome and proteome suggest functional independence of dual phloem transport systems in cucurbits. Proc Natl Acad Sci U S A 2010;107:13532–7.
[38] Zhang C, Yu X, Ayre BG, Turgeon R. The origin and composition of cucurbit "phloem" exudate. Plant Physiol 2012;158:1873–82.
[39] Sui X, Nie J, Li X, Scanlon MJ, Zhang C, Zheng Y, et al. Transcriptomic and functional analysis of cucumber (Cucumis sativus L.) fruit phloem during early development. Plant J 2018;96:982–96.
[40] Lu XH, Liu WQ, Wang T, Zhang JL, Li XJ, Zhang WN. Systemic long-distance signaling and communication between rootstock and scion in grafted vegetables. Front Plant Sci 2020:**11**.
[41] Warschefsky EJ, Klein LL, Frank MH, Chitwood DH, Londo JP, von Wettberg EJB, et al. Rootstocks: diversity, domestication, and impacts on shoot phenotypes. Trends Plant Sci 2016;21:418–37.
[42] Omid A, Keilin T, Glass A, Leshkowitz D, Wolf S. Characterization of phloem-sap transcription profile in melon plants. J Exp Bot 2007;58:3645–56.
[43] Ham BK, Brandom JL, Xoconostle-Cazares B, Ringgold V, Lough TJ, Lucas WJ. A polypyrimidine tract binding protein, pumpkin RBP50, forms the basis of a phloem-mobile ribonucleoprotein complex. Plant Cell 2009;21:197–215.
[44] Zhang WN, Thieme CJ, Kollwig G, Apelt F, Yang L, Winter N, et al. tRNA-related sequences trigger systemic mRNA transport in plants. Plant Cell 2016;28:1237–49.
[45] Ruiz-Medrano R, Xoconostle-Cazares B, Lucas WJ. Phloem long-distance transport of CmNACP mRNA: implications for supracellular regulation in plants. Development 1999;126:4405–19.
[46] Haywood V, Yu TS, Huang NC, Lucas WJ. Phloem long-distance trafficking of Gibberellic acid-insensitive RNA regulates leaf development. Plant J 2005;42:49–68.
[47] Huang NC, Yu TS. The sequences of Arabidopsis GA-INSENSITIVE RNA constitute the motifs that are necessary and sufficient for RNA long-distance trafficking. Plant J 2009;59:921–9.
[48] Liu WQ, Xiang CG, Li XJ, Wang T, Lu XH, Liu ZX, et al. Identification of Long-Distance Transmissible mRNA between Scion and Rootstock in Cucurbit Seedling Heterografts. Int J Mol Sci 2020:**21**.
[49] Wang YH, Bo KL, Gu XF, Pan JS, Li YH, Chen JF, et al. Molecularly tagged genes and quantitative trait loci in cucumber with recommendations for QTL nomenclature. Hortic Res-Engl 2020:**7**.
[50] Guan, DG, Yan, B, Thieme, C, Hua, JM, Zhu, HL, Boheler, KR, et al. (2020) PlaMoM: a comprehensive database compiles plant mobile macromolecules (vol 45, pg D1021, 2017). Nucleic Acids Res, 48, 7607–7607.
[51] Sun H, Wu S, Zhang G, Jiao C, Guo S, Ren Y, et al. Karyotype stability and unbiased fractionation in the paleo-allotetraploid cucurbita genomes. Mol Plant 2017;10:1293–306.
[52] Zheng Y, Wu S, Bai Y, Sun H, Jiao C, Guo S, et al. Cucurbit Genomics Database (CuGenDB): a central portal for comparative and functional genomics of cucurbit crops. Nucleic Acids Res 2019;47:D1128–36.
[53] Zhang SD, Sun L, Kragler F. The phloem-delivered RNA pool contains small noncoding RNAs and interferes with translation. Plant Physiol 2009;150:378–87.

[54] Chomczynski P. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. Biotechniques 1993;15:532.
[55] Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3 ' UTRs and near stop codons. Cell 2012;149:1635–46.
[56] Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J 2011;17:10–2.
[57] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–9.
[58] Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015;12:357–60.
[59] Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 2015;33:290–5.
[60] Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9:R137.
[61] Chen T, Chen X, Zhang S, Zhu J, Tang B, Wang A, et al. The genome sequence archive family: toward explosive data growth and diverse data types. Genom Proteom Bioinform 2021;19:578–83.
[62] Members C-N, Partners. Database resources of the national genomics data center, China National Center for bioinformation in 2022. Nucleic Acids Res 2022;50:D27–38.
[63] Bailey TL. DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics 2011;27:1653–9.
[64] Buels R, Yao E, Diesh CM, Hayes RD, Munoz-Torres M, Helt G, et al. JBrowse: a dynamic web platform for genome visualization and analysis. Genome Biol 2016:**17**.