

MoRNIING: A Database of RNA Modification Sites Associated with RNA Secondary Structure Dynamics

Yicen Zhou (周奕岑)^{1, #, }, Shanxin Lyu (吕善鑫)^{1, #, }, Shiau Wei Liew (刘晓薇)^{2, #, }, Xi Mou (牟希)^{2, },
Ian Hoffecker^{3, }, Jian Yan (严健)^{1, 4, 5, }, Yu Li (李煜)^{6, }, Chun Kit Kwok (郭骏杰)^{2, 7, *, },
Jilin Zhang (张继林)^{1, 4, 5, *, }

¹Department of Biomedical Sciences, College of Biomedicine, City University of Hong Kong, Hong Kong Special Administrative Region 999077, China

²Department of Chemistry and State Key Laboratory of Marine Environmental Health, City University of Hong Kong, Hong Kong Special Administrative Region 999077, China

³Department of Gene Technology, KTH Royal Institute of Technology, SciLifeLab, Solna 17165, Sweden

⁴Tung Biomedical Sciences Centre, City University of Hong Kong, Hong Kong Special Administrative Region 999077, China

⁵Department of Precision Diagnostic and Therapeutic Technology, City University of Hong Kong, Shenzhen Futian Research Institute, Shenzhen 518000, China

⁶Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong Special Administrative Region 999077, China

⁷Shenzhen Research Institute of City University of Hong Kong, Shenzhen 518000, China

*Corresponding authors: jilin.zhang@cityu.edu.hk (Zhang J), ckkwok42@cityu.edu.hk (Kwok CK).

#Equal contribution.

Handling Editor: Qiangfeng Cliff Zhang

Abstract

RNA structures are essential building blocks of functional RNA molecules. Profiling secondary structures *in vivo* and in real time remains challenging because RNAs exhibit dynamic structures and complex conformations. In addition to the canonical stem-loop secondary structure, the non-canonical RNA G-quadruplex (rG4) structure has attracted interest for its potential as a drug target. Early studies have demonstrated that RNAs can form distinct secondary structures. However, how distinct RNA structures formed from the same RNA sequence function within the transcriptome is poorly understood, and the factors that drive and regulate structural transitions remain to be investigated. Inspired by the ability of a *HOXB9* segment to form multiple structures, we found that many RNA segments across the transcriptome exhibit multi-faceted structure-forming potential. In the case of *HOXB9*, we demonstrated that *N*⁶-methyladenosine (m⁶A) modification influences RNA structure and binding to RNA-binding proteins (RBPs). Therefore, we collected RNA modification sites naturally occurring within the putative G-quadruplex-forming sequences (PQSs) of transcripts and developed MoRNIING, a database for RNA modifications in natural rG4 structures. MoRNIING is organized into reliability tiers determined by the resolution of RNA modification sites and is designed to accommodate various large datasets. We experimentally validated the influence of m⁶A, 5-methylcytosine (m⁵C), and adenosine-to-inosine (A-to-I) editing on rG4-forming sequences, providing evidence to support the modification switch concept. The diversity and transition of secondary structures from the same RNA segment offer valuable insights into the regulation of RNA structural dynamics. MoRNIING is freely accessible at <https://www.cityu.edu.hk/bms/morning>.

Key words: RNA G-quadruplex; RNA modification; RNA secondary structure; RNA-binding protein; Conformation dynamics.

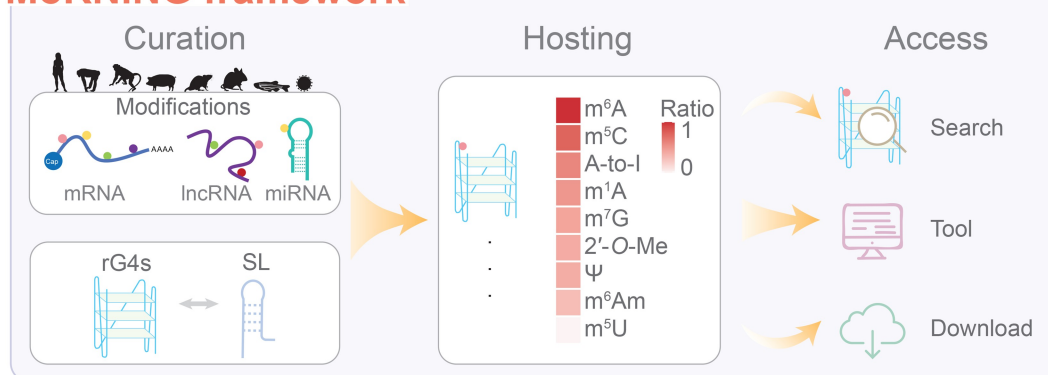
Received: 2 May 2025. Revised: 8 October 2025. Accepted: 11 November 2025.

© The Author(s) 2025. Published by Oxford University Press and Science Press on behalf of the Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation and Genetics Society of China.

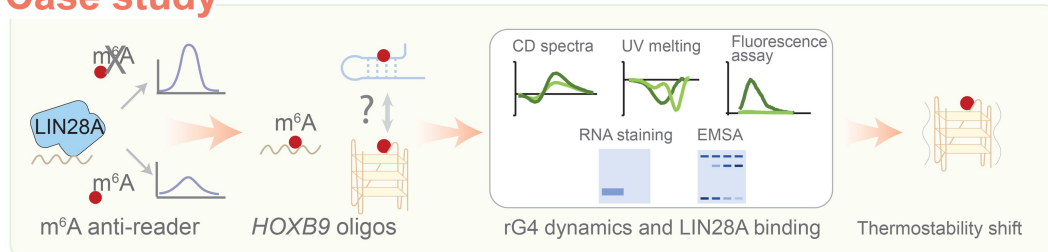
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Graphical abstract

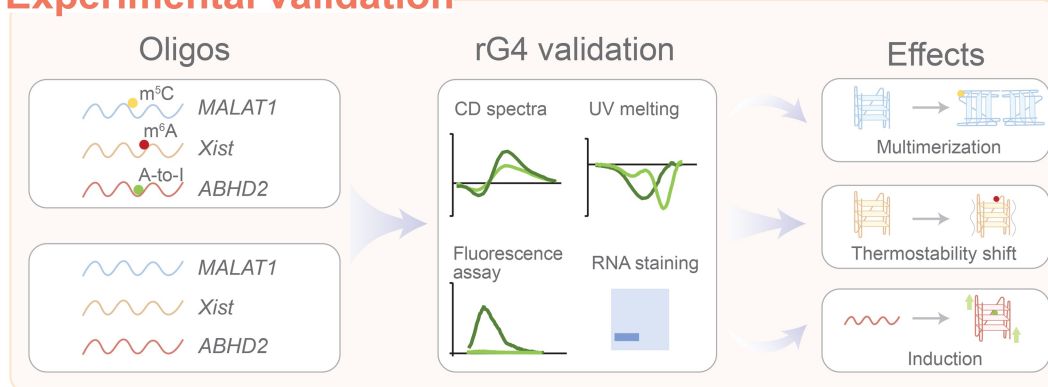
MoRNING framework



Case study



Experimental validation



Introduction

RNA molecules are instrumental in many biological processes [1]. Understanding their functional roles is crucial for constructing a comprehensive regulatory map. One of the significant challenges hindering this process is the elusive information that determines the transition between linear sequences and structural RNAs, as well as among different RNA structures.

RNAs form a variety of structures, and a growing volume of RNA profiling data has demonstrated that many bases participate in canonical base pairing [2,3]. Beyond canonical secondary RNA conformations, such as stem-loop (SL) structures, internal loops, and multi-branch junctions, the non-canonical structure RNA G-quadruplex (rG4) can form when RNA strands fold to stack more than two planar G-tetrads, with each tetrad consisting of four guanines (Gs) connected via Hoogsteen hydrogen bonds and stabilized by a monovalent cation [4,5]. Since its

discovery, the rG4 structure has gained recognition as a potential functional element in both transcriptional and post-transcriptional regulation. Numerous assays have been devised and refined to profile rG4s *in vitro* and *in vivo* [4,5], offering unprecedented transcriptome-wide insights into the distribution and function of rG4s. However, the rG4 landscape and dynamics require substantial exploration to gain a comprehensive understanding.

The flexible nature of single-stranded RNAs allows them to dynamically change conformation to interact with RNA-binding proteins (RBPs) with structural specificity [6,7]. A typical *in vivo* example is LIN28A, which binds to the pre-*let-7* terminal SL structure through its cold shock domain to inhibit the function of miRNA *let-7* [8]. This protein can also unwind rG4 in mRNAs to promote ribosomal scanning and enhance translation [9]. The polycomb repressive complex 2 (PRC2) interacts with rG4 in *Xist*, which is crucial to silence genes on chromosome X in a stepwise

manner [10]. These cases emphasize the importance of structural RNA in regulating gene expression. In addition to previously reported synthetic ligands, evidence suggests that a single RNA sequence can adopt multiple distinct structures, linking the biological function of RNA *in vivo* to structural specificity [11,12]. Moreover, the *N*⁶-methyladenosine (m⁶A) modification at the LIN28A-binding site on *HOXB9* mRNA can reduce the binding affinity of LIN28A, suggesting a potential alteration in RNA secondary structure. As RNA structures across cellular states remain incompletely characterized and the factors driving conformational changes are still poorly understood, further investigation is needed to elucidate transitions between RNA secondary structures.

RNA modifications, such as m⁶A and 5-methylcytosine (m⁵C), are chemical changes that can alter the properties of local RNA sequences and their corresponding RNA structures [13–16]. Indeed, these modifications play vital roles in regulating the formation and stability of rG4s [17,18]. In the case of miRNA *let-7*, *N*⁷-methylguanine (m⁷G) modification changes the equilibrium between SL and rG4, enabling DROSHA binding and promoting miRNA maturation [19]. Notably, the key component of the methyltransferase complex, METTL14, prefers binding to rG4s [18], underscoring the role of rG4s in regulating m⁶A-dependent biological processes. Yet, the impact of RNA modifications on RNA secondary structures remains poorly characterized and requires further investigation. Despite the need to increase the spectrum and resolution of RNA modification profiling, gaining a large volume of rG4 profiles still requires substantial effort. The emerging consequences of RNA modifications in reshaping RNA structures should also be considered to gain deeper insights into RNA function.

To bridge the knowledge gap between RNA epigenetics and structural dynamics, we present Modification of RNAs in Natural rG4 (MoRNING), a database with a curated compendium of rG4-forming sequences containing various cellular RNA modification sites, partially supported by experimental evidence. Although the diversity of rG4s is not clearly defined at this stage, we provide evidence supporting the observation that RNA modifications can modulate rG4 structures. The database design is structured to accommodate future expansion and ensure compatibility with diverse data types, laying the foundation for a more comprehensive understanding of RNA structural dynamics.

Data collection and processing

Annotation of rG4 and PQS

Three types of rG4 datasets were included to characterize transcriptome-wide rG4 features or identify potential rG4 loci. These datasets include *in vivo* experimental data such as rG4-seq and ultra-low-input rG4-seq [20–23], *in vitro* circular dichroism (CD) spectra [24], and rG4 prediction results generated by several computational tools, including pqsfinder [25], QGRS Mapper [26], G4Hunter (G4H) [27], G4RNA screener [28], RNAfold [29], G4 Neural Network (G4NN; score > 0.5), and the Quadparser-based G₂L₁₂ motif [defined as (G_{≥2}N_{1–12})₃G_{≥2}] [30]. As the detection capacity and scope of *in vivo* approaches are

usually limited, we applied a hierarchical prediction scheme to obtain consensus rG4 loci, thereby strengthening reliability and reducing false-positive loci. Following the primary predictions using pqsfinder [25] and QGRS Mapper [26], the extracted consensus putative G-quadruplex-forming sequences (PQSs) were further examined with additional prediction tools and validated by experimental evidence to enhance reliability. This process resulted in annotated tags based on the evidence sources (File S1). Additionally, putative GI-quadruplex-forming sequences (PIQSs), a type of G-quartet induced by the substitution of G with inosine (I) through RNA editing in the human transcriptome, were included [31]. As not all PIQSs predicted by pqsfinder (score ≥ 10) originated from the longest transcripts, a separate “modType” option was introduced on the search page.

The distribution of all PQSs within mRNAs and long non-coding RNAs (lncRNAs) was visualized using the Guitar R package [32], with the annotation of 5′ untranslated region (UTR), coding sequence (CDS), and 3′ UTR. The number of PQSs per transcript was normalized using the following formula: normalized count = (PQS count × 1000) / transcript length. Kernel density estimation was performed using a Gaussian kernel.

Prediction of SL structure embedded in PQSs

PQSs in the longest transcripts, including mRNAs and lncRNAs annotated in the Ensembl database, miRNAs shared between the Ensembl and miRBase databases [33], and all remaining RNA types (Others), were subjected to RNAfold [29] to predict the canonical SL structures.

Collection of RNA modification sites

To obtain a panoramic view of the RNA modification dataset associated with genes, nine types of modification sites were collected, including m⁶A, m⁵C, 5-methyluridine (m⁵U), *N*⁶,2′-O-dimethyladenosine (m⁶Am), m⁷G, *N*¹-methyladenosine (m¹A), 2′-O-methylation (2′-O-Me), adenosine-to-inosine (A-to-I) editing, and pseudouridine located within exons of the longest transcripts. Data were sourced from the Gene Expression Omnibus (GEO), RMBase, RMVar, MiREDiBase, m⁶A-Atlas, and other public datasets [34–37], covering nine species to construct a database illustrating the interplay between RNA modifications and rG4. Modification sites annotated on older genome assemblies were converted to the latest University of California, Santa Cruz (UCSC) or National Center for Biotechnology Information (NCBI) genome versions using LiftOver (Table 1), including human (UCSC hg38), mouse (UCSC mm39), rat (UCSC rn7), rhesus monkey (UCSC rheMac10), chimpanzee (UCSC panTro5), pig (UCSC susScr11), zebrafish (UCSC danRer11), human immunodeficiency virus type 1 (HIV-1; NCBI NC_001802.1), and simian virus 40 (SV40; NCBI NC_001669.1). Gene annotations of corresponding genomes were downloaded from the Ensembl database (release 109). Transcript annotations of mRNA and lncRNA from Ensembl and miRNA from miRBase (v22) were considered [33]. The modification sites of the two RNA viruses were based on their genomic RNA or mRNA.

RNA modification sites were categorized into three confidence levels: high, medium, and low, according to experimental or

Table 1 Data sources and summary of studied species

Species	Genome	Annotation	No. of modifications on transcripts	No. of unique transcripts	No. of unique PQSs on transcripts
Human	UCSC hg38	Ensembl (release 109)	m ⁶ A: 34,527 m ¹ A: 515 m ⁵ C: 5596 m ⁷ G: 208 m ⁵ U: 3 m ⁶ Am: 59 2'-O-Me: 163 Ψ: 103 A-to-I editing: 1165	mRNA: 12,400 lncRNA: 3057 pre-miRNA: 18 miRNA: 11	35,189
Mouse	UCSC mm39	Ensembl (release 109)	m ⁶ A: 25,153 m ¹ A: 3 m ⁵ C: 2041 m ⁷ G: 19 Ψ: 38 A-to-I editing: 123	mRNA: 10,304 lncRNA: 1102 pre-miRNA: 4 miRNA: 1	23,666
Rhesus monkey	UCSC rheMac10	Ensembl (release 109)	m ⁶ A: 2225	mRNA: 1474	1994
Chimpanzee	UCSC panTro5	Ensembl (release 109)	m ⁶ A: 2521	mRNA: 1665 lincRNA: 2	2254
Pig	UCSC susScr11	Ensembl (release 109)	m ⁶ A: 7899	mRNA: 4334 lncRNA: 11	7130
Rat	UCSC rn7	Ensembl (release 109)	m ⁶ A: 4320 m ¹ A: 1 m ⁵ C: 1 m ⁷ G: 1 Ψ: 2	mRNA: 2690 lncRNA: 12	3874
Zebrafish	UCSC danRer11	Ensembl (release 109)	m ⁶ A: 1373	mRNA: 1032 lincRNA: 6	1249
HIV-1	NCBI NC_001802.1		m ⁶ A: 2	Genomic RNA: 2	2
SV40	NCBI NC_001669.1		m ⁶ A: 1	Late mRNA: 1	1

Note: UCSC, University of California, Santa Cruz; NCBI, National Center for Biotechnology Information; PQS, putative G-quadruplex-forming sequence; m⁶A, N⁶-methyladenosine; m¹A, N¹-methyladenosine; m⁵C, 5-methylcytidine; m⁷G, N⁷-methylguanosine; m⁵U, 5-methyluridine; m⁶Am, N⁶,2'-O-dimethyladenosine; 2'-O-Me, 2'-O-methylation; Ψ, pseudouridine; A-to-I editing, adenosine-to-inosine RNA editing; HIV-1, human immunodeficiency virus type 1; SV40, simian virus 40; mRNA, messenger RNA; miRNA, microRNA; lncRNA, long non-coding RNA; lincRNA, long intergenic non-coding RNA.

prediction methods (Table S1). The high confidence level reflects modification sites detected at single-nucleotide resolution with experimental evidence, while the low confidence level denotes RNA modifications derived from either databases or predictions. Information regarding the source of modification sites and corresponding publications is incorporated into detailed pages in the MoRNING database (Table S2).

Data processing and analysis in the database

RNA modification sites and candidate rG4 loci were processed to obtain base-resolution RNA modification sites located within rG4 or potential rG4-forming sequences using BEDTools [38]. Each modification site was assigned a unique ID following the nomenclature format: species_modification_unique identifier. The locations of modification sites were further classified into

loop, G-tetrad, and bulge to reflect their positions within the structures.

Database implementation and content

Database implementation

The database was implemented using SQL Server to store RNA modifications and PQSs from nine species in separate tables at the back end. The web interface was developed using Hypertext Markup Language (HTML), Cascading Style Sheets (CSS), and JavaScript (JS). Active Server Pages (ASP) scripts were employed to query the SQL Server back-end database through JS by AJAX forms or Fetch API. For real-time statistical information, interactive plots were created by ECharts or D3.js. To display the detailed information of each genomic locus, an embedded

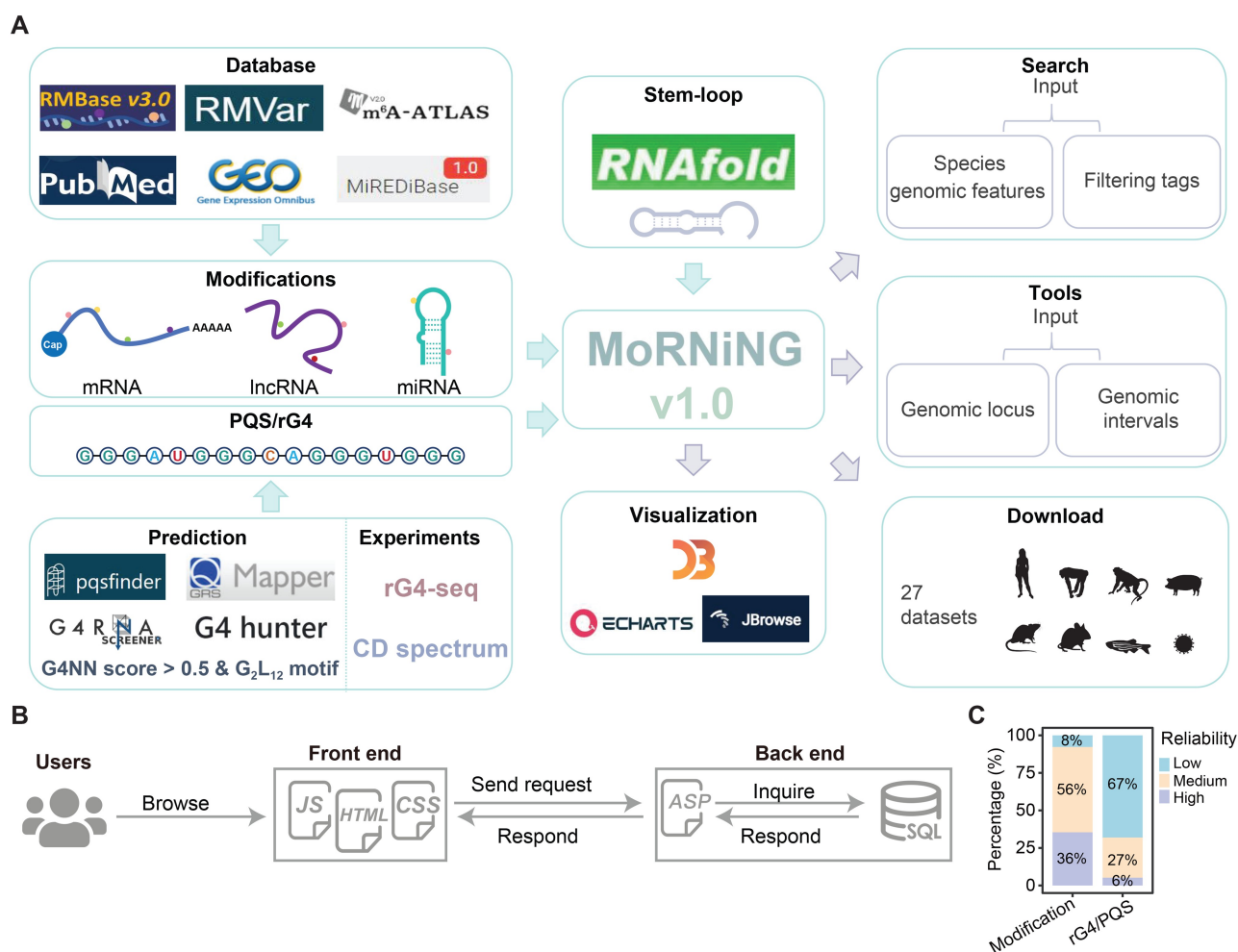


Figure 1 Architecture of MoRNING

A. Schematic illustration of the data sources, approaches, and functions of the MoRNING database. The sources of RNA modifications, tools, and experiments for rG4s and PQSs are shown in the left panel, while the major functions are presented in the center and right panels. The tools and datasets are summarized on the right. **B.** Architecture and modular implementation of the front and back ends. The major techniques used to implement front-end and back-end functions are depicted in groups. **C.** Reliability level of RNA modification sites and rG4s/PQSs cataloged in the database. MoRNING, Modification of RNAs in Natural rG4; lncRNA, long non-coding RNA; miRNA, microRNA; rG4, RNA G-quadruplex; PQS, putative G-quadruplex-forming sequence; CD, circular dichroism; JS, JavaScript; HTML, Hypertext Markup Language; CSS, Cascading Style Sheets; ASP, Active Server Pages; SQL, Structured Query Language.

JBrowse [39] was inserted into the corresponding subpages with a focused view at the selected locus.

Database content

In the database, we reported RNA modification loci within rG4s validated by experiments or PQSs predicted by several rG4 prediction tools. We collected nine types of RNA modifications from nine species to demonstrate the preference for modified sites in the transcriptome (Figure 1A and B; Table 1). The primary function of this database is to host RNA modification loci with tiered reliability, enabling researchers to explore the relationship between RNA modifications and rG4s and to reveal their potential biological impacts. To our knowledge, this dataset is a dedicated collection linking canonical and non-canonical RNA structures in natural RNAs. This work also proves that RNA

modifications can alter rG4-forming propensity, which may play a role in post-transcriptional regulation.

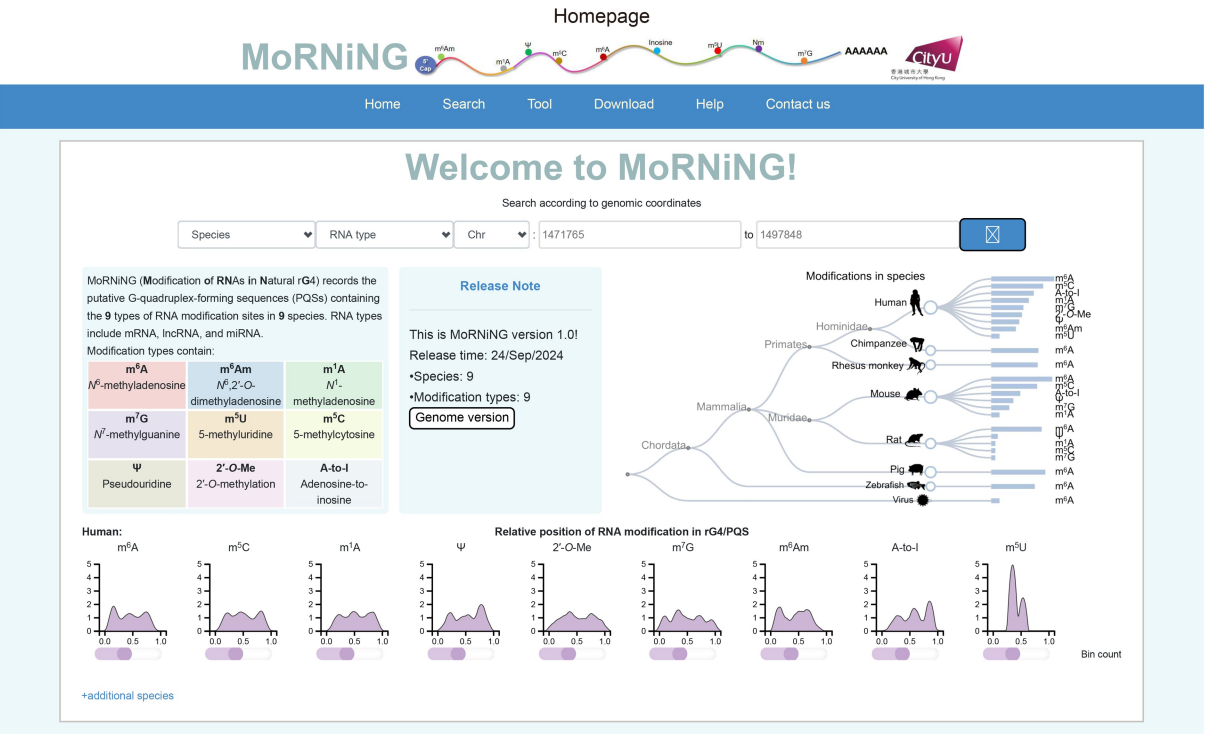
The MoRNING homepage features a real-time tree plot visualizing the constituent species and the number of cataloged RNA modification sites (Figure 2A). The positional preference plot of pronounced modification sites in each cataloged species is visualized. We implemented three primary pages, including search, download, and tool (Figure 2B, Figure S1), for users to navigate to distinct functions. A help page is offered to provide usage instructions, updates, and information on the corresponding data sources.

User interface

Browsing interface

The main page of the MoRNING database provides a summary of the total number of RNA modification sites in rG4s/PQSs across

A



B

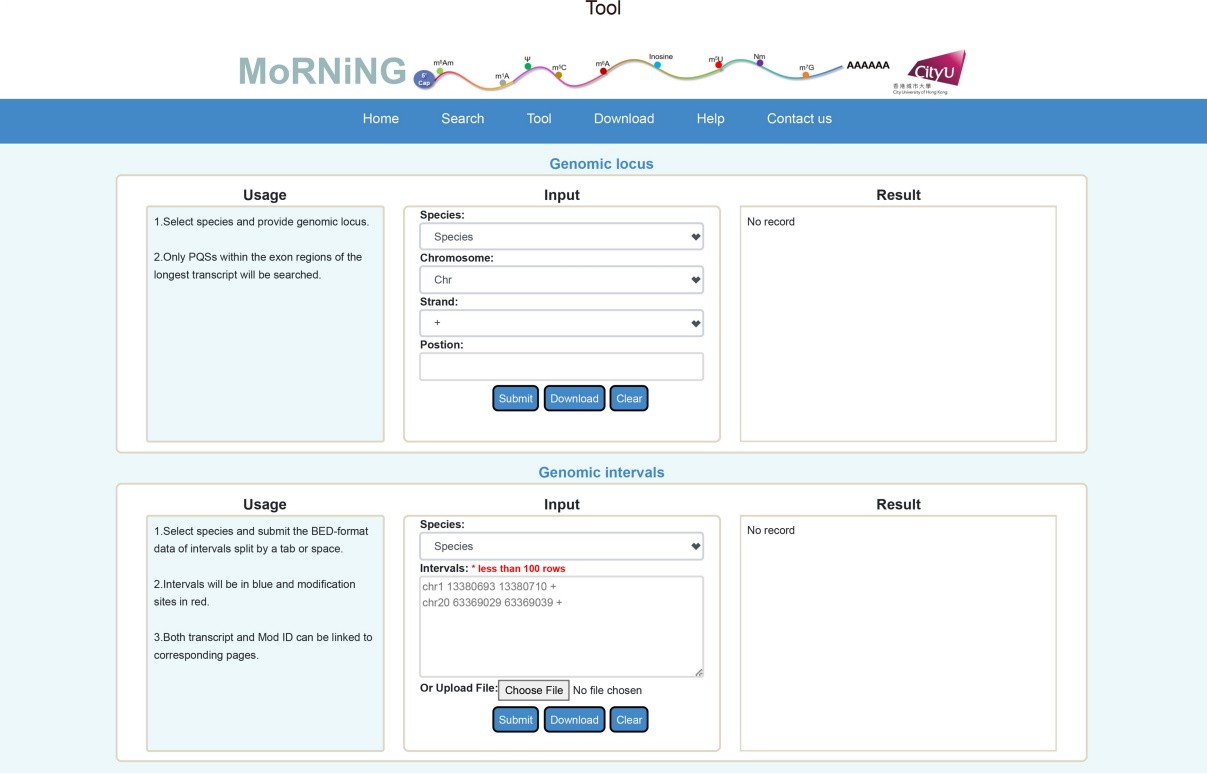


Figure 2 Interface and major functions of MoRNING
A. Snapshot of the main page. B. Snapshot of the tool function.

the studied species. A search engine is embedded within the interface, allowing users to access the database by specifying the species and genomic locus. The search page provides multiple

filters under eight categories, including species, chromosome, modification type, RNA type, and four additional options. Users can combine different filtering tags within each species to

retrieve desired RNA modifications associated with rG4s. The tool page offers two major functions that allow users to upload genomic loci or regions of interest with RNA modifications, either in locus-specific or batch mode. On the download page, formatted tables for each species can be downloaded.

Tiered reliability and interactive visualization

We classified RNA modification sites into three reliability tiers based on the resolution of corresponding experimental evidence or predictions. Loci in the high-reliability tier are supported by experiments at single-nucleotide resolution. Sites identified by methods that detect regional RNA modifications are classified as medium-reliability, while sites that are fully predicted or derived from the DARNED [40], RADAR [41], and MODOMICS [42] databases are assigned to the low-reliability tier. PIQs predicted by pqsfinder with scores ≥ 10 are also assigned to the low-reliability tier. The reliability information of sites supported by multiple sources is fully presented. Similarly, we introduced a system with tiered reliability to rG4s/PQs: loci with experimental evidence or supported by more than two prediction tools are labeled as high-reliability; loci supported by one additional prediction tool besides pqsfinder and QGRS Mapper are classified as medium-reliability; and loci in the low-reliability tier are those either supported by the consensus sequence predicted by pqsfinder and QGRS Mapper or obtained through the criteria of G4NN score > 0.5 and the G_2L_{12} motif (Figure 1C). Each RNA modification site includes modification type, genomic position, strand, gene symbol, Ensemble transcript ID, rG4/PQS coordinate, relative modification position within rG4/PQS, site type, detection method, and data source (Figure S2). An interactive genome browser, JBrowse, is embedded to visualize the rG4s within the annotated transcripts.

Tools for user-provided genomic loci or intervals

MoRNING provides web tools for preliminary analysis, including the extraction of RNA modification sites from experimentally validated rG4-forming loci and the identification of RNA modification-associated PQSs. These two tools allow users to obtain rG4s/PQs or RNA modification sites from the longest transcript of each gene (Figure S3). Each search job is executed instantly, and the results are visualized upon job completion.

Data download

All processed RNA modifications within rG4s of distinct RNA species can be downloaded through the browse page after applying selection criteria. Comprehensive modification sites and their information stored as zip files can be accessed on the download page for each species. The results generated by querying user-provided genomic loci can also be downloaded.

Case study

RNA modification affects the RNA secondary structure that interacts with LIN28A

As RNA modifications participate in regulating the dynamic interactions between RBPs and RNA structures, we investigated

the binding sites of RBP LIN28A, which acts as an m^6A anti-reader in mouse embryonic stem cells [17]. We discovered that one SL-forming RNA segment in *HOXB9*, documented in the MoRNING database and targeted by LIN28A, could also form an rG4, as predicted (Figure 3A and B). The prediction based on reactivity score, however, only indicated potential SLs (Figure S4A and B). Additionally, the profiles captured *in vivo* and *in vitro* did not align well within G-stretches, deviating from RNAfold prediction [29]. This shows that RNA structure prediction by RNAfold does not fully account for solution properties that could affect RNA folding, such as surrounding ions (Figure 3C). In the structure probing datasets, G-stretches are more frequently protected than other bases in PQSs, as Gs in G-tetrads show a statistically lower reactivity score than bases in the loop in several cell lines, including H9, HEK293, HeLa, HepG2, and K562 (Figure S4C). Although G tends to be more reactive in structure probing assays and the reactivity score reflects the structural average, G-tetrad scores are significantly lower than those of Gs in non-PQS regions (Figure S4D), implying that G-tetrads are generally protected in PQSs. To confirm the rG4 formation of this *HOXB9* RNA segment, synthesized RNA oligos were examined. A fully mutated sequence disrupting intramolecular rG4 formation was designed as a negative control. To demonstrate that rG4 can form from a sequence that also folds into an SL, we further designed two mutated sequences, Mut1_{GG2UU} and Mut2_{GG2UU}, based on the original oligo with the replacement of two bases, GG to UU, at the 5' end and 3' end, respectively (Figure 3A). Indeed, fluorescence assays using G4-specific ligands *N*-methylmesoporphyrin IX (NMM) and Thioflavin T (ThT) confirmed rG4 formation in the presence of K^+ (Figure 3D). In addition, the rG4 formed *in vitro* exhibited a complex composition according to the electrophoresis and fluorescence assays, as this RNA oligo contains G-tetrads capable of forming two types of rG4. The comparison of rG4s stained by G4-specific ligand ThT and total RNA stained with SYBR Gold provided strong evidence that both the original and Mut1_{GG2UU} sequences form rG4s (Figure 3E; Table S3). The original oligo likely forms several distinct rG4 structures in solution, given that several types of PQSs were predicted (Figure S4A), and the Mut1_{GG2UU} sequence yielded a much stronger signal in the fluorescence assay (Figure 3D). The CD and UV melting results for Mut2_{GG2UU} further confirmed the formation of compound rG4s (Figure S5). This finding is consistent with the observation that compound RNA conformations in solution exhibit lower rG4 absorption than pure rG4 conformation [43] (Figure S6).

Because m^6A modification occurs naturally in this RNA segment, we used a synthesized oligo carrying an m^6A modification to investigate rG4 formation. ThT staining demonstrated that m^6A reduces the rG4-forming potential under Li^+ conditions that destabilize rG4 (Figure 3E). This reduced rG4 stability is consistent with the results of UV melting experiment, which showed that m^6A decreases rG4 thermostability (Figure S7). The respective rG4 melting temperatures (Tms) of RNAs with and without m^6A were 55°C and 55.5°C in the forward scan, and 51°C and 53°C in the reverse scan. This suggests that RNA modifications affect rG4 folding.

Next, we examined the binding between rG4 and LIN28A to assess the influence of RNA modifications (Figure 3F, Figure S8). Despite their binding in the presence of K^+ , we discovered that the RNA carrying m^6A showed no significant difference in binding

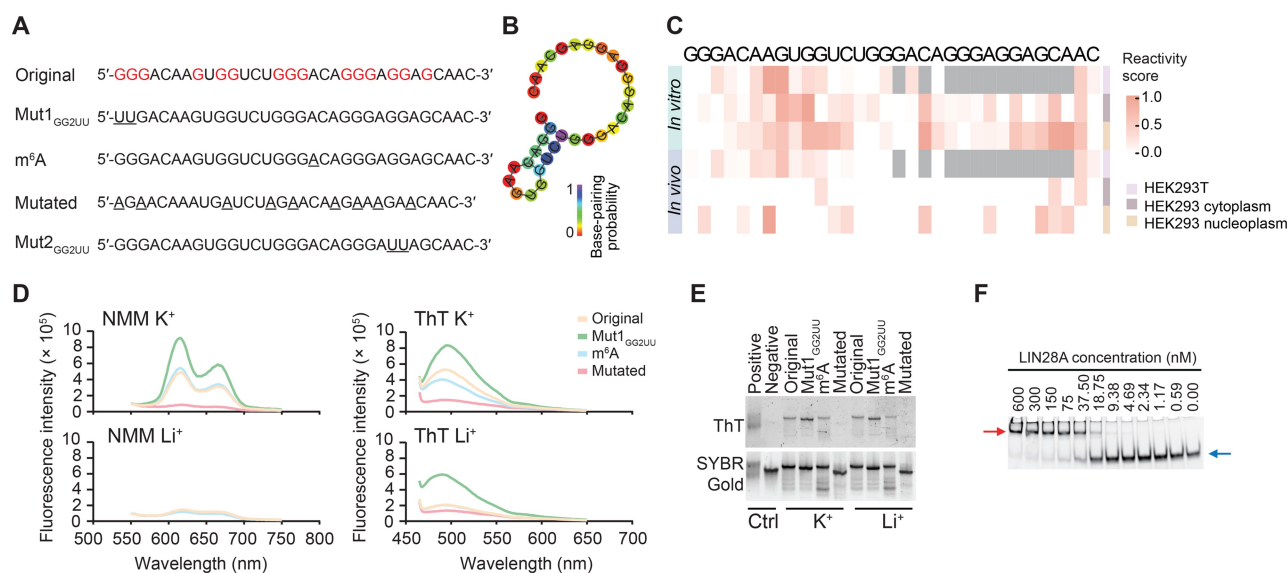


Figure 3 Structural diversity of the *HOXB9* mRNA segment

A. Sequences of the synthesized original *HOXB9* RNA oligo and its derivatives with mutations and m⁶A modification. Potential G-tetrads are highlighted in red in the original oligo, while UU replacement, m⁶A modification, and mutations in the fully mutated sequence are displayed with underlines. **B.** Predicted canonical SL structure of the original *HOXB9* RNA oligo by RNAfold. The base-pairing probability is visualized in color. **C.** Diverse icSHAPE reactivity scores of the original *HOXB9* RNA oligo. *In vivo* and *in vitro* profiles were captured from HEK293T cells and the cytoplasm and nucleoplasm of HEK293 cells. The reactivity score ranges from 0 to 1, with higher scores indicating greater single-strandedness. **D.** Fluorescence assay using G4-specific ligands ThT and NMM under K⁺ and Li⁺ conditions, respectively. **E.** rG4 formation of the original *HOXB9* RNA oligo and its derivatives stained by ThT and SYBR Gold. The conditions used to stabilize and destabilize rG4 are shown. The rG4-forming TERRA RNA and G-tetrad-mutated TERRA RNA were used as rG4 positive and negative controls, respectively. **F.** EMSA of the original *HOXB9* RNA oligo (10 nM, FAM-labeled) with increasing concentrations of LIN28A (0–600 nM) under K⁺ condition. Red and blue arrows indicate the LIN28A–RNA complex and free probe, respectively. SL, stem-loop; NMM, *N*-methylmesoporphyrin IX; ThT, Thioflavin T; TERRA, telomeric repeat-containing RNA; EMSA, electrophoretic mobility shift assay; FAM, fluorescein amidite; Ctrl, control.

affinity to LIN28A under the rG4 stabilizing condition compared to the RNA without m⁶A modification. This suggests that m⁶A potentially affects rG4 under conditions in which rG4 can easily unwind or compete with other conformations. The binding affinity between m⁶A-carrying RNA and LIN28A indeed declined under the Li⁺ condition (Figure S8). The comparison of electrophoretic mobility shift assay (EMSA) results under K⁺ and Li⁺ conditions indicates that rG4 enhances RNA binding to LIN28A. LIN28A could also bind to alternative RNA sequences or conformations, potentially influenced by K⁺. This finding aligns with the observation that m⁶A can slightly lower the rG4 T_m. This mild change enables the conformation to shift toward SL, as indicated by the ThT fluorescence assay (Figure 3D). The m⁶A-induced structural change can explain the decreased binding affinity between RNA and LIN28A when rG4 is the preferred binding conformation. Therefore, LIN28A can be captured as an m⁶A anti-reader in loci where rG4 and SL are nested. These results demonstrate that m⁶A can affect the conformation of rG4-containing sequences and their interactions with LIN28A.

Genome-wide existence of RNAs with multiple-structure forming potential

Rather than recording comprehensive maps of RNA structural dynamics, current RNA structure profiling assays, including SHAPE-Seq, icSHAPE, DMS–MaPseq, and rG4-seq [20,44–46], primarily provide static snapshots of structural landscapes in a few

experimental settings. It remains largely unknown how distinct secondary structures formed from the same locus undergo transition. We combined computational predictions and public datasets to investigate the potential transitions between distinct RNA structures. We focused on SL and rG4 structures because of their prevalence and increasingly available experimental datasets. We extracted PQSs from the longest transcripts of mRNA, lncRNA, and miRNA based on consensus PQSs predicted by pqsfinder [25] and QGRS Mapper [26], followed by an SL structure analysis using RNAfold [29] in six mammalian and one fish species (Figure 4A). A relatively loose threshold was applied to include non-canonical PQSs (File S1). By screening the human and mouse transcriptomes, we observed that at least 60% and 50% of PQSs, respectively, can form SLs across distinct RNAs. This was predicted by RNAfold [29] based on the minimum free energy (MFE) ensemble (Figure 4B and C), implying a widespread existence of embedded secondary RNA structures. The MFE distribution of canonical PQSs was significantly different from that of imperfect PQSs (Figure S9), indicating a relatively flexible nature of imperfect PQSs. Although G-tetrad demands G-stretches and G can pair with uracil (U) in nature [47], the stem of SL is more likely to appear in the loops of PQSs. By analyzing the number and occurrence of PQSs on mRNA and lncRNA in the human transcriptome, the normalized density of PQS counts based on relative positions revealed an uneven distribution of PQSs at both transcript ends. The number of PQSs in each transcript varied significantly, ranging from 1 to 346, suggesting that

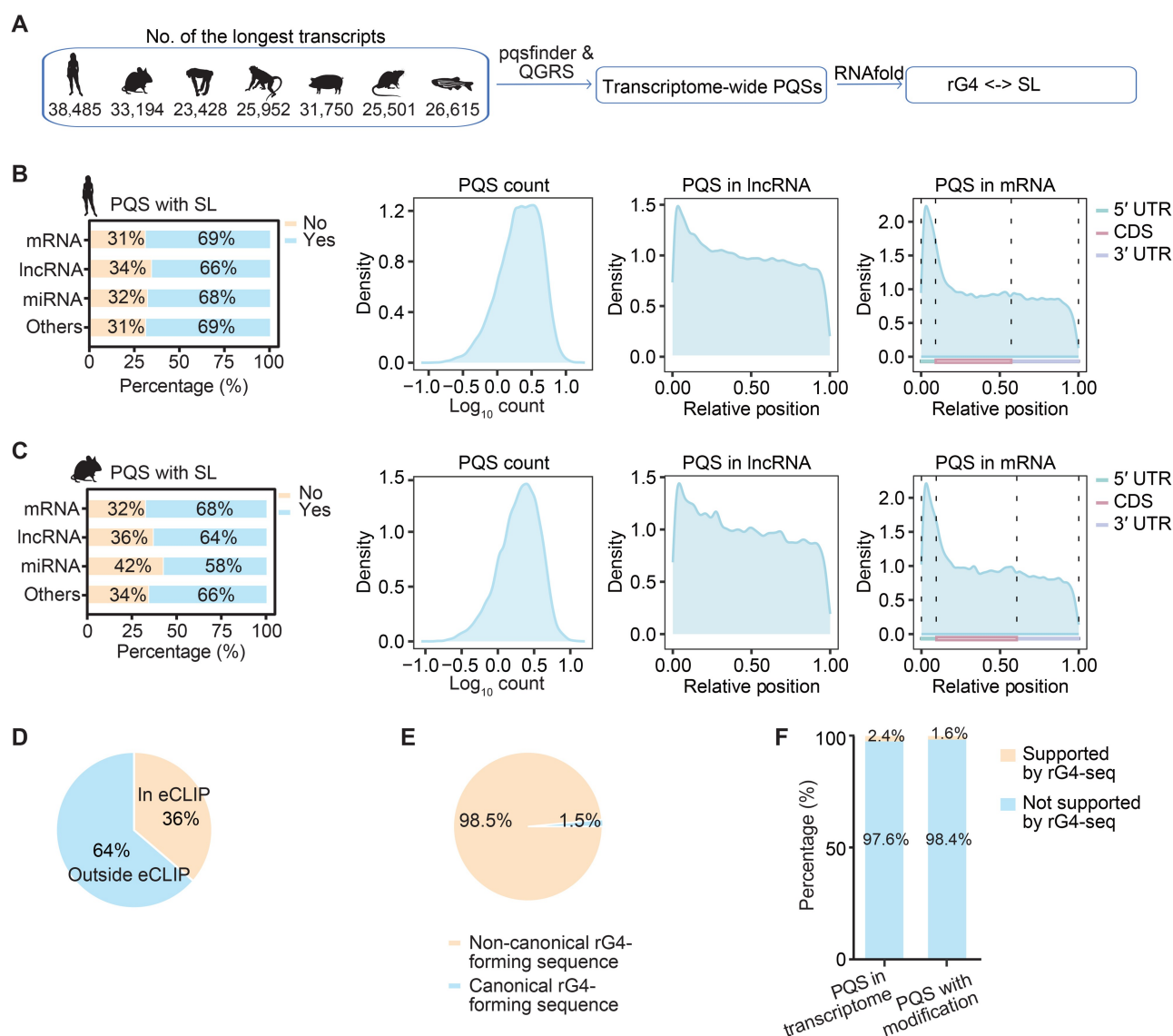


Figure 4 Dual structure-forming nature of RNA segments in transcripts

A. Workflow to collect PQSs with SL-forming potential in the transcriptome of vertebrate genomes, including human, chimpanzee, rhesus monkey, pig, mouse, rat, and zebrafish. **B.** SL-forming potential and distribution of PQSs in the human transcriptome. Left: Percentage of PQSs predicted to form SLs in mRNAs, lncRNAs, miRNAs, and the remaining RNAs in the human transcriptome. Middle-left: The distribution of PQS count per human transcript, normalized for length to counts per kilobase. Middle-right: The relative position of human PQSs in the longest lncRNAs. Right: The relative position of human PQSs in the longest mRNAs. **C.** SL-forming potential and distribution of PQSs in the mouse transcriptome. **D.** Percentage of PQSs supported by ENCODE eCLIP peaks in human sequences. **E.** Percentage of canonical and non-canonical rG4-forming sequences lacking SL structure in the human transcriptome. **F.** *In vivo* evidence for PQSs and PQSs with RNA modifications. Left: Percentage of PQSs supported by rG4-seq in the human transcript. Right: Percentage of PQSs with RNA modifications supported by rG4-seq in the human transcript. UTR, untranslated region; CDS, coding sequence.

transcripts use rG4 distinctly (Figure 4B). A similar pattern was observed in the transcriptomes of mice and other species, except for zebrafish (Figure 4C, Figure S10). These results imply the widespread existence of secondary structure equilibrium in transcriptomes.

PQSs and the scarcity of experimental rG4 profiles in the human transcriptome

About 36% of PQSs in the human transcriptome are supported by at least one eCLIP dataset from ENCODE [48], consistently

demonstrating that RBPs can interact with rG4s *in vivo* (Figure 4D). These PQS-overlapping peaks, which span at least half the length of the corresponding PQSs, were used to identify candidate rG4-binding RBPs. These rG4-binding RBPs were defined as proteins that bind PQS-overlapping peaks, which account for at least 15% of the total non-transposable element (TE) peaks in this study. The functional enrichment of rG4-binding RBPs showed significant overrepresentation in splicing, nucleoprotein granules, and stress granules (Figure S11). This result aligns with previous discoveries of co-localization between rG4s and ribonucleoproteins [49,50]. Notably, among the

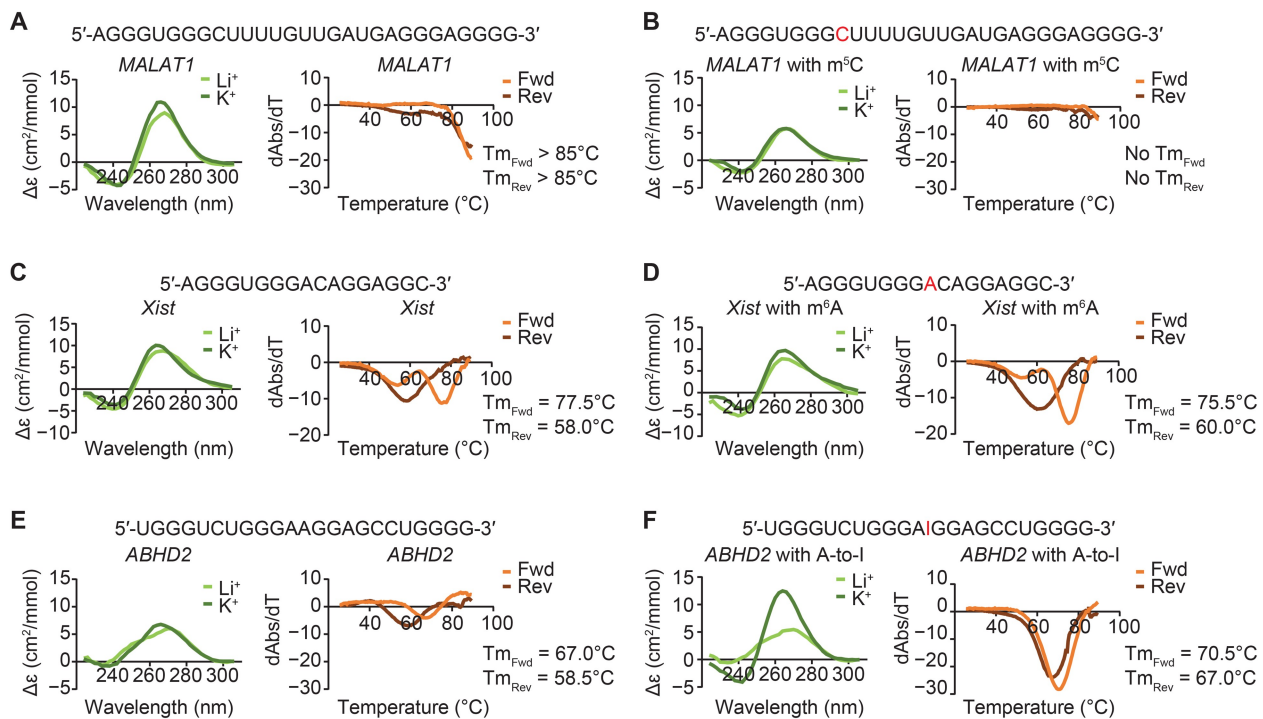


Figure 5 Biophysical characterization of PQSs in *MALAT1*, *Xist*, *ABHD2*, and their derivatives with RNA modifications

A. Left: CD spectra of *MALAT1* under K^+ and Li^+ conditions. The CD pattern was stronger in the presence of K^+ than in the presence of Li^+ , suggesting rG4 formation. The negative peak at 240 nm and the positive peak at 262 nm under K^+ conditions suggest a parallel topology of rG4. Right: UV melting results of *MALAT1*. The hypochromic shift at 295 nm is a strong sign of rG4 formation. The forward experiments were conducted from 20°C to 95°C, while the reverse experiments were conducted from 95°C to 20°C. The available T_m values for each ligand are indicated. The corresponding oligo sequence is shown at the top. **B.** Left: CD spectra of *MALAT1* with m^5C under K^+ and Li^+ conditions. Right: UV melting results of *MALAT1* with m^5C . The corresponding oligo sequence is shown at the top, with the modification site in red. **C.** Left: CD spectra of *Xist* under K^+ and Li^+ conditions. Right: UV melting results of *Xist*. **D.** Left: CD spectra of *Xist* with m^6A under K^+ and Li^+ conditions. Right: UV melting results of *Xist* with m^6A . **E.** Left: CD spectra of *ABHD2* under K^+ and Li^+ conditions. Right: UV melting results of *ABHD2*. **F.** Left: CD spectra of *ABHD2* with A-to-I editing under K^+ and Li^+ conditions. Right: UV melting results of *ABHD2* with A-to-I editing. T_m , melting temperature; Fwd, forward; Rev, reverse.

PQSs that RNAfold failed to predict SL structures, we noticed that canonical rG4s (G_3 , L_{1-7}) and non-canonical rG4s accounted for 1.5% and 98.5%, respectively (Figure 4E). This distribution significantly differed from that of PQSs predicted to form SL. Furthermore, based on rG4-seq and rG4-seq (v2.0) data from the human transcriptome, we analyzed PQSs that either completely overlap with rG4-seq peaks or have their last G at reverse transcriptase stalling (RTS) sites or 1 nt upstream, and found that only about 2.4% of PQSs were supported by rG4-seq (Figure 4F). Additionally, the current rG4-seq dataset solely includes data from HeLa and HEK293T cells, highlighting the necessity to generate more rG4 profiles to investigate the dynamics of distinct RNA secondary structures. To demonstrate the reliability of PQSs, CD-HIT-EST [51] was employed to cluster PQSs collected in this study with the reference PQSs documented in QUADTratlas [52] and G4Atlas [53]. Over 69% of the reference PQSs were clustered with the collected PQSs at a 90% sequence identity threshold; this increased to 87% when the identity threshold was set to 85% (Figure S12A). We additionally applied two large language models, RNA-FM [54] and RNAErnie [55], to compare feature similarity between PQS datasets using two PIWI-interacting RNA (piRNA) datasets as outgroups. The two PQS datasets showed good alignment in the projected space

after dimensionality reduction, indicating their high similarity (Figure S12B). In contrast, the piRNA sequences from *Mus musculus* and *Caenorhabditis elegans* formed distinct clusters as expected.

RNA modifications in rG4 loci

Emerging evidence has shown that rG4s are tightly coupled with RNA modifications. Our results further support the idea that RNA bases carrying chemical modifications can alter the base-pairing capability of nucleotides, thereby impairing RNA structure equilibrium, as many PQSs can also fold into SL structures. Therefore, we collected RNA modification sites in transcripts, including mRNAs, lncRNAs, and miRNAs, from resources cataloged in GEO, RMBase, RMVar, MiREDiBase, and m^6A -Atlas databases, as well as from published datasets. These modification sites were found within 15% and 13% of total PQSs in the human and mouse transcriptomes, respectively. The high-confidence m^6A sites within PQSs, supported by experimental evidence at single-base resolution, represent 41% and 32% of the total PQSs that contain m^6A sites in the human and mouse transcriptomes, corresponding to 14,137 and 7800 sites, respectively. Among

human PQSs with modifications, only 1.6% were supported by rG4-seq (Figure 4F).

Experimental validation

To comprehensively demonstrate the impact of RNA modifications on secondary structure dynamics, we selected four PQSs from human and mouse transcripts carrying three types of modifications that are relatively abundant *in vivo*. Sequences from these loci include m⁵C in human *MALAT1* (supported by rG4-seq [56]), m⁶A in mouse *Xist* [identified by K⁺-dependent reverse transcription (RT)-stop profiling [10]], an A-to-I editing site in human *ABHD2* (lacking rG4-seq support), and m⁵C in mouse *Kmt2d* (containing multiple m⁵C sites supported by ultra-low-input rG4-seq [23]). All rG4-forming sequences were documented in MoRNING. For the *MALAT1* oligo, we found that m⁵C can potentially promote the formation of rG4 multimers or intermolecular rG4s (Figure 5A and B). However, the T_m could not be reliably determined for *MALAT1* oligos, probably due to the low rG4 concentration. Compared to the *Xist* oligo without m⁶A, the m⁶A-containing oligo exhibited a significantly decreased T_m in the forward UV melting scan (Figure 5C and D). A-to-I editing significantly enhanced rG4 formation with considerable thermostability (Figure 5E and F). We further explored the effect of one m⁵C versus two m⁵Cs in mouse *Kmt2d* and found that all tested oligos could form rG4s. However, compared to oligos without m⁵C, those containing m⁵C appeared to enhance rG4 multimerization or promote intermolecular rG4 formation, likely due to changes in intermolecular interactions induced by m⁵C modifications (Figure S13). These findings demonstrate that modifications can influence RNA conformation, further confirming the utility of this database.

Discussion

The increasing volume of RNA modifications characterized at base resolution provides an opportunity to better understand the regulatory roles of RNAs, since chemical modifications to nucleotides alter base-pairing properties or the recognition of RBPs. However, the relationships between RNA structures and RNA modifications remain largely undefined. Beyond the known SL structure recognized by LIN28A, our experiments demonstrate that the non-canonical secondary structure, rG4, can also interact with this protein. rG4 can facilitate binding to LIN28A, regardless of the presence of potassium. For the existing structure profiling data of LIN28A, poor alignment between the reactivity score and SL can be explained by rG4 formation that competes with SL. Indeed, the UV melting experiment shows that m⁶A can reduce rG4 thermostability without drastically impairing the overall binding between RNA and LIN28A. This mechanism coincidentally explains LIN28A binding as an m⁶A anti-reader, with RNA modification acting as a switch [17]. Since the Mut1_{GG2UU} RNA exhibits a lower binding affinity for the protein than the original ligand, we could not rule out the possibility that LIN28A prefers binding to a certain form of rG4. Alternatively, a mixture of rG4s could somehow enhance the binding.

The compartmentalized distribution of K⁺ in cells [57] suggests that RNA conformation can dynamically change alongside ions within compartments in response to external cellular stress or stimuli, leading to a shift in the equilibrium of RNA conformation. In addition to the evidence that rG4s and RNA modifications are directly coupled [18], RNA modifications are commonly observed in neurons, where ion flux dictates cell excitability. Moreover, in several neurodegenerative diseases, including amyotrophic lateral sclerosis, frontotemporal dementia, and Alzheimer's disease, RNA modifications are directly associated with disease-associated RBPs and their molecular crowding [15,16]. The dysregulated interaction between RNA and RBP suggests modulation of RNA structure by RNA modifications. RNA modifications alter the structural stability of RNAs, representing a potential regulatory layer that facilitates RNA structural dynamics, as supported by evidence from several tRNA modifications [13,58,59]. Our findings support a model in which RNA modifications, together with ions, impact the equilibrium of RNA secondary conformations to alter binding affinity for RBPs, thereby influencing subsequent regulatory processes. It is also possible that RNA modifications influence intermolecular interactions, as observed in cases with m⁵C modifications. Although elucidating the detailed mechanism underlying RNA stability shifts remains challenging, modification-induced changes in RNA structure could guide future drug discovery and potentially be combined with single-nucleotide polymorphisms to empower personalized medicine strategies.

The large proportion of RNA sequences with dual RNA structure-forming capacity in the genome suggests that the regulation of RNA structures requires greater focus during profiling experiments. This also points out another aspect of RNA folding: current RNA folding algorithms require additional parameters to expand their capability to predict and distinguish structural forms derived from the same RNA segment. Our study is limited to *in vitro* assays, providing proof of concept that RNA modification acts as a structure switch by shifting the equilibrium of distinct secondary structures [12]. A similar rule can potentially be extended to the relationship between unstructured RNA segments and structures, as indicated by A-to-I editing. Therefore, large-scale investigations, *in vivo* profiling, and better RNA folding algorithms would be of great importance for gaining a systematic view. However, carefully designed experiments and corresponding simulations demand cross-disciplinary collaborations, which are beyond our current research capacity. Leveraging the accumulated experimental evidence and large-scale data, we established the MoRNING database to help the RNA research community explore the roles of RNA modifications at single-base resolution within their associated rG4s via web browsers. This development opens the door to targeting rG4 regulation through RNA modifications. However, due to the scope and throughput of existing experimental assays, many candidate rG4 loci remain to be validated, and low-confidence loci should be used with caution. Improved rG4-seq protocols and tailored rG4 capturing assays that boost resolution and throughput would be valuable resources [23]. Advanced models or algorithms for the accurate identification of PQSs are also on the horizon [54,55,60]. Nevertheless, our study lays the groundwork for targeting rG4 regulation through RNA modifications.

In follow-up work, we will focus on several areas: (1) expanding the database capacity by including more RNA modifications from newly sequenced species and data types from other experiments to annotate PQSs, such as icSHAPE [61] and HTR-SELEX [6]; (2) incorporating or developing state-of-the-art algorithms to improve the precision of rG4 prediction, and introducing new tools to integrate RNA structure data to strengthen the confidence level of experimental evidence; (3) automating the data processing pipeline to implement a robust data curation procedure that quickly incorporates the latest experimental evidence and allows users to submit published datasets. Overall, the MoRNING database will continue to provide insights into the interplay between RNA epigenetics and structural dynamics.

Data availability

MoRNING can be accessed at <https://www.cityu.edu.hk/bms/morning>. MoRNING has been submitted to Database Commons [62] at the National Genomics Data Center (NGDC), China National Center for Bioinformation (CNCB), which is publicly accessible at <https://ngdc.cncb.ac.cn/databasecommons/database/id/10241>.

CRediT author statement

Yicen Zhou: Data curation, Methodology, Investigation, Visualization, Software, Writing – original draft. **Shanxin Lyu:** Methodology, Investigation, Visualization, Writing – original draft. **Shiau Wei Liew:** Methodology, Investigation, Writing – review & editing. **Xi Mou:** Methodology, Investigation, Writing – review & editing. **Ian Hoffecker:** Investigation, Writing – review & editing. **Jian Yan:** Investigation, Writing – review & editing. **Yu Li:** Investigation, Writing – review & editing. **Chun Kit Kwok:** Conceptualization, Supervision, Investigation, Writing – review & editing. **Jilin Zhang:** Conceptualization, Investigation, Supervision, Investigation, Writing – original draft, review & editing. All authors have read and approved the final manuscript.

Competing interests

The authors have declared no competing interests.

Supplementary material

Supplementary material is available at *Genomics, Proteomics & Bioinformatics* online (<https://doi.org/10.1093/gpbjnl/qzaf106>).

Acknowledgments

We thank Dr. Lisheng Zhang for providing constructive suggestions. This work was supported in part by start-up grants for new faculty (Grant Nos. 7200735 and 9610580) to Jilin Zhang by City University of Hong Kong, China; the Research Grants Council of the Hong Kong Special Administrative Region, China (General Research Fund; Grant No. 11101022) to Jian Yan; the National Natural Science Foundation of China (Grant Nos.

32471343 and 32222089); the Research Grants Council of the Hong Kong Special Administrative Region, China (Grant No. RFS2425-1S02); the City University of Hong Kong projects, China (Grant Nos. 11101525, 11100123, 11100222, and 11100421); the Croucher Foundation Project, China (Grant No. 9509003); the State Key Laboratory of Marine Environmental Health Seed Collaborative Research Fund, China (Grant No. SCRF/0070); the City University of Hong Kong projects, China (Grant Nos. 9680376, 7030001, and 9678302) to Chun Kit Kwok; and the Hong Kong PhD Fellowship Scheme, China to Shiau Wei Liew.

ORCID

0009-0001-9650-649X (Yicen Zhou)
0009-0008-3827-7562 (Shanxin Lyu)
0000-0002-2424-9084 (Shiau Wei Liew)
0000-0002-0264-4702 (Xi Mou)
0000-0001-6941-4576 (Ian Hoffecker)
0000-0002-1267-2870 (Jian Yan)
0000-0002-3664-6722 (Yu Li)
0000-0001-9175-8543 (Chun Kit Kwok)
0000-0002-9976-1605 (Jilin Zhang)

References

- [1] Mortimer SA, Kidwell MA, Doudna JA. Insights into RNA structure and function from genome-wide studies. *Nat Rev Genet* 2014;15:469–79.
- [2] Baronti L, Guzzetti I, Ebrahimi P, Friebe Sandoz S, Steiner E, Schlagnitweit J, et al. Base-pair conformational switch modulates miR-34a targeting of *Sirt1* mRNA. *Nature* 2020; 583:139–44.
- [3] Cao X, Zhang Y, Ding Y, Wan Y. Identification of RNA structures and their roles in RNA functions. *Nat Rev Mol Cell Biol* 2024;25:784–801.
- [4] Lyu K, Chow EY, Mou X, Chan TF, Kwok CK. RNA G-quadruplexes (rG4s): genomics and biological functions. *Nucleic Acids Res* 2021;49:5426–50.
- [5] Dumas L, Herviou P, Dassi E, Cammas A, Millevoi S. G-quadruplexes in RNA biology: recent advances and future directions. *Trends Biochem Sci* 2021;46:270–83.
- [6] Jolma A, Zhang J, Mondragón E, Morgunova E, Kivioja T, Lavery KU, et al. Binding specificities of human RNA-binding proteins toward structured and linear RNA sequences. *Genome Res* 2020;30:962–73.
- [7] Orenstein Y, Ohler U, Berger B. Finding RNA structure in the unstructured RBPome. *BMC Genomics* 2018;19:154.
- [8] Nam Y, Chen C, Gregory RI, Chou JJ, Sliz P. Molecular basis for interaction of *let-7* microRNAs with Lin28. *Cell* 2011; 147:1080–91.
- [9] O'Day E, Le MTN, Imai S, Tan SM, Kirchner R, Arthanari H, et al. An RNA-binding protein, Lin28, recognizes and remodels G-quartets in the microRNAs (miRNAs) and mRNAs it regulates. *J Biol Chem* 2015;290:17909–22.
- [10] Lee YW, Weissbein U, Blum R, Lee JT. G-quadruplex folding in *Xist* RNA antagonizes PRC2 activity for stepwise regulation of X chromosome inactivation. *Mol Cell* 2024;84: 1870–85.e9.

- [11] Luo Y, Živković ML, Wang J, Ryneš J, Foldynová-Trantírková S, Trantírek L, et al. A sodium/potassium switch for G4-prone G/C-rich sequences. *Nucleic Acids Res* 2024;52:448–61.
- [12] Pandey S, Agarwala P, Jayaraj GG, Gargallo R, Maiti S. The RNA stem-loop to G-quadruplex equilibrium controls mature microRNA production inside the cell. *Biochemistry* 2015;54:7067–78.
- [13] Biedenbänder T, de Jesus V, Schmidt-Dengler M, Helm M, Corzilius B, Fürtig B. RNA modifications stabilize the tertiary structure of tRNA^{fMet} by locally increasing conformational dynamics. *Nucleic Acids Res* 2022; 50:2334–49.
- [14] Frye M, Harada BT, Behm M, He C. RNA modifications modulate gene expression during development. *Science* 2018; 361:1346–9.
- [15] McMillan M, Gomez N, Hsieh C, Bekier M, Li X, Miguez R, et al. RNA methylation influences TDP43 binding and disease pathogenesis in models of amyotrophic lateral sclerosis and frontotemporal dementia. *Mol Cell* 2023;83:219–36.e7.
- [16] Yang C, Hu Y, Zhou B, Bao Y, Li Z, Gong C, et al. The role of m⁶A modification in physiology and disease. *Cell Death Dis* 2020;11:960.
- [17] Sun L, Fazal FM, Li P, Broughton JP, Lee B, Tang L, et al. RNA structure maps across mammalian cellular compartments. *Nat Struct Mol Biol* 2019;26:322–30.
- [18] Yoshida A, Oyoshi T, Suda A, Futaki S, Imanishi M. Recognition of G-quadruplex RNA by a crucial RNA methyltransferase component, METTL14. *Nucleic Acids Res* 2022; 50:449–57.
- [19] Pandolfini L, Barbieri I, Bannister AJ, Hendrick A, Andrews B, Webster N, et al. METTL1 promotes *let-7* microRNA processing via m⁷G methylation. *Mol Cell* 2019;74:1278–90.e9.
- [20] Kwok CK, Marsico G, Sahakyan AB, Chambers VS, Balasubramanian S. rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. *Nat Methods* 2016;13:841–4.
- [21] Zhao J, Chow EY, Yeung PY, Zhang QC, Chan TF, Kwok CK. Enhanced transcriptome-wide RNA G-quadruplex sequencing for low RNA input samples with rG4-seq 2.0. *BMC Biol* 2022;20:257.
- [22] Chow EY, Lyu K, Kwok CK, Chan TF. rG4-seeker enables high-confidence identification of novel and non-canonical rG4 motifs from rG4-seq experiments. *RNA Biol* 2020; 17:903–17.
- [23] Meng TG, Yue W, Li C, Gao D, Du Z, Yi LT, et al. Ultra-low-input rG4-seq reveals the RNA G-quadruplex regulome in gene expression and genome integrity. *Nucleic Acids Res* 2026;54:gkag040.
- [24] Paramasivan S, Rujan I, Bolton PH. Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods* 2007;43:324–31.
- [25] Hon J, Martínek T, Zendulka J, Lexa M. Pqsfinder: an exhaustive and imperfection-tolerant search tool for potential quadruplex-forming sequences in R. *Bioinformatics* 2017;33:3373–9.
- [26] Kikin O, D'Antonio L, Bagga PS. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res* 2006;34:W676–82.
- [27] Brázda V, Kolomazník J, Lýsek J, Bartas M, Fojta M, Šťastný J, et al. G4Hunter web application: a web server for G-quadruplex prediction. *Bioinformatics* 2019; 35:3493–5.
- [28] Garant JM, Perreault JP, Scott MS. G4RNA screener web server: user focused interface for RNA G-quadruplex prediction. *Biochimie* 2018;151:115–8.
- [29] Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA package 2.0. *Algorithms Mol Biol* 2011;6:26.
- [30] Chan KL, Peng B, Umar MI, Chan CY, Sahakyan AB, Le MTN, et al. Structural analysis reveals the formation and role of RNA G-quadruplex structures in human mature microRNAs. *Chem Commun (Camb)* 2018;54:10878–81.
- [31] Hagen T, Laski A, Brümmer A, Pruška A, Schlösser V, Cléry A, et al. Inosine substitutions in RNA activate latent G-quadruplexes. *J Am Chem Soc* 2021;143:15120–30.
- [32] Cui X, Wei Z, Zhang L, Liu H, Sun L, Zhang SW, et al. Guitar: an R/Bioconductor package for gene annotation guided transcriptomic analysis of RNA-related genomic features. *Biomed Res Int* 2016;2016:8367534.
- [33] Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res* 2019; 47:D155–62.
- [34] Liang Z, Ye H, Ma J, Wei Z, Wang Y, Zhang Y, et al. m⁶A-Atlas v2.0: updated resources for unraveling the N⁶-methyladenosine (m⁶A) epitranscriptome among multiple species. *Nucleic Acids Res* 2024;52:D194–202.
- [35] Luo X, Li H, Liang J, Zhao Q, Xie Y, Ren J, et al. RMVar: an updated database of functional variants involved in RNA modifications. *Nucleic Acids Res* 2021;49:D1405–12.
- [36] Marceca GP, Distefano R, Tomasello L, Lagana A, Russo F, Calore F, et al. MiREDiBase, a manually curated database of validated and putative editing events in microRNAs. *Sci Data* 2021;8:199.
- [37] Xuan J, Chen L, Chen Z, Pang J, Huang J, Lin J, et al. RMBase v3.0: decode the landscape, mechanisms and functions of RNA modifications. *Nucleic Acids Res* 2024; 52:D273–84.
- [38] Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; 26:841–2.
- [39] Diesh C, Stevens GJ, Xie P, De Jesus Martinez T, Hershberg EA, Leung A, et al. JBrowse 2: a modular genome browser with views of synteny and structural variation. *Genome Biol* 2023;24:74.
- [40] Kiran A, Baranov PV. DARNED: a database of RNA editing in humans. *Bioinformatics* 2010;26:1772–6.
- [41] Ramaswami G, Li JB. RADAR: a rigorously annotated database of A-to-I RNA editing. *Nucleic Acids Res* 2014; 42:D109–13.
- [42] Boccaletto P, Stefaniak F, Ray A, Cappannini A, Mukherjee S, Purta E, et al. MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res* 2022; 50:D231–5.
- [43] Lyu K, Kwok CK. A G-quadruplex structure in microRNA interferes with messenger RNA recognition and controls gene expression. *Chem Commun (Camb)* 2023;59:8230–3.

- [44] Flynn RA, Zhang QC, Spitale RC, Lee B, Mumbach MR, Chang HY. Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat Protoc* 2016;11:273–90.
- [45] Lucks JB, Mortimer SA, Trapnell C, Luo S, Aviran S, Schroth GP, et al. Multiplexed RNA structure characterization with selective 2'-hydroxyl acylation analyzed by primer extension sequencing (SHAPE-Seq). *Proc Natl Acad Sci U S A* 2011;108:11063–8.
- [46] Lan TCT, Allan MF, Malsick LE, Woo JZ, Zhu C, Zhang F, et al. Secondary structural ensembles of the SARS-CoV-2 RNA genome in infected cells. *Nat Commun* 2022;13:1128.
- [47] Gabriel K, Schneider J, McClain WH. Functional evidence for indirect recognition of G·U in tRNA^{Ala} by alanyl-tRNA synthetase. *Science* 1996;271:195–7.
- [48] ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012;489:57–74.
- [49] Georgakopoulos-Soares I, Parada GE, Wong HY, Medhi R, Furlan G, Munita R, et al. Alternative splicing modulation by G-quadruplexes. *Nat Commun* 2022;13:2404.
- [50] Markmiller S, Soltanieh S, Server KL, Mak R, Jin W, Fang MY, et al. Context-dependent and disease-specific diversity in protein interactions within stress granules. *Cell* 2018;172:590–604.e13.
- [51] Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 2012;28:3150–2.
- [52] Bourdon S, Herviou P, Dumas L, Destefanis E, Zen A, Cammas A, et al. QUADRAtlas: the RNA G-quadruplex and rG4-binding proteins database. *Nucleic Acids Res* 2023;51:D240–7.
- [53] Yu H, Qi Y, Yang B, Yang X, Ding Y. G4Atlas: a comprehensive transcriptome-wide G-quadruplex database. *Nucleic Acids Res* 2023;51:D126–34.
- [54] Shen T, Hu Z, Sun S, Liu D, Wong F, Wang J, et al. Accurate RNA 3D structure prediction using a language model-based deep learning approach. *Nat Methods* 2024;21:2287–98.
- [55] Wang N, Bian J, Li Y, Li X, Mumtaz S, Kong L, et al. Multi-purpose RNA language modelling with motif-aware pretraining and type-guided fine-tuning. *Nat Mach Intell* 2024;6:548–57.
- [56] Mou X, Liew SW, Kwok CK. Identification and targeting of G-quadruplex structures in *MALAT1* long non-coding RNA. *Nucleic Acids Res* 2022;50:397–410.
- [57] Shen Y, Wu SY, Rancic V, Aggarwal A, Qian Y, Miyashita SI, et al. Genetically encoded fluorescent indicators for imaging intracellular potassium ion concentration. *Commun Biol* 2019;2:18.
- [58] Ruiz-Arroyo VM, Raj R, Babu K, Onolbaatar O, Roberts PH, Nam Y. Structures and mechanisms of tRNA methylation by METTL1-WDR4. *Nature* 2023;613:383–90.
- [59] Ohira T, Minowa K, Sugiyama K, Yamashita S, Sakaguchi Y, Miyauchi K, et al. Reversible RNA phosphorylation stabilizes tRNA for cellular thermotolerance. *Nature* 2022;605:372–9.
- [60] Zhuang F, Gutman D, Islas N, Guzman BB, Jimenez A, Jewell S, et al. G4mer: an RNA language model for transcriptome-wide identification of G-quadruplexes and disease variants from population-scale genetic data. *Nat Commun* 2025;16:10221.
- [61] Li P, Zhou X, Xu K, Zhang QC. RASP: an atlas of transcriptome-wide RNA secondary structure probing data. *Nucleic Acids Res* 2021;49:D183–91.
- [62] Ma L, Zou D, Liu L, Shireen H, Abbasi AA, Bateman A, et al. Database Commons: a catalog of worldwide biological databases. *Genomics Proteomics Bioinformatics* 2023;21:1054–8.