RNA G-quadruplex structure contributes to cold adaptation in plants

Xiaofei Yang1,2,3,4,9, Haopeng Yu1,4,9, Susan Duncan4, Yueying Zhang4, Jitender Cheema4, Haifeng Liu4,5, J. Benjamin Miller6, Jie Zhang4, Chun Kit Kwok7,8, Huakun Zhang1 & Yiliang Ding4

Nucleotide composition is suggested to infer gene functionality and ecological adaptation of species to distinct environments. However, the underlying biological function of nucleotide composition dictating environmental adaptation is largely unknown. Here, we systematically analyze the nucleotide composition of transcriptomes across 1000 plants (1KP) and their corresponding habitats. Intriguingly, we find that plants growing in cold climates have guanine (G)-enriched transcriptomes, which are prone to forming RNA G-quadruplex structures. Both immuno-fluorescence detection and in vivo structure profiling reveal that RNA G-quadruplex formation in plants is globally enhanced in response to cold. Cold-responsive RNA G-quadruplexes strongly enhanced mRNA stability, rather than affecting translation. Disruption of individual RNA G-quadruplex promotes mRNA decay in the cold, leading to impaired plant cold response. Therefore, we propose that plants adopted RNA G-quadruplex structure as a molecular signature to facilitate their adaptation to the cold during evolution.

The earth’s prodigious biodiversity was established in part by adaptation to diverse ecological habitats, driving speciation1. In particular, plants have become highly evolved to their specific environments, partly due to their sessile nature, with diversity across environments enabling their global colonization2–3. Varied nucleotide compositions were suggested to affect plant adaptation to specific habitats4–5. However, an understanding of the underlying molecular mechanisms remains unclear.

Results

Transcriptomic guanine (G) and RNA G-quadruplex frequencies exhibit climatic signatures of plant adaptation

The recent generation of transcriptome sequences of over 1000 plants (The 1000 plants initiative, 1KP) has allowed us to systematically assess the biological significance of nucleotide compositions across the plant kingdom1. We analyzed the frequency of the four nucleotides (A: adenine, U: uracil, C: cytosine, and G: guanine) across all major clades of land plants, including dicots, monocots, gymnosperms, ferns, lycophytes, and bryophytes (Fig.1a and Supplementary Data 1). Overall, the frequencies among the four nucleotides varied much more in the 3’UTR compared to 5’UTR and CDS (Fig. 1a). We then obtained the habitat locations of these plants from the Global Biodiversity Information Facility (GBIF) and extracted the corresponding 19 associated bioclimatic variables for these habitats from the WorldClim database (Fig. 1b and Supplementary Data 2, see Methods)7. We calculated the Pearson Correlation Coefficient (PCC) between the...
transcriptomic nucleotide frequencies and bioclimatic variables primarily concerned with temperature and precipitation (i.e., BIO1: annual mean temperature, BIO12: annual precipitation; full list in Methods). PCCs between nucleotide frequency and temperature bioclimatic variables (BIO1-BIO11) were overall stronger than those between PCCs between nucleotide frequency and temperature bioclimatic variables related to plant habitats. The frequencies of the four nucleotides were calculated in 906 land plants derived from the transcriptomes of the 1000 Plants (IKP Initiative). n = 556, 107, 80, 71, 71 for dicots, monocots, gymnosperms, ferns, lycophytes and bryophytes, respectively. b Heat plot showing the Pearson Correlation Coefficient (PCC) between transcriptomic nucleotide frequency and bioclimatic variables of plant habitats. The frequencies of the four nucleotides were calculated in 906 land plants derived from the transcriptomes of the 1000 Plants (IKP). The corresponding plant habitats were sourced from the Global Biodiversity Information Facility (GBIF). The bioclimatic variables of these plant habitats were derived from the WorldClim database, i.e., BIO1: annual mean temperature (for full details see Methods), as explained by WorldClim (https://www.worldclim.org/). c Bar plot showing the counts of significant PCCs between transcriptomic nucleotide frequency and temperature bioclimatic variables. PCCs with a P value less than a threshold of 0.01 were retained. d DNA G-quadruplex (RG4) frequencies in major clades of land plants. Land plants with over 100 occurrences in Global Biodiversity Information Facility (GBIF) were included (see Methods). n = 277, 43, 30, 10, 43 for dicots, monocots, gymnosperms, ferns, lycophytes and bryophytes, respectively. The corresponding annual mean temperatures for habitats of plant species were color-coded. e Heat plot showing the Pearson Correlation Coefficients (PCCs) between RG4 frequency in plant transcriptomes and associated bioclimatic variables related to plant habitats. PCCs with a P value less than a threshold of 0.01 were retained. Plant habitats were sourced from the GBIF, while associated bioclimatic variables were derived from the WorldClim database (full details see Methods).

Fig. 1 | Comparison of transcriptomic guanine (G) frequencies and RNA G-quadruplex frequencies exhibit climatic signatures of plant adaptation. a Scatter plot showing the frequency of each nucleotide (A: adenine, U: uracil, C: cytosine, and G: guanine) in transcriptomes of 906 land plants from the 1000 Plants (IKP Initiative). n = 556, 107, 80, 71, 71 for dicots, monocots, gymnosperms, ferns, lycophytes and bryophytes, respectively. b Heat plot showing the Pearson Correlation Coefficient (PCC) between transcriptomic nucleotide frequency and bioclimatic variables of plant habitats. The frequencies of the four nucleotides were calculated in 906 land plants derived from the transcriptomes of the 1000 Plants (IKP). The corresponding plant habitats were sourced from the Global Biodiversity Information Facility (GBIF). The bioclimatic variables of these plant habitats were derived from the WorldClim database, i.e., BIO1: annual mean temperature (for full details see Materials and Methods), as explained by WorldClim (https://www.worldclim.org/). c Bar plot showing the counts of significant PCCs between transcriptomic nucleotide frequency and temperature bioclimatic variables. PCCs with a P value less than a threshold of 0.01 were retained. d DNA G-quadruplex (RG4) frequencies in major clades of land plants. Land plants with over 100 occurrences in Global Biodiversity Information Facility (GBIF) were included (see Methods). n = 277, 43, 30, 10, 43 for dicots, monocots, gymnosperms, ferns, lycophytes and bryophytes, respectively. The corresponding annual mean temperatures for habitats of plant species were color-coded. e Heat plot showing the Pearson Correlation Coefficients (PCCs) between RG4 frequency in plant transcriptomes and associated bioclimatic variables related to plant habitats. PCCs with a P value less than a threshold of 0.01 were retained. Plant habitats were sourced from the GBIF, while associated bioclimatic variables were derived from the WorldClim database (full details see Methods).

Given that a G-rich region in an RNA molecule is capable of folding into a tertiary RNA structure called RNA G-quadruplex (RG4), involving the base pairs on both Hoogsteen and Watson-Crick faces (Supplementary Fig. 1a), we, therefore, hypothesized that G-enriched transcriptomes from plants prevalent in colder climates may be complemented with RG4 motifs. We then calculated the RG4 frequency across the 1KP dataset, yielding a median frequency value of 1.27, 2.35, and 0.99 RG4s/kb of RG4 for 5′-UTR, CDS, and 3′-UTR, respectively (Fig. 1d and Supplementary Data 3). Across all genic regions, the overall correlations between RG4 frequency and temperature bioclimatic variables were significantly negative, whilst generally less significant correlations between RG4 frequency and precipitation bioclimatic variables were observed (Figs. 1d, e).
sugest that RG4 is more strongly enriched in plant species growing in colder climates.

The folding state of RG4s is globally enhanced in response to cold

To understand the underlying mechanisms of higher RG4 enrichment in plant adaptation to cold, we first visualized RG4 in cells from the model plant Arabidopsis thaliana, using the BG4 antibody, which is able to detect RNA G-quadruplex in cells. We found a significant increase and enhancement of cytosolic BG4 foci in the cold at 4 °C, compared to the normal temperature control at 22 °C (Fig. 2a, P = 10−4, by Student’s t test). The increase of the BG4 foci in the cold mimicked RG4 stabilization using the well-characterized pyridostatin ligand (PDS) (Fig. 2a, P = 10−8 by Student’s t test)10, indicating that cold promotes RG4 folding in plant cells. Notably, when cold-treated plants were returned to 22 °C for 2 hours, BG4 signals recovered to a similar level to that before cold treatment (Fig. 2a, P = 0.17 by Student’s t test).

We then determined the in vivo folding status of individual RG4 motifs in the cold using SHALiPE-seq, a high throughput method for RG4 detection. SHALiPE-seq is based on the preferable motifs in the cold using SHALiPE-seq, a high throughput method for compared to the normal temperature control at 22 °C (Fig. 2b, P = 10−4, by Student’s t test). The increase of the BG4 foci in the cold mimicked RG4 stabilization using the well-characterized pyridostatin ligand (PDS) (Fig. 2a, P = 10−8 by Student’s t test)10, indicating that cold promotes RG4 folding in plant cells. Notably, when cold-treated plants were returned to 22 °C for 2 hours, BG4 signals recovered to a similar level to that before cold treatment (Fig. 2a, P = 0.17 by Student’s t test).

We then determined the in vivo folding status of individual RG4 motifs in the cold using SHALiPE-seq, a high throughput method for RG4 detection. SHALiPE-seq is based on the preferable motifs in the cold using SHALiPE-seq, a high throughput method for compared to the normal temperature control at 22 °C (Fig. 2b, P = 10−4, by Student’s t test). The increase of the BG4 foci in the cold mimicked RG4 stabilization using the well-characterized pyridostatin ligand (PDS) (Fig. 2a, P = 10−8 by Student’s t test)10, indicating that cold promotes RG4 folding in plant cells. Notably, when cold-treated plants were returned to 22 °C for 2 hours, BG4 signals recovered to a similar level to that before cold treatment (Fig. 2a, P = 0.17 by Student’s t test).

We then determined the in vivo folding status of individual RG4 motifs in the cold using SHALiPE-seq, a high throughput method for RG4 detection. SHALiPE-seq is based on the preferable motifs in the cold using SHALiPE-seq, a high throughput method for compared to the normal temperature control at 22 °C (Fig. 2b, P = 10−4, by Student’s t test). The increase of the BG4 foci in the cold mimicked RG4 stabilization using the well-characterized pyridostatin ligand (PDS) (Fig. 2a, P = 10−8 by Student’s t test)10, indicating that cold promotes RG4 folding in plant cells. Notably, when cold-treated plants were returned to 22 °C for 2 hours, BG4 signals recovered to a similar level to that before cold treatment (Fig. 2a, P = 0.17 by Student’s t test).

We then determined the in vivo folding status of individual RG4 motifs in the cold using SHALiPE-seq, a high throughput method for RG4 detection. SHALiPE-seq is based on the preferable motifs in the cold using SHALiPE-seq, a high throughput method for compared to the normal temperature control at 22 °C (Fig. 2b, P = 10−4, by Student’s t test). The increase of the BG4 foci in the cold mimicked RG4 stabilization using the well-characterized pyridostatin ligand (PDS) (Fig. 2a, P = 10−8 by Student’s t test)10, indicating that cold promotes RG4 folding in plant cells. Notably, when cold-treated plants were returned to 22 °C for 2 hours, BG4 signals recovered to a similar level to that before cold treatment (Fig. 2a, P = 0.17 by Student’s t test).

We then determined the in vivo folding status of individual RG4 motifs in the cold using SHALiPE-seq, a high throughput method for RG4 detection. SHALiPE-seq is based on the preferable motifs in the cold using SHALiPE-seq, a high throughput method for compared to the normal temperature control at 22 °C (Fig. 2b, P = 10−4, by Student’s t test). The increase of the BG4 foci in the cold mimicked RG4 stabilization using the well-characterized pyridostatin ligand (PDS) (Fig. 2a, P = 10−8 by Student’s t test)10, indicating that cold promotes RG4 folding in plant cells. Notably, when cold-treated plants were returned to 22 °C for 2 hours, BG4 signals recovered to a similar level to that before cold treatment (Fig. 2a, P = 0.17 by Student’s t test).

We then determined the in vivo folding status of individual RG4 motifs in the cold using SHALiPE-seq, a high throughput method for RG4 detection. SHALiPE-seq is based on the preferable motifs in the cold using SHALiPE-seq, a high throughput method for compared to the normal temperature control at 22 °C (Fig. 2b, P = 10−4, by Student’s t test). The increase of the BG4 foci in the cold mimicked RG4 stabilization using the well-characterized pyridostatin ligand (PDS) (Fig. 2a, P = 10−8 by Student’s t test)10, indicating that cold promotes RG4 folding in plant cells. Notably, when cold-treated plants were returned to 22 °C for 2 hours, BG4 signals recovered to a similar level to that before cold treatment (Fig. 2a, P = 0.17 by Student’s t test).
cold-responsive RG4s with higher folding scores upon cold treatment. To further explore how these cold-responsive RG4s may affect the plant response to cold, we assessed their molecular functions in regulating gene expression. Extensive studies have suggested that RG4 serves as a translational repressor. Thus, we assessed whether cold-responsive RG4s repress translation in the cold. We compared the translation efficiencies (TEs) of those genes containing cold-responsive RG4s at either 4 °C or 22 °C. Overall, we observed only a subtle variation (2%) and a high correlation between TE at 4 °C and TE at 22 °C (Fig. 3a, PCC = 0.86, difference = 2%, Supplementary Data S), suggesting cold-responsive RG4s are unlikely to contribute much to translation. We then examined the folding score differences between 4 °C and 22 °C of RG4s in different genomic regions and found RG4s in the 3′-UTR possessed the greatest folding score differences (Fig. 3b, 2P = 0.006, by Student’s t test). Thus, cold conditions are likely to strongly promote the folding of RG4s with a higher number of G-quartets and intermediate loop lengths.

To understand any biological relevance between RG4s and the cold, we performed gene ontology (GO) analysis on genes containing cold-responsive RG4s with higher folding scores upon cold treatment. We found specific transcript-related enrichment in biological functions, such as response to abiotic stimulus, response to temperature stimulus, and response to cold (Fig. 2d). Thus, those genes containing cold-responsive RG4s are also likely to function in plant response to cold.

Reinforcement of the notion that RG4 folding is responsive to cold in living plants.

RG4s have a range of G-quartets and loop lengths. To test whether different types of RG4s are favored in the cold, we compared the folding score differences between 4 °C and 22 °C on individual RG4s. We found greater folding score differences for RG4s with three G-quartets (G3-RG4) compared to those with two G-quartets (G2-RG4, Fig. 2d, 2P = 0.006, by Student’s t test). In addition, the folding status of RG4s with intermediate lengths from 4nt to 6nt was more easily enhanced in the cold, compared to RG4s with either a very short loop of 2nt or a very long loop of over 8nt (Fig. 2e, 2P < 0.05, by Student’s t test). Thus, cold conditions are likely to strongly promote the folding of RG4s with a higher number of G-quartets and intermediate loop lengths.

The cold-responsive RG4s in the 3′-UTR serve as mRNA stabilizers.

To further explore how these cold-responsive RG4s may affect the plant response to cold, we assessed their molecular functions in regulating gene expression. Extensive studies have suggested that RG4 serves as a translational repressor. Thus, we assessed whether cold-responsive RG4s repress translation in the cold. We compared the translation efficiencies (TEs) of those genes containing cold-responsive RG4s at either 4 °C or 22 °C. Overall, we observed only a subtle variation (2%) and a high correlation between TE at 4 °C and TE at 22 °C (Fig. 3a, PCC = 0.86, difference = 2%, Supplementary Data S), suggesting cold-responsive RG4s are unlikely to contribute much to translation. We then examined the folding score differences between 4 °C and 22 °C of RG4s in different genomic regions and found RG4s in the 3′-UTR possessed the greatest folding score differences (Fig. 3b, 2P = 0.006, by Student’s t test). Based on previous studies that suggested structural elements in 3′-UTRs are likely to regulate mRNA stability, we hypothesized that cold-responsive RG4s may impact mRNA stability. We measured the RNA stability of the Arabidopsis transcriptome at both 22 °C and 4 °C using the transcription arrest assay (Fig. 3c). Upon transcription arrest, the rate of mRNA abundance decline reveals RNA stability, whereby a rapid decrease indicates low stability whilst a slow decrease indicates high stability. Overall, we found a generally slower mRNA abundance decline at 4 °C compared to 22 °C (Fig. 3d). Notably, the cold effect of slowing down RNA abundance decline was found to be higher for transcripts with cold-responsive RG4s in 3′-UTRs, compared to those transcripts without RG4s (nonRG4), or with cold-responsive RG4s in 5′-UTR and/or CDS regions (Fig. 3d). The transcription arrest profiles for individual mRNAs are illustrated for AT1G33210, AT4G32020, and AT5G24930 along with their corresponding qRT-PCR validations (Supplementary Fig. S).

We then derived the mRNA decay rate for individual mRNAs with cold-responsive RG4s at both 4 °C and 22 °C (Supplementary Data S). The mRNA decay rates at 4 °C altered dramatically compared to those at 22 °C.
at 22 °C (Fig. 3d, e, PCC = 0.02), and generally showed slower mRNA decay rates in the cold. We subsequently calculated the decay rate differences between 4 °C and 22 °C and found a significantly greater difference in decay rate between 4 °C and 22 °C on transcripts with cold-responsive RG4s in 3′-UTR compared to that of transcripts without RG4s or with RG4s in 5′-UTR or CDS regions (Fig. 3f, P values = 0.02, 0.04 and 0.03 respectively, by Student’s t test). As mRNA stability directly affects steady-state mRNA abundance, we reasoned that cold enhances the steady-state mRNA abundance of transcripts with cold-responsive RG4s. Indeed, we found a strong increase in mRNA abundance in the cold for transcripts with cold-responsive RG4s in the 3′-UTR of transcripts with wtRG4 or mutRG4, as shown in the designs in a. Data are presented as mean values ± SE, n = 4. c Phenotypes of plants of different genotypes grown at 22 °C or 4 °C. wtRG4 or mutRG4 denotes corg1-1 mutant complemented with genomic DNA of CORG1 carrying wtRG4 or mutRG4 respectively, as illustrated in a. Comparison was performed on the plants grown vertically on plates at 22 °C for 1 week, or at 4 °C for 4 weeks. Significance tested by one-way ANOVA/Tukey HSD post hoc test (P < 0.05), n ≥ 20. Data are presented as mean values ± SE.

The cold-responsive RG4 modulates plant cold-responsive growth

Considering the strong molecular function of these cold-responsive RG4s in 3′-UTRs, we reasoned that these cold-responsive RG4s, as mRNA stabilizers, may affect plant phenotypic response to the cold. Of the cold-responsive RG4s present in 3′-UTRs, the one with the greatest folding score difference between 4 °C and 22 °C was found on AT1G13390 (Supplementary Fig. 6a and Supplementary Data 4), designated as CORG1 (Cold-responsive RNA G-quadruplex 1), with the annotation encoding a translocase subunit protein (courtesy of The Arabidopsis Information Resource). To further explore whether this CORG1 RG4 contributes to mRNA stabilization in the cold, we constructed the reporter gene, FIRELY-LUCIFERASE (FLUC) fused with the 3′-UTR of CORG1 containing the wild-type cold-responsive RG4 (wtRG4) or a disrupted RG4 (mutRG4, Fig. 4a). We then performed transcriptional arrest assays at both 4 °C and 22 °C to measure the corresponding mRNA stabilities. The FLUC mRNA with the 3′-UTR of CORG1 containing the wtRG4 degraded much more rapidly compared to that containing wtRG4 at 4 °C, while no significant differences were found between these two versions at 22 °C (Fig. 4b), indicating that this cold-responsive RG4 enhanced mRNA stabilization in cold conditions.

We then reintroduced Arabidopsis genomic DNA of CORG1 containing 3′-UTR with either wtRG4 or mutRG4 into the corg1-1 null mutant (Supplementary Fig. 6b, c). At 22 °C, the primary root growth of corg1-1 was similar to that of wild-type Col-0 (Fig. 4c). At 4 °C, the root length of corg1-1 was distinctively longer than that of Col-0, indicating CORG1 represses plant growth under cold conditions. When the corg1-1 mutant was complemented with mutRG4-CORG1, the root growth was further repressed (Fig. 4c).
length was significantly longer than \textit{utRG4-CORGI} complemented plants at 4 °C, while no significant differences were observed at 22 °C (Fig. 4c, P values < 0.01, one-way ANOVA/Tukey HSD post hoc test), suggesting impairment of \textit{mutRG4-CORGI} plants in sensing cold. The cold-sensitive defects in the \textit{mutRG4-CORGI} plants correspond to the more rapid RNA decay and decrease in steady-state abundance (Supplementary Fig. 6d, e). Additionally, the general plant sizes of the \textit{corGI-1} mutant complemented with \textit{mutRG4-CORGI} were overall bigger than \textit{utRG4-CORGI} complemented plants at 4 °C, while no significant differences were observed at 22 °C (Supplementary Fig. 6f). Furthermore, we found that the cold-responsive expression levels of the marker genes for \textit{Arabidopsis} cold response including \textit{CBF1}, \textit{CBF2}, \textit{CBF3}, and \textit{CORISA} were significantly lower in the \textit{corGI-1} mutant complemented with \textit{mutRG4-CORGI} than those in \textit{utRG4-CORGI} complemented plants upon cold treatments (Supplementary Fig. 7).

**Discussion**

Our broad and deep investigations of nucleotide composition across the plant kingdom promote the discovery that RNA G-quadruplex structure may serve as a molecular signature contributing to plant environmental adaptation. Correlations between RG4 frequency and temperature bioclimatic variables are not very high (Fig. 1e), suggesting other factors besides RG4 may also contribute to plant thermal adaptation. RNA G-quadruplexes embedded across the plant transcriptome might play an important role in stabilizing RNA molecules in response to cold (Fig. 4d). Hence, our results suggest that RNA structure, as a key regulator of gene expression\textsuperscript{21–25}, may have evolved due to selection pressures arising from distinct temperature environments. Our work, therefore, advances our fundamental understanding of the molecular mechanisms underlying species adaptation. Ultimately, the translation of this conceptual advance to molecular engineering strategies could help to address global climate change challenges indicated by continuing increases in the frequency, intensity, and duration of extreme low-temperature events\textsuperscript{26–28}.

**Methods**

**Statistics**

No statistical methods were used to predetermine the sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment. Sampling in all cases was performed by collecting materials independently from new plants for replicates.

**Plants and growth conditions**

\textit{Arabidopsis thaliana} ecotype Columbia (Col-0) and \textit{corGI-1} mutant (SALK_097604) with T-DNA insertion were obtained from the Nottingham Arabidopsis Stock Centre (NASC); the homozygous mutant was identified using PCR-based genotyping\textsuperscript{26}. \textit{Arabidopsis} seeds were sterilized using 70% ethanol for 10 min, washed with distilled water three times, and plated on half-strength Murashige and Skoog medium supplemented with 1% sucrose. After standing at 4 °C for 3 days, the plates were placed in a growth chamber at 22 °C. To complement the mutant, \textit{corGI-1} mutant plants were transformed using a floral dipping method\textsuperscript{27}. Transgenic plants were selected using GM media supplemented with 1% sucrose. After standing at 4 °C for 3 days, transgenic plants were harvested in a cold room and incubated with 400 mM NAI at 22 °C overnight before use. NAI probing at 22 °C was carried out as described in our previous study\textsuperscript{11}. For NAI probing at 4 °C, plates with \textit{Agrobacterium tumefaciens} GV3101, the plate was placed into a growth chamber MIR-254-PE (Panasonic) at 4 °C for 2 h. The NAI probing buffer was kept in an incubator at the desired temperature overnight before use. NAI probing at 4 °C was carried out as described in our previous study\textsuperscript{11}. For NAI probing at 4 °C, plates with \textit{Agrobacterium tumefaciens} GV3101, the plate was placed into a growth chamber MIR-254-PE (Panasonic) at 4 °C for 2 h.

**In vivo NAI probing of RNA**

The NAI probing buffer was kept in an incubator at the desired temperature overnight before use. NAI probing at 22 °C was carried out as described in our previous study\textsuperscript{11}. For NAI probing at 4 °C, plates with 5-DAG etiolated seedlings were treated at 4 °C for 2 h. Treated seedlings were harvested in a cold room and incubated with 400 mM NAI at 4 °C for 15 min. NAI was quenched using five times DTT to that of NAI. Seedlings were washed three times with distilled water, ground in liquid nitrogen, and applied to RNA extraction using RNeasy Plant Mini Kit (QIagen).

**Gel-based analysis of NAI probing**

RNA was extracted after NAI probing, 1 μg of total RNA was dissolved in 6 μL water, 1 μL of 5 μM Cy5 modified RT primer for 18 S RNA (listed in

![Image](https://example.com/image.png)
Supplementary Data 7) and 0.5 μL of 10 mM dNTPs were denatured at 95 °C for 3 min. The reaction was cooled down to 50 °C before adding 2 μL of 5× home-made RT buffer (100 mM Tris (pH 8.3), 500 mM LiCl, 15 mM MgCl₂, 5 mM DTT) and 0.5 μL reverse transcriptase Supernscript III (Invitrogen) and mixing quickly with a pipette. The RT reaction was incubated at 50 °C for 20 min followed by a stay at 85 °C for 10 min to inactivate reverse transcriptase. cDNA hybridized RNA was degraded by adding 0.5 μL of 2 M NaOH and incubating at 95 °C for 10 min. Equal volumes of 2× stopping dye (95% formamide, 20 mM EDTA (pH 8.0), 20 mM Tris (pH 7.5), orange G) were added and incubated at 95 °C for 5 min. The resultant reaction was kept at 65 °C and loaded to 5% Acrylamide: Bis-Acrylamide-Urea gel for electrophoresis. For the sequencing lanes, RNA was dissolved in 5 μL water, and 1 μL of corresponding ddNTP (Roche) was added at the beginning.

**Generation of the SHALIPE-seq libraries**

Libraries of SHALIPE-seq were prepared as described. Polya-selected RNA was recovered and reverse transcribed using super script III (Invitrogen) and RT primer (5′-CAGACGTTGCTTCGGATCTTCCCACACCTACCAATGGCTACGTAAGCTGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′, where NNNNNN denotes the barcodes, e.g. Index1 is CGTGAT for Illumina sequencing). The 3′-end of resultant cDNAs were ligated to an ssDNA linker (5′-PhosNNNAGATCGGGAGACGCTGCTGACCAAGCTACACGATGAGCTACATCCGAGATCCAGATGCTAGCCGGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT3′) using Circligase ssDNA Ligase (Epicentre) at 65 °C for 12 h. Product longer than 100 nt was recovered using QIAquick Gel Extraction Kit (Qiagen) after separation of the ligation products with TBE-Urea 10% Gel (Invitrogen). Purified cDNA was subjected to PCR amplification using KAPA Library Amplification Kits (Roche) with Forward Library primer (5′-CAGACGTTGCTTCGGATCTTCCCACACCTACCAATGGCTACGTAAGCTGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′) and Reverse Library primers (5′-CAGACGTTGCTTCGGATCTTCCCACACCTACCAATGGCTACGTAAGCTGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT3′) using Circligase ssDNA Ligase (Epicentre) at 95 °C for 2 min. The product was then amplified by 10 PCR cycles with primers (5′-CAGACGTTGCTTCGGATCTTCCCACACCTACCAATGGCTACGTAAGCTGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3′) and (5′-CAGACGTTGCTTCGGATCTTCCCACACCTACCAATGGCTACGTAAGCTGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-5′) using KAPA Library Amplification Kits (Roche). A low to high sucrose gradient was collected, taking fractions corresponding to 0 min of cordycepin incubation, RNA was extracted and subjected to either quantitative real-time PCR (qRT-PCR) or RNA sequencing.

**Quantitative real-time PCR**

RNA was digested using RNase-free TURBO™ DNase (Ambion). First-strand cDNA was synthesized using reverse transcriptase Superscript III (Invitrogen) and oligo dT primer with home-made RT buffer (20 mM Tris (pH 8.3), 100 mM LiCl, 3 mM MgCl₂, 1 mM DTT). QRT-PCR was performed with LightCycler® 480 SYBR Green 1 Master (Roche) using CFX96 Touch Real-Time PCR Detection System (BIORAD) according to the manufacturer’s protocol. PP2A (AT1G13320) was used as the internal control. All primers are listed in Supplementary Data 7.

**Phenotype assessment**

To measure primary root length, images of 1-week seedlings grown at 22 °C under short-day (SD, 8/16 h, light/dark) were captured using a digital camera. The root length of more than 20 seedlings was measured using ImageJ software (NIH). For cold treatment at 4 °C, plates were placed at 22 °C for 2 days for germination and transferred to 4 °C for further growth under the same light condition. Images were taken after 4 weeks’ growth at 4 °C.

**RNA G-quadruplex structure (RG4) prediction**

Plant transcriptomes were obtained from the resource data of the one thousand plant transcriptomes project. RNA G-quadruplex features were extracted according to our previous study, and further assigned to different structural subclasses of G2-RG4 (RG4 with two layers of G-quartet) or G3-RG4 (RG4 with three layers of G-quartet). The RG4 density was calculated by the count of predicted RG4s, normalized to the total amount of all four bases in the transcriptome for individual species.

**Bioclimatic variables for plants**

The latitude and longitude of plant geographic distribution were obtained from observations collected by the Global Biodiversity Information Facility (www.gbif.org, i.e. Arabidopsis thaliana: 10.15468/dl.pmy46j, full list see Supplementary Data 2). The 433 species with over 100 observations were included in further analysis. The nineteen bioclimatic variables at representative locations were extracted from the WorldClim database. BIO1: annual mean temperature; BIO2: mean diurnal range (mean of monthly (max temp–min temp)); BIO3: isotermality (BIO2/BIO7) ×100); BIO4: temperature seasonality (standard deviation ×100); BIO5: max temperature of the warmest month; BIO6: min temperature of the coldest month; BIO7: temperature annual range (BIO5-BIO6); BIO8: mean temperature of the wettest quarter; BIO9: mean temperature of the driest quarter; BIO10: mean temperature of the warmest quarter; BIO11: mean temperature of the coldest quarter; BIO12: annual precipitation; BIO13: precipitation of wettest month; BIO14: precipitation of driest month; BIO15: precipitation seasonality (coefficient of variation); BIO16: precipitation of wettest quarter; BIO17: precipitation of driest quarter; BIO18: precipitation of warmest quarter; BIO19: precipitation of coldest quarter. We obtained the corresponding climate features from the latitude and longitude of each plant observation and calculated the 10, 25, 50, 75, and 90 quartiles. The Pearson Coefficient Correlation analysis was performed by R scripts, and the corresponding P values were adjusted.
by the Benjamini–Hochberg method (FDR). The lowest FDR among the different quartiles of climate features was used to represent the corresponding features.

Alignment of the sequencing reads
For RNA-seq and polysome-seq libraries, reads were directly used for alignment. For SHALiPE-seq libraries, the first 3 bases at the 5’-end (random nucleotides on the adaptor for cDNA ligation) of the raw reads were cropped. The alignment was carried out against Arabidopsis transcriptome TAIR10 release using bowtie version 1.0.1 with iterative mapping procedure. The minimum read length allowed to map was fixed to 21 bases long. The resulting mapped bam files were converted to sam files and indexed using samtools-1.4.11. The stop counts were extracted using HTseq v0.7.2 and the code was written in Python v2.7.15. Reads counts of replicates were merged after observing high correlations among these.

Gini index and folding score calculation
Gini index was calculated from the SHALiPE-seq libraries with reads number of G residues in G-tract as described.

\[
Gini = \sum_{i=1}^{n} \frac{n_i}{2n^2} r_i - \sum_{i=1}^{n} \frac{n_i}{2n^2} r_j
\]

\(n\) denotes the number of G residues in the G-tracks, and \(r_i\) denotes the reads number in SHALiPE profiling at position \(i\). To calculate the folding score, the regions with \(G\) in vitro \(K^+\)/\(G\) in vitro \(Li^+\) ≥1.1 and \(G\) residues with average reads count ≥10 were included.

\[
\text{folding score} = \frac{\text{Gini in vivo} - \text{Gini in vitro } Li^+}{\text{Gini in vitro } K^+ - \text{Gini in vitro } Li^+}
\]

GO analysis
Enrichment analysis of GO categories was performed online by AgriGo.

Reporting summary
Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The raw sequencing data have been deposited in the Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/sra) under BioProject ID number PRJNA762705. SHALiPE-seq data in vitro and in vivo at 22 °C are available under BioProject ID number PRJNA561194. The processed data are provided in the Supplementary Data 1–7. Resources for the latitude and longitude of plant geographic distribution obtained from observations collected by the Global Biodiversity Information Facility (www.gbif.org) have been listed in Supplementary Data 2, i.e. Arabidopsis thaliana: https://doi.org/10.15468/dlf.pmy46j. WorldClim database could be accessed through the server https://worldclim.org/.

References
1. Smarda, P. et al. Ecological and evolutionary significance of genomic GC content diversity in monocots. Proc. Natl Acad. Sci. USA 111, E4096–E4102 (2014).
2. Xu, S., Wang, J., Guo, Z., He, Z. & Shi, S. Genomic convergence in the adaptation to extreme environments. Plant Commun. 1, 100117 (2020).
3. One Thousand Plant Transcriptomes. I. One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574, 679–685 (2019).
4. Fick, S. E. & Hjimans, R. J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315 (2017).
5. Savolainen, O., Lascoux, M. & Merila, J. Ecological genomics of local adaptation. Nat. Rev. Genet. 14, 807–820 (2013).
6. Verdu, M., Gomez, J. M., Valiente-Banuet, A. & Schob, C. Facilitation and plant–plant phenotypic evolution. Trends Plant Sci. 26, 913–923 (2021).
7. Lancaster, L. T. & Humphreys, A. M. Global variation in the thermal tolerances of plants. Proc. Natl Acad. Sci. USA 117, 13580–13587 (2020).
8. Kwok, C. K., Marsico, G. & Balasubramanian, S. Detecting RNA G-Quadruplexes (G4s) in the Transcriptome. Cold Spring Harb. Perspect. Biol. 10, a032284 (2018).
9. Biffi, G., Di Antonio, M., Tannahill, D. & Balasubramanian, S. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. Nat. Chem. 6, 75–80 (2014).
10. Rodriguez, R. et al. A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. J. Am. Chem. Soc. 130, 15758–15759 (2008).
11. Yang, X. et al. RNA G-quadruplex structures exist and function in vivo in plants. Genome Biol. 21, 226 (2020).
12. Kwock, C. K., Sahakyan, A. B. & Balasubramanian, S. Structural analysis using SHALiPE to reveal RNA G-quadruplex formation in human precursor microRNA. Angew. Chem. Int. Ed. Engl. 55, 8958–8961 (2016).
13. Spitale, R. C. et al. RNA SHAPE analysis in living cells. Nat. Chem. Biol. 9, 18–20 (2013).
14. Guo, J. U. & Bartel, D. P. RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria. Science 353, aa5371 (2016).
15. Bedrat, A., Lacroix, L. & Mergny, J. L. Re-evaluation of G-quadruplex propensity with G4Hunter. Nucleic Acids Res. 44, 1746–1759 (2016).
16. Cho, H. et al. Translational control of phloem development by RNA G-quadruplex-JULG1 determines plant sink strength. Nat. Plants 4, 376–390 (2018).
17. Wu, X. & Bartel, D. P. Widespread influence of 3′-end structures on mammalian mRNA processing and stability. Cell 169, 905–917 (2017).
18. Goodarzi, H. et al. Systematic discovery of structural elements governing stability of mammalian messenger RNAs. Nature 485, 264–268 (2012).
19. Khong, A. & Parker, R. mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction. J. Cell Biol. 217, 4124–4140 (2018).
20. Sorenson, R. S., Deshotel, M. J., Johnson, K., Adler, F. R. & Sieburth, L. E. Arabidopsis mRNA decay landscape arises from specialized RNA decay substrates, decapping-mediated degradation, and redundancy. Proc. Natl Acad. Sci. USA 115, E1485–E1494 (2018).
21. Cruz, J. A. & Westhof, E. The dynamic landscapes of RNA architecture. Cell 136, 604–609 (2009).
22. Cecch, T. R. The mRNA worlds in context. Cold Spring Harb. Perspect. Biol. 4, a006742 (2012).
23. O’Malley R.C., Berry C.C, Ecker J.R. Plant functional genomics. In: NatureMalley R.C., Berry C.C, Ecker J.R. Plant functional genomics. In: Nature
27. Zhang, X., Henriques, R., Lin, S. S., Niu, Q. W. & Chua, N. H. Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat. Protoc. 1, 641–646 (2006).
28. Heckman, K. L. & Pease, L. R. Gene splicing and mutagenesis by PCR-driven overlap extension. Nat. Protoc. 2, 924–932 (2007).
29. Richardson, L. L. & Dym, M. Improved adhesiveness of polyester wax sections for immunocytochemistry. Biotechniques 17, 846–848 (1994).
30. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675 (2012).
31. Bahry, E. et al. RS-FISH: Precise, interactive and scalable smFISH spot detection using radial symmetry. bioRxiv, https://doi.org/10.1101/2021.03.09.434205 (2021).
32. Ding, Y., Kwok, C. K., Tang, Y., Bevilacqua, P. C. & Assmann, S. M. Genome-wide profiling of in vivo RNA structure at single-nucleotide resolution using structure-seq. Nat. Protoc. 10, 1050–1066 (2015).
33. Li, W. et al. EIN2-directed translational regulation of ethylene signaling in Arabidopsis. Cell 163, 670–683 (2015).
34. Misra, A. & von Arnim, A. G. Analysis of mRNA translation states in Arabidopsis over the diurnal cycle by polysome microarray. Methods Mol. Biol. 1158, 157–174 (2014).
35. Langmead, B. Aligning short sequencing reads with Bowtie. Curr. Protoc. Bioinform. Chapter 11, Unit 11.17 (2010).
36. Norris, M. et al. FoldAtlas: a repository for genome-wide RNA structure probing data. Bioinformatics 33, 306–308 (2017).
37. Li, H. et al. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).
38. Du, Z., Zhou, X., Ling, Y., Zhang, Z. & Su, Z. agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. 38, W64–W70 (2010).

Acknowledgements

The study is supported by the National Natural Science Foundation of China (32170229) (H.Z.), the National Key Research and Development Program of China (2021YFF1000900) (H.Z.), United Kingdom Biotechnology and Biological Sciences Research Council (BBSRC: BBS/E/J/000PR9788, BB/L025000/1, and BB/N022572/1) (Y.D.), Norwich Research Park Science links seed fund (Y.D.), the European Research Council (ERC: 680324) (Y.D.), Human Frontier Science Program Fellowship (LTO10177/2021-L) (H.Y.), Shenzhen Basic Research Project (JCYJ201805071642811) (C.K.K.), Research Grants Council of the Hong Kong SAR, China Projects (CityU 11100421, CityU 11101519, CityU 11100218 and N_CityU110/17) (C.K.K.), Croucher Foundation Project (95090033) (C.K.K.), State Key Laboratory of Marine Pollution Director Discretionary Fund (C.K.K.). We thank professor Beverley Glover (U. Cambridge), Dr. Sam Brockington (U. Cambridge), professor Shiheng Tao (Northwest A&F University), professor Giles Oldroyd (SLCU, Cambridge), and Dame professor Caroline Dean (John Innes Centre), Dr. Antony Dodd (John Innes Centre), professor Cristobal Uauy (John Innes Centre), professor Alison Smith (John Innes Centre) and Dr. Desmond Bradley (John Innes Centre) for discussions with this work. We thank the John Innes Centre Bioimaging facility and staff for their contribution to this publication. This research was supported by the Norwich Biosciences Institutes Partnership’s Computing infrastructure for Science (CiS) group through the provision of a High-Performance Computing Cluster and the John Innes Centre Informatics team.

Author contributions

Y.D. conceived the study; X.Y., H.Y., J.B.M., C.K.K., H.Z., and Y.D. designed the study; X.Y., S.D., Y.Z., H. L., and J.Z. performed the experiments; H.Y., X.Y., S.D., and J.C. did the analyses; H.Z. and Y.D. supervised the analyses; X.Y., H.Y., S.D., J.C., H.Z., and Y.D. wrote the paper with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-34040-y.

Correspondence and requests for materials should be addressed to Huakun Zhang or Yiliang Ding.

Peer review information Nature Communications thanks Martin Crespi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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

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

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

© The Author(s) 2022