PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Review

Cite this article: Marondedze C. 2020 The increasing diversity and complexity of the RNA-binding protein repertoire in plants. *Proc. R. Soc. B* **287**: 20201397. http://dx.doi.org/10.1098/rspb.2020.1397

Received: 15 June 2020
Accepted: 1 September 2020

Subject Category:
Development and physiology

Subject Areas:
biochemistry, developmental biology, plant science

Keywords:
RNA-binding protein, RNA interactome capture, UV crosslinking, abiotic stress, mass spectrometry

Author for correspondence:
C. Marondedze
e-mail: cmarondedze@gmail.com

†Present address: Rijk Zwaan, De Lier 2678 ZG, Netherlands.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5113668.

THE ROYAL SOCIETY PUBLISHING

The increasing diversity and complexity of the RNA-binding protein repertoire in plants

C. Marondedze[1,2,3,†]

[1]Cambridge Centre for Proteomics, Department of Biochemistry, University of Cambridge, Cambridge CB2 1GA, UK
[2]Biological and Environmental Sciences and Engineering Division, 4700 King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia
[3]Department of Biochemistry, Midlands State University, P. Bag 9055, Gweru, Zimbabwe

CM, 0000-0002-2113-904X

Post-transcriptional regulation has far-reaching implications on the fate of RNAs. It is gaining increasing momentum as a critical component in adjusting global cellular transcript levels during development and in response to environmental stresses. In this process, RNA-binding proteins (RBPs) are indispensable chaperones that naturally bind RNA via one or multiple globular RNA-binding domains (RBDs) changing the function or fate of the bound RNAs. Despite the technical challenges faced in plants in large-scale studies, several hundreds of these RBPs have been discovered and elucidated globally over the past few years. Recent discoveries have more than doubled the number of proteins implicated in RNA interaction, including identification of RBPs lacking classical RBDs. This review will discuss these new emerging classes of RBPs, focusing on the current state of the RBP repertoire in *Arabidopsis thaliana*, including the diverse functional roles derived from quantitative studies implicating RBPs in abiotic stress responses. Notably, this review highlights that 836 RBPs are enriched as *Arabidopsis* RBPs while 1865 can be classified as candidate RBPs. The review will also outline outstanding areas within this field that require addressing to advance our understanding and potential biotechnological applications of RBPs.

1. Introduction

RNA–protein interaction is an imperative checkpoint to fine-tune gene expression at the RNA level. RNA-binding proteins (RBPs) interact with the untranslated regions of RNAs that have cis-acting regulatory functions forming dynamic ribonucleoprotein (RNP) complexes that control the fate of RNA. RBPs regulate the synthesis, editing, processing (including capping, splicing and polyadenylation), transport and localization, storage, translation and turnover of RNA (electronic supplementary material, figure S1) [1]. In plants, the diversity of RBPs is rather complex. A vast repertoire of RBPs exists in mitochondria and chloroplasts (for review see [2,3]); however, this review will focus more on the nuclear-cytosolic RBPs.

Typically, classical RNA-binding domains (RBDs) are used to annotate RBPs *in silico* based on the knowledge of conserved domain structure and function. In plants, the RNA recognition motifs (RRM) and K homology (KH) domain are among the most common classical domains. The RRM class is the most represented and dominant classical RBD across all interactomes identified even in yeast and animals [4,5]. Thus far and until recently, knowledge of RBPs in plants was acquired mainly from targeted studies on individual proteins or via bioinformatics predictions based on sequence homology with classical RBDs identified in other kingdoms. Of late, global identification of proteins binding *in vivo* to RNA has been made possible through RNA interactome capture (RIC), a method where proteins are fixed to target RNAs by UV crosslinking

and purified through affinity capture of polyadenylated (poly(A)) RNA [5]. Although this method is biased towards poly(A) RNA, it has set the basis for modifications to allow an unbiased global identification of proteins interacting with RNAs *in vivo*. The UV-crosslinking protocol is effective in selecting proteins that directly bind to RNA while discriminating against uncrosslinked proteins including proteins that associate as subunits of larger RNPs that have no direct contact with RNA [5,6]. This is because the protocol does not facilitate protein–protein crosslinking. In addition, the lysis buffer and stringent washes applied permit the dissociation of non-covalently associated protein–protein interactions. Also, stringent statistical criteria are applied to proteins identified in both UV-crosslinked and nUV samples in order to differentiate *bona fide* RBPs from non-specifically bound proteins. Furthermore, western blot analysis is performed to assess the sensitivity and selectivity of the UV-crosslinking technique.

In *Arabidopsis thaliana*, over 800 RBPs were enriched in UV crosslinked samples compared to control non-UV crosslinked samples (electronic supplementary material, file S1) [6–10], and recently, over 50 proteins were reported as enriched in the cereal model plant *Brachypodium distachyon* [11]. These results provide experimental evidence for numerous proteins identified via *in silico* algorithms. In addition, the data offer insights on the role of RBPs in stress responses [12]. Strikingly, as in mammals and yeast (*Saccharomyces cerevisae*), the RNA interactomes revealed novel outcomes with some surprises. For example, many of the proteins trapped bound on to RNA have not previously been linked to RNA-mediated processes, including proteins of the intermediary metabolism [6,7,9]. RBPs have been also reported to play roles in circadian rhythm, flowering transition and responses to biotic and abiotic stresses [8,12–18]. Overall, the studies on RNA–protein interactions afford exceptional insights in the composition of the proteins binding mRNA, underscoring the complexity of *in vivo* RNA-mediated processes.

Given such diverse important roles of RBPs in plants, it is imperative to foster the identification of the novel RBPs, which consequently will facilitate the understanding of plant molecular biology. Although a vast number of studies have discovered an increasingly growing number of classical and novel candidate RBPs, a lot remains to be uncovered from the hidden core of the RNA–protein interactome world. In particular, we need to understand the mechanisms regulating the function of metabolic enzymes in RNA recognition and interaction and their enzymatic activity. Deep insights on such post-transcriptional gene regulatory mechanisms will comprehensively enhance our understanding on the growth and developmental processes in organisms.

2. The RNA-binding protein repertoire studies in *Arabidopsis*

The existence of RBPs in plant cells and their significance in post-transcriptional gene regulation is a well-known phenomenon yet to be fully comprehended. Hitherto, the majority of studies only identified and functionally characterized a subset of the RBPs with targeted roles such as hormonal responses [19,20], pathogen defence [13,21,22], abiotic stress [23–27], flowering [16,28] and circadian clock [14,18,29]. These individual studies have demonstrated that mutations in specific RBPs can depict severe phenotypes or lethality indicating the crucial role of RBPs on the growth and development of plants [30]. Then, only a limited number of plant RBPs had experimental evidence, but this has significantly improved following attempts by various research laboratories to globally catalogue RBPs *in vivo*.

A limited number of RBPs were retrieved natively from cultured *Arabidopsis* cells by oligo(dT) capture and matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry [31]. Identified proteins comprised of a suite of RRM containing proteins including glycine-rich (GR) RBPs like AtGRP2, 7 and 8, cold-shock domain (CSD) proteins, WHIRLY 3 and members of the chloroplast RBPs such as chloroplast RBP29, a protein shown to interact with nuclear RNA [32].

Furthermore, the first global scale RBP repertoire was captured using the RIC approach employing UV crosslinking at 254 nm, oligo(dT) capture and tandem mass spectrometry performed on cultured *Arabidopsis* cells derived from root cells of Columbia-0 and Landsberg Erecta ecotypes, and 4-week-old *Arabidopsis* plants (figure 1). This approach identified 1145 proteins bound to mRNAs [6]. Moreover, quantitative RIC was employed on drought stress-treated cultured *Arabidopsis* cells identifying an additional 808 RNA-bound proteins [8]. In both independent studies, a total of 1953 proteins were identified, of which 550 proteins were linked to RNA biology and greater than 1000 were detected as novel candidates representing proteins not previously assigned an RNA-related function. Notably, 399 proteins were significantly enriched upon UV crosslinking when compared with the control non-UV crosslinked samples. In this review, proteins enriched in the UV crosslinked samples in comparison to the control non-UV crosslinked samples will be classified as *At*RBPs, while the remaining novel proteins only detected in the UV crosslinked samples will be termed candidate (cand)-*At*RBPs. RNA interaction for novel candidates such as clathrin heavy chain (At3G11130) and catalase 3 were validated using RNA electrophoretic mobility shift assay [6].

Moreover, only a limited number of RBPs (236 proteins) were identified from leaves, suggesting potential challenges associated with UV crosslinking efficiency in photosynthetic active tissues [6]. Similarly, of the 405 proteins identified from the *Brachypodium* RNA interactome, 203 were detected from the seedlings and 287 from leaf mesophyll protoplasts [11]. This limitation may be due to the waxy cuticle leaf layer interfering with the intensity of UV radiation reaching the cytoplasm at wavelength less than 400 nm [33], leading to insufficient crosslinking of RBP-RNA molecules. Recently, it has been reported that an improved version of RIC could overcome such difficulties. The improved RIC involves UV crosslinking the leaf twice on the adaxial side and once on the abaxial side unlike previously where the crosslinking was performed two to three times on the same side of the leaf. In both cases, 150 mJ cm^{-2} of UV light at 254 nm wavelength was applied. Using the improved RIC, 717 proteins were identified in *Arabidopsis* leaves, with about 75% of these linked to RNA biology [7]. Notably, a large number of chloroplast targeting ribosomal proteins and proteins linked to photosynthesis supracomplexes including photosystems I and II were identified. However, just like other novel RBPs, it is necessary to confirm the RNA-binding activity and target RNAs to shed light on their functional significance in RNA regulations.

Additional evidence of RBP diversity in *Arabidopsis* was revealed from etiolated *Arabidopsis* seedlings using the RIC

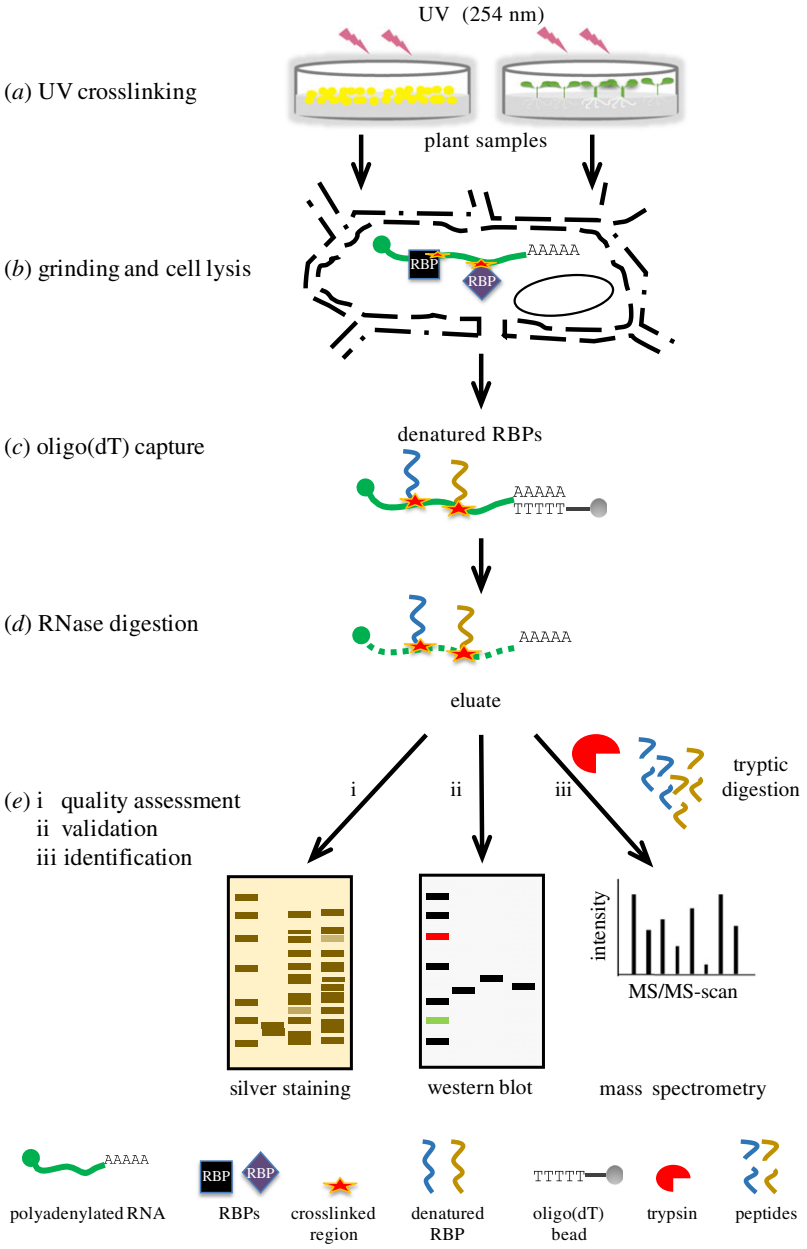

Figure 1. A schematic workflow of the plant RNA interactome capture technique. (*a*) Plant samples are irradiated with UV light at a wavelength of 254 nm to fix RNA and proteins that are in intimate contact. (*b*) Samples are ground in liquid nitrogen and cells are lysed in a buffer promoting denaturation of proteins. (*c*) polyadenylated RNAs are pulled down using oligo(dT) magnetic beads followed by stringent washes, (*d*) recovered RNA–protein complexes are subjected to RNAse digestion to release RNA-binding proteins (RBPs). (*e*) Proteins are quality checked and quantitatively analysed by (i) silver staining, (ii) western blot (also for validation of the technique) and (iii) identification of the RBPs by mass spectrometry following trypsin digestion. (Online version in colour.)

approach. A total of 746 proteins were identified and among these 299 were *At*RBPs [9]. Three-quarters of these *At*RBPs possessed known RNA biology-related functions, while only 46% of the cand-*At*RBPs were linked to RNA biology. Of interest to note is the limited capture of organelle localized or targeted RBPs in particular mitochondrial and chloroplast RBPs, e.g. only 18 pentatricopeptide repeat (PPR) proteins were detected in the crosslinked samples, which is far fewer than the 450 predicted PPRs in *Arabidopsis* [34]. Authors postulated that the low number of PPRs identified may be due to the developmental stage of the etioplasts that are yet to differentiate. The low number of PPRs can also be explained by the limitation of the RIC approach, that enriches poly(A) RNA, which are known to be limited in the mitochondria and chloroplast [35]. Perhaps an organelle targeted enrichment would enhance PPRs identification.

Another study using the RIC approach on *Arabidopsis* mesophyll protoplast detected 325 proteins of which 100 are *At*RBPs [10]. Similarly to the observations made in previous studies[6,9], ribosomal proteins were highly represented, accounting for 38% of the UV-enriched proteins. Besides, 70 proteins contained known RBDs, while 132 proteins constituted cand-*At*RBPs. The latter was dominated by metabolic enzymes (49 proteins) and photosynthesis-related proteins (29 proteins). Just like most novel cand-*At*RBPs from previous studies, their RNA-binding activity is yet to be confirmed. However, RNA-binding capacity of some enzymes such as the plant orthologue of yeast phosphoglycerate kinase was noted in both yeast and human cells [4] denoting that this enzyme and other plant enzymes could play a role in RNA metabolism although they lack conventional RBDs.

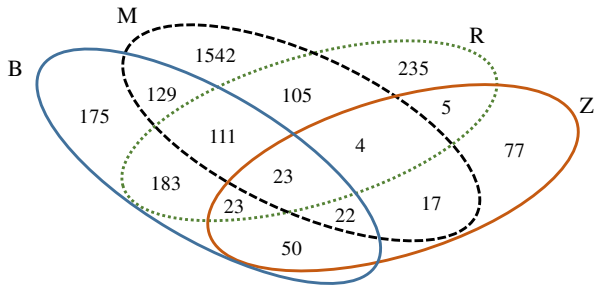

Figure 2. *Arabidopsis* RBP repertoire identified by RNA interactome capture. An overlap of all 2701 proteins (electronic supplementary material, file S1) identified bound to mRNA upon UV cross-linking in leaf tissue (B) [7], leaf and cultured *Arabidopsis* cells (M) [6,8], etiolated seedlings (R) [9] and leaf mesophyll protoplasts (Z) [10]. (Online version in colour.)

Overall, novel insights are gained from these diverse RNA interactomes generated from different plant tissues thus far. In particular, the complexity and variation among the diverse tissues or cells was depicted by the low overlap (figure 2). Furthermore, the impact of environmental cues denote that the scope and function of RNA–protein interactions is rather broad and informative about the developmental and physiological states within plant cells as is in other biological systems such as animals and yeast.

3. Drought stress-induced RNA interactome changes

As sessile organisms, plants are often subjected to variable environmental conditions including extreme cues like drought, heat and salinity that are unfavourable for their growth and productivity. Adaptation to such adverse effects of stress involves reprogramming of cellular events led by signalling networks driving modifications in gene expression and metabolism to turn on protective mechanisms. These activities partly rely on post-transcriptional modifications that in turn determine the ultimate fate of expressed genes. A proteomics analysis of oligo(dT)-bound messenger RNP revealed that peroxide-induced oxidative stress stimulated differential regulation of *Arabidopsis* glycine-rich RBPs, *At*GRP7 and *At*GRP8 [31]. Expression of these two proteins was rapidly upregulated upon oxidative stress. However, in response to drought stress, the abundance (in the affinity-purified fractions) of *At*GRP8 decreased [8].

A quantitative RIC approach was applied to investigate, at the systems level, the effect of drought stress on the RBP repertoire and its biological significance. Of the 567 RNA-bound proteins identified responding to the stress, abundance of 150 significantly ($p < 0.05$) changed in association with RNA and 417 were time-dependent transient changes, either detected only in the control (untreated) or treated samples [8], implying that drought stress induced new RNA–protein interactions or dissociations. Classical drought response proteins were detected including abscisic acid (ABA) hypersensitive 1 (or cap-binding protein 80), hyaluronan protein (*At*RGGA) and aldehyde dehydrogenase (ALDH) 7B4. The abundance of these proteins increased over time in correlation with the increase in ABA concentration [12]. In addition, intermediary metabolism proteins were also modified linking post-transcriptional gene regulation to stress-induced metabolic changes and this may be indicative of the regulatory role of these RBPs on their own mRNAs. Differential regulation of 44 spliceosome components was observed denoting their key role in orchestrating changes of the transcriptome in response to exogenous cues [36]. Identification of splicing factors may reflect a direct and/or indirect stress-induced splicing events that have a direct effect on transcriptome and possibly proteome changes under stress conditions. Thirty-two stress granule components were detected in consistent with a transcriptional arrest phenomenon. Transcriptional arrest has been observed to occur during stress exposure, inducing stress granule formation. Stress granules are supramolecular cytoplasmic foci resulting from cytoplasmic aggregates of non-translated mRNPs [37]. Stress granules have been detected in other stress responses such as low oxygen, oxidative and heat stress [38–40]. However, a comprehensive stress granule focused research is required to gain insights on their global composition and functional significance in stress adaptation. Identification of RBPs including stress granule components under various stresses is critical in determining common abiotic stress-induced targets that can have biotechnological applications. Moreover, RBPs have been proposed as targets to improve stress tolerance in crops, e.g. a recent study showed that in halotolerant sugar beet (*Beta vulgaris* (Bv)) expression of *BvSATO1*, an RNA metabolism associated RBP, was repressed by salt treatment, while in *Arabidopsis* *BvSATO1* increased salt tolerance [41].

Overall, various studies indicate that RBPs play crucial roles in stress adaptation and tolerance with functions including control and stability of metabolic process, RNA splicing and RNA metabolic processes. These works set the foundation for future mechanistic approaches to elucidate biological significance of RBPs and their target RNAs in abiotic stress responses and consequently towards crop improvement.

4. The contemporary RNA-binding protein repertoire in *Arabidopsis*

Although different criteria were used to classify *At*RBPs, various studies described in this review identified the largest documented *Arabidopsis* RBP repertoire to date. Here, a detailed view of the up-to-date state of the RBPome focusing on experiments that used the RIC approach will be discussed [6–10]. These five datasets are essential in understanding the complexity of the post-transcriptional gene regulatory processes in plants. It is also conceivable to speculate that some of the identified RBPs may be regulated by RNA [42]. Altogether, the studies identified 2701 unique RNA-bound proteins (figure 2), of which 31% (836 proteins) are classified as *At*RBPs (considering minimum enrichment of $\log_2$ fold-change greater than 1.5 and a FDR less than 5%) and 69% (1865 proteins) as cand-*At*RBPs (electronic supplementary material, file S1). Of the *At*RBPs identified, 456 proteins are present in at least two studies and only 18 are common in all the studies. Of the cand-*At*RBPs, 216 proteins are present in at least two studies. These small overlaps between datasets can be due to various reasons such as differences in the type and physiological states of the tissues used, growth conditions, environmental stress, variation in analysis pipelines including mass spectrometers, softwares and the statistical criteria that were applied. Moreover, some known RBPs were not enriched in the UV crosslinked samples

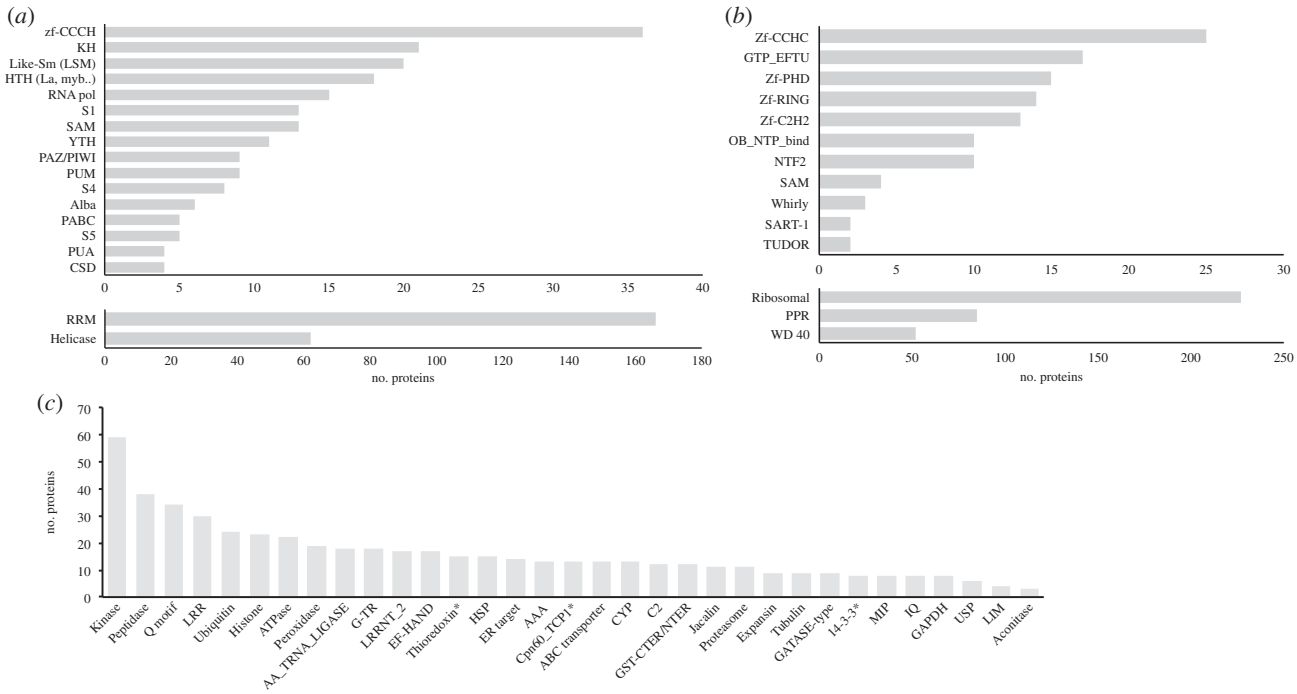

Figure 3. Domains detected in the *Arabidopsis* RBP repertoires. Number of proteins harbouring (*a*) classical RBDs, (*b*) non-classical RBDs and (*c*) unknown RNA interacting domains in plants. The domains are uncovered from the 2701 determined *Arabidopsis* mRNA interactors identified upon UV crosslinking [6–10]). Note that some domains such as ribosomal are a pool of individual domains (see electronic supplementary material, file S2). Classification of domains into classical, non-classical and unknown RBDs is based on [4,6].

although they were identified in the control samples denoting that they may be either involved in other molecular functions, not interacting with mRNA at the time of the experiment and possibly interacting with non-poly(A) RNAs or not efficiently crosslinked due to, for example, the geometry of the specific protein-RNA interactions.

5. RNA-binding protein repertoire reveals diverse proteins with a wide range of RNA-binding domains

In general, the data provide experimental evidence for RNA-protein association for numerous *in silico* determined proteins. For example, a total of 166 (greater than 80%) of 197 predicted RRM-containing proteins, 62 out of 182 helicases, 21 out of 28 predicted KH domain proteins, 11 out of 13 YT521-B homology and 10 out of 18 Nuclear Transport Factor 2 were described (figure 3*a*; electronic supplementary material, file S2). The most represented non-classical RBDs are the ribosomal proteins comprising of 227 out of 524 *Arabidopsis* ribosomal proteins based on AgriGO (http://bioinfo.cau.edu.cn/agriGO/). Numerous ribosomal proteins have been shown to interact directly with mRNA and playing extra-ribosomal functions in mRNA regulation [43].

Multiple zinc finger (zf) sub-types were noted such as zf-CCCH, zf-C2H2, zf-CCHC, zf-RING, zf-PHD and zf-RanBP (figure 3*b*). Various zf sub-types have been shown to interact with RNA but some represent novel RNA interactions [44,45]. Additionally, abiotic stress linked RBP families were also identified including the tudor-SN, *At*GRPs, Like-Sm (LSM) and CSD proteins [15,46,47]. Remarkably, some stress-linked domains are not evolutionarily conserved like the orthologues of CSD3 (At2G17870) in animal systems that lacked the CSD, suggesting that it may have evolved for plant optimal survival under abiotic stresses [8].

Several other domains were detected within the *Arabidopsis* RBP repertoire whose roles in RNA interaction are yet to be fully elucidated, suggesting the existence of new modes of RNA binding. Besides, the presence of uncharacterized RBDs is not just a phenomenon in plants but spans from yeast to human [4]. Essentially, the discovery of novel previously unknown RBDs unveils new ways of looking at RNA-protein research. Noteworthy is the presence of eight MIP domain containing plasma-membrane intrinsic proteins (aquaporins) (figure 3*c*). This class of proteins has not been directly linked to RNA interaction. However, as transmembrane channels transporting various substrates including small solutes, gases and water [9,48], it is tempting to speculate that RNA may be one of the substrates for aquaporins, contributing to cell-to-cell transport just like protein assisted cell-to-cell movement of RNA during virus infection [9,49].

Eight 14–3-3 containing proteins were identified that belong to the general regulatory factor family (figure 3*c*). 14-3-3 containing proteins have roles in intracellular signalling by directly regulating either catalytic activity of their interacting partners, interactions between bound proteins and other cellular molecules, or regulating subcellular localization. Although their role in RNA interaction is yet to be established, 14-3-3 proteins may regulate protein–RNA interactions within RNPs or establishing RNA localization.

Nineteen peroxidase domain-containing proteins were detected (figure 3*c*). Peroxidases are well-known hydrogen peroxidase scavenging enzymes that are essential in stress tolerance induced by oxidative stress [50,51]. A few peroxidases have been shown to interact with RNA such as thioredoxin peroxidase 1 from *Plasmodium knowlesi* [52]. Nevertheless, in plants the role of peroxidases in RNA interaction is yet to be elucidated.

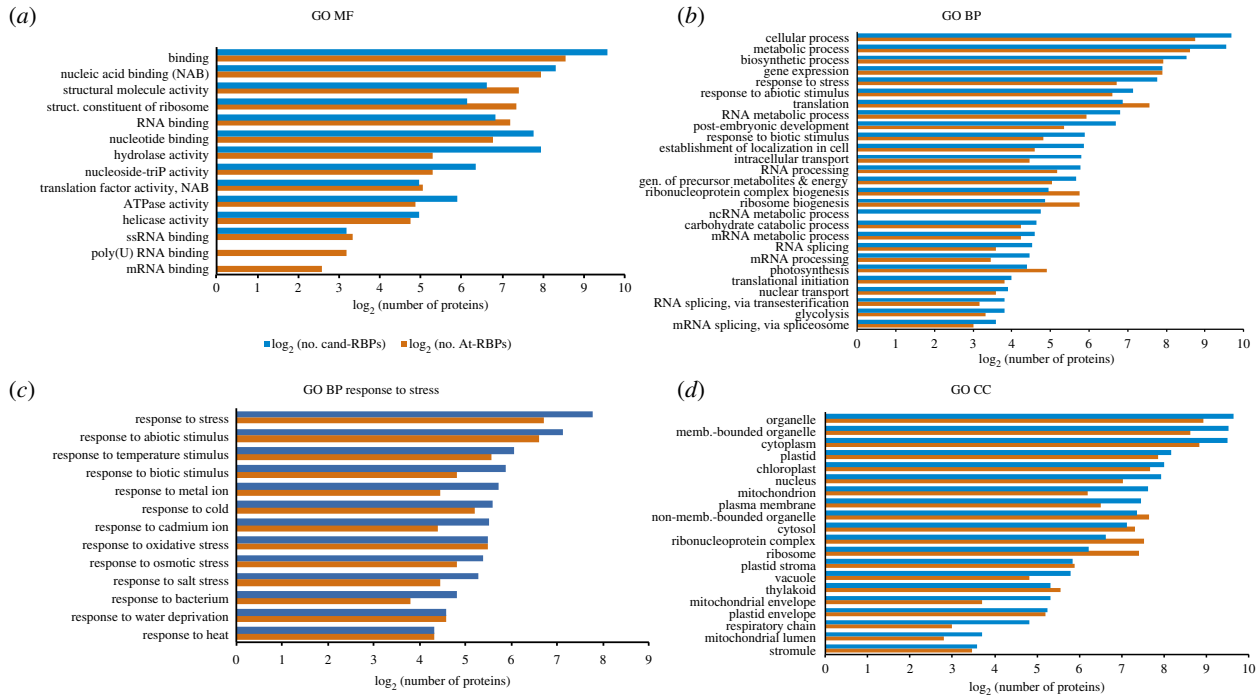

Figure 4. GO analysis of the *Arabidopsis* RBP repertoire enriched upon UV crosslinking. Number of proteins significantly enriched (adjusted $p < 0.05$) in the categories (*a*) 'molecular function', (*b*) 'biological process', (*c*) subcategory 'response to stress' and (*d*) 'cellular component'. Blue bars represent $\log_2$ of the number of candidate RBPs and orange bars $\log_2$ of *At*-RBPs. Complete GO list is in electronic supplementary material, file S3. (Online version in colour.)

Eight GAPDH NAD-binding and GAPDH C-terminal domain containing proteins were noted (figure 3*c*). Some of these NAD-binding domain-containing proteins were shown to interact with RNA during drought stress [8]. In non-plant systems, *in vivo* and *in vitro* evidence confirm the existence of RNA-binding activities within the NAD-binding pocket of GAPDH [53–55]. GAPDH has been shown to bind to diverse RNA species including AU-rich elements, tRNAs and TERC [56,57]. Based on these evidences, it is tempting to suggest that the same principle is conserved in plants, and that GAPDH potentially interact with RNA through its NAD-binding pocket.

6. Insights from the gene ontology analysis of the *Arabidopsis* RNA–binding protein repertoire

Gene ontology (GO) exploration using singular enrichment analysis in AGRIGO facilitated categorization of the *Arabidopsis* RBP repertoire (electronic supplementary material, file S3). *Arabidopsis* genemodel TAIR9 was used as a reference with advanced options: Fisher exact statistical test, Yekutieli (FDR under dependency) as multi-test adjustment and p-value $\leq$ 0.05. Molecular function analysis showed a bias towards categories such as 'RNA binding', 'mRNA binding', 'poly(U) RNA binding', 'single-stranded RNA binding' 'structural molecule activity' and 'helicase activity' (figure 4*a*). RNA binding functions are highly enriched in both *At*RBPs and cand-*At*RBPs, while catalytic activities are more enriched in the cand-*At*RBPs. Biological processes enriched included 'RNA metabolism processes' and 'response to stress' (figure 4*b*,*c*).

Notably, 51 proteins were enriched in the category 'photosynthesis', representing an essential plant specific gene regulatory network. Besides, a high number of proteins were enriched in the category 'catalytic activities' including various classes of enzymes such as ATPases, helicases, hydrolyases, lyases, oxidoreductases, peptidases, peroxidases and transferases, highlighting the diversity and complexity of RBPs (figure 4*b*; electronic supplementary material, file S3). Notably, 23 proteins with glycolytic functions were identified, highlighting the existence of dual functionality as reported in other systems [4,5]. Strikingly, some of these intermediary metabolism proteins such as phosphofructokinase, GAPDH, pyruvate dehydrogenase and ALDH7B4, are quantitatively regulated by drought stress [8].

Determining localization is important to understand the role of RBPs in intracellular trafficking of RNAs and localized protein biosynthesis and organelle biogenesis [58–61]. Although for years it has been proposed that translation is customarily cytosolic or associated with the ER, recent studies have reported the existence of nuclear encoded cytosolic mRNAs and ribosomes on the surface of organelles such as mitochondria and chloroplasts [62–67]. This essentially marks one of the primary roles of RBPs in RNA trafficking and sorting. Thus far, cellular compartments are significantly enriched in the *Arabidopsis* RBP repertoire including 'chloroplast' (205 *At*RBPs and 255 cand-*At*RBPs), 'nucleus' (130 *At*RBPs and 244 cand-*At*RBPs) and 'mitochondrion' (73 *At*RBPs and 197 cand-*At*RBPs) (figure 4*d*). These enrichments signal the existence of organelle targeted mRNA localization that could serve in organelle surface translation machinery or translational control. These enrichments partially denote the localization of certain proteins like for 73 mitochondria *At*RBPs, 7 and 13 proteins are enriched in the mitochondrial lumen and envelope, respectively (electronic supplementary material, file S3). This paves the way for targeted functional analysis and detection of their mRNA counterparts. However, globally, further elucidation is warranted. Besides, targeting of mRNAs to the surface of mitochondria for targeted translation would serve as an efficient mechanism for mitochondrial function in ATP

production in particular for the synthesis of metabolic enzymes to allow for a rapid adaptation of energy metabolism according to physiological and external environmental stimuli.

7. Amino acid motif enrichment in the *Arabidopsis* RNA-binding protein repertoire

Analysis of amino acid sequence motif enrichment was performed using the Discriminative Regular Expression Motif Elicitation (DREME) with an *E-value* threshold of 0.05 and allowing for motif widths between 3 and 12 amino acids [68]. DREME is part of the MEME suite (v. 5.1.1, available at http://meme-suite.org/index.html). Of the 12 amino acid motifs most significantly enriched within the *Arabidopsis* RBP repertoire were eight glycine (G)-rich motifs including previously detected motifs such as GGGY, FVGGL, GYGFV, GTGKT and GSGKT that have various roles in pre-mRNA processing (electronic supplementary material, figure S2A-H). Additionally, poly(D), poly(E), poly(S) and poly(P) motifs were detected, of which proline was previously significantly enriched in plant RBPomes [8,9] (electronic supplementary material, figure S2I-L). Poly(P) motif is involved in inter- and intra-molecular interactions [69]. Enriched amino acid motifs in the RBP repertoire comprise mostly of intrinsically disordered, low complexity sequences. These regions of low complexity play an essential role in phase separation in cells leading to the formation of RNA granules in cell-free systems and likewise *in vivo* [70,71].

8. Conclusion and future perspectives

Fundamental research on plant RNA-bound proteins has unveiled more than 2700 RBPs, representing a new order worldwide towards unraveling cellular complexity at the level of post-transcriptional gene regulation. The functions of RBPs emanates from their capacity to bind specific compendia of RNA molecules and direct several post-transcriptional RNA processing including RNA splicing, editing, transport and decay. Perhaps not surprisingly, a wealth of both expected and unanticipated proteins stemmed. Additional knowledge gained from RNA interactomes in various organisms dictates the presence of RBPs with dual functionality such as enzymatic activity suggesting that a large number of significant RBPs are yet to be discovered. In plants, the global interrogation of RBPs is still at its infancy with as yet many unknowns but the studies presented in this review provides a good starting point in understanding the regulatory roles of RBPs in the growth, development and stress responses of plants. In the latter, differential regulation of RBPs under environmental stress could offer an advantage for plants to cope with the stress and provide candidates for biotechnological applications that can be used to make crop plants tolerant to environmental stresses.

Besides these fascinating steps taken towards understanding the composition and system-wide functions of mRNA-binders, a lot still remains to be uncovered for a comprehensive understanding of the RNA-based regulation(s). This includes attending to the following:

(1) Advance our technical expertise in plant RNA interactome investigations through optimizing RIC approach or employing some of the recently developed non-mRNA-biased techniques from other systems, and in addition investigate organelle-specific RBP repertoires.
(2) Explore RNA interactomes to define structure and diversity of the RNA-binding sites on the RBPs and determine an RBDmap as described in mammalian cells [72].
(3) Characterize at the functional level the biological significance of RBPs, especially the novel cand-*At*RBPs and their relevance in stress responses.
(4) Dissect the RNP complex remodelling events under various developmental and stress stimuli including abiotic and biotic stresses (as described in [73] for example).
(5) Unveil the RBPs regulatory network impacting the fate of targeted RNAs. This can be exploited through genome-wide profiling of RBP *in vivo* target RNAs via advanced RNA immunoprecipitations such as individual nucleotide resolution crosslinking immunoprecipitation (iCLIP) [74] or high throughput sequencing (HITS)-CLIP [75].
(6) Determine the RNA-binding complement of a single RNA or an RNA regulon. Employing the MS2 stem-loop tagging described in yeast [76], RNA-protein interaction detection approach (RaPID) [77] or ascorbate peroxidase (APEX)-catalysed RNA biotinylation [78] could allow uncovering proteins interacting with a single RNA *in vivo*. This is essential for targeted engineering of a single RNA for biotechnological intervention towards improving crops for tackling food security.
(7) Unveil the organelle surface-targeting RNPs, organelle-specific translation platforms, and role in organelle biogenesis and localized protein synthesis.
(8) Deduce RBP networks to gain insights into the comprehensive scope of post-transcriptional RNA regulation in plants.

Overall, the plant-adapted UV crosslinking approach initiative will help facilitate answering the above and further unanswered questions in fostering our understanding of the post-transcriptional RNA regulation in plants.

Data accessibility. The datasets supporting this article have been uploaded as part of the electronic supplementary material.

Competing interests. The author declares no conflict of interest.

Funding. The work was supported by the KAUST Global Collaborative Research (grant no. CRG3-62140383).

Acknowledgements. The author wishes to thank Ludivine Thomas, Chris Gehring and Marco Chiapello for their critical reviewing of the manuscript and Chloé-Ann Marondedze for her assistance on the figures.

References

1. Glisovic T, Bachorik JL, Yong J, Dreyfuss G. 2008 RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* **582**, 1977–1986. (doi:10.1016/j.febslet.2008.03.004)
2. Barkan A. 2011 Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. *Plant Physiol.* **155**, 1520–1532. (doi:10.1104/pp.110.171231)
3. Hammani K, Bonnard G, Bouchoucha A, Gobert A, Pinker F, Salinas T, Giegé P. 2014 Helical repeats modular proteins are major players for organelle gene expression.

Biochimie **100**, 141–150. (doi:10.1016/j.biochi.2013.08.031)

4. Beckmann BM *et al*. 2015 The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat. Commun.* **6**, 10127. (doi:10.1038/ncomms10127)
5. Castello A *et al*. 2012 Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**, 1393–1406. (doi:10.1016/j.cell.2012.04.031)
6. Marondedze C, Thomas L, Serrano NL, Lilley KS, Gehring C. 2016 The RNA-binding protein repertoire of *Arabidopsis thaliana*. *Sci. Rep.* **6**, 29766. (doi:10.1038/srep29766)
7. Bach-Pages M, Homma F, Kourelis J, Kaschani F, Mohammed S, Kaiser M, Van Der Hoorn RAL, Castello A, Preston GM. 2020 Discovering the RNA-binding proteome of plant leaves with an improved RNA interactome capture method. *Biomolecules* **10**, 661. (doi:10.3390/biom10040661)
8. Marondedze C, Thomas L, Gehring C, Lilley KS. 2019 Changes in the *Arabidopsis* RNA-binding proteome reveal novel stress response mechanisms. *BMC Plant Biol.* **19**, 139. (doi:10.1186/s12870-019-1750-x)
9. Reichel M *et al*. 2016 In planta determination of the mRNA-binding proteome of *Arabidopsis* etiolated seedlings. *Plant Cell.* **28**, 2435–2452. (doi:10.1105/tpc.16.00562)
10. Zhang Z, Boonen K, Ferrari P, Schoofs L *et al*. 2016 UV crosslinked mRNA-binding proteins captured from leaf mesophyll protoplasts. *Plant Methods* **12**, 42. (doi:10.1186/s13007-016-0142-6)
11. Li M. 2019 *The effects of ambient temperature on* Brachypodium distachyon *development and mRNA binding proteins in flowering plants*. Leuven, Belgium: KU Leuven.
12. Marondedze C, Thomas L, Gehring C, Lilley KS. 2020 Drought stress causes specific changes to the spliceosome and stress granule components. *Front. Mol. Biosci.* **6**, 163. (doi:10.3389/fmolb.2019.00163)
13. Hackmann C *et al*. 2014 Salicylic acid-dependent and -independent impact of an RNA-binding protein on plant immunity. *Plant Cell Environ.* **37**, 696–706. (doi:10.1111/pce.12188)
14. Heintzen C, Nater M, Apel K, Staiger D. 1997 AtGRP7, a nuclear RNA-binding protein as a component of a circadian-regulated negative feedback loop in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **94**, 8515–8520. (doi:10.1073/pnas.94.16.8515)
15. Kim MH, Sasaki K, Imai R. 2009 Cold shock domain protein 3 regulates freezing tolerance in *Arabidopsis thaliana*. *J. Biol. Chem.* **284**, 23 454–23 460. (doi:10.1074/jbc.M109.025791)
16. Lim MH *et al*. 2004 A new Arabidopsis gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. *Plant Cell.* **16**, 731–740. (doi:10.1105/tpc.019331)
17. Lorkovic ZJ, Barta A. 2002 Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nucleic Acids Res.* **30**, 623–635. (doi:10.1093/nar/30.3.623)
18. Schmal C, Reimann P, Staiger D. 2013 A circadian clock-regulated toggle switch explains AtGRP7 and AtGRP8 oscillations in *Arabidopsis thaliana*. *PLoS Comput. Biol.* **9**, e1002986. (doi:10.1371/journal.pcbi.1002986)
19. Cruz TM, Carvalho RF, Richardson DN, Duque P. 2014 Abscisic acid (ABA) regulation of *Arabidopsis* SR protein gene expression. *Int. J. Mol. Sci.* **15**, 17 541–17 564. (doi:10.3390/ijms151017541)
20. Thatcher LF, Kamphuis LG, Hane JK, Onate-Sanchez L, Singh KB. 2015 The *Arabidopsis* KH-domain RNA-binding protein ESR1 functions in components of jasmonate signalling, unlinking growth restraint and resistance to stress. *PLoS ONE* **10**, e0126978. (doi:10.1371/journal.pone.0126978)
21. Lyons R *et al*. 2013 The RNA-binding protein FPA regulates flg22-triggered defense responses and transcription factor activity by alternative polyadenylation. *Sci. Rep.* **3**, 2866. (doi:10.1038/srep02866)
22. Qi Y *et al*. 2010 A putative RNA-binding protein positively regulates salicylic acid-mediated immunity in *Arabidopsis*. *Mol. Plant Microbe Interact.* **23**, 1573–1583. (doi:10.1094/MPMI-05-10-0106)
23. Daszkowska-Golec A, Wojnar W, Rosikiewicz M, Szarejko I, Maluszynski M, Szweykowska-Kulinska Z, Jarmolowski A. 2013 *Arabidopsis* suppressor mutant of abh1 shows a new face of the already known players: ABH1 (CBP80) and ABI4-in response to ABA and abiotic stresses during seed germination. *Plant Mol. Biol.* **81**, 189–209. (doi:10.1007/s11103-012-9991-1)
24. Huh SU, Paek KH. 2014 APUM5, encoding a Pumilio RNA binding protein, negatively regulates abiotic stress responsive gene expression. *BMC Plant Biol.* **14**, 75. (doi:10.1186/1471-2229-14-75)
25. Kim JS, Jung HJ, Lee HJ, Kim KA, Goh C-H, Woo Y, Oh SH, Han YS, Kang H. 2008 Glycine-rich RNA-binding protein 7 affects abiotic stress responses by regulating stomata opening and closing in *Arabidopsis thaliana*. *Plant J.* **55**, 455–466. (doi:10.1111/j.1365-313X.2008.03518.x)
26. Park HY, Kang IS, Han JS, Lee CH, An G, Moon Y-H. 2009 OsDEG10 encoding a small RNA-binding protein is involved in abiotic stress signaling. *Biochem. Biophys. Res. Commun.* **380**, 597–602. (doi:10.1016/j.bbrc.2009.01.131)
27. Park SJ, Kwak KJ, Oh TR, Kim YO, Kang H. 2009 Cold shock domain proteins affect seed germination and growth of *Arabidopsis thaliana* under abiotic stress conditions. *Plant Cell Physiol.* **50**, 869–878. (doi:10.1093/pcp/pcp037)
28. Schomburg FM, Patton DA, Meinke DW, Amasino RM. 2001 FPA, a gene involved in floral induction in *Arabidopsis*, encodes a protein containing RNA-recognition motifs. *Plant Cell.* **13**, 1427–1436. (doi:10.1105/tpc.13.6.1427)
29. Staiger D. 2001 RNA-binding proteins and circadian rhythms in *Arabidopsis thaliana*. *Phil. Trans. R. Soc. Lond. B* **356**, 1755–1759. (doi:10.1098/rstb.2001.0964)
30. Riehs-Kearnan N, Gloggnitzer J, Dekrout B, Jonak C, Riha K. 2012 Aberrant growth and lethality of *Arabidopsis* deficient in nonsense-mediated RNA decay factors is caused by autoimmune-like response. *Nucleic Acids Res.* **40**, 5615–5624. (doi:10.1093/nar/gks195)
31. Schmidt F, Marnef A, Cheung MK, Wilson I, Hancock J, Staiger D, Ladomery M. 2010 A proteomic analysis of oligo(dT)-bound mRNP containing oxidative stress-induced *Arabidopsis thaliana* RNA-binding proteins ATGRP7 and ATGRP8. *Mol. Biol. Rep.* **37**, 839–845. (doi:10.1007/s11033-009-9636-x)
32. Gosai SJ *et al*. 2015 Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the *Arabidopsis* nucleus. *Mol. Cell.* **57**, 376–388. (doi:10.1016/j.molcel.2014.12.004)
33. Krauss P, Markstadter C, Riederer M. 1997 Attenuation of UV radiation by plant cuticles from woody species. *Plant Cell Environ.* **20**, 1079–1085. (doi:10.1111/j.1365-3040.1997.tb00684.x)
34. Silverman IM, Li F, Gregory BD. 2013 Genomic era analyses of RNA secondary structure and RNA-binding proteins reveal their significance to post-transcriptional regulation in plants. *Plant Sci.* **205–206**, 55–62. (doi:10.1016/j.plantsci.2013.01.009)
35. Chang JH, Tong L. 2012 Mitochondrial poly(A) polymerase and polyadenylation. *Biochim. Biophys. Acta* **1819**, 992–997. (doi:10.1016/j.bbagrm.2011.10.012)
36. Köster T, Marondedze C, Meyer K, Staiger D. 2017 RNA-binding proteins revisited: the emerging *Arabidopsis* mRNA interactome. *Trends Plant Sci.* **22**, 512–526. (doi:10.1016/j.tplants.2017.03.009)
37. Kedersha N, Stoecklin G, Ayodele M, Yacono P *et al*. 2005 Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J. Cell Biol.* **169**, 871–884. (doi:10.1083/jcb.200502088)
38. Anderson P, Kedersha N. 2006 RNA granules. *J. Cell Biol.* **172**, 803–808. (doi:10.1083/jcb.200512082)
39. Ivanov PA, Nadezhdina ES. 2006 Stress granules: RNP-containing cytoplasmic bodies springing up under stress. The structure and mechanism of organization. *Mol Biol.* **40**, 937–944. (doi:10.1134/s0026893306060021)
40. Weber C, Nover L, Fauth M. 2008 Plant stress granules and mRNA processing bodies are distinct from heat stress granules. *Plant J. Cell Mol. Biol.* **56**, 517–530. (doi:10.1111/j.1365-313X.2008.03623.x)
41. Téllez SR, Kanhonou R, Alepuz P, Ros R. 2020 RNA-binding proteins as targets to improve salt stress tolerance in crops. *Agronomy* **10**, 250. (doi:10.3390/agronomy10020250)
42. Sanchez de Groot N, Armaos A, Grana-Montes R, Alriquet M, Calloni G, Vabulas RM, Tartaglia GG. 2019 RNA structure drives interaction with proteins. *Nat. Commun.* **10**, 3246. (doi:10.1038/s41467-019-10923-5)
43. Warner JR, McIntosh KB. 2009 How common are extraribosomal functions of ribosomal proteins? *Mol. Cell.* **34**, 3–11. (doi:10.1016/j.molcel.2009.03.006)
44. Ciftci-Yilmaz S, Mittler R. 2008 The zinc finger network of plants. *Cell. Mol. Life Sci.* **65**, 1150–1160. (doi:10.1007/s00018-007-7473-4)

45. Nguyen CD, Mansfield RE, Leung W, Vaz PM, Loughlin FE, Grant RP, Mackay JP. 2011 Characterization of a family of RanBP2-type zinc fingers that can recognize single-stranded RNA. *J. Mol. Biol.* **407**, 273–283. (doi:10.1016/j.jmb.2010.12.041)
46. Frei dit Frey N, Muller P, Jammes F, Kizis D, Leung J, Perrot-Rechenmann C, Bianchi MW. 2010 The RNA binding protein Tudor-SN is essential for stress tolerance and stabilizes levels of stress-responsive mRNAs encoding secreted proteins in *Arabidopsis*. *Plant Cell.* **22**, 1575–1591. (doi:10.1105/tpc.109.070680)
47. Kwak KJ, Kim YO, Kang H. 2005 Characterization of transgenic *Arabidopsis* plants overexpressing GR-RBP4 under high salinity, dehydration, or cold stress. *J. Exp. Bot.* **56**, 3007–3016. (doi:10.1093/jxb/eri298)
48. Maurel C, Boursiac Y, Luu DT, Santoni V, Shahzad Z, Verdoucq L. 2015 Aquaporins in plants. *Physiol. Rev.* **95**, 1321–1358. (doi:10.1152/physrev.00008.2015)
49. Pena EJ, Ferriol I, Sambade A, Buschmann H, Niehl A, Elena SF, Rubio L, Heinlein M. 2014 Experimental virus evolution reveals a role of plant microtubule dynamics and TORTIFOLIA1/SPIRAL2 in RNA trafficking. *PLoS ONE* **9**, e105364. (doi:10.1371/journal.pone.0105364)
50. Foyer CH, Mullineaux PM. 1998 The presence of dehydroascorbate and dehydroascorbate reductase in plant tissues. *FEBS Lett.* **425**, 528–529. (doi:10.1016/s0014-5793(98)00281-6)
51. Groden D, Beck E. 1979 H2O2 destruction by ascorbate-dependent systems from chloroplasts. *Biochim. Biophys. Acta* **546**, 426–435. (doi:10.1016/0005-2728(79)90078-1)
52. Hakimi H, Suganuma K, Usui M, Masuda-Suganuma H, Angeles JMM, Asada M, Kawai S, Inoue N, Kawazu S. 2014 *Plasmodium knowlesi* thioredoxin peroxidase 1 binds to nucleic acids and has RNA chaperone activity. *Parasitol. Res.* **113**, 3957–3962. (doi:10.1007/s00436-014-4060-0)
53. Carmona P, Rodriguez-Casado A, Molina M. 1999 Conformational structure and binding mode of glyceraldehyde-3-phosphate dehydrogenase to tRNA studied by Raman and CD spectroscopy. *Biochim. Biophys. Acta* **1432**, 222–233. (doi:10.1016/s0167-4838(99)00113-2)
54. Queiroz RML *et al*. 2019 Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nat. Biotechnol.* **37**, 169–178. (doi:10.1038/s41587-018-0001-2)
55. Singh R, Green MR. 1993 Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase. *Science* **259**, 365–368. (doi:10.1126/science.8420004)
56. Garcin ED. 2019 GAPDH as a model non-canonical AU-rich RNA binding protein. *Semin. Cell Dev. Biol.* **86**, 162–173. (doi:10.1016/j.semcdb.2018.03.013)
57. White MR, Garcin ED. 2016 The sweet side of RNA regulation: glyceraldehyde-3-phosphate dehydrogenase as a noncanonical RNA-binding protein. *Wiley Interdiscip. Rev. RNA.* **7**, 53–70. (doi:10.1002/wrna.1315)
58. Michaud M, Ubrig E, Filleur S, Erhardt M, Ephritikhine G, Marechal-Drouard L, Duchene A-M. 2014 Differential targeting of VDAC3 mRNA isoforms influences mitochondria morphology. *Proc. Natl Acad. Sci. USA* **111**, 8991–8996. (doi:10.1073/pnas.1402588111)
59. Saint-Georges Y, Garcia M, Delaveau T, Jourdren L, Le Crom S, Lemoine S, Tanty V, Devaux F, Jacq C. 2008 Yeast mitochondrial biogenesis: a role for the PUF RNA-binding protein Puf3p in mRNA localization. *PLoS ONE* **3**, e2293. (doi:10.1371/journal.pone.0002293)
60. Vincent T *et al*. 2017 A genome-scale analysis of mRNAs targeting to plant mitochondria: upstream AUGs in 5′ untranslated regions reduce mitochondrial association. *Plant J. Cell Mol. Biol.* **92**, 1132–1142. (doi:10.1111/tpj.13749)
61. Waltz F, Corre N, Hashem Y, Giege P. 2020 Specificities of the plant mitochondrial translation apparatus. *Mitochondrion* **53**, 30–37. (doi:10.1016/j.mito.2020.04.008)
62. Corral-Debrinski M. 2007 mRNA specific subcellular localization represents a crucial step for fine-tuning of gene expression in mammalian cells. *Biochim. Biophys. Acta* **1773**, 473–475. (doi:10.1016/j.bbamcr.2006.06.008)
63. Corral-Debrinski M, Blugeon C, Jacq C. 2000 In yeast, the 3′ untranslated region or the presequence of ATM1 is required for the exclusive localization of its mRNA to the vicinity of mitochondria. *Mol. Cell. Biol.* **20**, 7881–7892. (doi:10.1128/mcb.20.21.7881-7892.2000)
64. Marc P, Margeot A, Devaux F, Blugeon C, Jacq C. 2002 Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep.* **3**, 159–164. (doi:10.1093/embo-reports/kvf025)
65. Michaud M, Marechal-Drouard L, Duchene AM. 2010 RNA trafficking in plant cells: targeting of cytosolic mRNAs to the mitochondrial surface. *Plant Mol. Biol.* **73**, 697–704. (doi:10.1007/s11103-010-9650-3)
66. Lesnik C, Golani-Armon A, Arava Y. 2015 Localized translation near the mitochondrial outer membrane: an update. *RNA Biol.* **12**, 801–809. (doi:10.1080/15476286.2015.1058686)
67. Tian L, Chou HL, Fukuda M, Kumamaru T, Okita TW. 2020 mRNA localization in plant cells. *Plant Physiol.* **182**, 97–109. (doi:10.1104/pp.19.00972)
68. Bailey TL. 2011 DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653–1659. (doi:10.1093/bioinformatics/btr261)
69. Zarrinpar A, Bhattacharyya RP, Lim WA. 2003 The structure and function of proline recognition domains. *Sci. STKE.* **2003**, RE8. (doi:10.1126/stke.2003.179.re8)
70. Calabretta S, Richard S. 2015 Emerging roles of disordered sequences in RNA-binding proteins. *Trends Biochem. Sci.* **40**, 662–672. (doi:10.1016/j.tibs.2015.08.012)
71. Han TW *et al*. 2012 Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies. *Cell* **149**, 768–779. (doi:10.1016/j.cell.2012.04.016)
72. Castello A *et al*. 2016 Comprehensive identification of RNA-binding domains in human cells. *Mol. Cell.* **63**, 696–710. (doi:10.1016/j.molcel.2016.06.029)
73. Kilchert C *et al*. 2020 System-wide analyses of the fission yeast poly(A)(+) RNA interactome reveal insights into organization and function of RNA-protein complexes. *Genome Res.* **30**, 1012–1026. (doi:10.1101/gr.257006.119)
74. Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. 2010 iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* **17**, 909–915. (doi:10.1038/nsmb.1838)
75. Licatalosi DD *et al*. 2008 HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–469. (doi:10.1038/nature07488)
76. Slobodin B, Gerst JE. 2010 A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. *RNA* **16**, 2277–2290. (doi:10.1261/rna.2091710)
77. Ramanathan M *et al*. 2018 RNA-protein interaction detection in living cells. *Nat. Methods.* **15**, 207–212. (doi:10.1038/nmeth.4601)
78. Fazal FM, Han S, Parker KR, Kaewsapsak P, Xu J, Boettiger AN, Chang HY, Ting AY. 2019 Atlas of subcellular RNA localization revealed by APEX-Seq. *Cell* **178**, 473–490. (doi:10.1016/j.cell.2019.05.027)
79. Coppin L, Leclerc J, Vincent A, Porchet N, Pigny P. 2018 Messenger RNA life-cycle in cancer cells: emerging role of conventional and non-conventional RNA-binding proteins? *Int. J. Mol. Sci.* **19**, 650. (doi:10.3390/ijms19030650)