Uncovering the role of RNA-binding proteins in plant immunity

Marcel Bach Pages

St. Cross College
Department of Plant Sciences
University of Oxford

Thesis submitted for the degree of Doctor of Philosophy

Trinity 2019

Uncovering the role of RNA-binding proteins in plant immunity

Marcel Bach Pages

St. Cross College
Department of Plant Sciences
University of Oxford

Primary supervisor: Professor Gail M. Preston
Secondary supervisors: Alfredo Castello and Renier van der Hoorn

Trinity 2019

Abstract

RNA-binding proteins (RBPs) are essential regulators of RNA fate, from synthesis to decay. Hence, changes in the RBPome orchestrate the reprograming of RNA metabolism and gene expression that occur during plant responses to a myriad of cues, including pathogen attack. Although RBPs are known to play critical roles in plants, we lack comprehensive information on the identity and function of RBPs in different species. Moreover, it remains unknown how RBPs globally respond to different environmental, physiological and pathological cues.

Here, we have developed 'plant RNA-interactome capture' (ptRIC), a variant of RNA-interactome capture (RIC) optimised for plant tissues that allows efficient system-wide identification of RBPs in leaves. We used ptRIC to identify the most extensive Arabidopsis RBPome to date, which includes more than 1100 RBPs, and showed that it displays all the hallmarks expected from a high-quality RBPome. We have used ptRIC to uncover, for the first time, the RBPome of six additional plant species that span across the plant kingdom. This revealed hundreds of unconventional RBPs with no known RNA-binding domains or previously unlinked to RNA biology, including photosynthetic

components and metabolic enzymes. By comparing the RBPomes discovered in these different species we have determined the 'core plant RBPome', which is composed by RBPs conserved from bryophytes to angiosperms. Using ptRIC we discovered that the Arabidopsis RBPome is extensively reconfigured during plant immune responses. Almost 300 RBPs displayed altered association with RNA in response to the bacterial elicitor flg22. Moreover, these alterations that cannot be explained due to changes in protein abundance, indicating that they are due to changes in RNA-binding activity. Through immunity-based mutant screens, we have discovered that several of these flg22-responsive RBPs play a role in immunity against bacterial pathogens.

Taken together, the results from this thesis provide a significant advance in the knowledge of RBPs in plants and their regulation during immunity. Hence, my results provide a framework for future studies on the role of RBPs in regulating important processes in plants.

Acknowledgements

Primer de tot vull agrair a tota la meva família, sense la qual arribar fins aquí i culminar aquest projecte hagués sigut impossible. Mil gràcies mama, papa, germà, tiets (els que hi són i els que malauradament ja no), cosins, avis (allà on sigueu) i a tots els demés, us estimo. Aquesta tesis va dedicada a vosaltres. També vull mencionar a tots els meus amics, que per sort en son molts (i que així continuï!). Lleida, amics de Sant Cugat, Biofamília, i amics d'arreu, aquesta tesis també es en part vostra.

I would really like to acknowledge Gail Preston for her unconditional supervision and help. I would also like to thank you for allowing me to develop this project in your lab and for always being supportive and give valuable help. Without you this thesis would not have been possible! Quiero agradecer también especialmente a Alfredo Castello por su excelente supervisión y por enseñarme tanto. No menos importante, quiero agradecerte por crear tan buen clima de trabajo y por estar ahí cada vez que necesitaba ayuda aun cuando no eras oficialmente mi supervisor. Sin ti esta tesis no hubiera sido posible!

I also want to thank Renier van der Hoorn for his support and feedback on the project, and for his useful advices. I would like to thank St Cross College (Oxford) and the DTP for allowing me to do my DPhil here and give me support.

I would also like to thank all the members of the Preston, Castello and van der Hoorn labs, I have learned a lot from you and we had lots of fun! Thanks to Malaka (Jiorgos), Judy, mama (Daniela) Sueldo, Pollito (Laura), Cochiwanda (Flo), Tee, Maria, Emma, Friederike, Mariana, Kyoko, Shivani, Philippe, Ricardo, Julianna, Pierre, Lauren, Rose, Felix, Izzy, Alice, Manuel, Marko, Aino,

Sam, Honglin and all the past members of the labs, you made me enjoy every minute in the lab! Thanks a lot as well to Urszula, Sarah and Caroline for making my work easier every day!

Vull també agrair a tots els meus companys de laboratori de Barcelona, en especial a la Lidia per ensenyar-me molt del que sé al laboratori. Gràcies Blanca per acollir-me y gràcies a tots els membres del lab i en especial a Pat i Mireia, ens ho vam passar molt bé. Gràcies també als meus amics del CRAG.

Muchas gracias también a mis compañeros de laboratorio de Argentina (Andrés, Franco, Vani y Leo) y a toda la demás gente de Chascomús, pasamos muy buenos momentos y fue un placer conocerlos a base de mate.

Last but not least. I would like to finish my acknowledgements by thanking all my friends in Oxford, especially my good friends from DTP and my Zenos family. You have made me enjoy every bit of Oxford! E a te Niky, grazie per farmi sorridere.

Thank you all,

Marcel Bach Pages

Table of contents

CHAPTER 1

General introduction

OVERVIEW

The aim of this thesis is to unravel the RNA-binding proteomes (RBPomes) of important model and crop species across the plant kingdom and to profile the dynamic modulation of the plant RBPome during plant defence responses.

In this introduction I will describe recent advances in the role of RNA-binding proteins (RBPs) in plant immunity. I will start with a broad overview about our current knowledge on plant immunity followed by a general introduction to RBPs and to RNA interactome capture (RIC) as a valuable technique to detect RBPs. The main body of the introduction will focus on explaining the function of specific RBPs in plant immunity. Parts of this introduction were previously published by Bach-Pages and colleagues (Bach-Pages et al., 2017).

1. The plant immune system

Plants are sessile organisms that face attacks by pathogens, pests and herbivores. Hence, it is not surprising that plants have evolved an innate immune system to defend themselves from these attacks.

Plants are armed with pattern recognition receptors (PRR) that are generally surface-localized receptor-like kinases (RLKs) and receptor-like proteins (RLPs; similar to RLKs but lack the kinase domain and require a co-receptor; **Fig. 1**) (Zipfel, 2014). Plants can perceive the presence of pathogens using PPRs that recognise conserved pathogen-derived features termed pathogen/microbe-associated molecular patterns (PAMPs/MAMPs) or plant-derived molecules that are released during pathogen invasion termed damage-associated molecular patterns (DAMPs; **Fig. 1**) (Zipfel, 2014). Perhaps the most studied PRRs are the flagellin-sensing2 (FLS2) and the elongation factor tu receptor (EFR) that recognise the bacterial PAMPs flg22 (a flagella-derived peptide) and EF-Tu (elongation factor Tu) and require the co-receptor brassinosteroid

insensitive 1 - associated receptor kinase 1 (BAK1) for immune signalling (Zipfel, 2014). PRR activation elicits a signalling cascade that involves mitogen-activated protein kinases (MAPKs) and calcium-dependent protein kinases (CDPKs) and results in a defence response called PAMP-triggered immunity (PTI; **Fig. 1**) (Zipfel, 2014).

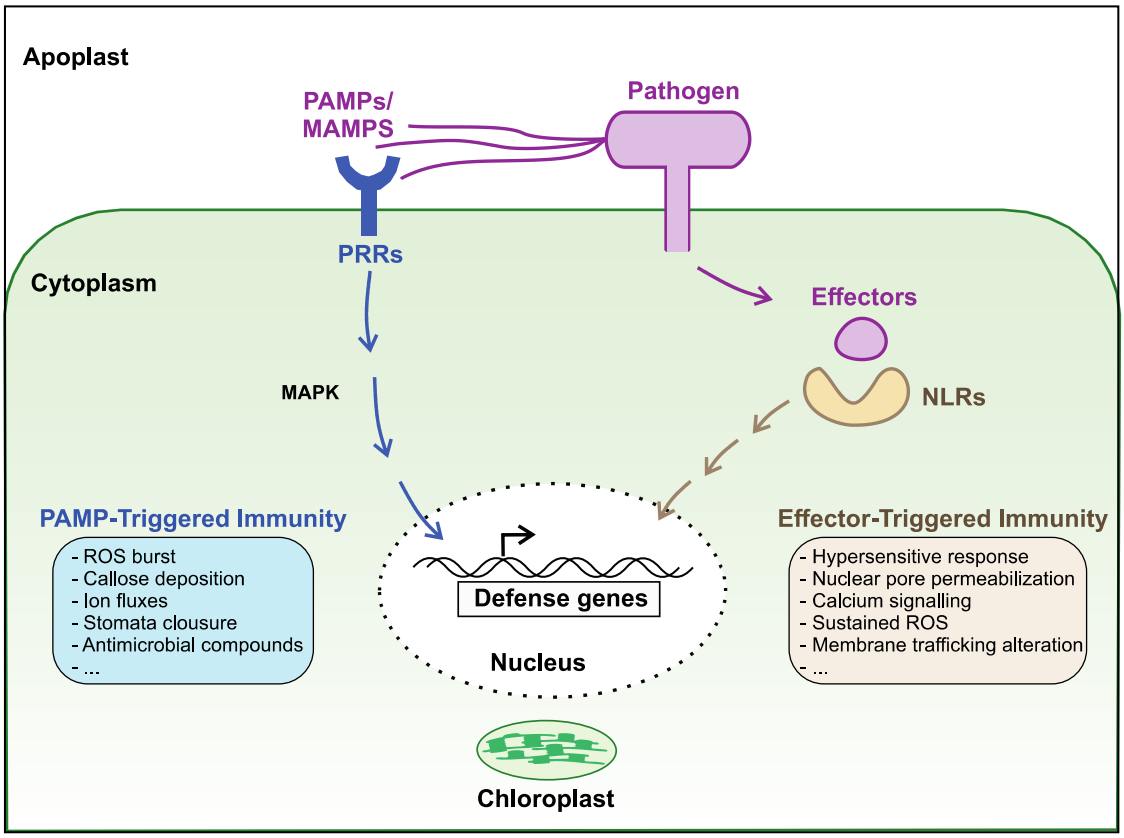


Fig. 1. Simplified schematic representation of the plant immune system. Plants can perceive pathogens of diverse nature (bacteria, fungi, oomycetes, etc) through recognition of conserved pathogen/microbe-associated molecular patterns (PAMPs/MAMPs) using pattern recognition receptors (PRR). This recognition triggers a signalling cascade that leads to an immune response termed PAMP-triggered immunity (PTI). PTI generally involves the production of reactive oxygen species (ROS), callose deposition, stomata closure, etc. Pathogens can deliver virulent effector proteins that target diverse host processes and is able to disrupt defence responses, aid at nutrient acquisition and facilitate dispersal. Plants can perceive the effectors or their activity using intracellular receptors that are often NLRs and trigger a defence response called effector-triggered immunity (ETI). ETI responses typically involve hypersensitive response, nuclear pore permeabilization and calcium signalling amongst others. Both PTI and ETI the defence responses may occur simultaneously, involve a pervasive reprogramming of the plant transcriptome and are aimed to limit pathogen growth.

Some pathogens have evolved virulence effector proteins that are able to suppress PTI, help in nutrient acquisition and facilitate pathogen dispersal (**Fig. 1**). A wide range of pathogens of

different nature are armed with suites of effectors that are delivered to the plant cell or extracellular space using diverse mechanisms (Toruño et al., 2016). For instance, bacteria can use multiple secretion systems to deliver effectors and perhaps the most studied one is the type III secretion system (T3SS; Galán et al., 2014). The T3SS is a protein complex that forms a needle-like structure through which effectors are delivered into the host cell (Galán et al., 2014). Fungi and oomycetes generally secrete effectors in specialised structures dedicated to feeding and infection termed haustoria and appressoria (Petre and Kamoun, 2014). Moreover, nematodes and aphids can deliver effectors through the stylet when feeding (Goverse and Smant, 2014).

The cellular host targets of effectors are diverse (reviewed in Toruño et al., 2016). Some effectors impede PAMPs detection by promoting degradation of PPRs (AvrPtoB), by inhibiting their kinase activity (AvrPto) and by targeting co-receptors (HopF2, AvrPtoB and AvrPto) (Macho and Zipfel, 2015). Other effectors such as the Pep1 effector from *Ustilago maydis* suppress defence responses by inhibiting maize peroxidases (Hemetsberger et al., 2012). In addition, other effectors promote entry inside the host. For example, *Pseudomonas syringae* pv. *tomato* produces coronatine that mimics jasmonic acid (JA) isoleucine and induces stomatal opening to facilitate pathogen entry (Melotto et al., 2006). Other *P. syringae* effectors also induce stomatal opening such as HopX1, HopZ1a and HopM1 (Gimenez-Ibanez et al., 2014; Jiang et al., 2013b; Lozano-Durán et al., 2014). In addition, effectors can also target negative regulators of host immunity (susceptibility factors) to promote disease, some of which are RNA-binding proteins and are discussed below.

Plants possess intracellular receptors that are able to recognise these effectors directly, or indirectly through recognition of the effector's activity, and trigger the induction of an additional layer of defence termed effector-triggered immunity (ETI; **Fig. 1**) (Chiang and Coaker, 2015). These receptors are generally nucleotide-binding site leucine-rich repeat (NBS-LRR or NLRs)

receptors, are highly variable, and normally specific effectors activate specific NLRs (**Fig. 1**; Chiang and Coaker, 2015). Moreover, there is an evolutionary arms race between pathogens and hosts whereby effectors can mutate to evade NLR recognition and plants can evolve new NLRs that recognise these effectors. In addition, the 'sensor' NLRs involved in effector perception are generally paired with 'helper' NLRs for immune signalling. Hence, NLRs are part an intricate network with 'sensor' NLRs that confer resistance to pathogens associating with specific 'helper' NLRs (Wu et al., 2017). If plants fail to recognise PAMPs or effectors, or the pathogen suppresses the immune responses at any level, this normally results in pathogen growth and disease.

Generally, PTI responses include the production of reactive oxygen species (ROS) by both NADPH oxidases called respiratory burst oxidase homologs (RBOHs) and the apoplastic peroxidases (**Fig. 1**). The ROS burst is thought to limit pathogen growth due to its antimicrobial actions, its signalling activity and by strengthening of the cell wall (Qi et al., 2017). Other defence outputs are the callose deposition that acts to limit pathogen penetration, ion fluxes, expression of defence-related genes, stomatal closure, decrease of photosynthesis, and production of antimicrobial compounds (**Fig. 1**; Melotto et al., 2006; Zeng and He, 2010; Bolton, 2009; Brown et al., 1998; Meng and Zhang, 2013; Boller and Felix, 2009). However, ETI typically induces the hypersensitive response (HR), a localized cell death that may limit pathogen growth (Chiang and Coaker, 2015). Other ETI responses include nuclear pore permeabilization, calcium signalling, modulations in membrane trafficking and increased and sustained ROS (**Fig. 1**; Gu et al., 2016; Cui et al., 2015). Furthermore, in order to protect themselves from apoplast colonizing pathogens, plants secrete different proteins and hydrolytic enzymes into the apoplast, which effectively limit disease (Buscaill et al., 2019). Although historically PTI and ETI have been described as separate events with different defence outputs, increasingly the division between

PTI and ETI has become blurred and both immune responses often occur simultaneously upon pathogen attack (Thomma et al., 2011; van der Burgh and Joosten, 2019).

Most of these responses - both PTI and ETI - involve extensive re-programming of the plant transcriptome (**Fig. 1**). Thus, upon pathogen recognition defence-related genes are transcribed in the nucleus and are eventually translated in the cytoplasm to mount a defence response against pathogens (Buscaill and Rivas, 2014). Moreover, the chloroplasts have an important role in immunity since they are involved in the synthesis of defence-related compounds such as the hormones jasmonic acid (JA) and salicylic acid (SA) (Serrano et al., 2016). This complex re-programming of the plant transcriptome can occur at many levels, from transcription to translation, and in different plant compartments. In recent years, it has become evident that post-transcriptional gene regulation plays crucial roles in plant adaptation to environmental cues, including biotic stress. RNA-binding proteins (RBPs) are critical players in posttranscriptional gene regulation (Glisovic et al., 2008). Hence, RBPs are orchestrators of the re-programming of the immunity transcriptome through regulation of RNA metabolism or localization.

2. RNA-binding proteins

2.1. Structure and role of RNA-binding proteins

The dynamic interactions between RNAs and RBPs to form ribonucleoprotein (RNP) complexes determine the fate of cellular RNAs (Glisovic et al., 2008). RBPs are of paramount importance for cell functioning since they regulate RNA synthesis, processing (capping, splicing and polyadenylation), post-transcriptional modifications, transport, storage, translation and degradation (Glisovic et al., 2008). Moreover, the composition of RNP complexes is highly dynamic, coordinating transcriptional changes according to different developmental processes

or in response to external cues. Thus, the RNA-binding proteome (RBPome) of multiple species have been reported to be modulated in response numerous cues. For example, the human RBPome is profoundly modulated upon viral infection (Garcia-Moreno et al., 2019) and in response to treatment with arsenite (Trendel et al., 2019). The Drosophila RBPome is remodelled to control the maternal-to-zygotic transition (Sysoev et al., 2016). In addition, RBPome of Arabidopsis cell cultures has been recently described to respond to drought stress (Marondedze et al., 2019). Taken together these studies evidence the dynamic nature of the RBPome in response to different physiological or pathological conditions.

RBPs have a modular structure that includes one or more RNA-binding domains (RBDs) that confer the ability to bind RNA, and auxiliary domains that modulate these properties or exert RNA-related molecular activities (Lunde et al., 2007). By combining individual RBDs that bind to RNAs, RBPs possess high specificity, versatility and affinity for target RNAs (Lunde et al., 2007). Another layer of regulation is conferred by the auxiliary domains, which can be post-translationally modified and enable RBDs to interact with other proteins or biomolecules (Lunde et al., 2007; Nagai et al., 1995). Furthermore, the RNP configuration is thought to be dependent on each individual target mRNA and to be dynamically altered throughout RNA processing and maturation. Therefore, the RBPome of a given cell is composed of unique combinations of different RBD and auxiliary domains within individual RBPs (Lunde et al., 2007). These modular structures confer different binding activities and protein-protein interactions to each specific RBP, and thus, determine RNA metabolism.

Despite the high diversity of RNA targets, it was generally believed that the specificity of RBPs relied on the combination of a limited subset of RNA-binding modules in the appropriate order (Lunde et al., 2007). RBDs have been classified into classical and non-classical according to their evidence to bind RNA (Castello et al., 2012). Classical RBDs include well characterised domains

that have been validated in multiple organisms. These include the RNA recognition motif (RRM), K-homology (KH) domain, zinc finger (ZnF), PIWI, etc (Lunde et al., 2007). Other non-classical domains have been linked to RNA-binding activity in at least one publication including ribosomal, pentatricopeptide repeat (PPR) and helicase domains. However, it has been recently reported that hundreds of novel experimentally identified RBPs do not harbour known RBDs, indicating that more domains or modes of binding to RNA await to be discovered (Castello et al., 2016; Hentze et al., 2018).

2.2. The plant RNA-binding proteome

The RBPome of plants is composed of a large repertoire of RBPs. It has been estimated that the Arabidopsis genome encodes approximately 1600-1800 RBPs based on the presence of bioinformatically predicted RBDs (Van Ruyskensvelde et al., 2018; Cho et al., 2019). Some of these are plant-specific, indicating the functional diversity of RBPs in several plant-specific processes, including development and responses to both abiotic and biotic stresses (Van Ruyskensvelde et al., 2018; Cho et al., 2019; Staiger et al., 2013). Moreover, RBPs are active not only in the cytoplasm or nucleus, but they can also be located in mitochondria and chloroplasts, where they exert crucial roles in organelle-RNA metabolism (de Longevialle et al., 2007; Hammani and Giegé, 2014; Stern et al., 2010; del Campo, 2009).

2.3. RNA-Interactome capture (RIC)

Recently, a technique termed RNA interactome capture (RIC) has been developed that allows global determination of the RBPs actively bound to polyadenylated RNAs *in vivo* (Castello et al., 2012; Baltz et al., 2012; Castello et al., 2013b). In brief, cells are UV-irradiated to covalently crosslink the proteins that are in intimate contact with the RNAs. Then, the cells are lysed and the RNAs together with the associated proteins are isolated using oligo(dT) magnetic beads under stringent conditions. The stringency of the washing steps is key to dissociate any non-

specific interactors and allow isolation of just proteins that directly bind RNA. The RNAs are degraded by RNase treatment, and the protein identified and quantified by mass spectrometry (Castello et al., 2012, 2013b). RIC has represented a breakthrough in the study of RBPs since it has allowed discovery of hundreds of RBPs that were previously unknown in multiple species (Hentze et al., 2018). Of note, RIC allows isolation a subset of the total RBPome, that which binds to mRNA (the mRNA-binding proteome – mRBPome – hereinafter referred as RBPome).

The development of RIC has represented a great technical advance, since RIC has many improvements over previous methods. Cells are irradiated with UV light *in vivo*, thus allowing identification of native and physiological RNA-protein interactions. UV light promotes formation of covalent bonds between RNAs and proteins that are in very close proximity (≤2Å), but it does not promote protein-protein crosslinking (Pashev et al., 1991). This allows the use of stringent buffers to remove non-specific binders, and the capture of only proteins that are covalently bound to RNA. Since UV crosslinking relies on quantitative parameters of protein-RNA interactions (i.e. time and strength of interaction) and proteins are identified by quantitative proteomics, RIC allows discrimination between active (RNA bound) and inactive (unbound) RBPs in different physiological contexts (Castello et al., 2012, 2013a). This offers the unprecedented opportunity to study the dynamic behaviour of RBPs under different cellular or environmental contexts at the system-wide scale. RIC also allows identification of non-canonical RBPs that do not harbour classical RBDs, and thus allows potential identification of novel RBDs or novel modes of RNA binding.

However, RIC has some limitations. Because RIC relies of the capture of poly(A) RNAs using oligo(dT) beads, RIC fails to isolate RBPs bound to non-polyadenylated RNAs. Moreover, RIC will not likely identify proteins which do not efficiently UV-crosslink and RBPs interacting with the phosphate-ribose backbone, since UV light only acts on the nucleotide bases and not the

backbone. Likewise, RIC will probably fail to identify RBPs that do not efficiently crosslink due to RNA-protein geometry such as double stranded RBPs (Hentze et al., 2018). Finally, RIC will not identify RBPs that are not expressed in the cell line or tissue analysed.

RIC has greatly contributed to unveil the RNA-binding proteomes (RBPomes; RNA interactomes) in multiple different organisms (Hentze et al., 2018). Recently, four studies applied RIC to different tissues of the model plant *Arabidopsis thaliana*: i) etiolated seedlings (Reichel et al., 2016); ii) cell cultures (Marondedze et al., 2016, 2019); iii) leaves (Marondedze et al., 2016) and iv) protoplasts derived from leaf mesophyll (Zhang et al., 2016b). This has allowed experimental identification, for the first time, of the Arabidopsis RBPome, which included approximately 700 RBPs. However, the application of RIC to Arabidopsis leaves was not successful, since only 27 RBPs were identified from leaf tissue (Marondedze et al., 2016), numbers remarkably low as compared to the RBPomes identified in other organisms (Hentze et al., 2018). Therefore, the leaf RBPome remains largely unexplored. In addition, the majority of these studies have used plant materials that are easy-to-crosslink, but that may lack relevance for certain areas of plant biology since they are rather artificial systems. Moreover, the number of RBPs identified is low as compared to the almost 2000 RBPs identified in humans (Hentze et al., 2018). To our knowledge RIC has not been applied to any other plant species. Hence, our knowledge of the plant RBPome is limited.

3. RBP-mediated immunity in plants: RBPs acting in plant immunity

In recent years, there have been numerous studies that have provided evidence of RBPs playing important roles in plant immunity. In this introduction we review the evidence for what we term 'RBP-mediated immunity in plants', whereby immunity against pathogens requires the re-programming of the transcriptome mediated by RBPs that regulate RNA metabolism and/or localization. Moreover, RBPs can also be targeted by pathogens, and therefore, represent one

of the plant's Achilles' heels. We therefore also review evidence for 'RBP-mediated susceptibility' in plants.

RNAs are mainly transcribed in the nucleus and before being translated into proteins they undergo different processes including capping, splicing and alternative splicing, polyadenylation and alternative polyadenylation, editing, export, stabilization and translation, amongst others. During all these processes RNAs are bound to RBPs, which play crucial roles in regulating the fate of RNAs. For each of the steps of the RNA lifecycle we summarise the main evidence for RBP-mediated immunity, focusing mainly on RBPs that are known or hypothesised to bind RNA directly (**Fig. 2, Supplemental table 1**). We also discuss important RBPs such as helicases or EV-localised RBPs. Some RBPs for which an immune mechanism is known are described in the main text. Additional RBPs with links to immunity are summarised in **Supplemental table 1**. In this section we describe the phenotypes of multiple RBP mutant lines, which are loss of function mutants unless stated otherwise.

A number of RBPs are involved in transcription of RNA from DNA (Xiao et al., 2019). However, this section will focus on RBPs that regulate the RNAs downstream of transcription since there are many reviews that describe transcriptional regulation of plant immunity (please see Buscaill and Rivas, 2014; Tsuda and Somssich, 2015; Li et al., 2016). Several RBPs have also been described to play roles in viral infection and immunity (Huh and Paek, 2013; Musidlak et al., 2017). This is not surprising since viruses require the cellular machinery in order to replicate and survive. However, here we will focus on the RBPs implicated in immunity against bacteria, fungi and oomycetes. See Huh and colleagues (Huh and Paek, 2013) and Musidlak and colleagues (Musidlak et al., 2017) for recent reviews about RBPs and viral infection.

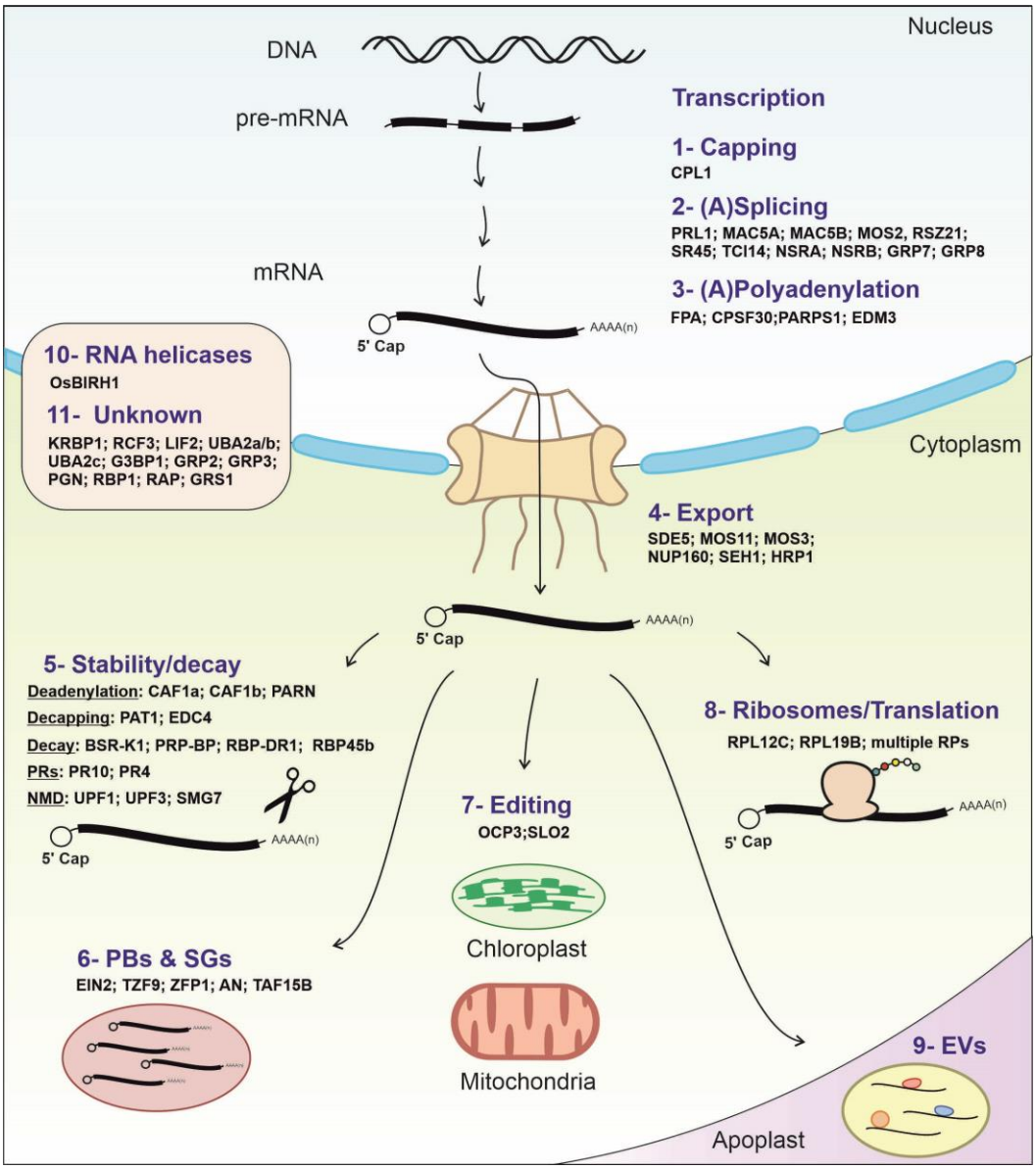


Fig. 2. RNA-binding proteins involved in plant immunity. RBPs involved in immunity play important roles in all the steps of RNA metabolism and localization including transcription, [1] capping, [2] (alternative) splicing, [3] (alternative) polyadenylation, [4] mRNA export, [5] RNA stability/decay, [6] localization to P-bodies (PBs) or stress granules (SGs), [7] RNA editing mainly in organelles, [8] ribosomes and translation or [9] extracellular vesicles (EVs). [10] RNA helicases are involved in multiple processes in different cellular compartments. [11] A number of RBPs have known roles in immunity, but their precise role in RNA metabolism and/or localization awaits to be discovered.

3.1. RNA capping

The first step after transcription is the co-transcriptional addition of the 7-methylguanosine (m^7G) cap at the 5' end of the pre-mRNA. This increases the stability and translational efficiency of the transcript and mediates some functions such as RNA processing or nuclear export (Ramanathan et al., 2016). RNA capping occurs mainly in the nucleus, although recapping of

decapped mRNAs has also been reported to take place in the cytoplasm (Trotman and Schoenberg, 2019). However, to our knowledge this has not been reported in plants. It has been described that addition of a methyl group on the ribose 2'-*O* position of the cap (2'-O-methylation) can also be identified as a sign to distinguish self RNA from non-self RNA (e.g. viral RNAs) in immunity (Züst et al., 2011).

RNA POLYMERASE II C-TERMINAL DOMAIN (CTD) PHOSPHATASE-LIKE1 (CPL1)

One of the RBPs that is involved in capping and plant immunity is the RNA polymerase II carboxyl terminal domain (CTD) phosphatase-like1 (CPL1), a protein that contains two double-stranded RNA (dsRNA) binding domains and a phosphatase domain. However, although CPL1 possesses two putative dsRNA domains, to our knowledge it has not been experimentally demonstrated to directly bind RNA.

CPL1 has been described to have important roles in transcription and RNA regulation (for example RNA capping and decay), by dephosphorylation of the Carboxyl Terminal Domain (CTD) of RNA polymerase II (Jiang et al., 2013a; Chen et al., 2013; Cui et al., 2016; Jeong et al., 2013; Bang et al., 2008; Manavella et al., 2012). In eukaryotes, diverse reversible modifications of the CTD of RNA Pol II recruit different factors that regulate both transcription and some processes of RNA metabolism such as capping, splicing and polyadenylation (Kerk et al., 2008; Hsin and Manley, 2012). Kinases and phosphatases like CPL1 regulate the phosphorylation state of serine residues within the CTD, although how these interactions translate in recruitment of factors and regulation of gene expression it is not fully understood.

CPL1 has been found to be a negative regulator or susceptibility factor in plant immunity since transcriptomic analyses of *cpl1* mutants revealed upregulation of genes involved in biotic stresses (Thatcher et al., 2018). Accordingly, *cpl1* mutants are more resistant to the fungi

Fusarium oxysporum and *Alternaria brassicicola* and the aphid *Myzus persicae* (Thatcher et al., 2018). However, *cpl1* mutants have no altered resistance to the bacterium *Pseudomonas syringae*. In addition, CPL1 regulates other important processes such as flowering, abiotic stresses (oxidative, cold, drought, salt, ABA, nutrient deficiency/toxicity), wounding responses amongst others (Jiang et al., 2013a; Matsuda et al., 2009; Koiwa et al., 2002; Aksoy et al., 2013). Therefore, it seems that by dephosphorylating the CTD of RNA Pol II, CPL1 plays essential roles in transcriptional regulation and RNA processing that affect many different biological processes.

3.2. Splicing and alternative splicing

Most eukaryotic pre-mRNAs contain introns, which are removed by the spliceosome in a process termed splicing (Labadorf et al., 2010; Sharp, 2005). The spliceosome is a large molecular complex composed of five small nuclear ribonucleoproteins (snRNPs) and multiple spliceosome-associated proteins (Plaschka et al., 2018). In a normal splicing event, the splice sites located at the 5' and 3' ends of the introns are successfully recognised by snRNPs, and the intron is spliced out. However, the splicing sites' recognition can be perturbed or regulated and give rise to multiple splice sites that yield multiple mRNA isoforms from a single pre-mRNA. This is called alternative splicing (AS) (Chaudhary et al., 2019). Hence, AS increases the complexity of the proteome and is believed to occur in about 60% of the genes that contain introns in Arabidopsis (Marquez et al., 2012; Laloum et al., 2018). AS can give rise to alternative spliced isoforms in response to environmental cues (Gulledge et al., 2012; Tanabe et al., 2007; Palusa et al., 2007), thus, AS has important roles in post-transcriptional regulation of gene expression in response to stresses. There are different forms of AS that include intron retention, exon skipping or alternative splice sites (5' and 3'). The most common form of AS is intron retention (Marquez et al., 2012; Vélez-Bermúdez and Schmidt, 2014), which can result in the degradation of the transcript by the nonsense-mediated decay (NMD) pathway due to premature stop codons, or nuclear retention and storage (Nicholson and Mühlemann, 2010; McGlincy and Smith, 2008).

Exon skipping can lead to differences in protein function or subcellular localization by re-arranging domains within proteins (Syed et al., 2012). A compendium of proteins governs AS including serine/arginine-rich (SR) proteins (Reddy 2004, Shen 2004) and heterogeneous RNPs (hnRNPs).

Many of the spliceosome and splicing components have been related to immune-related functions, highlighting the importance of splicing and alternative splicing in plant defence responses. For a recent review of (alternative) splicing in plant immunity see Rigo et al., 2019.

MOS4-ASSOCIATED COMPLEX (MAC) COMPLEX

In Arabidopsis, the MOS4-associated complex (MAC) has been linked to immunity and development through functions in miRNA biogenesis and splicing (see Johnson et al., 2011 for review). MAC is a conserved complex in eukaryotes, and has been associated with the spliceosome in humans, yeast and plants (Monaghan et al., 2009; Koncz et al., 2012; Deng et al., 2016; Johnson et al., 2011). In Arabidopsis, MAC is composed of multiple proteins and some of these are RBPs with roles associated to immunity (Johnson et al., 2011). For instance, pleiotropic regulatory locus 1 (PRL1) is a nuclear RBP containing seven WD-40 repeats that has additional roles in sRNAs accumulation (Palma et al., 2007; Zhang et al., 2014b). However, the mechanism by which PRL1 contributes to immunity is still unclear. Other components of MAC are MAC5A and MAC5B, two RBPs that possess a CCCH-type zinc finger domain and a RRM and are highly homologous and partly redundant (Monaghan et al., 2010). MAC5A is involved in plant immunity and *mac5a* mutant lines partially supress the autoimmune phenotypes of *suppressor of npr1-1, constitutive1* (*snc1*) gain of function mutant (Monaghan et al., 2010).

MODIFIER OF SNC1 2 (MOS2)

Another protein connected to the spliceosome, although it is not part of the MAC complex, is modifier of snc1 2 (MOS2). MOS2 encodes a nuclear and evolutionarily conserved protein with a G-patch domain and two KOW motifs that are predicted to bind RNA (Zhang et al., 2005). MOS2 is important in both basal defence as well as *R*-gene-mediated resistance, and *mos2* mutants are more susceptible to virulent and avirulent *P. syringae* pv. *maculicola* ES4326 and to *P. syringae* pv. *tomato* DC3000 *AvrRPS4*. Although MOS2 has been linked to the spliceosome (Koncz et al., 2012) and interacts with the miRNA biogenesis machinery (Wu et al., 2013), the exact role of *MOS2* in plant immunity awaits to be discovered (Zhang et al., 2005; Copeland et al., 2013)

GLYCINE RICH PROTEIN 7 (GRP7)

Interestingly, some studies have identified that some pathogens can target components of the splicing machinery to disrupt host immunity. One protein that has been particularly well studied is the Arabidopsis glycine rich protein 7 (AtGRP7), which is an RBP with a N-terminal RRM domain and a C-terminal glycine-rich domain. It has been described that perturbations of GRP7 levels lead to changes in alternative splicing and polyadenylation, suggesting that GRP7 might play a role in alternative splicing (specifically at the alternative 5' splice site choice) and polyadenylation (Streitner et al., 2012; Meyer et al., 2017). GRP7 is also involved in pri-miRNA processing (Köster et al., 2014).

Importantly, GRP7 plays a critical role in immunity and is targeted by a type III secretion system effector of *P. syringae* (Fu et al., 2007; Jeong et al., 2011a). This effector, named HopU1, is a mono-ADP-ribosyl transferase that ADP-ribosylates the arginine 49 residue within the RRM of GRP7. Upon ADP-ribosylation, GRP7 is unable to bind to its usual target transcripts of *FLS2* and *EFR* through its RRM domain, and therefore, the accumulation of these immune receptors is

reduced (Nicaise et al., 2013). Accordingly, *grp7* mutant lines are more susceptible to *P. syringae* pv. *tomato* DC3000 (Fu et al., 2007; Jeong et al., 2011a; Nicaise et al., 2013). Moreover, the effector HopU1 is required for pathogen virulence (Fu et al., 2007). Interestingly, the Arabidopsis glycine rich protein 8 (AtGRP8), which is also involved in alternative splicing and can be targeted by the effector HopU1, also binds the *FLS2* and *EFR* transcripts *in vivo* (Fu et al., 2007; Streitner et al., 2012). The targeting of a host RBP by ADP-ribosylation represents a novel virulence mechanism whereby pathogens are able to suppress immunity and cause disease. Moreover, GRP7 seems to be involved in defence against pathogens in a species-specific manner: it is a positive regulator of immunity against *Pectobacterium carotovorum* and tobacco mosaic virus (TMV), whereas it is a negative regulator of immunity against *Botrytis cinerea* (Lee et al., 2012b).

Other proteins involved in splicing are targeted by pathogen effectors

Other proteins belonging to the splicing machinery are targeted by pathogens to subvert plant defence responses. For example, a *Phytophthora sojae* effector targets GmSKRP (Huang et al., 2017) and a *Heterodera schachtii* effector targets SMU2 (Verma et al., 2018), although these two proteins do not directly bind RNA.

Therefore, splicing and alternative splicing are key in reprograming the transcriptome to modulate the defence responses in plants and some pathogens, in turn, have evolved strategies to target the host splicing machinery to disrupt the immune responses.

3.3. Polyadenylation and alternative polyadenylation

Before leaving the nucleus, virtually all eukaryotic mRNAs are polyadenylated by addition of a poly(A) tail at the 3' end that protects them from degradation and helps to recruit the translational machinery (Eckmann et al., 2011). Moreover, about 70% of the Arabidopsis genes can undergo alternative polyadenylation (APA), altering the stability, translation or functionality

of those genes (Wu et al., 2011). APA has been shown to be an important mechanism to regulate gene expression in different organisms. Recent studies have evidenced the link between APA and immunity in multiple organisms beyond plants (Kondrashov et al., 2009).

FLOWERING TIME CONTROL PROTEIN (FPA)

One of the best studied RBPs involved in APA is the Arabidopsis flowering time control protein (FPA), for which a mechanistic link to immunity is known. FPA is an RBP that regulates the 3' mRNA polyadenylation of the immunity-related transcriptional repressor Ethylene Response Factor 4 (ERF4) (Lyons et al., 2013). PTI elicitation induces the expression of regular as well as APA isoforms of ERF4. The regular isoform positively regulates the ROS burst, whereas the alternative isoform lacks the ERF-associated amphiphilic repression (EAR) motif and functions as a transcriptional activator that suppresses the ROS burst. FPA inhibits the induction of APA isoforms of ERF4 by regulating poly(A) site choice, therefore, FPA represses the formation of EAR-lacking ERF4 (Lyons et al., 2013). However, both EAR-lacking ERF4 and FPA suppress ROS. Hence, it has been speculated that FPA also contributes to regulating the ROS burst by regulating APA of other defence-related genes that contribute to ROS regulation (Lyons et al., 2013). It has been shown that FPA negatively regulates basal resistance, since FPA overexpressing plants are more susceptible to *P. syringae* (Lyons et al., 2013).

Therefore, FPA represses the production of an APA form of ERF4 induced by PTI responses, quantitatively regulating the defence responses and negatively regulating resistance to *P. syringae*.

3.4. Nuclear mRNA export

When mRNAs are fully processed, a set of RBPs bind to them and the large mRNP complexes are exported to the cytoplasm via the nuclear pore complexes (NPCs). Since the export rates of

mRNA influence their translation, mRNA export is involved in the control of many biological processes including responses to stress (Van Ruyskensvelde et al., 2018).

SILENCING DEFECTIVE 5 (SDE5)

Silencing defective 5 (SDE5) encodes an RBP that has been described to have a role in transgene silencing and production of ta-siRNAs (Hernandez-Pinzon et al., 2007). SDE5 is the homolog of human TAP (NXF1), which is one of the principal mRNA nuclear export factors (Katahira, 2015). This, together with the fact that *sde5* mutants accumulate more polyadenylated RNAs in the nuclei, led to the recent assumption that SDE5 is also likely to be involved in mRNA export (Uddin et al., 2017). Moreover, SDE5 is involved in plant immunity and contributes to ETI while supressing PTI. SDE5 acts as a positive regulator of the SA-mediated defence response, and a negative regulator of the JA-mediated immunity, and possibly participates in cross-talk between these pathways (Uddin et al., 2017). Accordingly, SDE5 contributes to increased resistance to the biotrophic pathogen *P. syringae* pv. *tomato* DC3000 and to increased susceptibility to the necrotrophic bacterium *Erwinia caratovora* pv. *caratovora* (Uddin et al., 2017)**.** However, the precise mechanism by which SDE5 contributes to immunity by modulating mRNA export is not fully understood.

3.5. RNA stability/decay

The pool of a given RNA that is available for translation depends not only on its synthesis, processing and export rates, but also on its stability or degradation. Therefore, the rate of degradation, which depends on transcript stability or inhibition/stimulation of degradation pathways, largely determines the abundance of RNA that is available to be translated. Many different factors affect the stability of a given mRNA, including presence/absence of introns, sequence elements in the UTRs, UTR and ORF length and secondary structures (Narsai et al., 2007). Therefore, the control of the mRNA pool through mRNA stability or degradation is

regulated to quickly adjust to different cellular states and in responses to different environmental conditions, including biotic stresses (Jiao et al., 2008; Yu et al., 2019).

a) Deadenylation

mRNA needs to be deadenylated before it can either be degraded in a 3' – 5' direction or be decapped and degraded in the 5' – 3' direction (Garneau et al., 2007). Deadenylation involves shortening or removing the poly(A) tail at the 3' of the mRNA. There are multiple deadenylase complexes in plants such as the poly(A) ribonuclease PARN, the poly(A) nuclease PAN and CCR4/CAF1 deadenylase complex. The PARN and the CCR4/CAF1 deadenylase complex have both been linked to plant immunity (Johnson et al., 2018; Walley et al., 2007; Liang et al., 2009; Walley et al., 2010).

CCR4-ASSOCIATED FACTOR 1 (CAF1s)

The Arabidopsis CCR4-associated factor 1 a and b (AtCAF1a and AtCAF1b) have both been shown to be involved in mRNA deadenylation and immunity *in vivo* (Liang et al., 2009). Their expression is induced upon multiple treatments such as stress-related hormones (SA and ABA) and biotic stresses such as bacteria, indicating that they may respond to a broad range of stimuli (Walley et al., 2007; Liang et al., 2009; Walley et al., 2010). It has been speculated that upon elicitation of immunity, *CAF1a* and *CAF1b* are transcriptionally upregulated and participate in the deadenylation and thus degradation of a repressor of *PR* gene transcription (Liang et al., 2009). Consequently, in Arabidopsis CAF1a and CAF1b overexpressing lines, the repressor is highly deadenylated and degraded, allowing high expression of *PR1* and *PR2* genes and increased resistance to *P. syringae* pv. *tomato* DC3000, and the opposite is true for *caf1a* and *caf1b* mutants (Liang et al., 2009).

Moreover, the role of CAF1 in immunity is not exclusive to Arabidopsis, since the chilli pepper CAF1a (CaCAF1a) is also involved in immunity (Sarowar et al., 2007). Overexpression of CaCAF1a in tomato leads to increased resistance against *Phytophthora infestans*, to the upregulation of many defence-related genes (such as polyamine biosynthesis or *PR* genes) and to thickening of the cell wall and cuticle (Sarowar et al., 2007). In agreement, virus-induced gene silencing (VIGS) of CaCAF1a in pepper results in increased susceptibility to the bacterial pathogen *Xanthomonas axonopodis* pv. vesicatoria (Sarowar et al., 2007).

b) Decapping

After deadenylation, RNA needs to be decapped prior to 5' to 3' degradation. Decapping involves removing the m^7G cap at the 5' end of the mRNA and is mainly carried out by the decapping enzyme decapping 2 (DCP2), the decapping activators DCP1/5, varicose (VCS) and PAT1, and exoribonuclease 4 (XRN4) that degrades the RNA. Many of the members of the decapping complex are involved in plant immunity.

DECAPPING 1 and 2 (DCP1 and DCP2)

The decapping 2 (DCP2) enzyme functions together with the co-activator DCP1 (and other factors) in a complex involved in decapping and both have been recently described to be involved in immunity-related mRNA decay in the processing bodies (P-bodies; Yu et al., 2019). Upon flg22 perception, DCP1 is phosphorylated by two MAPKs (MPK3/6), which promotes dissociation from DCP2 and association with XRN4 (Yu et al., 2019). DCP1-XRN4 association stimulates XRN4 exonuclease activity or allows access of XRN4 to mRNAs, thus resulting in degradation of decapped RNAs (Yu et al., 2019). The DCP1-DCP2 complex contributes to immunity by positively regulating PTI since *DCP1* and *DCP2* silencing results in decreased PTI-induced defence gene expression and reduced callose deposition (Yu et al., 2019). Moreover,

DCP1 and *DCP2* silenced plants have increased susceptibility to *P. syringae* pv *tomato* and *P. syringae* pv. *maculicola* ES4326 (Yu et al., 2019).

Therefore, PTI activation triggers a DCP1-dependent mRNA decay of a set of immunity-downregulated genes. This mechanism could be involved in the downregulation of a set of genes as part of the transcriptomic reprogramming that occurs during immune responses (Yu et al., 2019).

PROTEIN-ASSOCIATED WITH TOPOISOMERASE 1 (PAT1)

The protein-associated with topoisomerase 1 (PAT1) is a decapping enhancer that together with LSM1-7 links deadenylation and decapping by binding the 3' of deadenylated mRNAs and promoting decapping of specific transcripts (Tharun, 2009). Moreover, PAT1 also plays a role in translational inhibition and P-body formation (Roux et al., 2015). Consistently, *pat1* mutants accumulate more capped mRNAs, confirming its role in decapping (Roux et al., 2015).

The involvement of PAT1 in immunity has recently been described. Upon flg22 treatment PAT1 interacts and is phosphorylated by the MPK4 (and to a lesser extent MPK6), after which it re-localises to the P-bodies (Roux et al., 2015). Moreover, PAT1 is not regulated by flg22 treatment at the transcriptional level since PAT1 mRNA levels remain unaltered upon flg22 treatment, indicating post-transcriptional regulation of PAT1 (Roux et al., 2015). *pat1* mutants exhibit autoimmunity, have increased resistance to *P. syringae* pv. *tomato* DC3000 and a high constitutive expression of *PR1* and *PR2* genes (Roux et al., 2015). PAT1 provides a clear example of a mechanism whereby MPKs regulate mRNA decay machinery during immune responses, although the downstream RNAs regulated by PAT1 remain unknown.

c) RNA stability/turnover

BROAD-SPECTRUM RESISTANCE KITAAKE-1 (BSR-K1)

Another example of an RBP involved in RNA stability/turnover is the broad-spectrum resistance kitaake-1 (BSR-K1). BSR-K1 is a rice cytoplasmic tetratricopeptide repeat (TPR)-containing protein that has been recently described to be involved in immunity (Zhou et al., 2018). *bsk-k1* mutants possess enhanced resistance to different races of the fungus *Magnaporthe oryzae* and the bacteria *Xanthomonas oryzae* pv *oryzae*. BSR-K1 negatively regulates plant immunity by binding to multiple defence-related *OsPAL* transcripts (*OsPAL1-7*) and promoting their turnover. Therefore, in the absence of BSR-K1, the *OsPAL* transcripts are accumulated and rice is resistant to bacterial and fungal pathogens (Zhou et al., 2018).

PVPRP7 MRNA BINDING PROTEIN (PRP-BP)

PVPRP7 mRNA binding protein (PRP-BP) is a bean cytoplasmic RBP that binds specifically to the 3' UTR of *PvPRP1* (Zhang and Mehdy, 1994). The RNA-binding activity of PRP-BP is increased upon treatment of bean cells (*Phaseolus vulgaris*) with elicitors from the fungus *Colletotrichum lindemuthianum* (Zhang and Mehdy, 1994). Treatment with elicitors also causes concomitant downregulation of the target gene *PvPRP1*, which encodes for a cell wall protein rich in proline but poor in tyrosine (Sheng et al., 1991; Zhang et al., 1993; Zhang and Mehdy, 1994). It has been proposed that PRP-BP mediates downregulation of *PvPRP1* upon immune activation because the PvPRP1 protein has low concentration of tyrosine, and therefore does not contribute to cell wall strengthening through isodityrosine crosslinking (Zhang et al., 1993; Zhang and Mehdy, 1994). Interestingly, the RNA-binding activity of PRP-BP was found to be regulated by the redox state of the sulfhydryl groups, and redox changes typically occur after pathogen infection (Zhang and Mehdy, 1994). Hence, upon elicitor treatment PRP-BP is post-translationally activated and binds to the target *PvPRP1* mRNA to promote its degradation.

PATHOGENESIS-RELATED PROTEINS (PRs)

An important group of RBPs that function in immunity are the pathogenesis-related 10 (PR10) family. A number of PR-10 proteins from different species have been described to be ribonucleases and to have important roles against multiple species of bacteria and fungi. For example, CaPR10 is a RNase from hot pepper (*Capsicum annuum*) that has been shown to play a role in defence against multiple pathogens including *Xanthomonas campestris* pv. *vesicatoria*, *Phytophthora capsici* and TMV (Park et al., 2004). CaPR10 was shown to interact with LRR1 (LEUCINE-RICH REPEAT 1) protein, leading to HR-like cell death in *N. benthamiana* and pepper and activation of defence signalling, although the mechanism is still not completely understood (Choi et al., 2012). Silencing of PR10 together with LRR1 in pepper results in increased susceptibility to *X. campestris*, whereas overexpression in Arabidopsis leads to increased resistance to *P. syringae* pv. *tomato DC3000* and *Hyaloperonospora arabidopsidis* (Choi et al., 2012).

GaPR-10 is a cotton (*Gossypium arboretum*) cytoplasmic PR-like protein with RNase activity (Zhou et al., 2002). Its expression is gradually induced after treatment of cotton with elicitors, *Verticillium dahliae* or jasmonic acid (Zhou et al., 2002). It has been proposed that GaPR-10 might control cellular RNA homeostasis by selectively degrading some target RNAs that are not required after pathogen attack, allowing plants to return to normal physiological conditions (Zhou et al., 2002). TcPR10 from cacao (*Theobroma cacao*), SsPR-10 from *Solanum surattense* and AhPR-10 from peanut (*Arachis hypogaea*) are additional PR proteins with ribonuclease activity that are implicated in immunity against different pathogens (Pungartnik et al., 2009; Chadha and Das, 2006; Liu et al., 2006). Likewise, PR-4 proteins from different species also possess ribonuclease activity and are involved in immunity (Filipenko et al., 2013). For reviews on the role of PR10 and PR-4 in biotic stress please see Jain and Kumar, 2015 and Filipenko et al., 2013.

d) mRNA surveillance pathways

Eukaryotic cells possess a range of mRNA surveillance mechanisms to ensure that the mRNAs that are to be translated are adequate. These include nonsense-mediated mRNA decay (NMD), nonstop mRNA decay (NSD), and no-go mRNA decay (NGD) (Doma and Parker, 2007).

Nonsense-mediated decay (NMD) machinery

Perhaps the best studied pathway for mRNA surveillance is the NMD. NMD targets mRNAs with aberrant translation termination such as alternative splicing variants containing premature termination codons, long (≥ 300–350 nt) 3' UTRs, introns ≥ 50–55 nt downstream of termination codons, and uORFs (Shaul 2015). The NMD machinery is highly conserved in eukaryotes, although there are some differences between different species. The plant NMD machinery includes UPF1/2/3, SMG1/7 and exon-junction complex (EJC). NMD has been extensively linked to immunity against bacteria (described below) and viruses in plants (Garcia et al., 2014).

Several NMD components have been recently linked to plant immunity and mutations in some NMD proteins such as UFP1, UPF5 and SMG7 lead to increased resistance to *P. syringae* pv. *tomato* DC3000 (Jeong et al., 2011b; Riehs-Kearnan et al., 2012; Shi et al., 2012; Rayson et al., 2012). Moreover, NMD mutants have autoimmune phenotypes and display high constitutive *PR* gene expression and higher SA content (Jeong et al., 2011b; Riehs-Kearnan et al., 2012; Shi et al., 2012; Rayson et al., 2012). Furthermore, upon infection the mRNA levels of UPF1 and UPF3 are down-regulated, indicating that repression of NMD might be part of the immune response in order to stabilise genes important for the defence response that might contain premature termination codons (Jeong et al., 2011b). It has recently been described that NMD controls the turnover of *R* genes since *TNL* (TIR domain-containing, nucleotide-binding, leucine-rich repeat) immune receptor transcripts have increased half-life in *smg7* mutants (Gloggnitzer et al., 2014).

In summary, some NMD components have been shown to negatively regulate defence responses. Upon mutation of NMD proteins or upon host-programmed inhibition of NMD in response to *P. syringae* infection, TNL transcripts are stabilized, thus enhancing immune responses (Gloggnitzer et al., 2014). For more detailed information about the function of NMD in plants please see Shaul, 2015 or Ohtani and Wachter, 2019.

e) sRNA biogenesis machinery

It is well known that sRNAs have important roles in plant immunity (for review see Brant and Budak, 2018, Hua et al., 2018 and Zhu et al., 2019). A number of proteins involved in the biogenesis of sRNA have also been demonstrated to be involved in plant immunity through their role in the production of small RNAs. However, here we will not focus on these RBPs since this topic has been reviewed extensively (Staiger et al., 2013; Pumplin and Voinnet, 2013).

3.6. Stress granules (SGs) and processing bodies (P-bodies)

Upon stress conditions including biotic stresses, mRNAs can be re-localised to stress granules (SGs) and processing bodies (P-bodies) where they can be stored or degraded (Weber et al., 2008; Chantarachot and Bailey-Serres, 2017). Many processes of mRNAs turnover occur in P-bodies, including deadenylation, decapping, miRNA-targeted gene silencing and NMD. However, it has also been recently reported that mRNAs can also be stabilized and stored in P-bodies without undergoing turnover (Li et al., 2015; Merchante et al., 2015; Horvathova et al., 2017; Hubstenberger et al., 2017). Some recent studies link P-bodies or proteins within the P-bodies, and the re-localisation of RNAs to P-bodies and SGs, with immune responses (Petre et al., 2016; Roux et al., 2015; Maldonado-Bonilla et al., 2014; Li et al., 2015; Merchante et al., 2015; Yu et al., 2019). Additionally, some pathogens may also target P-bodies to interfere with RNA metabolism and disrupt defence responses in plants (Petre et al., 2016) and in other organisms

such as mammals (Ariumi et al., 2011; Pérez-Vilaró et al., 2015). For a recent review about SGs and P-bodies, please see Chantarachot and Bailey-Serres, 2017.

ETHYLENE INSENSITIVE 2 (EIN2)

One example of a protein associated with P-bodies and plant immunity is ethylene insensitive 2 (EIN2). EIN2 has been described to play a role in the response to various stimuli, including biotic stress in different species (Gazzarrini and Mccourt, 2003; Salvador-Guirao et al., 2018; Rin et al., 2017). EIN2 is an evolutionarily conserved ethylene signalling component that contains a transmembrane domain anchored to the ER, and a cytoplasmic domain (Ju et al., 2015). Upon ethylene perception, a fraction of the EIN2 cytoplasmic domain termed CEND is cleaved and shuttles to the nucleus where it acts by stabilizing two transcription factors (EIN3/EIL1) that positively regulate ethylene responses (Ju et al., 2012; Qiao et al., 2012; Wen et al., 2012; **Fig. 3**). This has been termed the 'cleave and shuttle model'. However, it was recently described that after cleavage the CEND can remain cytoplasmic, where it binds the 3'UTR of ethylene-responsive mRNAs (EBF1/2) and together with the NMD machinery promotes translational repression and localization to P-bodies of those mRNAs (Li et al., 2015; Merchante et al., 2015) (**Fig. 3**). EBF1/2 act to promote degradation of EIN3/EIL1 transcription factors. Therefore, by promoting translational repression of EBF1/2, EIN2 promotes the expression of ethylene responsive genes. Support of the CEND functioning as an RBP came from Reichel and colleagues, since they identified EIN2 in the RBPome of etiolated seedlings (Reichel et al., 2016).

EIN2 has been shown to regulate the flg22 receptor FLS2, and *ein2* mutants accumulate less *FLS2* mRNA and protein than wild-type plants (Tintor et al., 2013; Mersmann et al., 2010; Boutrot et al., 2010). However, the exact mechanism whereby EIN2 directly or indirectly regulates FLS2 transcripts and impacts on plant immunity remains unknown. It has been shown that that the *FLS2* promoter has nine potential EIN3/EIL1 binding sites, and EIN3 binds to two positions

(Boutrot et al., 2010). The fact that *ein3/eil1* mutants mirror the immune responses and disease phenotypes of some *ein2* mutants indicates that the effect of EIN2 on FLS2 may be indirectly through EIN3/EIL1 regulation (Boutrot et al., 2010). It is tempting to speculate that, in addition to regulating EBF1/2, EIN2 could directly regulate the metabolism of other mRNAs promoting their P-bodies re-localization and thus contribute to other outputs of plant immunity.

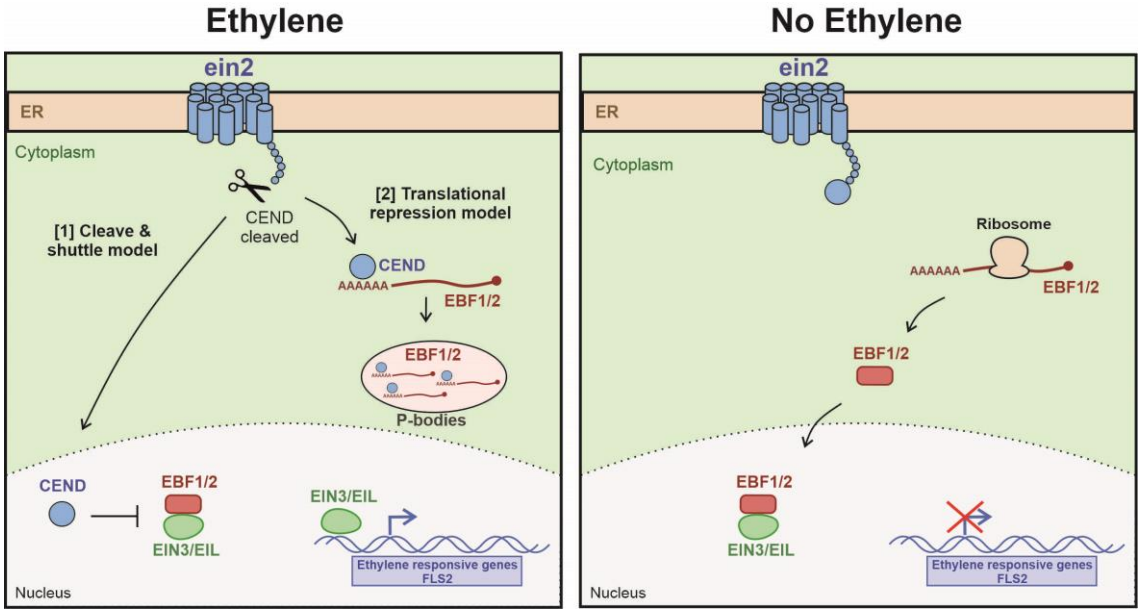


Fig. 3. Dual mode of action of EIN2. Upon ethylene perception, the CEND domain is cleaved and can play two different roles: [1] CEND can shuttle to the nucleus to stabilize two transcription factors (EIN3/EIL1) that positively regulate ethylene responses; [2] CEND can bind the 3'UTR of the negative regulators EBF1/2 in the cytoplasm and together with the NMD machinery re-localise them to the P-bodies to promote translational repression. EBF1/2 promotes degradation of EIN3/EIL1 transcription factors. Hence, translational repression of EBF1/2, results in the expression of ethylene responsive genes.

Several independent groups have described that *ein2* mutants have reduced immunity since they have reduced *PR* induction, abolished stomatal closure and reduced ROS burst and callose deposition (Mersmann et al., 2010; Tintor et al., 2013; Boutrot et al., 2010). Consistently, *ein2* mutant plants are more susceptible to *P. syringae* pv. *tomato* DC3000, especially at early time points (Tintor et al., 2013; Mersmann et al., 2010; Clay et al., 2009; Washington et al., 2016), suggesting an important role of EIN2 in regulating the early stages of immunity. However, a number of *ein2* mutants have been isolated during the past years, and there has been controversy with respect to their immune responses and disease phenotypes. For example,

Mersmann and colleagues found that EIN2 is required for some immune responses such as ROS burst, callose deposition and stomatal closure and for early *P. syringae* pv. *tomato* DC3000 resistance, but not for MAPK activation, seedling growth arrest or late resistance to *P. syringae* pv. *tomato* DC300 (Mersmann et al., 2010). Contrarily, Boutrot and colleagues identified that EIN2 is required for flg22-induced MAPK activation but not required for callose deposition (Boutrot et al., 2010). Regarding the disease phenotype, contrarily to the previously described, Chen and colleagues and Boutrot and colleagues described that *ein2* mutants possess an increased resistance to *P. syringae* pv. *tomato* DC3000 due to higher SA levels (Chen et al., 2009; Boutrot et al., 2010). Clearly, a proper identification and characterization of different mutants involved in ethylene perception and signalling (e.g. *etr1*, *ctr1*, *ein2*, *ein3*, *eil1*) is required to be able to draw solid conclusions.

TANDEM CCCH ZINC FINGER PROTEINS (TZFs)

The tandem CCCH zinc finger proteins (TZFs) typically contain an arginine-rich (RR) region followed by two CCCH-type zinc-finger (zf) motifs arranged in tandem (Bogamuwa and Jang, 2014). Mammalian TZFs have been described to have crucial roles for assembly of P-bodies and SGs (Carballo, 1998). Recently, some plant TZFs such as TZF9 and GhZFP1 have also been associated with P-bodies and SGs (Maldonado-Bonilla et al., 2014; Guo et al., 2009).

The Arabidopsis protein TZF9 has been postulated to play a central role in post-transcriptional regulation of PTI responses. *tzf9* mutants have altered immune responses such as decreased MPKs activation, decreased ROS burst and defective induction of early defence genes (Maldonado-Bonilla et al., 2014). Accordingly, *tzf9* mutants are more susceptible to infection by *P. syringae* pv. *tomato* DC3000 (Maldonado-Bonilla et al., 2014). TZF9 interacts and is phosphorylated by MPK3 and MPK6 upon immune activation, is able to bind RNA through the zf motifs, and its cytoplasmic (P-bodies) localization is RNA-dependent (Maldonado-Bonilla et

al., 2014). However, the precise role of TZF9 in post-transcriptional RNA regulation remains largely unknown.

3.7. RNA Editing

Organellar (i.e. plastid and mitochondrial) RNAs can undergo post-transcriptional RNA editing before translation. One common type of RNA editing is deamination from cytidine (C) to uridine (U), which is performed by the RNA editing complex. The RNA editing complex includes both PPR and RRM domain-containing proteins, RNA editing factors/RNA editing interacting proteins and zinc-finger proteins (Lu, 2018; Ichinose and Sugita, 2017). RNA editing alters the RNA sequence, thus potentially affecting the characteristics of the transcript. Hence, RNA editing can affect the start and stop codons, splicing and alternative splicing sites, miRNA maturation sites or can result in changes at the amino acid level, potentially changing the protein function or interaction with other proteins (Rodrigues et al., 2017). Because RNA editing can cause such dramatic changes to the protein function, it is not unexpected that some of them may be involved in important processes such as plant defence responses. One example is OCP3, which is involved in disease resistance regulation by controlling the editing efficiency of plastid transcripts (García-Andrade et al., 2013).

3.8. Ribosomes and translation

Recent studies have provided evidence that one of the major determinants of cellular protein levels is the control of translation (Guo, 2018). Traditionally the ribosomes were seen as static macromolecules involved in protein synthesis. However, recent findings have led to the identification of ‘specialised ribosomes’, with heterogeneous composition of ribosomal RNAs and proteins or different post-translational modifications of the ribosomal components (Xue and Barna, 2012; Mauro and Edelman, 2002). Moreover, these heterogeneous ribosomes can preferentially translate different subsets of mRNAs (Guo, 2018). This hypothesis is supported by

the presence of multiple paralogs of ribosomal proteins (RP) coding genes which are produced under certain conditions (Parenteau et al., 2011). For example, in Arabidopsis each RP can be encoded by 2-7 genes (Barakat et al., 2001). Ribosomes have been described to vary in response to different factors such as cellular status, environmental conditions, developmental stage or pathological condition (Venezia et al., 2019). For instance, several ribosomal proteins are altered in human cells infected with virus (Garcia-Moreno et al., 2019). Therefore, it not unexpected that ribosomes are also involved in translational control of immune-related proteins.

It has been described that RPL12C and RPL19B are involved in the defence response against bacterial pathogens in both *N. benthamiana* and Arabidopsis (Nagaraj et al., 2016). Fakih and colleagues identified different accumulation of ribosomal and translation-related proteins in plants treated with the elicitor chitosan (Fakih et al., 2016). Moreover, Eskelin and colleagues analysed the ribosome profiles of *N. benthamiana* plants untreated or treated with *Agrobacterium tumefaciens* either carrying infectious cDNA of Potato virus A (PVA) or a luciferase control by affinity-tag ribosome purification and revealed substantial changes in the ribosome composition (Eskelin et al., 2019).

Therefore, it is clear that the composition of ribosomes and the proteins associated with them is dynamically regulated. This is likely to result in distinct translation patterns to regulate the plant responses to different environmental cues, including biotic stresses.

3.9. Extracellular vesicles

Extracellular vesicles (EVs) are important for intercellular RNA transport and are especially produced in response to pathogen infection (Rutter and Innes, 2017). Moreover, EVs are known to be important for plant immune responses (Rybak and Robatzek, 2019). Recently, the protein contents of apoplastic EVs from Arabidopsis were analysed (Rutter and Innes, 2017). Many of

the identified proteins were related to defence responses, and some of them were involved in RNA metabolism such as GRP7 or ribosomal proteins (Rutter and Innes, 2017). Because EVs contain RNAs (Baldrich et al., 2019), it would not be unexpected to identify RBPs that accompany and regulate EV-localised RNAs. However, the role of RBPs within the EV and whether they have any specific role in RNA metabolism during immune responses remains unexplored.

3.10. RNA Helicases

RNA helicases are a large family of proteins involved in the unwinding of RNA secondary structure in virtually all steps during the RNA lifecycle (Chen and Shyu, 2014). For instance, the genome of Arabidopsis encodes for 58 RNA helicases (Mingam et al., 2004). Because changes in RNA helicase activity can result drastic remodelling of RNPs (Chen and Shyu, 2014), it is not unexpected that RNA helicases are critical players in different plant processes, including plant immunity. Multiple RNA helicases have been described to be involved in viral infection in multiple organisms, including humans and plants (Garcia-Moreno et al., 2019; Ahmad and Hur, 2015).

However, RNA helicases have been largely unstudied in plants and only a few proteins such as the rice BIRH1 has been characterised in response to biotic stress. OsBIRH1 has been described to be important in responses to biotic and oxidative stress (Li et al., 2008). Overexpression of OsBIRH1 in Arabidopsis results in up-regulated expression of defence-related genes and confers increased resistance to *A. brassicicola* and *P. syringae* pv. *tomato* DC3000 (Li et al., 2008). However, the mechanism through which it contributes to immunity is currently unknown.

Moreover, some RNA helicases have been described to be targeted by pathogens. For example, the P*hytophthora* RXLR effector suppressor of RNA silencing 1 (PSR1) targets the host evolutionarily conserved PSR1-interacting protein 1 (PINP1) and PINP1 silencing results in hyper-

susceptibility to *Phytophthora* (Qiao et al., 2015). PINP1 is a nuclear protein containing an RNA helicase domain that regulates accumulation of microRNAs and small interfering RNAs. Moreover, PINP1 has also been described to be involved in splicing (Tsugeki et al., 2015).

3.11. Unknown role

A number of RBPs that have not been assigned to the categories discussed above have been described to play important roles in plant immunity, although their implication in RNA metabolism is currently unknown. Below we discuss one particular interesting RBP, and additional examples can be found in **Supplemental table 1**.

K-HOMOLOGY (KH) RNA-BINDING PROTEIN 1 (StKRBP1)

StKRBP1 is a putative potato K-homology (KH) RNA-binding protein that is targeted by a *P. infestans* RXLR effector and re-localises to nuclear speckles upon *P. infestans* infection (Wang et al., 2015). StKRBP1 has been postulated to be a susceptibility factor since both overexpression of StKRBP1 or of the effector in *N. benthamiana* leads to increased susceptibility to *P. infestans* (Wang et al., 2015). The GxxG motifs within the three KH domains of StKRBP1 are key for the interaction with the effector, and mutation of these motifs disrupts the interaction, localization to nuclear speckles and increased susceptibility to *P. infestans* (Wang et al., 2015). Although the precise function of StKRBP1 in immunity awaits to be discovered, it was postulated that the binding of StKRBP1 to the effector could lead to stabilization of this susceptibility factor (Wang et al., 2015). Despite StKRBP1 having 3 KH domains, the function of StKRBP1 in RNA metabolism not yet known.

SUPPLEMENTAL DATA

APPENDIX I – Supplemental data Chapter 1

Supplemental table 1. RNA-binding proteins involved in plant immunity

AIMS AND RATIONALE OF THE THESIS

RNA-binding proteins (RBPs) regulate RNAs in all the steps of RNA lifecycle (Glisovic et al., 2008). Thus, RBPs are crucial players controlling gene expression and RNA regulation not only during cellular homeostasis but also in response to different environmental cues (Staiger et al., 2013; Yu et al., 2019).

The development of RNA interactome capture (RIC) has permitted researchers to uncover the RNA-binding proteome (RBPome; mRNA interactome) of multiple species from yeast to humans (Hentze et al., 2018). However, for plants only the RBPome of Arabidopsis has been experimentally investigated to date. Hence, the RBPomes of economically and agriculturally important crops await to be discovered. Moreover, due to the technical challenges plant tissue imposes, RIC has only been applied to Arabidopsis tissues that are not relevant for some areas of plant biology, especially for plant pathology - etiolated seedlings (Reichel et al., 2016), protoplasts (Zhang et al., 2016b) and cell cultures (Marondedze et al., 2016). We hypothesise that by modifying several key steps and parameters within the RIC protocol it can be extended to a large number of plant species and tissues.

Hence, we hypothesise that RIC could be key at unveiling the RBPomes of different plant species, thus allowing studies on the conservation of RBPs across the plant kingdom. We also envisage that this could lead to the discovery of a large number of novel RBPs, similarly to what has been reported in other species (Castello et al., 2012).

Plants require reprogramming of the transcriptome to mount a defence response against pathogens. RBPs are known to be key players in post-transcriptional gene regulation (Glisovic et al., 2008), and indeed, multiple RBPs play important roles during immunity (reviewed above). Moreover, the composition and activity of the RBPome has been shown to be tightly regulated

to orchestrate the changes in RNA metabolism that occur under different cellular contexts (Garcia-Moreno et al., 2019; Sysoev et al., 2016; Trendel et al., 2019; Perez-Perri et al., 2018; Marondedze et al., 2019). However, the evidence for the role of RBPs in plant immunity has been collected stepwise and we lack comprehensive approaches to determine the identity and regulation of RBPs that function during plant immune responses. RIC has allowed the unprecedented opportunity to investigate the dynamic behaviour of the RBPome in response to different cues and in wide range of species (Hentze et al., 2018). Hence, we hypothesise that RIC could be instrumental in uncovering the scope of plant RBPs that orchestrate transcriptome reprograming during plant immunity.

Therefore, in this thesis I have the following aims:

- In **Chapter 2** I aim to optimise RNA-interactome capture to efficiently identify RBPs actively bound to RNA from plant leaves.
- In **Chapter 3** I aim to identify the RBPomes of different plant species of agricultural and economical importance across the plant kingdom. By comparing the different plant RBPomes I also aim to unravel the core plant RBPome and to identify the potential novel RBPs and RBDs in plants.
- In **Chapter 4** I aim to profile the dynamics of the Arabidopsis RBPome in response to the immune elicitor flg22 and to validate some of flg22-responsive RBPs by immunity-based mutant screen.
- In **Chapter 5** I discuss the results and perspectives of the thesis.

CHAPTER 2

Plant RNA-Interactome Capture (ptRIC):

an Improved Method to Identify RBPs from Plant Leaves

Plant RNA-Interactome Capture (ptRIC): an Improved Method to Identify RBPs from Plant Leaves

Marcel Bach-Pages[1], Felix Homma[1], Jiorgos Kourelis[1], Farnusch Kaschani[2], Shabaz Mohammed[3], Renier A. L. van der Hoorn[1], Alfredo Castello[3]*, Gail M. Preston[1]*

[1]Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, UK

[2]Fakultät für Biologie, *Universität Duisburg-Essen, Essen, Germany*

[3]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK

* Co-corresponding authors

Keywords

RNA-binding proteins, RBP // RNA-binding proteome // RBPome // RNA-interactome capture // RIC // ptRIC // Arabidopsis

ABSTRACT

RNA-binding proteins (RBPs) are important post-transcriptional regulators of gene expression. Hence, unravelling the identity and activity of RBPs is necessary to understand the regulation of RNA function and gene expression in different cellular contexts. The development of RNA interactome capture (RIC) has aided in unveiling the RNA-binding proteomes (RBPomes) in multiple species and has recently been applied to plants. We have developed an improved RIC protocol termed 'plant RNA-interactome capture' (ptRIC), which is optimised for plant leaves. We have used ptRIC to identify large numbers of *bona fide* RBPs, including organellar RBPs, from mature *Arabidopsis* leaves. ptRIC has allowed identification of 722 RBPs co-purifying with RNA, which represents the first extensive Arabidopsis leaf RBPome to date. Interestingly, while 74%

of the high-confidence leaf RBPome was already linked to RNA biology, the remaining 26% represents putative novel RBPs. In addition, a large proportion of the leaf RBPs identified lack known RNA-binding domains (RBDs), providing scope for discovery of novel RBDs. ptRIC represents a significant advance on current methods, since deeper resolution of the RBPome allows a more robust study of the dynamics of the RBPome. We anticipate that ptRIC will greatly benefit the understanding of plant RBPomes and their dynamic regulation upon environmental changes or in different cellular states.

1. INTRODUCTION

1.1. RNA-binding proteins (RBPs)

RBPs interact with RNAs to form ribonucleoprotein (RNP) complexes and regulate the fate of RNAs at each of the steps of RNA life cycle. Therefore, RBPs are key players in the control of gene expression by regulating the synthesis, processing (capping, splicing and polyadenylation), post-transcriptional modifications, transport, storage, surveillance/quality control, translation and turnover of RNA (Glisovic et al., 2008). Because RBPs have a critical role for cellular function, the RNA-binding proteome (RBPome) is tightly regulated, adapting to alterations in environmental conditions and cellular states (Sysoev et al., 2016; Perez-Perri et al., 2018; Garcia-Moreno et al., 2019; Trendel et al., 2019). Therefore, elucidating the composition and activity of the RBPome is crucial to understand how RNA function and gene expression is regulated *in vivo*.

1.2. Historical study of RNA-protein complexes

In the 1950s, one of the first studies of RNA-protein (ribonucleoprotein) complexes visualised transcripts associated with proteins following transcription by RNA polymerase II. Subsequently, researchers have studied RNPs using a variety of techniques in a wide range of organisms (reviewed in Singh et al., 2015). For example, Pramanik and Bewley isolated and characterised

RNPs in alfalfa embryos by combining UV crosslinking and affinity chromatography on an oligo(dT)-cellulose column (Pramanik and Bewley, 1996). Other studies have used other techniques such as velocity sedimentation, gel filtration, or formaldehyde fixation and isopycnic centrifugation to purify RNP complexes (Jenkins et al., 1978; Grainger and Winkler, 1987; Spirin et al., 1965).

In the last decade, many efforts have been undertaken to identify RBPs in a comprehensive manner. For example, some studies have used fluorescent RNA probes incubated on protein arrays (Scherrer et al., 2010; Tsvetanova et al., 2010). Although these reported hundreds of proteins able to interact with RNA *in vitro*, they fail to identify physiological, *in vivo* RNA-protein interactions (Castello et al., 2013b). This is because *in vitro* studies can capture interactions between RNAs and proteins that may not occur in the cellular context due to, for example, different cellular localisation. Moreover, *in vitro* studies cannot distinguish *bona fide* RNA-protein interactions from the non-physiological interactions that occur *in vitro* due to the biochemical properties of proteins. Many *in silico* algorithms have been developed to identify novel RBPs by searching for domains with homology to well-established RNA-binding domains (RBDs). This has helped to classify of hundreds of proteins as putative RBPs in different species (Anantharaman et al., 2002). However, these computational approaches are based on homology and thus fail to identify RBPs with an unconventional architecture.

1.3. RNA-Interactome capture (RIC)

A new technique termed RNA interactome capture (RIC) was recently described that can systematically determine the proteins bound to polyadenylated RNAs in living cells (Castello et al., 2012; Baltz et al., 2012; Castello et al., 2013b). RIC employs ultraviolet (UV) irradiation of cells to promote RNA-to-protein crosslinking and capture of the poly(A) RNAs to isolate the RNA-bound proteins. RIC has many advantages over previous methods. Because UV light acts on the

nucleotide bases, it promotes selective crosslinking between RNA and proteins but does not promote protein-protein crosslinking (Pashev et al., 1991); thus, RIC identifies RBPs actively bound to RNA, preventing the capture of proteins associated with RNA through non-covalent interactions or protein-protein interactions. Moreover, UV light is directly applied to cells or tissues, thus "freezing" native and physiological protein-RNA interactions. RBP identification by RIC is not biased towards 'classic' RNA-binding domains, thus allowing the discovery of unconventional RBPs. Moreover, RIC can be coupled to quantitative proteomic analyses, thus enabling comparative analyses to profile RBP activities in different physiological contexts (Sysoev et al., 2016; Garcia-Moreno et al., 2019; Perez-Perri et al., 2018). This offers the unprecedented possibility to study the dynamics of RBPs under different developmental stages, cellular states and in response to changes in environmental conditions.

Since its publication in 2012, RIC has been applied to multiple organisms including *Trypanosoma brucei* (Lueong et al., 2016), *Leishmania donovani* (Nandan et al., 2017; De Pablos et al., 2019), *Plasmodium falciparum* (Bunnik et al., 2016), *Saccharomyces cerevisiae* (Beckmann, 2017; Beckmann et al., 2015; Mitchell et al., 2013; Matia-González et al., 2015), *Caenorhabditis elegans* (Matia-González et al., 2015), *Drosophila melanogaster* (Sysoev et al., 2016; Wessels et al., 2016), *Danio rerio* (Despic et al., 2017), *Mus musculus* (Kwon et al., 2013; Liao et al., 2016; Liepelt et al., 2016; Boucas et al., 2015) and different *Homo sapiens* cell lines (Baltz et al., 2012; Castello et al., 2012; Kramer et al., 2014; Beckmann et al., 2015; Castello et al., 2016; Perez-Perri et al., 2018; Conrad et al., 2016). Therefore, RIC has made an important contribution to the RNA field by enabling systematic identification of RBPomes in different biological systems (Hentze et al., 2018).

More recently, RIC has also been applied to different tissues of the model plant *Arabidopsis thaliana*: etiolated seedlings (Reichel et al., 2016), cell cultures (Marondedze et al., 2016, 2019)

and protoplasts derived from the mesophyll (Zhang et al., 2016b, 2017c). Although a first attempt to apply RIC to leaves was made (Marondedze et al., 2016), it barely identified 27 proteins. These are extremely low numbers when compared with RBPomes generated in other systems (~350-1000 RBPs). Hence, the leaf RBPome awaits to be uncovered.

1.4. Rationale and aim of the study

The RBPomes of multiple species have been uncovered using RIC (Hentze et al., 2018). Although three studies have recently applied RIC to analyse the RBPome of Arabidopsis (Reichel et al., 2016; Zhang et al., 2016b; Marondedze et al., 2016), none of them have successfully uncovered the RBPome of leaves. This is probably because applying RIC to plant leaves is challenging since: i) the UV crosslinking efficiency can be reduced due to the thickness of leaves and the presence of UV-absorbing pigments such as chlorophyll; ii) the composition of leaves is more complex than human cell lines due to the presence of the cell wall, multiple secondary metabolites, etc. and thus different aspects of the protocol such as tissue homogenisation and cell lysis are challenging.

Leaves are a tissue that is broadly used in plant biology and is critical for understanding plant physiology and functioning. We aimed to optimise RIC to unravel for the first time the RBPome of mature leaves. Here we describe an improved RIC protocol especially tailored to plant mature leaves, referred to here as 'plant RNA-interactome capture' (ptRIC).

2. RESULTS

2.1. Development of ptRIC protocol

To comprehensively determine the leaf RNA-binding proteome, we decided to further optimise RIC to be efficiently applied to leaves. In our initial attempt, we applied the RIC settings

established by Reichel and colleagues (Reichel et al., 2016), but using mature leaves of *Arabidopsis thaliana*. In brief, leaves were detached from the plants and irradiated with UV light at 254 nm to induce covalent bonds between RNAs and proteins that are in intimate contact (≤2Å). Next, tissue was disrupted and cells were lysed, followed by the capture of polyadenylated RNA-protein complexes using oligo(dT) magnetic beads. This isolation was performed under stringent denaturing conditions using ionic detergents and high LiCl concentrations to exclusively

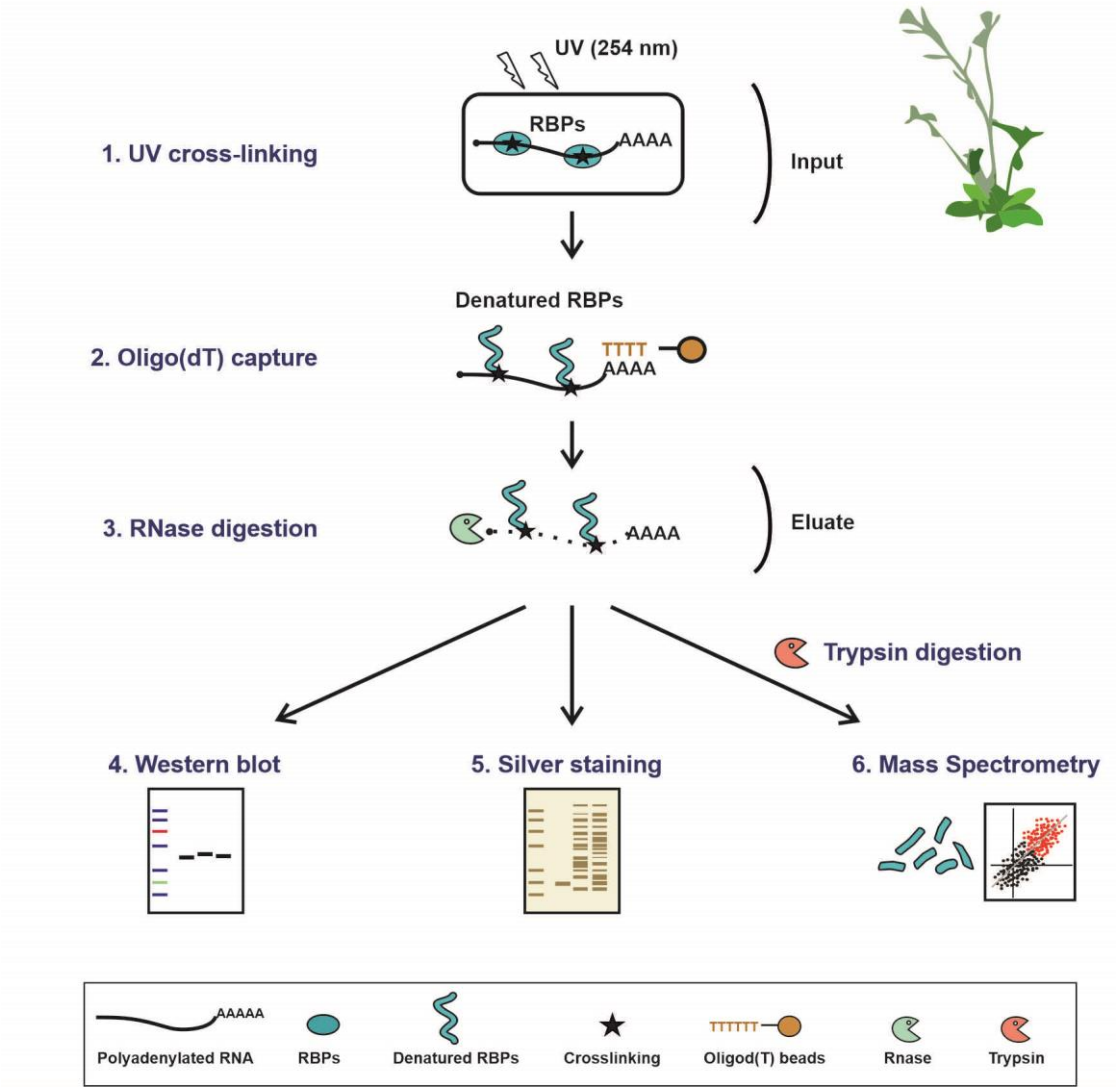


Fig.1 Schematic representation of ptRIC. Mature plants (leaves) are irradiated with UV light at 254 nm to promote crosslinking between RNAs and proteins that are in intimate contact (1). Next, cells are lysed and mRNAs pulled-down using oligo(dT) magnetic beads (2). After stringent washes, the RNA-protein complexes are recovered and the RBPs released by RNase digestion (3). The proteins can be quantitatively analysed by (4) western blot, (5) silver staining or (6) mass spectrometry after trypsin digestion.

isolate RNAs together with their directly bound proteins. The RNAs were subsequently degraded using RNases and the proteins associated with RNAs were analysed by silver staining and mass spectrometry (**Fig. 1**). A non-UV irradiated sample (NoCL) was processed in parallel as negative control. In our pilot experiment, although proteomic analysis revealed near two thousand proteins, only 40% of those were categorised as RBPs (log_2FC [CL/NoCL] ≥ 2) and almost 60% were not enriched over the non-UV irradiated negative control, which are likely to be contaminants (**Fig. 2,** MS1). The high incidence of potential contaminants present in both NoCL and CL resulted in a higher sample complexity that may reduce the depth of the proteomic analysis. Moreover, the presence of 'contaminants' can compromise the downstream analyses, leading to a higher incidence of false positives. The fact that ~60% of the identified proteins may be contaminants strongly called for further optimisation of the protocol.

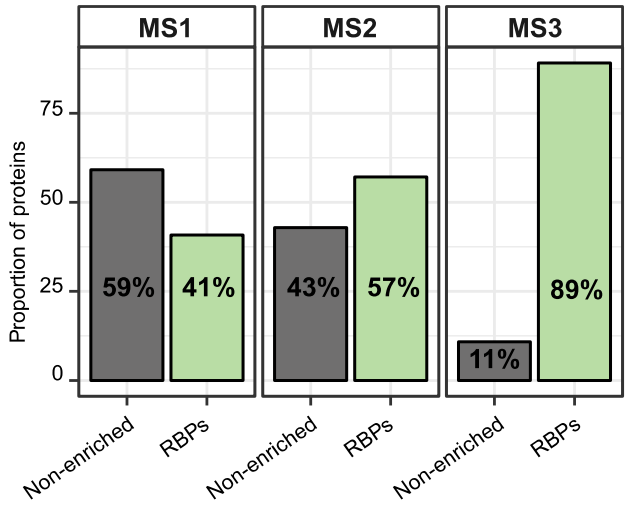


Fig.2. Development of ptRIC, an improved method to isolate RBPs actively bound to RNA from leaf tissue. Proportion of proteins classified as +UV/-UV 'non-enriched' or 'RBPs' in each of the mass spectrometry (MS) experiments. In MS1 and MS2 pilot experiments, proteins with missing values in the NoCL sample or log_2FC [CL/NoCL] ≥ 2 are classified as 'RBPs', otherwise they are classified as 'non-enriched'. In the MS3 experiment proteins with missing values in all the NoCL samples or log_2FC [CL/NoCL] ≥ 1.5 and FDR ≤ 0.05 are classified as 'RBPs', otherwise are were classified as 'non-enriched'.

To reduce the incidence of contaminants we increased the volume of lysis buffer to tissue ratio and introduced an extra washing step with a buffer that we termed 'harsh buffer' (**Table 1**), which includes higher concentrations of the ionic detergent LiDS (1% w/v) and LiCl (2M). The

addition of this buffer was intended to increase the stringency of the washes to dissociate proteins interacting with RNA via non-covalent interactions or protein-protein interactions, and to remove proteins sticking to the oligo(dT) beads. This second pilot experiment resulted in a decrease in the proportion of proteins with non-significant +UV/-UV ratio (from ∼60% to 43%; **Fig. 2,** MS2).

We suspected that some of the non-enriched contaminants could be associated with cellular debris trapped in the beads, resulting from incomplete homogenization of the tissue. To circumvent this, we increased the lysis buffer to tissue ratio and added an additional step of homogenisation with a Potter-Elvehjem homogenizer. We passed the lysates through a narrow syringe to shear the gDNA that could stick to the beads and increase protein background (DNA-binding proteins also crosslink with DNA upon UV exposure). Moreover, we further optimised multiple different parameters that we suspected were affecting RBP recovery. UV crosslinking can be inefficient in leaves due to the thickness of the tissue and the presence of UV-absorbing pigments. Hence, we reasoned that the processing of the tissue and the UV dose could be important factors determining UV-dependent protein recovery. We tested UV irradiation of either intact or pulverised leaves and determined that intact detached leaves were the most optimal starting material (data not shown). We next tested different doses of UV light (from 150 mJ/cm^2 to triple irradiations with 800 mJ/cm^2) and determined that the most optimal UV crosslinking was three irradiations at 150 mJ/cm^2 spaced by a 30 seconds pause (data not shown). Furthermore, we hypothesised that the amount of oligo(dT) beads could influence RNA recovery and, if in excess, promote unspecific capture of contaminants. Indeed, while higher amounts of oligo(dT) resulted in an increase in recovered RBPs, it also caused a concomitant increase in contaminants (data not shown). We observed that 250 µl of magnetic beads kept the optimal balance between signal and noise. By applying these modifications, we increased the

proportion of RBPs to 89% of the proteins identified, and as a consequence, the proportion of +UV/-UV non-enriched proteins dropped to 11% (**Fig. 2,** MS3; **Fig. 3A**).

Silver staining analyses of the proteins eluted with these ptRIC settings revealed a specific protein pattern in CL samples that resembled previously established RNA interactomes (**Fig. 3B**). Importantly, the non-irradiated control was devoid of proteins, confirming the stringency of our purification (**Fig. 3B**). Moreover, the banding pattern of the CL lane differed from the whole cell lysate, indicating the isolation of a specific subset of proteins, likely to be RBPs (**Fig. 3B**).

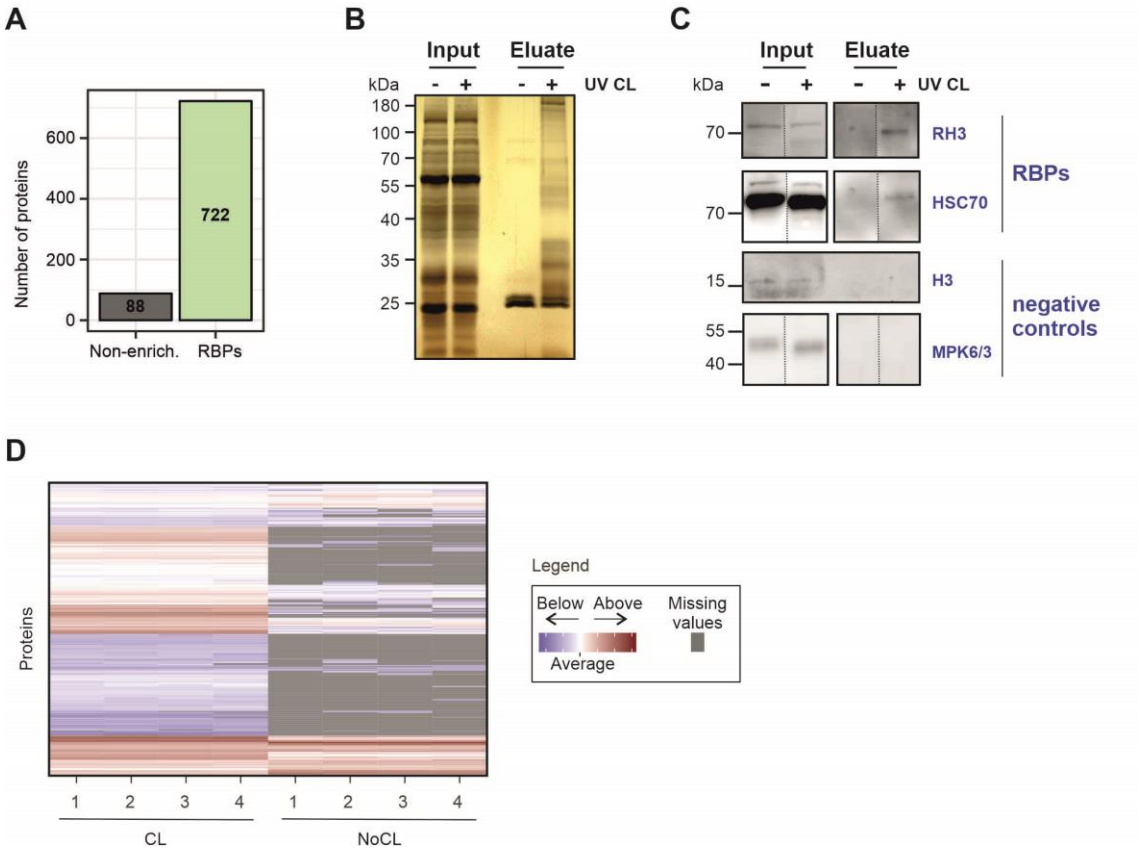


Fig.3. Identification and validation of Arabidopsis leaf RBPs using ptRIC. A) Classification of the proteins identified in the MS3 experiment. B) Silver staining analyses of the inputs (whole cell lysates) and eluates (RBPs) of the ptRIC. '+' and '-' indicate '+ UV crosslinking' or '- UV crosslinking'. C) Western blotting analyses of the inputs (whole cell lysates) and eluates (RBPome) of the ptRIC. '+' and '-' indicate '+ UV crosslinking' or '- UV crosslinking'. D) Heat map representing the profile of intensities of the proteins identified by MS in the MS3 experiment. The raw intensities were re-scaled and intensities above the average are coloured in red, whereas intensities below the average are coloured in blue. Missing values are coloured in grey. Proteins are clustered based on intensity profiles similarity.

2.2. **Validation of ptRIC**

To further validate these results, we used ptRIC followed by western blot focusing on specific proteins. We confirmed that RH3 and HSC70, which were classified as RBPs by ptRIC, are enriched in eluates only when UV is applied (**Fig. 3C**). However, the antibody used to validate HSC70 recognise three members of Arabidopsis HSC70 family (HSC70-1-3) and all three were identified as RBPs by mass spectrometry analyses. All three HSC70 members but HSC70-1 are around 70 kDa, hence we cannot distinguish whether the western blot signal corresponds to HSC70-2, HSC70-3 or both. As negative controls we used MAPK6 and MAPK3 which are absent in the RBPome, and histone 3 (H3) as marker of DNA contamination. Both proteins were detected in the whole cell lysates (inputs) but not in the eluates, indicating that the incidence of DNA contamination in ptRIC experiments is negligible (**Fig. 3C**).

2.3. **Building a high confidence leaf RBPome for Arabidopsis**

To establish the Arabidopsis leaf RBPome, protein $\log_2$ fold change between UV crosslinked and non-crosslinked samples ($\log_2$FC [CL/NoCL]) was calculated and moderate t-test and false discovery rate (FDR) estimation were applied using information from the four biological replicates. We classified proteins as high-confidence RBPs when the $\log_2$FC [CL/NoCL] ≥ 1.5 and the adjusted p-value (FDR) ≤ 0.05. Following this method, 432 proteins were classified as high-confidence Arabidopsis RBPs. However, a large number of proteins identified had missing values (i.e. zero intensity) in all the NoCL controls (**Fig. 3D**), which impedes the generation of ratios because it results in 'infinite' values for the $\log_2$FC [CL/NoCL]. The high incidence of 'zero' intensity values in the negative control highlights the high stringency of our experimental approach. To analyse the proteins with missing values in the negative control, we used a modification of the semiquantitative method described by Sysoev (Sysoev et al., 2016), which classifies as RBP proteins with intensity values in (at least) three out of four CL samples and

lacking signal in the non-irradiated controls. This analysis resulted in 290 additional proteins classified as RBPs.

Overall, our improved ptRIC protocol and the combination of quantitative and semi-quantitative data analyses led to the identification 722 high-confidence RBPs in leaves of Arabidopsis (**Fig. 3A**). The full list of proteins included in the high-confidence leaf RBPome (high-confidence leaf RNA interactome) can be found in **Supplemental digital table 1 (Appendix III)**. This represents the first comprehensive RBPome of mature leaves, since Marondedze and colleagues only identified 27 proteins as RBPs (Marondedze et al., 2016).

2.4. Insights into the leaf RBPome

2.4.1. Function and localization of leaf RBPs

We interrogated whether the high-confidence leaf RBPome had known links to RNA biology using gene ontology (GO) annotations as defined by Beckmann and colleagues (Beckmann et al., 2015). Importantly, approximately 74% of the leaf RBPs were already linked to RNA biology (**Fig. 4D**), indicating that ptRIC yields a high confidence RBPome. Interestingly, the remaining 26% RBPs had no known or predicted function in RNA biology and thus represent novel RBPs. 'RNA binding', 'mRNA binding' or 'nucleic acid binding' were amongst the most enriched GO terms, further confirming the validity of our approach (**Fig. 4A**). Other statistically enriched GO terms included other RNA-related functions such as 'structural constituent of ribosome', 'translation initiation factor activity' and 'translation regulator activity'. We found amongst the statistically underrepresented terms RNA-unrelated functions such as 'phosphotransferase activity', 'protein kinase' and 'DNA-binding transcription factor'. The depletion of the 'DNA-binding transcription factor' GO term in the leaf RBPome indicates that our eluates are largely free of DNA contamination. These results are similar to those previously observed in mammalian RIC

experiments (Castello et al., 2012). Leaf RBPs were annotated to be localised not only in nucleus (147 RBPs) or cytoplasm (304 RBPs), but also in the chloroplast (311 RBPs) and mitochondria (85 RBPs), indicating that ptRIC allows identification of RBPs from different cellular compartments, including subcellular organelles. Moreover, 'chloroplast' and 'plastic' were statistically enriched GO cellular component terms (**Fig. 4B**).

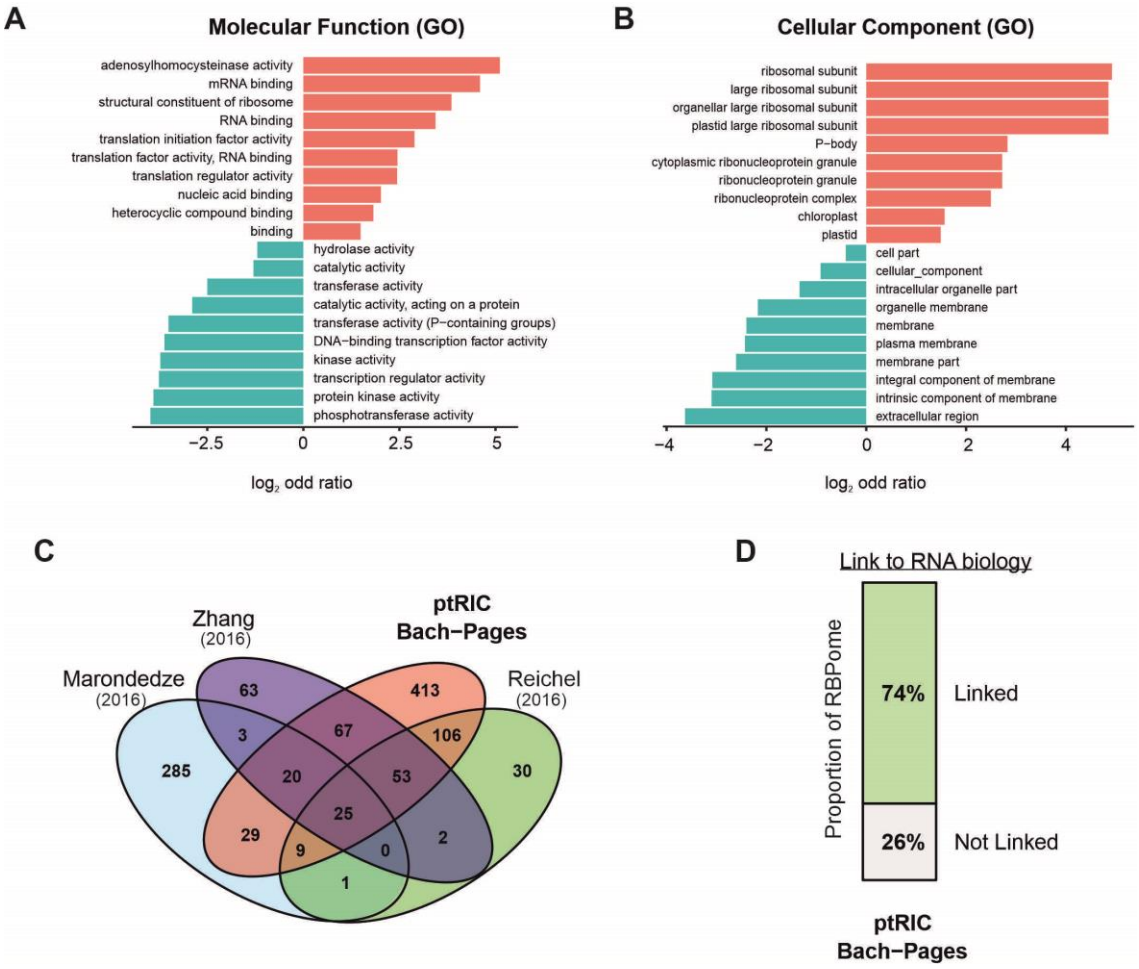


Fig.4. Insights into the Arabidopsis leaf RBPome identified by ptRIC. GO analysis with ten of the most significant enriched (red) or underrepresented (blue) molecular function (A) or cellular component (B) GO terms of the leaf RBPome. C) Venn diagrams depicting the overlap between the three publically available Arabidopsis RBPomes and the ptRIC leaf RBPome. D) Proportion of the leaf RBPome linked to RNA biology based on GO annotations.

Remarkably, out of the 201 ribosomal proteins (RP) identified by ptRIC, 18 were mitochondrial and 83 were chloroplastic RPs. This high proportion of identified organelle RPs contrasts with the RBPome of etiolated seedlings, where just one chloroplastic and one mitochondrial RP were

identified (Reichel et al., 2016). Hence, ptRIC allows deeper analyses of the Arabidopsis RBPs, including organellar RBPs.

2.4.2. RNA-binding domains in plant RBPs

We expected a strong enrichment in RNA-binding domains (RBDs) if our RBPome is enriched in *bona fide* RBPs. In agreement, about 60% of the leaf RBPs harbour known RBDs (classical or non-classical RBDs), whereas the remaining 40% harbour no recognisable RBD (**Fig. 5A**). These proportions are similar in RNA interactomes of other species including humans and plants (Castello et al., 2012; Reichel et al., 2016; Marondedze et al., 2016). One hundred and sixty RBPs harbour classical RBDs, which are domains well-characterised at the biochemical and structural level (Lunde et al., 2007). The most prominent of these domains are, as expected, the RNA recognition motif (RRM), followed by DEAD-box, zinc finger (zf)-CCCH and K-homology domains (KH) (**Fig. 5B**), which are the most abundant RBDs in plants and mammals (Castello et al., 2012; Reichel et al., 2016). We have identified 110 RRM-containing proteins out of the 197 predicted in the Arabidopsis genome (Silverman et al., 2013); 10 out of 28 predicted KH-containing proteins; 4 out of 5 predicted CSD-containing proteins (Silverman et al., 2013) and 6 out of 25 PUF-containing proteins (Wang et al., 2018). To date, the only Arabidopsis members of the PUF-containing protein family that have been experimentally validated to bind RNA are PUM2 and PUM5 (Francischini and Quaggio, 2009; Huh and Paek, 2014; Huh et al., 2013). We observed that PUM1-6 co-purify with RNAs *in vivo*, and these proteins are also identified as RBPs in the RNA interactome of etiolated seedlings (Reichel et al., 2016), providing strong evidence that multiple members of the PUF protein family bind to RNA in plants.

We refer to protein domains that have been described to bind to RNA in literature at least once, although their interaction with RNA are not well-established biochemically and structurally as non-classical RBDs. We identified 277 RBPs harbouring non-classical RBDs, with ribosomal

domains and pentatricopeptide repeat (PPR), helicase C-terminal, zf-CCHC and like-Sm (LSM) as the most prominent ones (**Fig. 5C**). The fact that a large proportion of the proteins harbouring non-classical RBDs contained ribosomal domains has been observed in the RNA interactomes in other species (Castello et al., 2012; Liao et al., 2016; Reichel et al., 2016). This is not surprising since it is known that some ribosomal proteins possess roles in mRNA regulation (Warner and McIntosh, 2009) while others interact directly with mRNAs at the ribosomal RNA channel (Pisarev et al., 2008). We have also identified 10 of the 13 YT521-B homology (YTH)-containing proteins in Arabidopsis (Li et al., 2014), which are known to 'read' methyl 6 adenosine modifications within RNA (Liao et al., 2018). YTH-containing proteins were identified in the RNA interactome of Arabidopsis seedlings (Reichel et al., 2016) and we confirmed their RNA-binding activity here. Moreover, we identify a total of 161 RBPs containing PPRs. This contrasts with the low numbers identified in other plant RNA interactomes ranging from 8 - 50 PPR-containing proteins (Reichel et al., 2016; Marondedze et al., 2016, 2019; Zhang et al., 2016b). There is evidence that some PPR subclasses are involved in interactions with RNA (Barkan and Small, 2014) and are thus classified as 'non classical RBD'. However, other subclasses lack links to RNA biology and are thus classified here as 'putative RBD'.

A large proportion of the identified RBPs do not possess any recognisable RBD (285 RBPs; ∼40% of the leaf RBPome; **Fig. 5D**). We evaluated whether these proteins are genuine RBPs by performing a GO enrichment analysis. Strikingly, the most enriched GO terms were related to 'RNA biology', including 'mRNA binding', and 'translation initiation factor activity' (**Fig. 5E**). This indicates that although 40% of the leaf RBPs do not possess known RBDs, they are likely to be

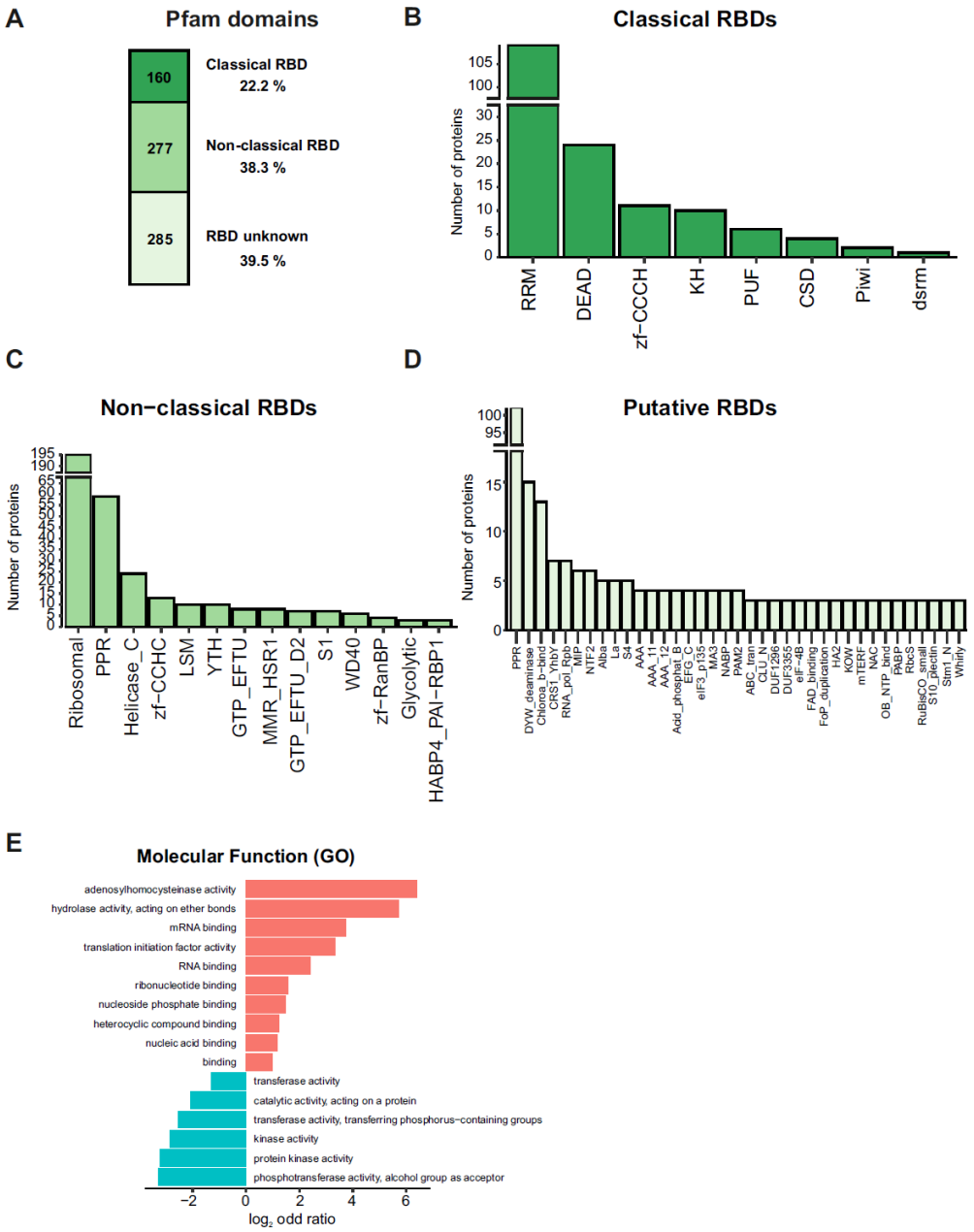


Fig.5 Domain architecture of the Arabidopsis leaf RBPome. A) Number of proteins harbouring classical, non-classical or no known RBDs in the leaf RBPome. (B) Number of proteins annotated as possessing classical RBDs, (C) non-classical RBDs or (D) putative RBDs. For non-classical and putative RBD, only RBDs with at least three counts are shown. E) GO analysis with the most significant enriched (red) molecular function GO terms of the proteins that do not contain known RBDs.

bona fide RBPs. For instance, we identify 5 out of 6 predicted Alba-containing proteins (Yuan et al., 2019), and 4 were also present in the RNA interactome of etiolated seedlings (Reichel et al., 2016). Moreover, recently two Alba-containing proteins have been shown to interact with RNA in plants (Gosai et al., 2015; Yuan et al., 2019). We identified 7 CRS1/YhbY (CRM)-containing

proteins and this domain is present in Archaea, Bacteria and plants and has been linked to RNA (Barkan et al., 2007). Strikingly we identified 14 different photosynthesis-related domains such as the domains Chlorophyll A-B binding, PsaA_PsaB, PsaD, PsaL and PsbH in our leaf RBPome (**Fig. 5D**). There are multiple lines of evidence that support the hypothesis that photosynthetic components can associate with RNA and these are discussed further in **Chapter 3**. Hence, these domains that do not have previous links to RNA biology but are enriched in RBPs could represent novel RBDs.

2.5. RBPs uniquely identified by ptRIC

Next we compared our high-confidence RBPome with those previously published for Arabidopsis (Reichel et al., 2016; Marondedze et al., 2016; Zhang et al., 2016b). Together, the previously established plant RNA interactomes identified around 700 RBPs from different Arabidopsis tissues. Remarkably, we identified 722 RBPs from Arabidopsis leaves using ptRIC, of which 413 RBPs (57.2%) are uniquely identified in our dataset (**Fig. 4C**). These RBPs may represent leaf- or developmental stage-specific RBPs. Alternatively, these RBPs may be present in our dataset simply because ptRIC allows deeper interrogation of the plant RBPome than any of the other published plant RIC protocols. Surprisingly, there are just 25 RBPs that are common between the four RBPomes. Although this number may seem low *a priori*, this could reflect the different nature of the starting materials and, more likely, the distinct stringency and reproducibility of the different RIC variants: cell cultures, mesophyll protoplasts, etiolated seedlings and leaves of mature plants. Moreover, the RNA interactome of cell cultures (Marondedze et al., 2016) shows the smallest overlap with the other datasets (**Fig. 4C**) and therefore decreases the number of commonly identified RBPs. This is further discussed below.

To get a deeper inside into the 413 RBPs uniquely identified using ptRIC (ptRIC-specific leaf RBPome), we performed gene set enrichment analysis using GO (**Fig. 6A, B**). Similarly to what

was observed for the entire high-confidence leaf RBPome, the most enriched GO terms in the ptRIC-specific leaf RBPome were related to RNA including 'mRNA binding', 'ribonucleoside binding' or 'RNA binding' (**Fig. 6A**). Other RNA-related functions were also statistically enriched such as 'structural constituent of ribosome'. Amongst the statistically underrepresented terms were 'phosphotransferase activity', 'protein kinase' activity and 'hydrolase activity' (**Fig. 6A**). This indicates that the proteins specifically identified using ptRIC are indeed *bona fide* RBPs and not contaminants.

We examined whether the large number of ptRIC-specific RBPs were due to a better capture of RBPs from any cellular compartment such as organelles. To test this, we constructed a superset of Arabidopsis RBPs comprising those identified in previous papers (Hentze et al., 2018) together with our identified leaf RBPs and performed gene set enrichment analyses assessing the GO cellular component. We identified no statistical enrichment or depletion of GO cellular component (data not shown), indicating that we do not differentially isolate RBPs from any cellular compartment, including organelles. However, overall we observe a large number of mitochondrial and chloroplastic RBPs, indicating that we obtain a good coverage of these subcellular organelles.

Importantly, 67% of the leaf ptRIC-specific RBPs were already linked to RNA biology (**Fig. 6C**), which highlights the excellent quality of our dataset. This contrasts with the set of 285 proteins exclusively identified by Marondedze and colleagues (Marondedze et al., 2016), of which only 18% are linked to RNA biology (**Fig. 6C**). Moreover, gene set enrichment analyses against the total proteome revealed that the Marondedze-specific 'RBPome' is enriched in GO terms largely unrelated to RNA biology such as 'aconitase hydratase', 'L-aspartate:2-oxoglutarate aminotransferase activity' or 'hydro-lyase activity' (data not shown). Gene set enrichment analyses of Marondedze-specific 'RBPome' against the superset of Arabidopsis RBPs revealed

depletion of GO terms related to RNA biology such as ‘RNA binding’, ‘nucleic acid binding’ and ‘structural constituent of ribosome’ (data not shown). This indicates that the Marondedze-specific proteins that we fail to identify by ptRIC are largely not related to RNA biology, so further analyses should be performed to confirm whether they are *bona fide* RBPs.

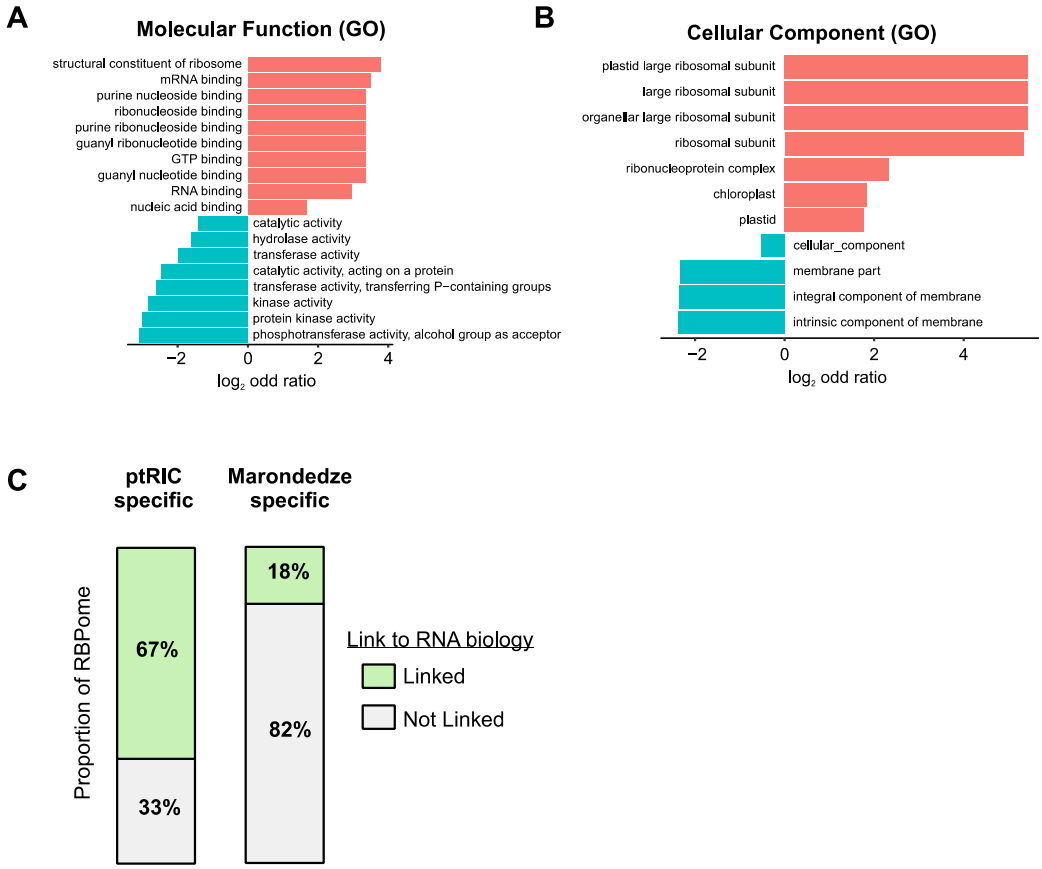


Fig.6. Insights into the ptRIC-specific leaf RBPs. GO analysis with ten of the most significant enriched (red) or underrepresented (blue) molecular function (A) or cellular compartment (B) GO terms of the ptRIC-specific leaf RBPome. C) Proportion of the ptRIC-specific leaf RBPome or the Marondedze-specific RBPome linked to RNA biology based on GO annotations.

We analysed the domain architecture of the ptRIC-specific leaf RBPome (**Fig. 7**) and identified that it yields a similar RBD composition to the complete leaf RBPome (**Fig. 5**). This indicates that ptRIC-specific proteins are not biased towards a particular type of RBD or molecular function.

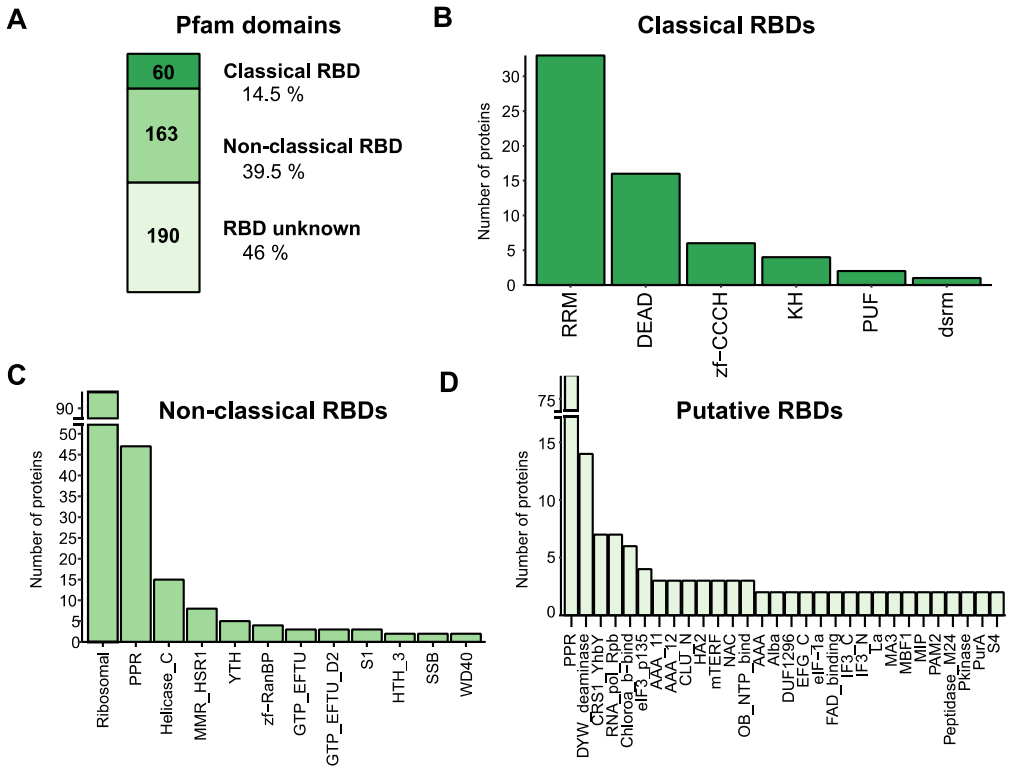


Fig.7 Domain architecture of the RBPs identified uniquely by ptRIC. A) Number of proteins harbouring classical, non-classical or no known RBDs in the ptRIC-specific leaf RBPome. (B) Number of proteins annotated as possessing classical RBDs, (C) non-classical RBDs or (D) putative RBDs in the ptRIC-specific leaf RBPome. For non-classical and putative RBD, only RBDs with at least three protein counts are shown.

3. DISCUSSION

3.1. Advantages and limitations of ptRIC

RIC has multiple advantages over other published techniques to isolate RBPs. For instance, RIC allows: i) identification of proteins directly associated with RNA and not proteins interacting with RNA though protein-protein interactions or non-covalent interactions since UV light does not promote protein-protein crosslinking; ii) identification of native and physiological protein-RNA interactions because UV crosslinking is applied to living cells; iii) discovery of unconventional RBPs since it is unbiased towards specific classical domains; iv) comparative studies to analyse the RBP dynamics because it is coupled to quantitative mass spectrometry (Castello et al., 2012, 2013b).

We have improved RIC to be applied efficiently to plant leaves. Using this modified protocol, referred to here as 'ptRIC', we have identified first comprehensive leaf RBPome to date, comprising 722 RBPs in Arabidopsis leaves (**Fig. 3**). This represents a significant advance, since deeper scrutiny of the RBPome allows identification of low abundance RBPs that would be missed by currently available protocols. Moreover, ptRIC can be applied now to any environmental, physiological or pathological condition, allowing identification of the RBPs that regulate RNA metabolism in any particular cellular state.

Although four recent works have attempted to determine the RBPome of *A. thaliana*, these studies used plant tissues that are, in principle, easy-to-crosslink (i.e. etiolated seedlings (Reichel et al., 2016), mesophyll protoplasts (Zhang et al., 2016b) or cell cultures (Marondedze et al., 2016, 2019)). Marondedze and colleagues carried out the first attempt to apply RIC to Arabidopsis leaves with very limited success, since only 27 proteins were enriched in an UV-dependent manner (Marondedze et al., 2016). We have successfully optimised ptRIC to leaves of mature plants, which is a more relevant tissue for many areas of plant biology research since it reflects natural tissue physiology. We believe that our protocol not only outperforms published protocols, but has broader applicability, as it can be adapted for other plant tissues such as roots or flowers. However, application to other tissues may need additional optimisation of certain steps such as the UV irradiation strategy.

RIC and ptRIC have some limitations that should be considered. RIC/ptRIC fail to identify: i) RBPs that only bind to non-polyadenylated RNAs because RIC/ptRIC relies on the oligo(dT) capture of poly(A) RNA-RBP complexes; ii) RBPs that are not expressed or active in the tissue studied and under the experimental conditions used; iii) RBPs that do not efficiently crosslink using UV light; iv) RBPs whose UV crosslinking is not facilitated by the RNA-protein geometry, as is the case for many double stranded RBPs (Hentze et al., 2018); iv) RBPs interacting with the phosphate-ribose

backbone. Moreover, the peptide abundance and sequence influence the mass spectrometry analyses.

3.2. ptRIC: a RIC protocol optimised for leaf tissue

We optimised several parameters to circumvent the many challenges that leaf tissue impose for the isolation of RBPs. One of the critical parameters optimized was the UV dosage. The duration and intensity of UV light determine the efficiency of crosslinking between RNAs and RBPs; however, excessive exposure leads to RNA degradation and therefore, reduction in the yield of RBP (Urdaneta et al., 2019). Leaves are particularly challenging to UV-crosslink due to their thickness and the presence of pigments able to absorb UV light. We determined that the optimal UV dosage for leaves was three UV irradiations of 0.15J/cm^2 on top of ice pads separated by 30 seconds pause in between treatments. Moreover, we irradiated both adaxial and abaxial sides of the leaves to increase the crosslinking efficiency. The optimal amounts of oligo(dT) magnetic beads was experimentally determined to be 250 µl of beads/sample. Although higher amounts of beads can increase the amount of pulled-down mRNA (associated with RBPs), this results in an increase in the background noise. Importantly, we also identified that the buffer to tissue ratio and an increased homogenization are crucial parameters to achieve high-quality RBPomes. Incomplete homogenization can cause co-purification of cellular debris or large chains of gDNA (associated with crosslinked DNA-binding proteins) along with the poly(A) RNA samples, thus resulting in high background noise. We determined that the most optimal homogenization method for leaf tissue consisted in grinding the tissue using mortar and pestle (in liquid N_2) followed by potter tube homogenization in lysis buffer and shearing of gDNA by passing lysates through a syringe with a narrow needle. To increase the stringency of the washes we included an additional wash with a 'harsh buffer' with high concentrations of ionic detergent (LiDS) and high salt (LiCL). This leads to the elimination of contaminant proteins that stick to the oligo(dT) beads or bind to the RNA via protein-protein interactions or non-covalent interactions.

3.3. RBPome of mature Arabidopsis leaves

To our knowledge using ptRIC we have unravelled the first complete Arabidopsis leaf RBPome to date, comprising 722 leaf RBPs (**Fig. 3**). Most of the identified leaf RBPs (74%) were linked to RNA biology (**Fig. 4D**), indicating that ptRIC is a suitable technique to uncover high-quality RBPomes. Notably, the rest of leaf RBPs (26%) had no predicted or known role in RNA biology and are novel RBPs. The proportion of potential novel RBPs is similar to that in the RNA interactome of etiolated seedlings and protoplasts (42%) (Reichel et al., 2016; Zhang et al., 2016b), yeast (42%) or humans (31%) (Beckmann et al., 2015). Stronger evidence of the novel role of proteins as RNA binders can be gained when comparing the candidate RBPs identified in different Arabidopsis RBPomes. For instance, catalase-3 has been identified in our high-confidence leaf RBPome and in the three Arabidopsis RBPomes (Reichel et al., 2016; Zhang et al., 2016b; Marondedze et al., 2016), providing convincing evidence that this enzyme has an RNA-related moonlighting function in plants. Moreover, Clerch and colleagues also hinted that catalase could bind RNA in cows (Clerch et al., 1996).

A large proportion of the high-confidence leaf RBPs possess known RBDs (60%; **Fig. 5**), a proportion similar to other Arabidopsis RBPomes (Reichel et al., 2016). The remaining 40% of the leaf RBPome does not possess domains related to RNA biology and represents an excellent dataset to discover novel RBDs. By comparing motifs enriched in RBPs but with unknown RNA-binding function, novel RBDs may be identified. Moreover, extending these analyses to multiple species will increase the support for the novel role of certain motifs in RNA binding. However, the ability of these putative novel RBDs to bind RNA should be experimentally tested to exclude the possibility that we identify certain motifs as RBDs only because they coexist with RBDs in the same protein. Traditionally the RNA-binding ability of domains has been tested in single proteins by deletion and mutagenesis of the domains coupled to techniques such as electrophoretic

mobility shift assays (EMSA) or crosslinking and immunoprecipitation assays (CLIP). Recently a technique termed RBDmap has been developed that allows identification of the RNA-binding sites of proteins in a system-wide scale (Castello et al., 2016, 2017). RBDmap is similar to RIC but uses an extra step of controlled proteolysis and a second oligo(dT) capture to allow system-wide identification of RNA-binding regions within proteins (Castello et al., 2016, 2017). The use of RBDmap would be instrumental to determine the scope of RBDs in plants and might reveal novel plant-specific RBDs.

Notably, when comparing the leaf RBPs isolated by ptRIC with those from previous RIC studies in Arabidopsis (Reichel et al., 2016; Zhang et al., 2016b; Marondedze et al., 2016) we identified that 413 are unique to our high-confidence leaf RBPome. This highlights the efficiency of ptRIC in isolating RBPs from leaf tissue. The ptRIC-specific RBPs may represent tissue- or developmental-specific RBPs, thus highlighting the dynamism and flexibility of the plant RBPome.

3.4. Conclusion

We have developed 'plant RNA-interactome capture' (ptRIC), which is a RIC variant optimised to efficiently isolate RBPs from plant leaves. Using ptRIC we have uncovered, to our knowledge, the first Arabidopsis leaf RBPome that includes 722 leaf RBPs. Interestingly, a large proportion of the RBPome had no previous links to RNA biology and may represent novel RBPs. Moreover, many identified RBPs had no known RBDs, which represents an excellent opportunity for novel RBD discovery. Because ptRIC allows a deep resolution of the plant RBPome, we envisage that ptRIC will be instrumental to the study of the dynamic modulation of the plant RBPome in response to different stimuli.

4. MATERIALS AND METHODS

4.1. Plant material

We used mature Arabidopsis plants (5-6 weeks old) of Col-0 ecotype for all our experiments. Plants were grown in soil at neutral day conditions (12 h light, 12 h dark) at 20 °C and light intensity of approximately 100 µmol/m^2/s.

4.2. Plant RNA interactome capture (ptRIC)

4.2.1. UV crosslinking

For UV crosslinking, leaves of mature Arabidopsis plants were excised and placed on a plastic sheet on top of ice pads to prevent sample overheating. The leaves were crosslinked three times (twice on the adaxial side of the leaves and once in the abaxial) with 150 mJ/cm^2 of UV light at 254 nm wavelength. We allowed 30 seconds pause in between irradiations. Crosslinking both sides of the leaves ensures that both adaxial and abaxial sides of the leave receive a UV dose, thus increasing crosslinking efficiency. For the non-crosslinked (NoCL) negative control, we placed the leaves on ice for approximately the same time the crosslinked (CL) samples were maintained on the ice pads during irradiation (∼3 minutes). After irradiation the leaves were immediately frozen in liquid nitrogen to preserve the molecular interactions and sample integrity. Both CL and NoCL samples were processed in parallel following the same protocol.

4.2.2. Buffer preparation

Solutions were prepared, filtered and stored at 4°C for up to 3 months. IGEPAL, Polyvinylpyrrolidone (PVP40), ß-mercaptoethanol (B-ME), DTT (Dithiothreitol), and the Protease and RNase inhibitors were immediately added before use (**Supplemental table 1, Appendix II**).

4.2.3. Cell lysis

Leaf tissue was ground to a fine powder in liquid nitrogen using a mortar and pestle. 1.2 g of tissue was mixed with 12 ml of lysis buffer in a 50 ml tube. The lysates were kept on ice to minimise RNA degradation. To further homogenise the lysates a Potter-Elvehjem homogenizer was used. Samples were homogenised for 1 minute, while kept on ice to avoid sample overheating. The lysates were cleared by centrifugation (4000 rpm, 10 min, 4 °C) and filtration of the supernatant through miracloth (Merk, cat. no. 475855). To shear the gDNA, lysates were passed through a narrow needle (27G) five times. The lysates were cleared again by centrifugation (4000 rpm, 10 min, 4 °C) and filtration of the supernatant through miracloth.

4.2.4. RNA interactome capture

Before starting RNA interactome capture, aliquots of 200-500 µl of the inputs (whole cell lysates / total proteomes) were taken and stored at – 80 °C. Oligo(dT) beads (250 ul/sample; NEB, cat. no. S1419S) were activated by washing them three times with lysis buffer, using a magnet to trap the magnetic beads followed by supernatant removal. Next, 250 µl of beads were added to each of the lysates and incubated at 4 °C for 1 h in a rotator (10 rpm) to allow hybridisation of the oligo(dT) to the poly(A) tail of the RNAs. The beads were captured on a magnet for ∼20-30 minutes at 4 °C, making sure the supernatant was cleared. The supernatant was collected and stored on ice for a second round of capture. The beads were washed three times with 1.5 ml of lysis buffer with an incubation of 5 min on ice - inverting the tube every 30-60 seconds - followed by magnet capture and supernatant removal. The stringency of the lysis buffer allows removal of contaminant proteins sticking to the oligo(dT) beads or bound to the RNAs non-covalently or via protein-protein interactions. The beads were washed with 1.5 ml of harsh buffer at room temperature (5 minutes). This washing step is performed at room temperature to increase the removal of contaminants and avoid precipitation of the LiDS. The harsh buffer contains higher concentrations of the ionic detergent LiDS (1%, w/v) and LiCL (2M) to further increase the

stringency of the capture. The beads were then washed two times with buffer I, two times with buffer II, one time with buffer III with detergent (0.02% IGEPAL) and one time with buffer III without detergent as this detergent interferes with downstream mass spectrometry analyses. To elute the RNA-protein complexes, beads were resuspended and incubated with 300 µl of elution buffer for 3 minutes at 55 °C. The beads were pelleted using a magnet and the supernatant containing the RNA-protein complexes was transferred to a new tube. Typically, using our settings we obtain RNA concentrations of about 400 - 1000 ng/µl.

For the second round of capture, the beads were washed three times with lysis buffer and added to the supernatants recovered from the first capture that were kept at 4 °C. Beads were only reused for the same condition (i.e. the beads of treatment 1 were reused for the second round of capture of samples of treatment 1).

4.2.5. RNA quantification and normalization

For each of the samples, the two eluates from different rounds of capture were pooled and quantified using a NanoDrop spectrophotometer (Thermo Fisher scientific). Elution buffer was used to adjust the volume of the samples so that the same amount of RNA was present in each sample. An aliquot of 50 µl of the pooled normalised eluates can be separated for RT-qPCR or RNAseq analyses.

4.3. Downstream applications of isolated RNA-protein complexes

After isolation of the RNA-protein complexes, analyses can be performed on the proteins or the RNAs. For protein analyses, we treated each of the eluates with 4 µl of RNase A (Sigma-Aldrich, cat. no. R4642) and T1 (Sigma-Aldrich, cat. no. R1003) mix (RNase A and RNase T1 mixed at equal proportions and diluted 1/100) for 1 h at 37 °C followed by incubation for 15 min at 50 °C. Samples were then analysed as following: i) western blotting with specific antibodies; ii) silver

staining for total protein analysis; and iii) quantitative mass spectrometry to identify and quantify the isolated proteins. Out of the 600 µl of the eluates, 50 µl were used for western blotting and silver staining analyses, and remaining material (550 µl) was used for mass spectrometry (stored at – 80 °C).

4.3.1. Protein concentration and western blot or silver staining

For western blot or silver staining analyses, proteins were concentrated using an Amicon centrifugal filter of 3 KDa cut-off (Merck, cat. no. UFC500324) by centrifugation (15000 rpm, 4 °C, 1-2h). The volume recovered from each filter (typically between 20 – 40 µl) was measured and normalised using elution buffer to proceed with the same volume in each sample.

The inputs (whole cell lysates / total proteome) and the eluates (RBPs) were mixed with protein loading buffer (1/4, v/v) and incubated for 4 min at 95 °C. Input samples were diluted 1/20 with water and the marker was diluted 1/15 to avoid signal saturation in the silver staining. Proteins were separated in a 10-12% acrylamide gel followed by western blot or silver staining. Western blots were performed following standard procedures (Bass et al., 2017) using the following antibodies: anti-RH3 (Agrisera, AS132714), anti-HSC70 (Agrisera, AS08371), anti-H3 (abcam, ab1791) and anti-phospho-p44/42 MAPK (Cell signalling technology, 4370). Silver staining was performed using Silver quest (Invitrogen, cat. no. LC6070) following the manufacturer's instructions.

4.3.2. Sample preparation for MS

Samples were prepared for mass spectrometry using the standard FASP (filter aided sample preparation) with minor modifications (Wiśniewski et al., 2009). Briefly, proteins were mixed with urea buffer (8 M urea (Sigma-aldrich cat. no. U1250), 100 mM TEAB (triethylammonium bicarbonate buffer; Sigma-aldrich cat. no. T7408)), loaded in a 10 KDa cut-off Amicon filter unit (Millipore) and centrifuged at 15,000g for 30 min at 20 °C. Proteins were washed six times with

200 µl urea buffer to remove any detergents from the sample. Protein reduction was performed with 200 µl urea buffer supplemented with 10mM TCEP (Tris-(2-Carboxyethyl)phosphine)) for 30 min at room temperature. Protein alkylation was performed by adding CA (Chloracetamide) to a final concentration of 50 mM for 30 min, at room temperature and in the dark. The samples were washed three times with 200 µl a buffer containing 6 M urea, 50 mM TEAB and centrifuged (15,000g, 30 min, 20 °C). Proteins were digested inside the filter unit by adding 1 µg of Lysyl endopeptidase (Wako, cat. 125-05061) in 99 µl of a buffer containing 1 M urea and 50 mM TEAB for 4 hours at 37 °C. Polypeptides were further digested with 1 µg trypsin (Promega, cat. no. V5111) in 99 µl of a buffer containing 50 mM TEAB overnight at 37°C. The digested peptides were harvested by centrifugation (15,000g, 25 min, 20 °C). To increase the peptide recovery, the filters were washed once with 0.1 % TFA (Trifluoroacetic acid; Thermo Fisher Scientific, cat. no. 85183), followed by a wash with 50 % acetonitrile (ACN; Honeywell, cat. no. 14261) and 0.1 % TFA. Peptides were harvested by centrifugation (15,000g, 25 min, 20 °C) and eluates were dried using a speed vac. The acidified tryptic digests were desalted on home-made 2 disc C18 StageTips as described (Rappsilber et al., 2007). After elution from the StageTips, samples were dried using a vacuum concentrator (Eppendorf) and the peptides were taken up in 10 µL 0.1 % formic acid solution.

4.3.3. Mass spectrometry analysis

Experiments were performed on an Orbitrap Elite instrument (Thermo Fisher Scientific) (Michalski et al., 2012) that was coupled to an EASY-nLC 1000 liquid chromatography (LC) system (Thermo Fisher Scientific). The LC was operated in the one-column mode. The analytical column was a fused silica capillary (75 µm × 35 or 50 cm) with an integrated PicoFrit emitter (New Objective) packed in-house with Reprosil-Pur 120 C18-AQ 1.9 µm resin (Dr. Maisch). The analytical column was encased by a column oven (Sonation) and attached to a nanospray flex ion source (Thermo Fisher Scientific). The column oven temperature was adjusted to 45 °C during

data acquisition. The LC was equipped with two mobile phases: solvent A (0.1% formic acid (FA) in water) and solvent B (0.1% FA in ACN). All solvents were of UPLC grade (Sigma-Aldrich). Peptides were directly loaded onto the analytical column with a maximum flow rate that would not exceed the set pressure limit of 980 bar (usually around 0.6 – 1.0 µL/min). Peptides were subsequently separated on the analytical column by running a 140 min gradient of solvent A and solvent B (start with 7% B; gradient 7% to 35% B for 120 min; gradient 35% to 100% B for 10 min and 100% B for 10 min) at a flow rate of 300 nl/min. The mass spectrometer was operated using Xcalibur software (version 2.2 SP1.48). The mass spectrometer was set in the positive ion mode. Precursor ion scanning was performed in the Orbitrap analyser (FTMS; Fourier Transform Mass Spectrometry) in the scan range of m/z 300-1800 and at a resolution of 60000 with the internal lock mass option turned on (lock mass was 445.120025 m/z, polysiloxane) (Olsen et al., 2005). Product ion spectra were recorded in a data dependent fashion in the ion trap (ITMS) in a variable scan range and at a rapid scan rate. The ionization potential (spray voltage) was set to 1.8 kV. Peptides were analysed using a repeating cycle consisting of a full precursor ion scan (3.0×10^6 ions or 50 ms) followed by 15 product ion scans (1.0×10^4 ions or 50 ms) where peptides are isolated based on their intensity in the full survey scan (threshold of 500 counts) for tandem mass spectrum (MS2) generation that permits peptide sequencing and identification. Collision induced dissociation (CID) energy was set to 35% for the generation of MS2 spectra. During MS2 data acquisition dynamic ion exclusion was set to 120 seconds with a maximum list of excluded ions consisting of 500 members and a repeat count of one. Ion injection time prediction, preview mode for the FTMS, monoisotopic precursor selection and charge state screening were enabled. Only charge states higher than 1 were considered for fragmentation.

RAW spectra were submitted to an Andromeda (Cox et al., 2011) search in MaxQuant (1.5.3.30) using the default settings (Cox and Mann, 2008). Label-free quantification and match-between-runs was activated (Cox et al., 2014). We downloaded the reference *A. thaliana* database from

NCBI (Thale cress, ARA_GCF_000001735.4_TAIR10.1_protein.fasta, downloaded 20/02/2019). All searches included a contaminants database search (as implemented in MaxQuant). The contaminants database contains known MS contaminants and was included to estimate the level of contamination. Andromeda searches allowed oxidation of methionine residues (16 Da) and acetylation of the protein N-terminus (42 Da) as dynamic modifications and the static modification of cysteine (57 Da, alkylation with iodoacetamide). Enzyme specificity was set to "Trypsin/P" with two missed cleavages allowed. The instrument type in Andromeda searches was set to Orbitrap and the precursor mass tolerance was set to ±20 ppm (first search) and ±4.5 ppm (main search). The MS/MS match tolerance was set to ±0.5 Da. The peptide spectrum match FDR and the protein FDR were set to 0.01 (based on target-decoy approach). Minimum peptide length was 7 amino acids. For protein quantification unique and razor peptides were allowed. Modified peptides were allowed for quantification. The minimum score for modified peptides was 40. Label-free protein quantification was switched on, and unique and razor peptides were considered for quantification with a minimum ratio count of 2. Retention times were recalibrated based on the built-in nonlinear time-rescaling algorithm. MS/MS identifications were transferred between LC-MS/MS runs with the "match between runs" option in which the maximal match time window was set to 0.7 min and the alignment time window set to 20 min. The quantification is based on the "value at maximum" of the extracted ion current. At least two quantitation events were required for a quantifiable protein.

4.4. Data analysis

4.4.1. Statistical analysis of the RBPome

All statistical analyses were conducted in R (R Core Team, 2014). The contaminant protein groups, including reversed sequences and protein groups identified by site, were filtered from our dataset. Raw intensities were used when comparing NoCL and CL because LFQ intensities

are normalised and thus are not suitable for experiments in which one condition is rich in proteins (CL) while the other is devoid in proteins (NoCL) (Cox et al., 2014). The proteins from the CL samples were only considered for analysis if quantified in at least 3 out of 4 replicates. Some of the identified peptides were not unique and could not be assigned to a single protein. For these, we conducted analyses using the protein groups, although we cannot confidently assign the peptide to one unique protein.

The identified proteins were separated into two groups and analysed using different methods as following. The proteins with values in the CL samples, which were not detected in the NoCL sample (negative control) were analysed using a modified version of the semi-quantitative approach described by Sysoev and colleagues (Sysoev et al., 2016). We classified as RBPs proteins quantified in at least 3 out of 4 replicates in the CL replicates but not detected in the NoCL replicates. For proteins that had intensity values in both CL and NoCL, we used a quantitative method. Briefly, the missing values were imputed using the impute.minDet function (R Core Team, 2014) and the ratios of CL/NoCL were calculated. Statistical analysis of CL/NoCL enrichment samples was performed using a moderate t-test implemented in the R/Bioconductor package limma (Smyth, 2004). The resulting p-values were corrected for multiple testing using Benjamini-Hochberg (Benjamini and Hochberg, 1995) to calculate the false discovery rate (FDR; p-adjusted). Proteins with a $\log_2$ fold change [CL/NoCL] ≥ 1.5 and an adjusted p-value (FDR) ≤ 0.05 were defined as 'RBPs', whereas the remaining proteins ('non-enriched in +UV/-UV') were discarded to avoid high incidence of false positives. During the optimisation process of the ptRIC, we analysed by mass spectrometry two pilot experiments (MS1 and MS2). For each of these two early experiments we analysed one CL and one NoCL sample. Therefore, in these two pilot experiments we classified a protein as RBPs if $\log_2$FC [CL/NoCL] ≥ 2.

The proteins defined as RBPs by either semi-quantitative or quantitative method were pooled to define a high confidence Arabidopsis leaf RBPome.

4.4.2. Bioinformatic analysis of the RBPome

The reference Arabidopsis proteome was downloaded (July 2019) from Uniprot (www.uniprot.org) and was used for GO and Pfam annotations. A superset of Arabidopsis RBPs was generated by adding the leaf RBPs identified by ptRIC to the Arabidopsis RBPs identified in previous studies (Reichel et al., 2016; Zhang et al., 2016b; Marondedze et al., 2016) and standardised by Hentze and colleagues (Hentze et al., 2018).

Gene set enrichment analyses were performed by comparing frequencies of gene ontology (GO) terms in the leaf RBPome with the frequency of the same GO terms in either the reference Arabidopsis proteome or the superset of Arabidopsis RBPs. Statistical testing was performed applying Fisher's exact test and Bonferroni corrected p-values were used to account for multiple testing. We allowed for propagation of GO terms to account for the intrinsic hierarchy of GO terms. All statistical testing was performed using GOATOOLS (Klopfenstein et al., 2018) with the basic GO file version 1.2 (released on 01.07.2019). Significantly enriched GO terms were divided into the different GO categories 'cellular component and 'molecular function'. For each category, $\log_2$ odds ratios of some of the most significantly enriched GO terms were calculated. The occurrence of significantly enriched GO terms that were absent in the RBPome was artificially set to the lowest biologically sensible value (1) to enable calculation of $\log_2$ odds ratios.

The Pfam domains were categorised into classical RBDs, non-classical RBDs and domains not known to be linked to RNA-binding (putative RBDs) by manual annotation based on data from previous studies (Castello et al., 2012, 2016) as previously described (Castello et al., 2012). Two analyses were performed. First, hierarchical analysis of Pfam distribution whereby proteins were

analysed to determine whether they contained a Pfam code that corresponded to a classical, non-classical or putative RBD (RBD unknown). Classical was dominant over non-classical and non-classical was dominant over unknown. If one protein was associated with more than one RBD, the most dominant category was selected for the protein. Second, the distribution of Pfam names in each of the three groups (classical, non-classical and putative) was analysed. The following Pfam names were merged into groups. Within classical RBDs, all Pfam names beginning with 'KH' were merged. Within non-classical RBDs, all Pfam names beginning with 'Ribosomal' were merged. Within putative RBDs, all Pfam names beginning with 'RNA pol Rbd' and 'PPR', 'FAD binding' were merged.

The links to RNA biology of identified RBPs were determined as described by Beckmann and colleagues (Beckmann et al., 2015). To analyse the overlap between the different Arabidopsis RICs, the RBPomes were obtained from the supplemental dataset of Hentze and colleagues (Hentze et al., 2018). Because Hentze and colleagues classified a given protein as RBP when FDR < 0.01 (without considering the log2FC[CL/NoCL]), we applied the same criteria to our leaf RBPome to be able to compare it with available datasets.

Most of the graphs were generated using the ggplot2 package within R (Wickham, 2009). The Venn diagram was generated using the VennDiagram package within R (Chen and Boutros, 2011) and the heat map was generated using the ggdendro and grid packages within R (R Core Team, 2014).

ACKNOWLEDGEMENTS

We want to thank all the members of van der Hoorn, Castello and Preston labs for their fruitful discussions and inputs on the manuscript. We also want to thank Nattapong Sanguankiattichai and Honglin Chen for their advice and help in the bioinformatics pipeline. We thank Urszula

Pyzio, Sarah Rodgers and Caroline O'Brien for excellent technical support. We acknowledge Marlene Reichel and Anthony Millar for advice on the original plant-adapted RIC protocol.

FUNDING

Marcel Bach-Pages is supported by Biotechnology and Biological Sciences Research Council (BBSRC, grant BB/M011224/1) and by the Lorna Casselton Memorial Scholarship at St. Cross College, Oxford. Alfredo Castello is supported by MRC Career Development Award MR/L019434/1 and MRC grant MR/R021562/1.

AUTHOR CONTRIBUTIONS

MBP designed the experiments with feedback from RvdH, AC and GMP. MBP performed all the experiments except for the MS run, which was performed by SM and FK. FH and JK assisted with the bioinformatics analyses. MBP analysed all the data and wrote the manuscript with feedback from GMP and AC.

SUPPLEMENTAL DATA

APPENDIX II – Supplemental data Chapter 2

Supplemental table 1. Buffers used in ptRIC

CHAPTER 3

Identification of the RBPome of the Plant Kingdom

Identification of the RBPome of the Plant Kingdom

Marcel Bach-Pages[1], Felix Homma[1], Jiorgos Kourelis[1], Farnusch Kaschani[2], Renier A.L. van der Hoorn[1], Alfredo Castello[3]*, Gail M. Preston[1]*

[1]Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, UK

[2]Fakultät für Biologie, *Universität Duisburg-Essen, Essen, Germany*

[3]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK

* Co-corresponding authors

Key words

RNA-binding proteins // RBPs // RNA-binding proteome // RBPome // ptRIC

ABSTRACT

RNA-binding proteins (RBPs) are critical for normal cellular homeostasis and to adapt to environmental changes. Moreover, many RBPs are unique to plants and play roles in plant-specific processes. Therefore, identifying the composition of the RNA-binding proteome (RBPome) of plants is critical for our understanding of plant physiology. Recently, RNA-interactome capture (RIC) has allowed the systematic identification of RBPs in multiple organisms, from yeast to humans. However, only the model *Arabidopsis thaliana* has been studied by RIC and agriculturally and economically important plant species remain unstudied. To determine the RBPome of different plant species, we have applied an improved and adapted version of RIC for plants (plant RNA interactome capture; ptRIC) to seven plant species from bryophytes to angiosperms. Our datasets are enriched in *bona fide* RBPs, display all the molecular features expected from high quality RBPomes and contain hundreds of putative novel

plant RBPs. Moreover, our analyses revealed hundreds of protein domains that are enriched in RBPs but lack previous links to RNA biology, thus representing potential novel RNA-binding domains (RBDs). By comparing the RBPome of the different species, we have identified 304 conserved RBPs that represent the 'core plant RBPome', many of which were previously unknown to bind RNA. Our study provides an unprecedented resource that will be of great use to the plant biology community.

1. INTRODUCTION

1.1. The study of RNA-binding proteins (RBPs)

The RNA-binding proteome (RBPome) has been estimated to comprise about 7.5 % of the total cellular proteome in human cells (Gerstberger et al., 2014; Järvelin et al., 2016), highlighting the importance of RNA-binding proteins (RBPs) for cellular functioning. In the recent years, prediction of RBPs has been mostly done *in silico* based on the identification of domains with similarity to known RNA-binding domains (RBDs). This has contributed to identification of hundreds of RBPs in different species (Anantharaman et al., 2002). However, these predictions are based on domain homology, hence, are unable to identify RBPs that do not possess known RBDs. Moreover, some studies have tried to experimentally identify RBPs in a comprehensive way by using fluorescently labelled RNA probes and protein arrays (Scherrer et al., 2010; Tsvetanova et al., 2010). These studies have allowed identification of hundreds of RBPs interacting with RNA *in vitro.* However, these studies also yielded non-physiological RNA binders that do not bind RNA *in vivo*.

In 2012 Castello and colleagues and Baltz and colleagues described a technique termed RNA interactome capture (RIC) that enables systematic identification of proteins actively bound to mRNAs *in vivo* (Castello et al., 2012; Baltz et al., 2012). To date RIC been applied to multiple organisms including the parasites *Trypanosoma brucei* (Lueong et al., 2016), *Leishmania*

donovani (Nandan et al., 2017; De Pablos et al., 2019), and *Plasmodium falciparium* (Bunnik et al., 2016), but also other organisms including yeast (Beckmann, 2017; Beckmann et al., 2015; Mitchell et al., 2013; Matia-González et al., 2015), worm (Matia-González et al., 2015), fly (Sysoev et al., 2016; Wessels et al., 2016), zebrafish (Despic et al., 2017), mouse (Kwon et al., 2013; Liao et al., 2016; Liepelt et al., 2016; Boucas et al., 2015) and different human cell lines (Baltz et al., 2012; Castello et al., 2012; Kramer et al., 2014; Beckmann et al., 2015; Castello et al., 2016; Perez-Perri et al., 2018; Conrad et al., 2016). Therefore, RIC has uncovered the RBPomes of multiple organisms (Hentze et al., 2018). Recent studies have provided the first experimental evidence that RNA interactome capture can be applied to plants, enabling the determination for the first time of the Arabidopsis RBPome (Reichel et al., 2016; Zhang et al., 2016b; Marondedze et al., 2016). However, the RBPome of the rest of the plant species remain experimentally unexplored.

1.2. Conservation of RNA-binding proteins

Cells rely on essential processes for cellular RNA homeostasis, and some of these pathways are highly conserved throughout evolution. For instance, 230 RBP orthogroups are common between humans and yeast (Beckmann et al., 2015), whereas 148 RBPs are conserved amongst humans, mice and flies (Sysoev et al., 2016). These conserved RBPs are enriched in core machinery elements of common eukaryotic processes such as splicing or translation (Sysoev et al., 2016). Therefore, it is logical to expect that RBPs involved in essential steps of post-transcriptional regulation such as RNA helicases or ribosomal proteins will be conserved amongst plant species and even eukaryotes (Lorković and Barta, 2002). However, a remarkable proportion of the conserved RBPs have no links to RNA biology or known RBDs and represent non-canonical RBPs (Beckmann et al., 2015; Matia-González et al., 2015). These non-canonical RBPs include different metabolic enzymes such as phosphoglycerate kinase or proteins of the proteasome, some of which have already been validated to bind RNA (Beckmann et al., 2015;

Castello et al., 2016; Matia-González et al., 2015). Other conserved non-canonical RBPs are involved in a wide range of processes such as oxidoreductases or ATP-binding and represent proteins and metabolic enzymes that moonlight as RNA binders (Beckmann et al., 2015).

On the other hand, proteomes have changed dramatically across evolution due to the changes in complexity and new functionality of proteins (Schad et al., 2011; Baer and Millar, 2016). Hence, one can anticipate that some of these changes can also occur at the RBPome level. Some authors have proposed that due to their sessile nature, plants require a more precise post-transcriptional regulation to adapt to the everchanging environment (Prall et al., 2019). Plants would have achieved that by expansion of RBP families leading to functional redundancies but also to subfunctionalization (Prall et al., 2019). Perhaps the most illustrative example is the pentatricopeptide repeat (PPR) domain containing family, which has expanded dramatically in plants when compared to humans (more than 50-fold) (Schmitz-Linneweber and Small, 2008; Silverman et al., 2013). Moreover, many plant RBPs do not possess orthologs in other kingdoms and might be plant-specific (Reichel et al., 2016; Prall et al., 2019). Two examples are SR45, a serine/arginine-rich protein involved in alternative splicing during plant development and stress and EIN2, an ethylene signalling component that has been recently described to bind RNA (McCarty and Chory, 2000; Li et al., 2015; Merchante et al., 2015; Reichel et al., 2016).

1.3. Rationale and aims of the study

To date only the RBPome of *Arabidopsis thaliana* has been experimentally determined and the RBPomes of economically and agriculturally relevant plant species remain undiscovered. We have previously developed 'ptRIC' (**Chapter 2**), an improved version of RIC that allows identification of RBPs from leaves of mature plants. We envisioned that ptRIC could be used to determine the RBPome of different plant species, thus offering an unprecedented opportunity to uncover the conserved and divergent RBPs across plant species.

2. RESULTS

2.1. ptRIC can be applied to multiple plant species

We aimed to apply ptRIC to different plant species to determine their RBPome (**Fig. 1A**). To do so we selected a panel of 10 species (**Fig. 1B**) using the following criteria: i) they should represent different groups within the plant phylogenetic tree; ii) their genomes and proteomes must be characterised; iii) they should either possess an agronomical interest or/and be broadly used model plants.

Thallus (for *Marchantia polymorpha*), protonemal tissue (for *Physcomitrella patens*) or leaves (for the rest of plant species) were irradiated with UV light at 254 nm to promote covalent bonds between RNAs and proteins at zero distance (**Fig. 1A**). After cell lysis, the polyadenylated RNAs (together with the RBPs) were isolated using oligo d(T) magnetic beads under stringent conditions to remove non-specific binders (**Fig. 1A**). The RBPs co-purified with RNA were released by RNase treatment and analysed by gel electrophoresis followed by protein analysis (**Fig. 1B)**. Silver staining of eluates for all species analysed revealed a specific protein pattern similar to that observed in other RIC studies (**Fig. 1B**). In absence of UV light almost no protein was detected, reflecting the high stringency of the approach. Importantly, the banding pattern of the eluates differed from the inputs (whole cell lysate), which indicates the isolation of a specific subset of proteins, likely RBPs (**Fig. 1B**). However, protein recovery was unequal across the different species (**Fig. 1B**). For example, *Physcomitrella patens* and *Nicotiana benthamiana* yielded more proteins than *Marchantia polymorpha* and *Zea mays*.

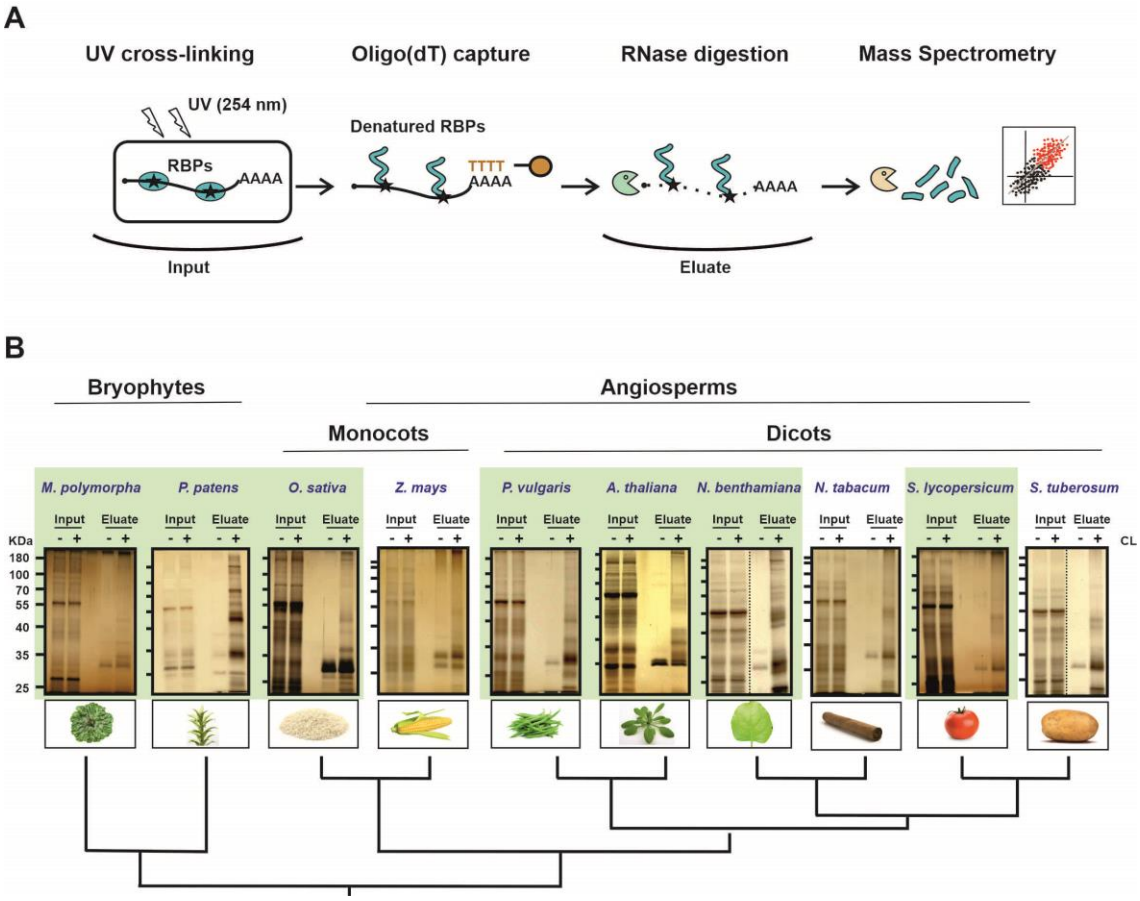


Fig.1. Application of ptRIC to different species across the plant kingdom.
A) **Schematic representation of ptRIC.** Leaves/thallus/protonemal tissue are irradiated with UV light (254 nm) to promote crosslinking between RNAs and proteins that are in close proximity. Cells are lysed and the mRNAs pulled down using oligo(dT) magnetic beads. After stringent washes, the RNA-protein complexes are recovered and the RBPs released by RNase digestion. The proteins are quantitatively analysed by mass spectrometry after trypsin digestion. B) **Silver staining analyses of the inputs** (total proteome) and **eluates** (RBPs) from different plant species. '+' and '-' refer to '+UV' and '-UV' crosslinking, respectively.

This could be due to several reasons: i) differential UV crosslinking efficiency due to distinct characteristics of the leaf/thallus/protonemal tissue such as thickness, amounts of pigments, morphology, etc; or ii) differential lysis or efficiency of isolation of RNA-RBP complexes due to particular features of each tissue. For instance, *Physcomitrella patens* may have the highest efficiency of RBPs isolation because the protonemal tissue is composed of a cell monolayer. Together, these results suggest that ptRIC can be applied to study the RBPome of these plant species.

2.2. The high-confidence RBPomes of different plant species

We further selected seven plant species to determine their RBPome by proteomics (**Fig. 1B,** species marked in green). We selected the two bryophytes *Marchantia polymorpha* (liverwort) and *Physcomitrella patens* (moss) as two model species at the base of the plant phylogenetic tree. We selected one monocot species, *Oryza sativa* (rice), since it is a major staple crop worldwide and is an important model species with amenable genetic manipulation. We also selected *Phaseolus vulgaris* (bean) as an important crop and model for leguminous plants, and *Arabidopsis thaliana* (Thale cress) since it is a broadly used plant model. Finally, we selected two Solanaceae species, *Nicotiana benthamiana* as a model plant used for transient expression studies, and *Solanum lycopersicum* (tomato) as an economically important crop and model species. We used the same methods as previously described (**Chapter 2**) to define the high-confidence RBPomes for each of the species (**Fig. 2**). This allowed us to generate a phylogenetically diverse compendium of high-confidence plant leaf/protonema/thallus RBPomes (herein after leaf RBPome).

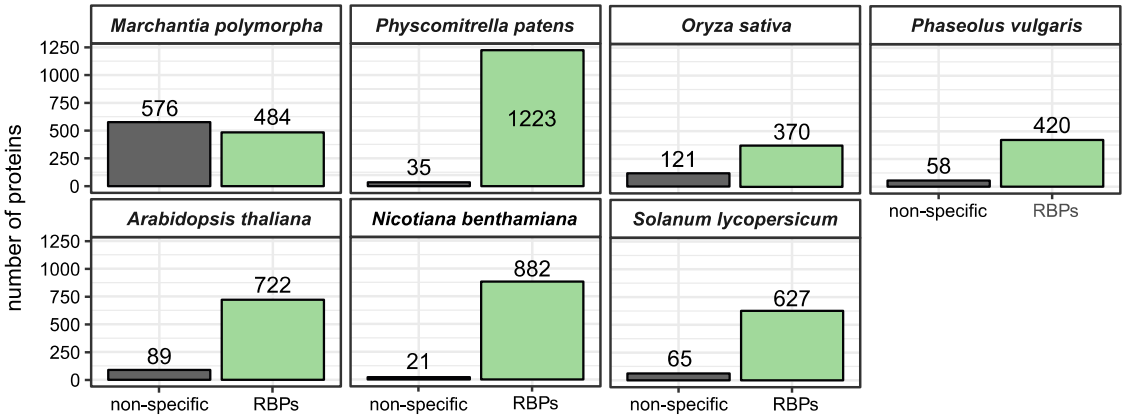


Fig.2 Determination of the high-confidence leaf RBPome for seven plant species. Proteins with missing values in all the NoCL replicates or log_2FC [CL/NoCL] ≥ 1.5 and FDR ≤ 0.05 are classified as 'RBPs' (in green), otherwise are classified as 'non-enriched in +UV/-UV' (in grey).

Jointly, these experiments revealed a total of 4727 RBPs across all species (**Supplemental digital table 1**). The species with the highest number of identified RBPs was *P. patens* (1223 RBPs),

followed by *N. benthamiana* (882 RBPs), *A. thaliana* (722 RBPs), *S. lycopersicum* (627 RBPs), *M. polymorpha* (484 RBPs), *P. vulgaris* (420 RBPs) and *O. sativa* (370 RBPs) (**Fig. 2**). Both MS and silver staining analyses (**Fig. 1B**) revealed different efficiency of RBP recovery across the different species. The number of proteins non-enriched in +UV/-UV varied across the different species, with *M. polymorpha* having the highest incidence (**Fig. 2**). To our knowledge only the RBPome of Arabidopsis has been previously unravelled by RIC (**Chapter 2** and **4**; Reichel et al., 2016; Marondedze et al., 2016; Zhang et al., 2016; Marondedze et al., 2019), so we cannot compare the RBPomes of the remaining six species with any established reference RBPome. In **Chapter 2** we examine the Arabidopsis RBPome obtained by ptRIC and compare it to those obtained using the other published RIC protocols.

2.3. Insights into the high-confidence plant RBPomes

We examined the high-confidence leaf RBPomes of each of the plant species by performing a gene set enrichment analysis using gene ontology (GO). For all seven species, the most enriched GO terms were related to RNA, including 'RNA binding', 'rRNA binding' and 'nucleic acid binding' (**Fig. 3**). Other RNA-related functions were also statistically enriched such as 'structural constituent of ribosome' and 'translation'-related GO terms (**Fig. 3**). Interestingly, in *M. polymorpha* the GO term 'glucose-6-phosphate isomerase' was statistically enriched. Likewise, in *O. sativa* and *P. vulgaris* the GO terms 'fructose-bisphosphate aldolase' and 'carbonate dehydratase activity' were statistically enriched, respectively (**Fig. 3**). It is known that many metabolic enzymes (especially within glucose metabolism) can bind RNA (Castello et al., 2015; Beckmann et al., 2015; Marondedze et al., 2016; Chang et al., 2013) and this feature is conserved from yeast to human (Beckmann et al., 2015), thus it is not unexpected that glucose metabolism is present in the plant RBPomes. In fact, many metabolic enzymes have been identified in the

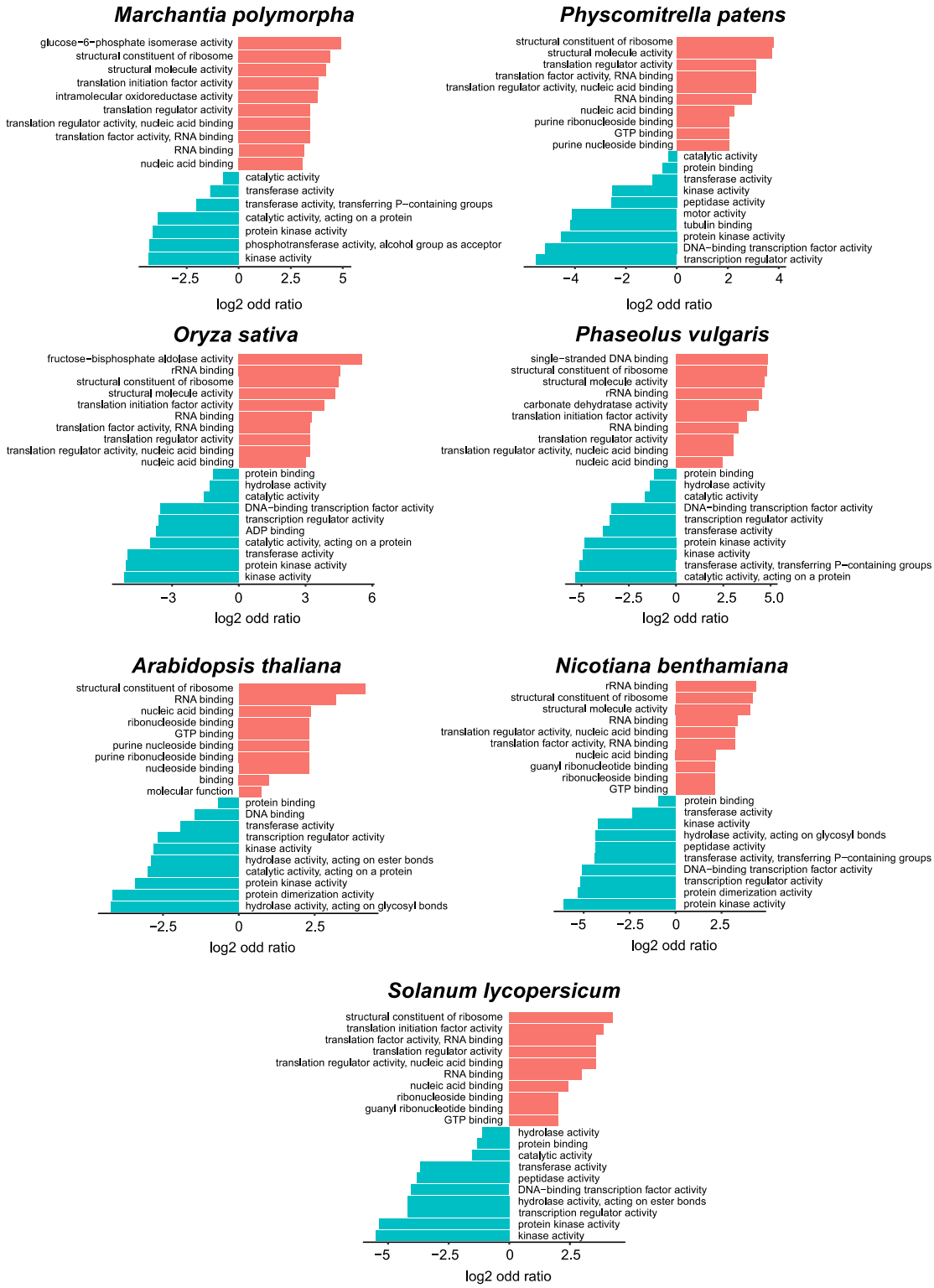


Fig. 3 Gene ontology analyses reveal enrichment of RNA-related functions. Ten of the most significantly enriched (red) or underrepresented (blue) molecular functions identified by GO analysis of the RBPomes.

RBPome of Arabidopsis cell cultures (Marondedze et al., 2016). Amongst the statistically underrepresented terms were 'protein kinase', 'DNA-binding transcription factor', 'peptidase' and 'phosphotransferase activity' (**Fig. 3**). The fact that 'DNA-binding transcription factor' is underrepresented indicates that our eluates are mostly free of DNA contamination. These results indicate that ptRIC enriches for *bona fide* RBPs. Hence, ptRIC is proven here to be an effective technique to isolate high-confidence leaf RBPomes in a wide variety of plant species.

To benchmark our inter-specie plant RBPome, we investigated the domains present in our datasets using Interpro scan annotations. *A priori*, we predict a large proportion of known RBDs (classical and nonclassical) if the plant RBPomes are selectively enriched in RBPs. Accordingly, about 50-75% of the RBPs identified harboured known RBD (**Fig. 4**). These numbers are in agreement with previous studies of RBPomes in other species. For example, 53% of the HeLa RBPome harbour known RBDs (Castello et al., 2012). Moreover, the incidence of classical and non-classical RBDs or of proteins lacking known RNA-binding signatures differed slightly between species, possibly reflecting the different depth of the different RBPomes. The plant species with the highest proportion of known RBDs (classical and nonclassical) was *O. sativa*, whereas the species with the highest proportion of proteins with no known RBD was *P. patens* (**Fig. 4)**. The fact that deepest interrogation of the RBPome, yields larger proportion of RBPs with unknown RBDs (*P. patens*) has also been described in other RNA interactomes (Hentze et al., 2018).

As expected, the most highly represented classical RBDs were RNA recognition motif (RRM), DEAD-box, zinc finger (zf)-CCCH and K-homology domain (KH) (**Supplemental table 1**). These domains are biochemically and structurally well-characterised and are the most prominent RBDs in plants and mammals (Castello et al., 2012; Reichel et al., 2016). Of the non-classical RBDs, the most abundant domains were ribosomal domains followed by pentatricopeptide repeat (PPR)

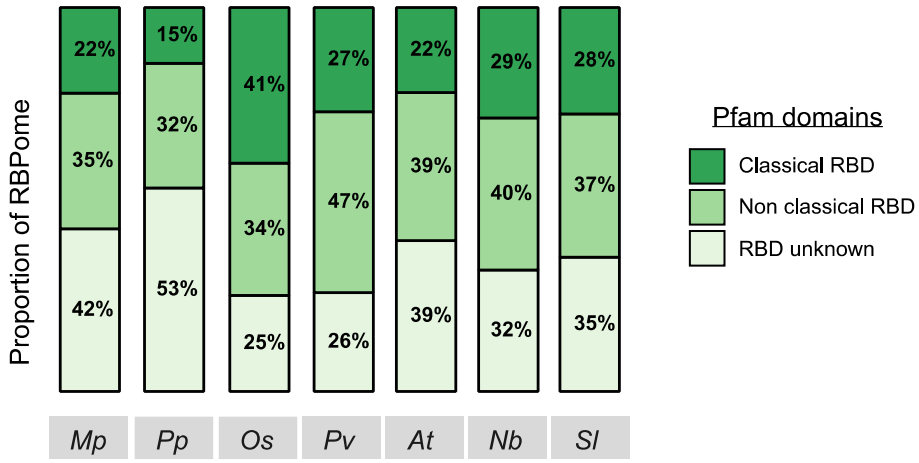


Fig. 4 Insights into the domain architecture of the plant RBPomes. Proportion of each RBPome harbouring classical, non-classical or no known RBDs. Abbreviations: Mp, *Marchantia polymorpha*; Pp, *Physcomitrella patens*; Os, *Oryza sativa*; Pv, *Phaseolus vulgaris*; At, *Arabidopsis thaliana*; Nb, *Nicotiana benthamiana*; Sl, *Solanum lycopersicum*.

domains, helicase C-terminal, zf-CCHC and YT521-B homology (YTH) domains (**Supplemental table 2**). These numbers align well with previous RIC experiments (Castello et al., 2012; Liao et al., 2016; Reichel et al., 2016). Ribosomal domains are the most predominant nonclassical RBDs (1253 proteins in all the species) and comprise cytoplasmic as well as chloroplastic and mitochondrial ribosomal proteins (**Supplemental table 2**). This is not unexpected since ribosomes are large macromolecular complexes that are composed of RNA and proteins and interact with mRNA during translation (Pisarev et al., 2008). Some ribosomal proteins also have extra-ribosome functions and interact directly with mRNA (Warner and McIntosh, 2009; Arif et al., 2018; Shi et al., 2017). We also identified a large number of PPR-containing proteins in the different high-confidence RBPomes (**Supplemental table 2**). The PPR-containing protein family has expanded in plants as compared to metazoans (e.g. 450 PPR-containing proteins in Arabidopsis vs 8 in humans; Silverman et al., 2013) and many plant PPR-containing RBPs seem to have evolved from expansion of a small subset of eukaryotic conserved PPR-containing proteins (Lurin et al., 2004). Moreover, many PPR-containing proteins are conserved from bryophytes to angiosperms and are exclusive to plants (Lurin et al., 2004). Some PPR subclasses have been described to have functions as RNA regulators and are classified as 'non classical

RBD', whereas for other subclasses there are no established links to RNA biology yet, so they are classified as 'RBD unknown/putative RBDs' (Barkan and Small, 2014). We identified large numbers of PPR-containing proteins in both categories (**Supplemental table 2, Supplemental digital table 2**). Moreover, other domains such as DYW or C-terminal E domains are often present with the PPR domains in proteins (Lurin et al., 2004). Indeed, we also identified a number of DYW-containing RBPs in most of the species (**Supplemental digital table 2**). The DYW domain contributes to specific recognition of the RNA editing sites (Ichinose and Sugita, 2018) and has been proposed to be the catalytic domain for RNA editing (Salone et al., 2007; Boussardon et al., 2014). A large proportion of the RBPs identified did not possess any known RBD (25-50% of the high-confidence leaf RBPome depending on the species; **Fig. 4**). Thus, some common domains found in these RBPs might represent novel RBDs (**Supplemental table 3, Supplemental digital table 2**). These numbers are similar to those previously observed for other species. For example, 39% and 47% of the RBPomes of etiolated Arabidopsis seedlings and HeLa contain proteins with no recognisable RBD, respectively (Reichel et al., 2016; Castello et al., 2012).

2.4. Uncovering potential novel RBDs in plants

We hypothesised that if a group of RBPs lacking known RBDs share a particular domain, this would likely contribute to RNA binding. We thus analysed the domains that are not annotated as RBDs but are enriched in RBPs from most of the high-confidence plant RBPomes (**Supplemental table 3, Supplemental digital table 2**). We also analysed those domains sparsely connected to RNA interactions.

Whirly (WHY)

We identified 13 RBPs containing the Whirly (WHY) domain in the RBPomes of most of the plant species (**Supplemental table 3**). For *A. thaliana* and *O. sativa*, we identified all the WHY

members encoded in their genomes (3 and 2, respectively; **Supplemental table 3**). These proteins are mainly found in plants and constitute a family of DNA-binding proteins that act as transcription factors primarily in organelles (Cappadocia et al., 2013; Desveaux et al., 2005; Krause et al., 2009). Recent evidence has linked members of this family to RNA binding. For instance, the *Zea mays* WHY1 has been demonstrated to bind both DNA and RNA *in vitro* (Prikryl et al., 2008), and the three Arabidopsis WHY-containing proteins were identified by RIC in etiolated seedlings (Reichel et al., 2016). Interestingly, we identified WHY-containing RBPs in all the species but the two bryophytes (i.e. *M. polymorpha* and *P. patens*). It is known that WHY-containing proteins have expanded in seed plants (Wilhelmsson et al., 2017), and *M. polymorpha* genome encodes for only one WHY-containing protein while *P. patens* genome encodes for none (Moriyama and Sato, 2014). Thus, it not unexpected that we identify WHY-containing RBPs exclusively in angiosperms.

ALBA

We found a total 40 Alba-containing proteins in the RBPomes of all the different plant species (**Supplemental table 3**). Alba domain-containing proteins bind DNA and play roles in transcriptional repression in Archaea, but have also been described to bind RNA (Bell et al., 2002; Jelinska et al., 2005; Forterre et al., 1999; Guo et al., 2003). However, the role of Alba-containing proteins remains largely unstudied. Recent publications have shown that two Alba-containing proteins can bind RNA in plants (Gosai et al., 2015; Yuan et al., 2019; Goyal et al., 2016), and four Alba-containing proteins have been identified by RIC in etiolated Arabidopsis seedlings (Reichel et al., 2016). Here we reveal that Alba domain-containing proteins can bind RNA not only in Arabidopsis, but multiple other plant species.

DUF1296/GBF-interacting protein 1

With the exception of *P. vulgaris*, all the identified plant RBPomes contained RBPs harbouring the DUF1296/GBF-interacting protein 1 (GIP1) domain (**Supplemental table 3**). This domain has been postulated to be a novel RBD in plants, since some proteins harbouring DUF1296/GIP1 were identified by RIC in etiolated Arabidopsis seedlings (Reichel et al., 2016). Using ptRIC we identified three out of the eight predicted DUF1296/GIP1-containing proteins in Arabidopsis, and additional 11 DUF1296/GIP1-containing proteins in other plant species. Our results therefore support a role for DUF1296/GIP1 as a novel RBD across the plant kingdom.

Multiprotein Bridging Factor 1 (MBF1)

Globally, we found the Multiprotein Bridging Factor 1 (MBF1) domain in 10 different RBPs isolated from the RBPomes of all the plant species (**Supplemental table 3**). MBF1 is found in a wide range of eukaryotes, from fungi to metazoans and plants (Nishioka et al., 2018). To our knowledge, this domain has not been previously linked with RNA binding, hence MBF1 may represent a novel RBD.

Proteasome component domain (PCI)

A total of 23 proteins harbouring the Proteasome component (PCI) domain were identified in the RBPomes of all the tested species (**Supplemental table 3**). The PCI domain is found in different proteasome subunits, the COP9 signalosome (CSN) and some eIF3 (Hofmann and Bucher, 1998). Accordingly, 11 identified PCI-containing proteins have an additional eIF-3c_N (eukaryotic translation initiation factor 3 subunit 8 N-terminus) domain and two have an additional eIF3_N (eIF3 subunit 6 N terminal) domain (**Supplemental digital table 2**). Moreover, the PCI domain has been shown to bind RNA in yeast (Kouba et al., 2012), and it has been proposed to bind RNA in plants (Dessau et al., 2008). Here we suggest that PCI domain may be a novel RBD across different plant species.

Peptidase_M24

We identified the domain Peptidase_M24 in 15 different RBPs from the RBPomes of all the tested species (**Supplemental table 3**). This domain is present in various peptidases, but also in some proteins without catalytic activity such as the yeast Spt16, which is part of the FACT complex involved in transcription initiation and elongation (Belotserkovskaya et al., 2003; Biswas et al., 2005; Stuwe et al., 2008). We also identified RBPs harbouring other peptidase domains (**Supplemental digital table 2**). Therefore, one could envisage that the Peptidase_M24, which is normally involved in proteolytic activity, could have lost the catalytic activity in some instances and acquired a novel RNA-binding function.

Phosphoglycerate kinase (PGK)

Nine RBPs identified in the RBPomes of all the tested species contained the Phosphoglycerate kinase (PGK) domain (**Supplemental table 3**). This domain is present in phosphoglycerate kinases, which are enzymes involved in the glucose metabolism. It has already been reported that multiple metabolic enzymes can bind RNA (Beckmann et al., 2015; Marondedze et al., 2016), and the involvement of metabolic enzymes in RNA regulation is discussed below. Moreover, PGK is an non-canonical RBP that has been experimentally validated to bind RNA and its RNA-binding moonlighting function is conserved in humans and yeast (Beckmann et al., 2015; Castello et al., 2016). However, to our knowledge it is currently unknown how phosphoglycerate kinases bind RNA, thus, the PGK domain could be a possible motif through which phosphoglycerate-RNA interactions occur.

S10_plectin

Overall, we found 14 RBPs from the RBPomes of all the tested species that harbour the S10_plectin domain (**Supplemental table 3**). This domain has been identified in the ribosomal

S10 protein as well as some cytoskeletal muscle protein plectin. It was postulated that the S10_plectin domain may be involved in RNA binding (El-Gebali et al., 2019). Here we suggest that S10_plectin functions as novel RBD in different plant species

Domains coexisting with known RBDs

Some domains with unknown RNA-binding function that are enriched in RBPs frequently co-occur with known RBDs within the same protein. For these, it is worth investigating whether the putative novel RBDs binds RNA, or if they are only identified by ptRIC because they co-exist with *bona fide* RBDs. One example is Friend of Prmt1 (FoP_duplication) domain. Overall, 23 FoP_duplication-domain containing RBPs were identified in the RBPomes of all the tested species (**Supplemental table 3**). FoP_duplication is a domain conserved from fungi to plants and is present in nuclear proteins associated with heterochromatin (van Dijk et al., 2010). Moreover, FoP_duplication is also present in known RBPs involved in different aspects of RNA metabolism such as Alyref, Alyref2, Mlo3 or CRE_20669 (van Dijk et al., 2010; Rodrigues et al., 2001). However, based on Interpro annotations, all the FoP_duplication domain-containing proteins identified except for two contain an additional RRM domain (**Supplemental digital table 2**). Hence, we cannot confidently classify FoP_duplication as a novel RBD.

Another example is the Nuclear transport factor 2 (NTF2) domain. We identified 32 NTF2-domain containing proteins in the RBPomes of all the plant species (**Supplemental table 3**). The NTF2 domain has been characterised to be involved in protein-protein interactions (Ribbeck, 1998; Fribourg et al., 2001), but recently several NTF2-containing RBPs were identified in etiolated seedlings (Reichel et al., 2016). Moreover, the Nuclear transport factor 2-like (NTF2L) domain was described to bind RNA (Katahira et al., 2015). However, based on Interpro annotations, all the identified proteins containing NTF2 domain also contain an additional RRM domain (**Supplemental digital table 2**). Thus, we cannot exclude that we identified NTF2-

containing RBPs that bind RNA through the RRM-RNA interaction instead of NTF2-RNA interaction.

In addition, a total of 23 RBPs harbouring the Suppressor of Tom1 (Stm1_N) domain were identified in the RBPomes of all plant species tested (**Supplemental table 3**). Stm1_N domain is present in various proteins that interact with RNA such as the yeast protein STM1 and the Arabidopsis protein AtRGGA (Van Dyke et al., 2004; Ambrosone et al., 2015). We identified AtRGGA and 22 additional RBPs harbouring the Stm1_N domain in plants, indicating that Stm1_N could be important for RNA binding. However, all the Stm1_N-containing proteins harbour an additional Hyaluronan/mRNA-binding protein domain (**Supplemental digital table 2**), so we cannot exclude that Stm1_N domain proteins are captured by ptRIC due to the presence of other RBDs within the same protein.

2.5. Identification of potential novel RBP families in plants

Using ptRIC we identified a number of proteins that were not previously known to bind RNA and may represent novel RBPs (**Supplemental digital table 1**). We also identified a number of proteins with links to RNA biology in other organisms or for which a role as RBP had been previously suggested but not experimentally demonstrated (**Supplemental digital table 1**). Therefore, our data confirm their roles as RBPs in different plant species.

Photosynthesis-related proteins

We identified a large number of RBPs that are components of the photosynthetic machinery in different species. These include proteins from the rubisco small subunit (SSU) in all the plant species and rubisco large subunit (LSU) in all the species with the exception of *M. polymorpha* (**Supplemental table 3, Supplemental digital table 1**). In addition, we identified a total of 56 RBPs in the RBPomes of all the species harbouring the Chlorophyll A-B binding domain (**Supplemental table 3, Supplemental digital table 1**), which are proteins that belong to the

light-harvesting complex (LHC). We also identified RBPs that belong to photosystem I (PSI) and contain several domains related to PSI (PsaA_PsaB, PsaD, PsaL, PsaE, PsaF, PSAK) and to the photosystem II (PSII) harbouring different PSII-related domains (PSII, PsbH, PsbP, PsbQ, PsbR, Psb28, BNR, Cytochrom_B559) (**Supplemental table 3, Supplemental digital table 1**). Other proteins related to photosynthesis were identified such as proteins from the cytochrome b6-f complex. This could indicate a dual function of these RBPs in regulating RNAs as well as in photosynthesis and are further discussed below.

Aquaporins

Remarkably, we identified a total of 22 proteins in the RBPomes of all of the species that contain the Major intrinsic protein (MIP) domain (**Supplemental table 3, Supplemental digital table 1**). MIP is present in proteins belonging to a large family of transmembrane channels (including aquaporins) that transport water, small molecules, ions and gases (Gaspar et al., 2003; Holm et al., 2005; Quigley et al., 2001). Although aquaporins have been postulated to bind RNA in Arabidopsis (Reichel et al., 2016), their ability to bind RNA remains to be experimentally validated by orthogonal approaches. Here we provide evidence that MIP-containing proteins, including aquaporins, bind RNA in multiple plant species.

Heat shock proteins

We identified multiple members of the HSP70 and HSP90 families in the RBPomes of most of the plant species (**Supplemental digital table 1**). Heat shock proteins (HSPs) are molecular chaperones that are involved in protein folding during stress responses. Some HSPs have also been linked to RNA metabolism and translation (Iwasaki et al., 2010; Willmund et al., 2013). Interestingly, different members of the HSP90 and HSP70 families bind RNA in human cells, and RBDmap identified their RNA-binding regions (Castello et al., 2016). In addition, several members of HSP70 and HSP90 have been identified to bind RNA in etiolated Arabidopsis

seedlings (Reichel et al., 2016). Here, we provide extensive evidence that HSP70 and HSP90 family proteins can bind RNA in different plant species.

Peptidyl-prolyl cis-trans isomerases

Interestingly, we identified several peptidyl-prolyl cis-trans isomerases (PPIs) in the RBPomes of different plant species (**Supplemental digital table 1**). PPIs are known to play important roles in transcription, pre-mRNA processing and mRNA decay by mediating ribonucleoprotein (RNP) complexes remodelling (Thapar, 2015). Although only a few PPIs possess known RBDs (Wang et al., 2008), recent evidence indicates that 11 human PPIs lacking known RBDs function as RBPs and the RNA-binding regions were identified by RBDmap (Castello et al., 2016). Our results provide evidence that PPIs might function as RNA binders in different species across the plant kingdom.

Metabolic enzymes

We identified multiple metabolic enzymes in the RBPomes of different plant species (**Supplemental digital table 1**). Several metabolic enzymes have been shown to moonlight as RBPs (Castello et al., 2015) and are further discussed below.

ABC transporters

Interestingly, we identified several proteins harbouring ATP-Binding Cassette (ABC) transporter-related domains (ABC_tran, ABC_tran_Xtn, ABC_trans_N, ABC2_membrane, ABC_membrane_3) in the RBPomes of most of the tested species (**Supplemental digital table 1, Supplemental digital table 2**). ABC transporters translocate different substrates across membranes using ATP. However, some members of the ABC transporter family do not possess transport-related functions and are involved in mRNA translation and DNA repair (Davidson et al., 2008; Chakraburtty, 2001). Moreover, one ABC transporter has been validated to bind RNA

in humans (Urdaneta et al., 2019). Therefore, we propose that a subset of ABC transporters may be involved in binding RNA and that their RNA-binding function is conserved across different species.

2.6. Discovery of the core plant RBPome

To determine the RBPs conserved across species, we cross-referenced all plant RBPomes and classified as 'core plant RBP' when a protein was present in at least five out of seven species. For this, we defined the orthologous RBPs within our set of seven high-confidence RBPomes and searched for conserved RBPs. Using these criteria, we identified 304 orthogroups that could represent the core plant RBPome (**Supplemental table 4).** We analysed the annotation of the members of the core RBPome (Beckmann et al., 2015). Notably, 80% of the core plant RBPome had links to RNA biology (**Fig. 5C**), while the remaining 20% lacked them, representing core unconventional RBPs.

We also interrogated the core RBPome by performing a gene set enrichment analysis using GO annotation. 'RNA binding', 'rRNA binding' and 'nucleic acid binding' were amongst the most enriched GO terms (**Fig. 5A**). Other RNA-related functions statistically enriched were 'structural constituent of ribosome', 'translation initiation factor activity' and 'translation regulation activity'. Amongst the statistically depleted GO terms were 'Protein kinase', 'DNA-binding transcription factor', 'peptidase' and 'transferase' (**Fig. 5A**). These results are in agreement with previously published RIC experiments (Castello et al., 2012), and with the high confidence Arabidopsis RBPome analysed in **Chapter 2**. The fact that 'DNA-binding transcription factor' is underrepresented indicates that ptRIC eluates do not contain DNA contamination. These analyses indicate that the core RBPome of plants is enriched in proteins with functions in RNA biology. Moreover, analyses of the cellular component GO terms associated with the identified RBPs indicated that the core plant RBPome is statistically enriched in RBPs from 'ribosomes',

'RNP complexes', 'cytoplasm' and 'organelles', amongst others (**Fig. 5B**). Interestingly, the core plant RBPome is also enriched in 'photosystem I' proteins and depleted in 'membrane proteins' and 'proteins associated with chromosomes' (**Fig. 5B**).

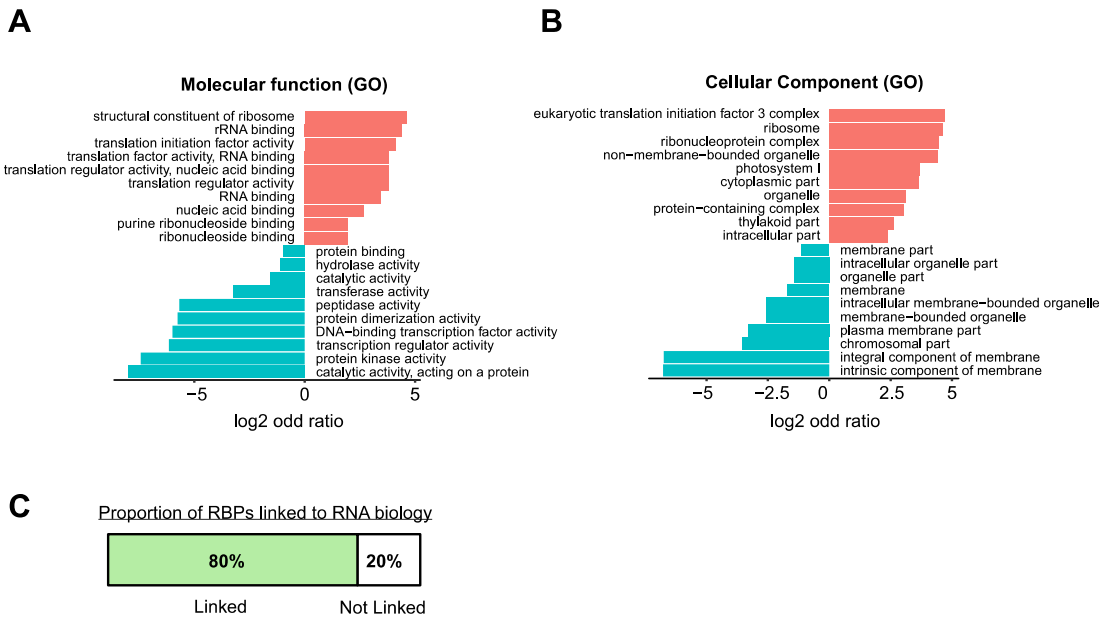


Fig. 5 The core plant RBPome is enriched in proteins linked to RNA biology. Ten of the most significant enriched (red) or underrepresented (blue) GO terms of the core plant RBPome classified by A) Molecular function and B) Cellular compartment. C) Proportion of the core plant RBPome linked to RNA biology or not based on GO annotations as described previously Beckmann and colleagues (Beckmann et al., 2015).

We analysed the domain architecture of the core plant RBPome and determined that 78 orthogroups possessed classical RBDs, 133 non-classical RBDs and 93 did not harbour domains known to bind RNA (RBD unknown) (**Fig. 6A**). Therefore, about 69% of the core plant RBPs harbours known RBDs (classical or nonclassical). This is in line with what has been observed for humans (53%; Castello et al., 2012) and in Arabidopsis (60%; **Chapter 2**)**.** The most represented classical RBD was RRM, followed by zf-CCCH, DEAD and KH (**Fig. 6B**), which are known to be predominant classical RBDs (Castello et al., 2012). Of the non-classical RBDs, the most abundant domains were ribosomal followed by PPR, helicase, MMR_HSR1 and S1 (**Fig. 6C**). These results are very similar to those obtained when analysing the individual high-confidence RBPomes of the different plant species and are discussed above. Moreover, a proportion of the core plant

RBPs do not possess any known RBD (30%) indicating that putative novel RBDs await to be discovered (**Fig. 6D**).

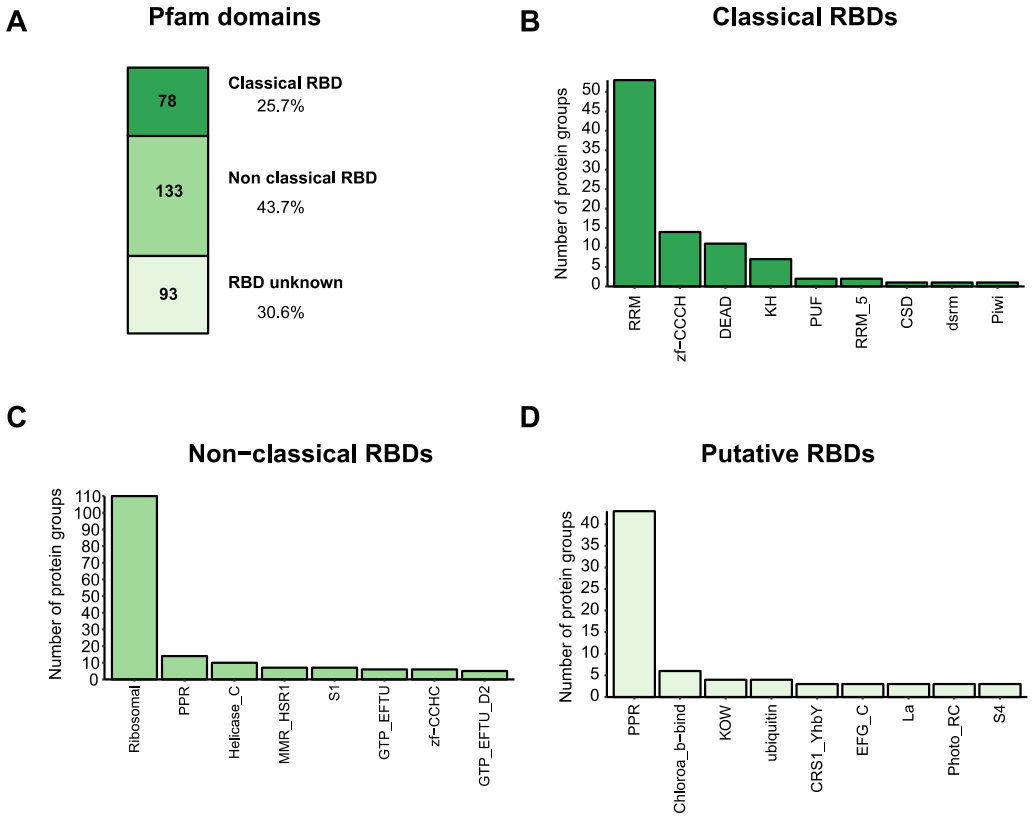


Fig. 6 Domains identified in the core plant RBPome. A) Number of proteins harbouring classical, non-classical or no known RBDs in the core plant RBPome. (B) Number of proteins annotated as possessing classical RBDs, (C) non-classical RBDs or (D) putative RBDs.

2.7. RBP pathways of the core plant RBPome

We wanted to determine the RBP networks that are conserved across the plant kingdom. Towards this aim, we performed STRING analyses of the orthogroups within the core RBPome of plants searching for enriched RNA-related functions (**Fig. 7**). Below we discuss the results.

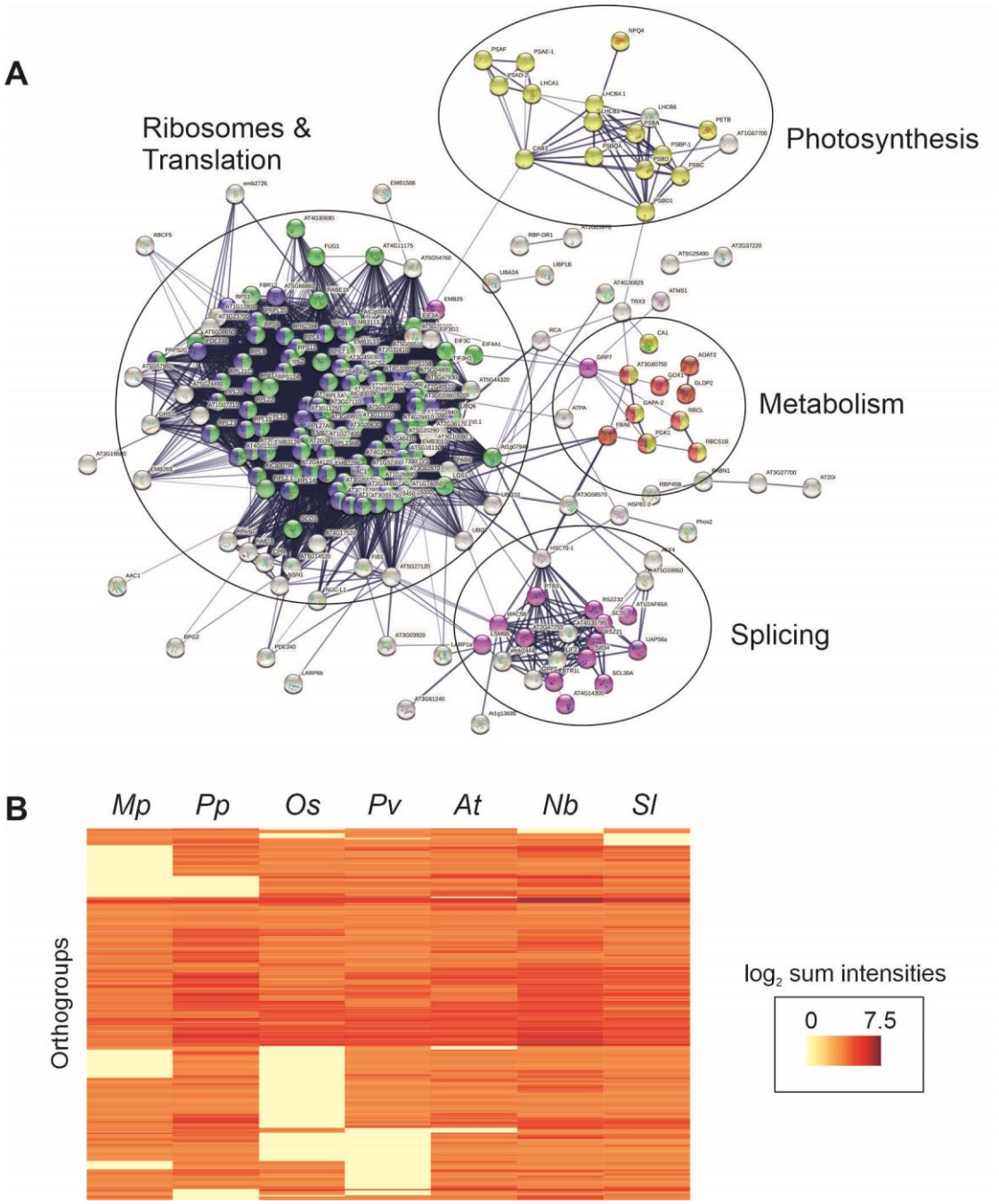


Fig. 7 The core plant RBPome. A) STRING was used to display the connectivity between RBPs of the core plant RBPome. Proteins involved in ribosomes are coloured in blue, translation in green, splicing in purple, photosynthesis in yellow and metabolism in red. B) Heat map representing the sum of the averaged intensities of all the proteins within each of the orthogroups (log_2). Abbreviations: Mp, *Marchantia polymorpha*; Pp, *Physcomitrella patens*; Os, *Oryza sativa*; Pv, *Phaseolus vulgaris*; At, *Arabidopsis thaliana*; Nb, *Nicotiana benthamiana*; Sl, *Solanum lycopersicum*.

Translation

Approximately 32% of the core plant RBPome (97 orthogroups) was related to translation and ribosomes (**Fig. 7A, Supplemental table 4**) including initiation and elongation factors and large

and small subunit ribosomal proteins (RP) from cytoplasm and organelles. It is well known that translation is tightly regulated, allowing plants to quickly respond to the everchanging environment (Xu et al., 2017). This regulation is done via adjustment of translation components and ribosomes. Hence, the large number of conserved RBPs involved in translation and ribosomes highlights the importance of translation control for normal cell functioning.

Splicing

Moreover, 20 orthogroups of the core plant RBPome were linked to splicing (**Fig. 7A, Supplemental table 4**). These included RBPs with well-known roles in splicing such as U2 snRNP auxiliary factors, helicases, ribonucleases, members of the exon junction complex (EJC) and serine/arginine-rich (SR) splicing factors. Most eukaryotic pre-mRNA contain introns, which are removed by splicing (Labadorf et al., 2010; Sharp, 2005). In addition, pre-mRNA can also be alternatively spliced, which increases the complexity of the proteome and has an important role during plant responses to environmental stimuli (Reddy, 2004). With some exceptions, the core splicing machinery is conserved amongst eukaryotes, and this is also observed in the core plant RBPome.

Photosynthesis

A large number of RBPs linked to photosynthesis were detected in the RBPome of all the different species (**Supplemental digital table 1**). Importantly, our data reveal that association with RNA of members of photosynthetic machinery such as rubisco small subunit (SSU), light-harvesting complex (LHC), photosystem I (PSI) and photosystem II (PSII), is conserved across the plant kingdom, from bryophytes to angiosperms (**Fig. 7A, Supplemental table 4**). This is further discussed below.

Metabolic enzymes (not involved in photosynthesis)

We identified a set of metabolic enzymes as part of the core plant RBPome, indicating that they are conserved moonlighting enzymes that bind RNA (**Fig. 7A, Supplemental table 4**). Moonlighting functions been described for a number of metabolic enzymes in other species, including plants (Castello et al., 2015; Matia-González et al., 2015; Beckmann et al., 2015; Marondedze et al., 2016). The core plant RBPome included glyceraldehyde-3-phosphate dehydrogenase (GAPDH), fructose-bisphosphate aldolase and transketolase, which have been previously identified to bind RNA in other organisms by RIC (Castello et al., 2012; Marondedze et al., 2016; Reichel et al., 2016; Hentze et al., 2018). Moreover, we also identified phosphoglycerate kinase and thioredoxin, which have been validated to bind RNA in humans and yeast (Beckmann et al., 2015; Castello et al., 2016). Thus, these metabolic enzymes have conserved RNA-binding activities from yeast to humans and plants. We also identified additional enzymes in the core plant RBPome such as cobalamin-independent methionine synthase, (S)-2-hydroxy-acid oxidase, glycine dehydrogenase and glutamate-glyoxylate aminotransferase, which, to our knowledge, have not previously been linked to RNA biology (**Supplemental table 4**). Importantly, all these moonlighting RNA-binding metabolic enzymes are conserved throughout the plant kingdom. Moreover, proteins that bind ATP were also found in the core plant RBPome such as ATPase E1, plasma membrane ATPase and ATP synthase (**Supplemental table 4**). Multiple RNA-binding metabolic enzymes can bind mono or di-nucleotides (Castello et al., 2015, 2016) and some conserved RBPs between yeast and humans have functions related to ATP-binding (Beckmann et al., 2015). Therefore, this RNA binding mode could be extended to the ATPases and ATP synthases detected in the core plant RBPome.

Other processes also meet RNA

In addition, we identified proton pump interactor 1 (PPI1) in the core plant RBPome (**Supplemental table 4**), which is a protein that promotes the activity of a plasma membrane

ATPase (Morandini et al., 2002; Viotti et al., 2005; Anzi et al., 2008). Interestingly, PPI1 has been previously described to bind RNA in etiolated seedlings (Reichel et al., 2016). Our results indicate that PPI1 can bind RNA beyond Arabidopsis. Other proteins previously unknown to bind RNA were found in the core plant RBPome and in the RBPome of Arabidopsis etiolated seedlings. These included the Apoptosis inhibitory protein 5 (API5) and the Multiprotein-bridging factor 1A (MBF1A), the latter one being a transcriptional coactivator (Tsuda et al., 2004; **Supplemental table 4**). Finally, the core RBPome contained two aquaporins and multiple RBPs known to be involved in stress responses such GRP7, GRP2, RBP-DR1, FCA, HSP70 and HSP90, amongst others (**Supplemental table 4**).

2.8. Lineage-specific RBPs

Next, we wanted to determine if any RBPs from the high-confidence plant RBPomes was unique to a specific plant lineage (e.g. specific to Solanaceae species). Thus, we analysed the orthogroups that were unique to the different plant lineages and searched for RBP families that could be enriched. There are 706 orthogroups shared between two or more plant RBPomes (**Fig. 8B**)**.** Of those, 75 orthogroups were bryophyte-specific and were only present in *P. patens* and *M. polymorpha* (**Fig. 8B, Supplemental table 5**). Moreover, 17 orthogroups were angiosperm-specific since they were present in all the angiosperm species but absent in the bryophytes (**Fig. 8B, Supplemental table 5**). These included some members of the WHY and PUMILIO families. In addition, 15 orthogroups were dicotyledonous-specific (mainly PPR) and 26 were Solanaceae-specific (**Fig. 8B, Supplemental table 5**). We did not observe an enrichment of specific RBP families in the putative lineage-specific RBPs, so we conclude that a deeper experimental analysis coupled to bioinformatics prediction should be performed to obtain solid conclusions of lineage-specific RBPs. A limitation of this analysis is the differential depth of the different RBPomes, thus we decided to avoid comparison between individual datasets.

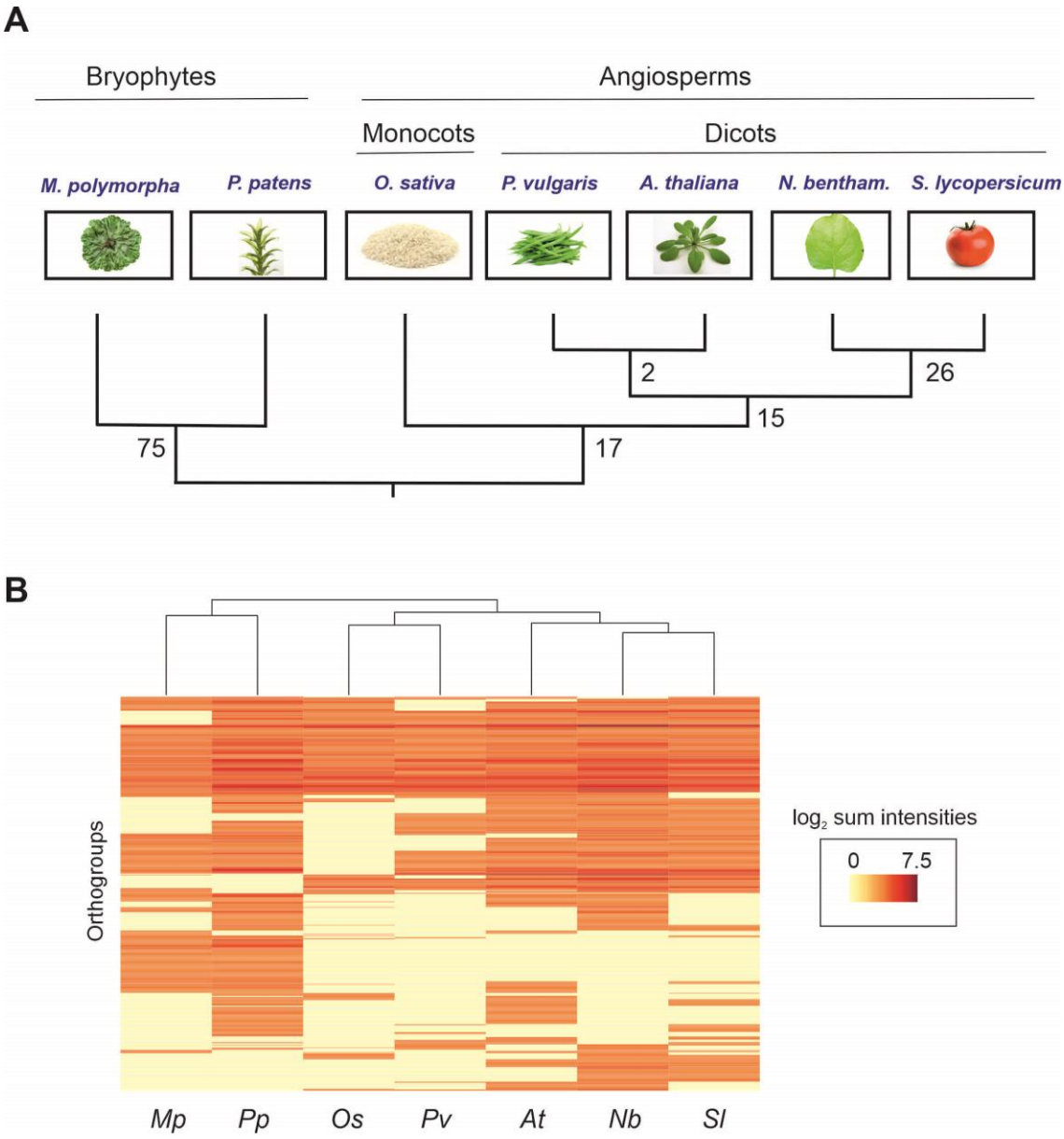


Fig. 8 Putative lineage-specific RBPs. A) Phylogenetic relationship between the species analysed. Numbers of putative lineage-specific orthogroups are depicted at the node of each lineage. B) Heat map representing the sum of the averaged intensities of all the proteins within each of the orthogroups (log_2). Orthogroups shared by at least two plant species are shown. Species are clustered based on similarity and reflect the phylogenetic relationship between the species.

3. DISCUSSION

3.1. Unravelling the RBPome of different plant species

We have shown that our variant of RIC, referred to here as ptRIC, can be applied to a wide range of plant species beyond the model plant *Arabidopsis thaliana*, providing comprehensive RBPomes (**Fig. 1B**). This robustness has allowed the unprecedented discovery of the leaf RBPomes of seven plant species ranging from bryophytes to angiosperms (**Fig. 2**). Our findings are of broad relevance and applicability, since the plant species used here are agriculturally and economically important and/or are model systems. However, ptRIC also possesses some limitations that should be considered. Variations in composition or structure between the species or plant organs may result in differences in the efficiency of UV crosslinking and RNA-RBP isolation, thus yielding different depths of RBPs identification as observed in by silver staining and mass spectrometry analyses (**Fig. 1**, **Fig. 2**).

3.2. Discovery of putative novel RBDs

We observed that a large proportion of the plant RBPs (25-50% depending on the plant species) does not harbour recognisable RBDs (**Fig. 4**), which agrees with what has been observed in mammalian RBPomes (Castello et al., 2012). However, several protein domains previously unlinked to RNA biology are reproducible identified in RBPs from different species (**Supplemental table 3, Supplemental digital table 2**). The robustness of these identifications suggests that these domains may be engaged in RNA binding in plant tissues. Moreover, RBDs with suggested RNA-binding activity are confirmed here to interact with RNA due to their high incidence in the RBPomes of the different plant species. These putative novel RBDs include different photosynthesis-related domains (Chlorophyll A-B binding domain, Rubisco_small, etc), MIP, PCI, DUF1296, and others (**Supplemental table 3, Supplemental digital table 2**). Here, we only discuss a small subset of the 862 domains previously unreported to bind RNA that were

enriched in RBPs of multiple plant species. Moreover, other domains are identified in some but not all the species and could represent lineage-specific RBDs (**Supplemental digital table 2**). Although exciting, a key limitation of this approach is that protein domains that do not interact directly with RNA can co-occur within the same protein with RBDs. Hence, we encourage to validate the RNA-binding ability of these individual domains prior to in depth characterisation. A number of *in vitro* techniques can be used to study the RNA binding activity of such RBDs, including electrophoretic RNA mobility shift assay (EMSA; Ryder et al., 2003) and systematic evolution of ligands by exponential enrichment (SELEX; Zhuo et al., 2017). Alternatively, application of RBDmap (Castello et al., 2016, 2017) to comprehensively determine RBDs in plants would help to validate some of the identified domains as novel RBDs in plants.

3.3. The RBPome of the plant kingdom reveals many non-canonical RBP families

Since it was developed in human cells, RIC has allowed discovery of hundreds of novel RBPs in multiple species (Hentze et al., 2018). Using ptRIC we have identified a large number of RBPs that were previously unlinked or sparsely related to RNA biology in Arabidopsis (**Chapter 2**). Here, we have expanded these discoveries by uncovering hundreds of novel RBPs across different plant species, many of which are non-canonical RBPs conserved across the plant kingdom (**Supplemental digital table 1, Supplemental table 4**). Some of these unconventional RBPs include photosynthetic machinery components, transporters and metabolic enzymes and are discussed below.

Photosynthesis-related proteins moonlighting as RBPs

A large number of proteins linked to photosynthesis were detected in the RBPomes of all the plant species and many of them are present in the core plant RBPome (**Supplemental digital table 1, Supplemental table 4**). *A priori* this may seem striking since photosynthesis has not been extensively linked to RNA metabolism. However, there are multiple lines of evidence that

support that photosynthesis-related proteins are involved in RNA binding. Firstly, we identified different proteins from rubisco large subunit (LSU) in the RBPomes of the different plant species (**Supplemental digital table 1**) and some of them are conserved in the core plant RBPome (**Supplemental table 4**, **Supplemental digital table 1**). Moreover, LSU can bind RNA in *Chlamydomonas reinhardtii* under stress conditions (Yosef et al., 2004; Cohen et al., 2006, 2005) and its N-terminal domain contains an RRM. In normal conditions this RRM is buried inside the protein, but it becomes exposed under oxidizing conditions, triggering its capacity to bind RNA (Yosef et al., 2004; Cohen et al., 2005). The structure of this N-terminal domain is highly conserved throughout evolution, despite low sequence similarity between species, and its RNA-binding activity has been experimentally validated in different photosynthetic organisms (Cohen et al., 2006). Therefore, our data confirms that rubisco LSU is an RBP that interacts with RNA in multiple plant species (**Supplemental digital table 1**). Moreover, our data suggest that a fraction of rubisco LSU is bound to RNA under steady-state conditions as otherwise would not have been detected in our experimental conditions. Secondly, we identified different components of the rubisco small subunit (SSU) in the RBPomes of the different plant species (**Supplemental digital table 1**) and several are present in the core plant RBPome (**Supplemental table 4**), indicating that, in addition to LSU, SSU also interacts with RNA. Furthermore, components of the light-harvesting complex (LHC), photosystem I (PSI) and photosystem II (PSII) are identified in the plant RBPomes and are conserved in the core RBPome (**Supplemental digital table 1, Supplemental table 4**). These proteins harbour Chlorophyll A-B binding domains or domains related to PSI and PSII, respectively (**Supplemental table 3)**. Other proteins related to photosynthesis were identified such as proteins from the cytochrome b6-f complex that transfer the electrons between PSII and PSI (**Supplemental digital table 1**). Importantly, Choquet and colleagues demonstrated that cytochrome f regulates its own RNA translation by binding the 5'UTR (Choquet et al., 2003). The binding is through the C-termini of the protein, which is shielded in the complex, but that upon disassembly the C-termini is exposed and binds RNA.

Other studies have also identified photosynthesis-related proteins in the RBPome of Arabidopsis (Zhang et al., 2016b; Marondedze et al., 2016). However, strikingly none of the RIC studies in plants have examined the possibility that photosynthetic enzymes act as RNA binders. It is tempting to speculate that photosynthesis-related proteins may exploit their RNA-binding activity to regulate gene expression in particular cellular states. Alternatively, these proteins could be allosterically regulated by RNA binding in a new case of 'riboregulation' as it has been described by the vault1-1 RNA on p62 (Horos et al., 2019). Further functional studies to identify the RNAs bound by these proteins and the biological consequences of these interactions will shed light on how photosynthesis regulates or is regulated by RNA.

Unconventional RNA transport?

Several proteins involved in transport of molecular cargos across membranes were robustly identified as RBPs across the studied plant species (**Supplemental digital table 1**). These include aquaporins, ABC transporters and NTF2-containing proteins, several of which are present in the core RBPome (**Supplemental table 4**). Interestingly, several aquaporins were also identified as RBPs in etiolated seedlings and were postulated to be involved in cell-to-cell transport of RNA (Reichel et al., 2016). In addition, some NTF2-containing proteins are involved in importing cargoes to the nucleus, hence, we hypothesise that NTF2-containing proteins could be involved in retrograde RNA transport. Although the function of these transporters as RBPs still needs validation, it is tempting to speculate that these proteins might contribute to RNA transport and localisation. This would open the possibility that unconventional pathways participate in the transport of RNAs through the cell or from cell to cell.

Metabolic enzymes moonlight as RNA binders

Multiple metabolic enzymes were identified to bind RNA in the RBPomes of the different species (**Supplemental digital table 1**), and some of them were conserved in the core RBPome (**Supplemental table 4**). One example is glyceraldehyde-3-phosphate dehydrogenase (GAPDH), which is well studied metabolic enzyme that moonlights as RBP. GAPDH catalyses the conversion of glyceraldehyde-3-phosphate to D-glycerate-1,3-bisphosphate and generates NADH. Moreover, GAPDH binds to a variety of RNAs from mRNAs to rRNAs and viral RNAs (Castello et al., 2015). For example, GAPDH regulates T-cell effector function by moonlighting between glycolysis and binding to the 3' UTR of cytokine mRNA (Chang et al., 2013). GAPDH has also been identified by RIC in Arabidopsis cell cultures (Marondedze et al., 2016). As part of the core plant RBPome we also identified several additional metabolic enzymes that are known to bind RNA in plants and humans such as fructose-bisphosphate aldolase, phosphoglycerate kinase and transketolase (Castello et al., 2012; Beckmann et al., 2015; Marondedze et al., 2016), but also a number of metabolic enzymes with unknown links to RNA biology and that might represent novel RNA-binding moonlighting enzymes (**Supplemental table 4**).

With the use of RIC it has become evident that the moonlighting function of some metabolic enzymes as RBPs is a widespread phenomenon, and that the RNA-binding activity of some metabolic enzymes is crucial for cell functioning (Castello et al., 2015). However, additional functional studies are required to understand the crosstalk between metabolism and gene regulation. It will be particularly interesting to study the RNAs regulated by metabolic enzymes and whether enzyme-RNA binding alters enzymatic activity. It will also be crucial to determine whether metabolic enzymes interact with RNAs in a cellular state-dependent manner, as has been described for the human aconitase (Volz, 2008).

3.4. Conclusion

Using ptRIC we have unravelled the RBPome of different species across the plant kingdom. These important datasets have allowed the discovery of hundreds of RBPs that were previously unknown to interact with RNA. Many particular domains previously unlinked to RNA binding are enriched in RBPs and may be endowed with RNA binding. We have defined the 'core plant RBPome' that compresses more than 300 RBPs conserved from bryophytes to angiosperms Amongst the conserved RBPs, photosynthesis-related proteins emerge as prevalent group of novel RNA interactors.

4. MATERIALS & METHODS

4.1. Plant growth conditions

Arabidopsis thaliana ecotype Col-0 and *Nicotiana benthamiana* plants were grown in soil for 6 weeks at neutral day conditions (12 h light, 12 h dark) at 20 °C and light intensity of approximately 100 µmol/m^2/s. *Phaseolus vulgaris* cv Canadian Wonder (bean) and *Zea mays* var. B73 (maize) were germinated on wet tissue, transplanted to soil and grown for 2/3 weeks at long day conditions (16 h light, 8 h dark) at 21°C (*P. vulgaris*) or 29°C day/20°C night (*Z. mays*) and light intensity of approximately 150-200 µmol/m^2/s. *Oryza sativa* subsp. *japonica* cv Nipponbare (rice) seeds were peeled and surface sterilized with ethanol 70 % (1 minute), 30% sodium hypochlorite (30 minutes) followed by 5 washes in sterile water. Seeds were germinated on a wet tissue, transplanted to soil and plants were grown for 3/4 weeks in long day conditions at 30°C day/25 °C night and light intensity of approximately 100 µmol/m^2/s. *Solanum lycopersicum* cv Moneymaker (tomato), *Solanum tuberosum* (potato) and *Nicotiana tabacum* (tabaco) were grown in soil for 4 weeks at long day conditions, 21°C and light intensity of approximately 150-200 µmol/m^2/s. *Marchantia polymorpha* Tak1 (male) and Tak2 (female) were grown from gemmae in Johnson's medium 1% agar for 2/3 weeks at short day conditions

(8h light /16h dark) at 20 °C day /17 °C night. *Physcomitrella patens* genotype Grandson was grown from protonemal tissue on top of cellophane in moss media (Based on Essential Moss Methods, from the University of Leeds) at 24 °C at long day conditions for 1 week.

4.2. Plant RNA interactome capture (ptRIC) and MS analyses

ptRIC and was performed as described in **Chapter 2** with minor modifications. For *P. patens*, protonemal tissue was used for ptRIC analyses. For *M. polymorpha* equal amounts (weight) of thallus from Tak1 (male) and Tak2 (female) were pooled for ptRIC analyses. For the rest of the species, leaves of mature plants were used.

Sample preparation for MS and MS analyses were performed as described in **Chapter 2**. For all the species except for *N. benthamiana*, the reference proteomes were downloaded from NCBI (https://www.ncbi.nlm.nih.gov) on February 2019: *M. polymorpha* (GCA_003032435.1_Marchanta_polymorpha_v1_protein.faa), *P. patens subsp. patens* (GCF_000002425.4_Phypa_V3_protein.faa), *O. sativa* (GCF_001433935.1_IRGSP-1.0_protein.faa), *Phaseolus vulgaris* (GCF_000499845.1_PhaVulg1_0_protein.faa), *A. thaliana* (GCF_000001735.4_TAIR10.1_protein.faa) and *S. lycopersicum* (GCF_000188115.4_SL3.0_protein.faa). For *N. benthamiana* we used the recently compiled reference database (ACE_0319_Sec7.23.Niben101_plus.fasta, 74812 entries; Kourelis et al., 2019).

4.3. SDS-PAGE and silver staining analyses

The inputs and the eluates of ptRIC were mixed with protein loading buffer and incubated for 4 min at 95 °C. The inputs and the marker were diluted 1/20 and 1/15 with water, respectively. The denatured proteins were separated in a 10-12% acrylamide gel followed by silver staining using Silver quest (Invitrogen) following the manufacturer's instructions.

4.4. Statistical analysis of the RBPomes

MS data was analysed as described in **Chapter 2** with minor modifications. Briefly, after processing the data, proteins were classified as 'RBPs' or 'non-enriched in +UV/-UV' using a dual criteria approach. Proteins with missing values in all the NoCL replicates (negative control) and intensity values in at least n-1 of the CL replicates were defined as 'RBPs' (semi-quantitative approach). For proteins with values in both CL and NoCL samples, the $\log_2$ fold change [CL/NoCL] and false discovery rate (FDR) were calculated and used for statistical enrichment of +UV/-UV (quantitative approach). Proteins with $\log_2$FC[CL/NoCL] ≥ 1.5 and an FDR ≤ 0.05 were defined as 'RBPs' whereas the remaining were considered as 'non-enriched in +UV/-UV'. For *Oryza sativa*, *Phaseolus vulgaris* and *Solanum lycopersicum* one of the samples was lost during the processing. Therefore, for those species the statistical analyses were performed using three biological samples for CL and NoCL. All statistical analyses were conducted in R (R Core Team, 2014). Some of the peptides identified could not be assigned to a unique protein, hence, the analyses were performed using the protein group. Protein groups are specified in the tables.

4.5. Bioinformatics analyses of the RBPomes

The reference proteomes were obtained from NCBI (https://www.ncbi.nlm.nih.gov). GO terms and PFAM and Interpro codes were added to the reference proteomes of each of the species using Interproscan (InterProScan-5.36-75.0; Mitchell et al., 2019). The superset of plant proteomes was generated by adding all the individual proteomes (NCBI) for each of the seven plant species. Gene set enrichment analyses were performed as described in **Chapter 2** with the following minor modification. The frequencies of GO terms of each individual RBPome or the core RBPome were compared against the reference proteome (NCBI) or to the superset of plant proteomes, respectively. Analyses of protein domains were performed as described in **Chapter 2**. Analyses of links to RNA biology of the identified RBPs were performed as described

previously (**Chapter 2**; Beckmann et al., 2015). Graphs were generated using the ggplot2 package within R (Wickham, 2009).

4.6. Conservation analyses

The RBPs identified for each of the plant species were grouped into orthogroups based on sequence homology using Orthofinder (Emms and Kelly, 2015). Orthogroups contain orthologs and paralogs. An orthogroup was defined to belong to the core RBPome if it contained at least one protein from five or more of the seven species analysed. In order to visualise differences and similarities between different species, the python (3.6.8) implemented seaborn clustermaps (seaborn 5.0) was used (Waskom et al., 2014). Clustermaps show the sum of averaged absolute intensities of proteins of the displayed orthogroup. All quantitative data was $\log_2$ transformed.

STRING (Szklarczyk et al., 2017) was used to display connectivity between RBPs of the core plant RBPome. Protein networks were generated using the following parameters: display – confidence; interaction sources – experiments, databases; interaction score – medium confidence (0.400). Disconnected nodes were hidden from display and nodes coloured based on functional enrichment within the network as determined by STRING.

ACKNOWLEDGEMENTS

We want to thank all the members of Preston, Castello and van der Hoorn lab for their fruitful discussions and inputs on the manuscript. We also want to specially thank Nattapong Sanguankiattichai for his valuable advice and help in the bioinformatics pipeline. We would like to acknowledge several people for providing us with different plant species: Susanna Streubel for *M. polymorpha*, Laura Moody for *P. patens,* Blanca San Segundo for *O. sativa* and Tom Hughes for *Z. mays*. We would like to thank Peng Wang for his advice on growing *O. sativa***.** We thank Urszula Pyzio, Sarah Rodgers and Caroline O'Brien for excellent technical support.

FUNDING

Marcel Bach-Pages is supported by Biotechnology and Biological Sciences Research Council (BBSRC, grant BB/M011224/1) and by the Lorna Casselton Memorial Scholarship at St. Cross College, Oxford. Alfredo Castello is supported by MRC Career Development Award MR/L019434/1 and MRC grant MR/R021562/1.

AUTHOR CONTRIBUTIONS

MBP designed the experiments with feedback from RvdH, AC and GMP. MBP performed all the experiments except for the MS run, which was performed by FK. FH and JK assisted with the bioinformatics analyses. MBP analysed all the data and wrote the manuscript with feedback from AC and GMP.

SUPPLEMENTAL DATA

APPENDIX III – Supplemental data Chapter 3

Note: supplemental table 4 is a summary of the results; the full data can be found in supplemental digital table 3.

Printed supplemental files

Supplemental Table 1. Classical RBDs

Supplemental table 2. Non-classical RBDs

Supplemental table 3. Selected putative novel RBDs

Supplemental table 4. Summary table of the core plant RBPome

Supplemental table 5. Putative lineage-specific RBPs

Digital supplemental files

Supplemental digital table 1. Superset of plant RBPs for all the species

Supplemental digital table 2 Domains not classified as RBDs and associated with RBPs

Supplemental digital table 3. Core RBPome

CHAPTER 4

Proteome-wide Profiling of RNA-Binding Protein Responses to flg22 Reveals Novel Regulators of Plant Immunity

Proteome-wide Profiling of RNA-Binding Protein Responses to flg22 Reveals Novel Regulators of Plant Immunity

Marcel Bach-Pages[1], Honglin Chen[2], Farnusch Kaschani[3], Renier A.L. van der Hoorn[1], Alfredo Castello[2]*, Gail M. Preston[1]*

[1]Department of Plant Sciences, University of Oxford, South Parks Road, Oxford, UK

[2]Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK

[3]Fakultät für Biologie, *Universität Duisburg-Essen, Essen, Germany*

* Co-corresponding authors

Key words

RNA-binding proteins // RNA-binding proteome // ptRIC // Immunity // Arabidopsis // flg22

ABSTRACT

Plants respond to pathogens with a myriad of immunity mechanisms that require transcriptome reprogramming. RNA-binding proteins (RBPs) play critical roles in post-transcriptional gene regulation and are known to contribute to these immunity regulatory mechanisms. However, the complement of RBPs required for these processes remains largely unknown. To profile the modulation of the RNA-binding proteome (RBPome) during immune responses we have applied 'ptRIC' to plants treated with the immune elicitor flg22. We have identified 1135 leaf RBPs actively bound to RNA, which is the most comprehensive Arabidopsis RBPome to date and includes many putative novel RBPs. Importantly, we have uncovered that the plant RBPome dynamically responds to immune activation by flg22. Nearly 300 RBPs display differential association with RNA upon elicitor perception, which cannot be explained by matching changes

in protein abundance. This suggests that flg22 elicits alterations in the RNA-binding affinity of responsive RBPs. Analyses of the flg22-responsive RBPome uncovers distinct RBP networks functioning during immune responses. These include well-characterised RBPs as well as unconventional RBPs lacking known RNA-binding domains, including peptidyl-prolyl cis-trans isomerases (PPIs), heat shock proteins (HSP) and proteins involved in photosynthesis. Importantly, by phenotypic analyses of selected mutants we have discovered that flg22-responsive RBPs modulate plant resistance to bacterial pathogens. In summary, the use of ptRIC has uncovered hundreds of RBPs that respond to flg22, and which may regulate plant immunity. Therefore, this study provides a unique resource that will be of great interest for the plant pathology community.

1. INTRODUCTION

1.1. RBPs regulate plant immunity

Plants have evolved a sophisticated two-branched immune system able to effectively restrict and counteract pathogens. Plants can perceive pathogens through recognition of pathogen-associated molecular patterns (PAMPs) or damage-associated molecular patterns (DAMPs) using pattern recognition receptors (PRR; Zipfel, 2014). One such PAMP is flg22, a 22-amino acid peptide derived from the bacterial flagella, which is recognised by the plant flagellin-sensing 2 (FLS2) receptor (Chinchilla et al., 2006), and elicits PAMP-triggered immunity (PTI; Zipfel, 2014). Multiple pathogens have evolved effectors able to target host processes, including PTI. However, some plants can detect these effectors using nucleotide-binding site leucine-rich repeat (NBS-LRR) receptors and elicit effector-triggered immunity (ETI; Chiang and Coaker, 2015). Plant defence responses typically involve production of reactive oxygen species (ROS), callose deposition, inhibition of photosynthesis, stomatal closure and production of

antimicrobial compounds, amongst others (for more detailed information about the plant immune system please see **Chapter 1**).

However, for most of the immune responses to occur, an extensive reprogramming of defence-related genes is required (Buscaill and Rivas, 2014). Recently, RBPs have emerged as critical players in post-transcriptional regulation of gene expression by regulating RNA localization and metabolism. Therefore, RBPs are likely to play important roles as orchestrators of the re-programming of the plant transcriptome during plant immunity. Indeed, a number of studies have demonstrated the importance of RBPs in immunity. For instance, multiple plant RBPs have been shown to contribute to defences against bacteria (Roux et al., 2015), fungi (Zhou et al., 2018) and other pathogens (Lee et al., 2012b), whereas others are targeted by pathogen effectors (Fu et al., 2007; Wang et al., 2015). For more details on specific RBPs involved in plant immunity please see **Chapter 1**. Although the mechanism of action of some of these RBPs is known, the scope of immunity RBPs and their mode of action remain largely unknown. Moreover, immunity-related RBPs have been identified stepwise, and global approaches to determine the identity and activity of RBPs participating in plant defences are lacking.

1.2. Profiling the dynamics of the RBPome by RNA interactome capture

For the last decades we lacked a method to study RBP dynamics in cellular systems and organisms. However, recently 'RNA interactome capture' (RIC) was developed that exploits the UV irradiation of living material to crosslink protein-RNA interactions, opening the opportunity to explore RBPomes under different pathological and physiological conditions (Castello et al., 2012; Baltz et al., 2012). RIC has dramatically expanded our census of RBPs, identifying hundreds of unconventional RBPs that were previously unknown to bind RNA (Hentze et al., 2018). Since then, RIC has rapidly expanded to a wide range of organisms including yeast, worm and human cell lines, opening new horizons for the study of RBPs in live tissues (Hentze et al., 2018).

Because RIC identifies RBPs bound to RNA in a proteome-wide scale, it can be applied to decipher how the complement of RBPs changes in response to physiological cues or environmental stimuli. Multiple recent studies have shown that the RNA-binding proteome (RBPome) is dynamically regulated to govern the changes that occur during adaptation to physiological and environmental cues. For example, the RBPome of Arabidopsis cell cultures has been described to respond to PEG-induced drought stress (Marondedze et al., 2019) and more than 200 human RBPs differentially associate with RNA upon viral infection (Garcia-Moreno et al., 2019). In addition, the human RBPome has also been found respond to treatment with dimethyloxalylglycine, which is an inhibitor of RNA demethylases (Perez-Perri et al., 2018). Moreover, the *Drosophila* RBPome is dynamically modulated during maternal-to-zygotic transition to orchestrate the post-transcriptional changes whereby the maternal mRNAs are replaced by zygotic mRNAs (Sysoev et al., 2016).

1.3. Rationale and aim of the study

RIC has allowed the study of the dynamic responses of the RBPome to different cellular contexts (Hentze et al., 2018). Hence, RIC represents an excellent tool to identify the scope of RBPs functioning in immunity and to study the dynamics of the plant RBPome during immune responses.

To identify the complement RBPs contributing to plant immunity, we applied a plant-optimised version of RIC (ptRIC; see **Chapter 2**) to plants treated with the immunogenic peptide flg22. Flg22 was used as a well-characterised elicitor of plant immunity (PTI), avoiding responses that occur when multiple elicitors (e.g. heat-killed bacteria/fungi) are used or the manipulation of the RBPome by pathogens. Moreover, the immune response to flg22 has been extensively studied in plants, offering substantial publically available data to compare our data to. We focused on

plant leaves as this is a widely used tissue for plant-pathogen studies and the relevant tissue to study foliar pathogens.

2. RESULTS

2.1. Elicitation of immune responses in Arabidopsis

To capture the dynamics of the plant RBPome during immune responses, Arabidopsis plants were infiltrated with either H_2O (mock) or the immune elicitor flg22 and leaf tissue was harvested at 2 h and 12 h post treatment (hpt). RBPs bound to RNA in each condition were isolated by ptRIC and quantified by quantitative proteomics (**Fig. 1A, B**). These time points were chosen as early (2 hpt) and late (12 hpt) responses to flg22. We reasoned that the early responses would mainly include rapid post-translational stimulation/inhibition of RBPs whereas the late responses could additionally be affected by changes in protein levels or RNA availability. To assess that the flg22 treatment was efficient at eliciting plant immunity and to ensure that the flg22 peptide was not degraded during plant infiltration, we analysed activation of immune signalling by mitogen-activated protein kinases (MAPKs). MAPKs are important players in immune signalling and are activated by phosphorylation upon pathogen perception (Suarez Rodriguez et al., 2010). Western blotting (WB) analyses indicated that MPK6 and MPK3 were phosphorylated 30 min after flg22 treatment, indicating that plant immunity was rapidly and efficiently elicited (**Fig. 1C**).

2.2. Building the Arabidopsis leaf RBPome

Before studying the dynamics of the RBPome upon flg22 treatment, the individual leaf RBPomes for each of the 4 different conditions (mock 2 hpt, flg22 2 hpt, mock 12 hpt and flg22 12 hpt) were determined by analysing the enrichment of the identified proteins in CL (+UV) over NoCL (-UV) samples. For proteins with intensity values in both CL and NoCL samples, the $\log_2$ fold

change between CL and NoCL samples (log_2FC [CL/NoCL]) was calculated and moderate t-test and false discovery rate estimation (FDR) was used to calculate the statistical enrichment of +UV/-UV. We classified proteins as RBPs when FDR ≤ 0.01 and log_2FC [CL/NoCL] ≥ 1.5 (**Fig. 2A-D**) or FDR ≤ 0.1 and log_2FC [CL/NoCL] ≥ 3.3 (**Fig. 2A-D**). Proteins undetected in the NoCL samples and with intensity values in all the CL samples were also considered RBPs using a modification of the semiquantitative method (Sysoev et al., 2016) as described in **Chapter 2**.

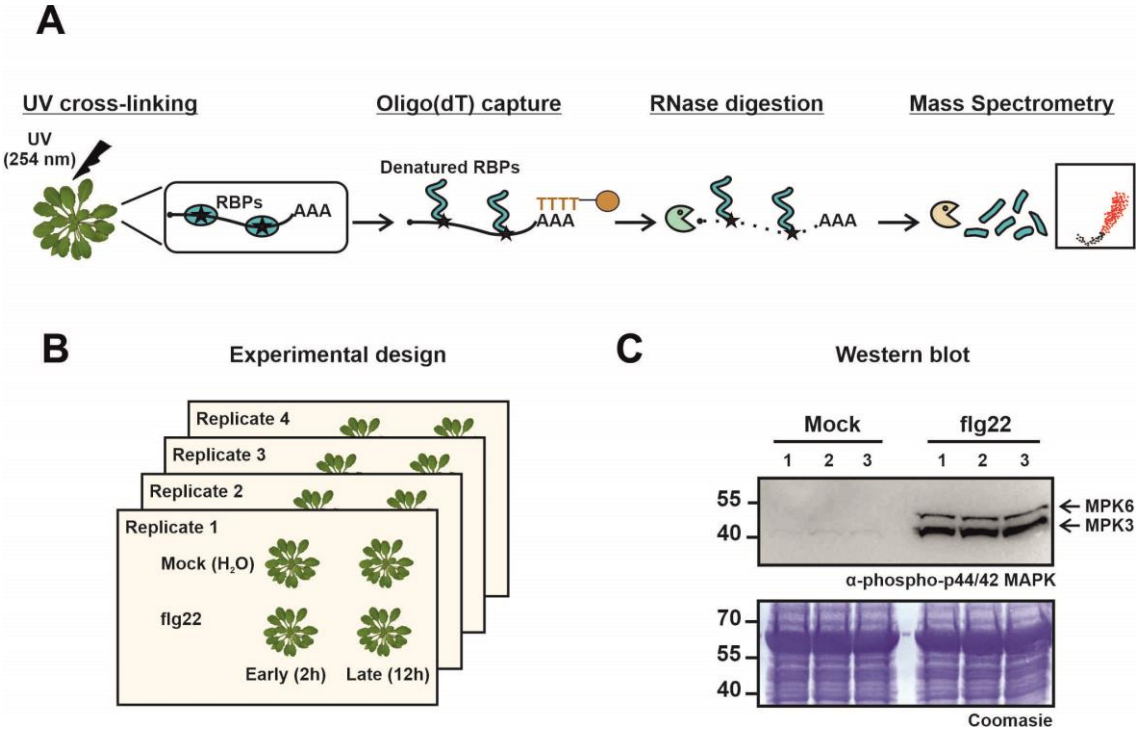


Fig.1. A) Overview of plant RNA Interactome Capture (ptRIC). Plants (leaves) are irradiated with UV light at 254 nm to promote crosslinking (CL) between RNAs and proteins that are in intimate contact. Next, cells are lysed and mRNAs pulled-down using oligo(dT) magnetic beads. After stringent washes, the RNA-protein complexes are recovered and the RBPs released by RNase digestion. Proteins are quantitatively analysed by mass spectrometry after trypsin digestion. **B) Experimental design of flg22 treatment for MS analysis.** Mature Arabidopsis plants (5-6 weeks old) were infiltrated with either H_2O (mock) or the immune elicitor flg22 (1 µM) and leaf tissue was harvested at 2 hpt and 12 hpt. Four biological replicates per time point and treatment were performed. **C) Western blotting analyses of phosphorylated MAPK.** Leaves of Arabidopsis plants were infiltrated with either H_2O or the immune elicitor flg22 (1 µM). Leaves were harvested 30 minutes after infiltrating, the proteins extracted and separated by SDS-PAGE. Proteins were transferred onto a nitrocellulose membrane and analysed by western blot using an antibody recognising phosphorylated MAPK (anti-phospho-p44/42 MAPK antibody; CST #9102). For each treatment, three biological replicates were performed (1-3). Abbreviations: UV, ultraviolet; RBPs, RNA-binding proteins; CL, crosslinking.

Using these criteria, a total of 1135 proteins were identified as +UV/-UV enriched (RBPs) in the 4 conditions, constituting the Arabidopsis leaf RBPome (**Supplemental digital table 1**). Approximately 58.5% of the leaf RBPome was linked to RNA biology (**Fig. 2F**) as determined by gene ontology (GO) annotations, which highlights the quality of our leaf RBPome. Moreover, 890 of the 1135 RBPs overlapped between the 4 datasets (**Fig. 2E**), supporting the robustness of

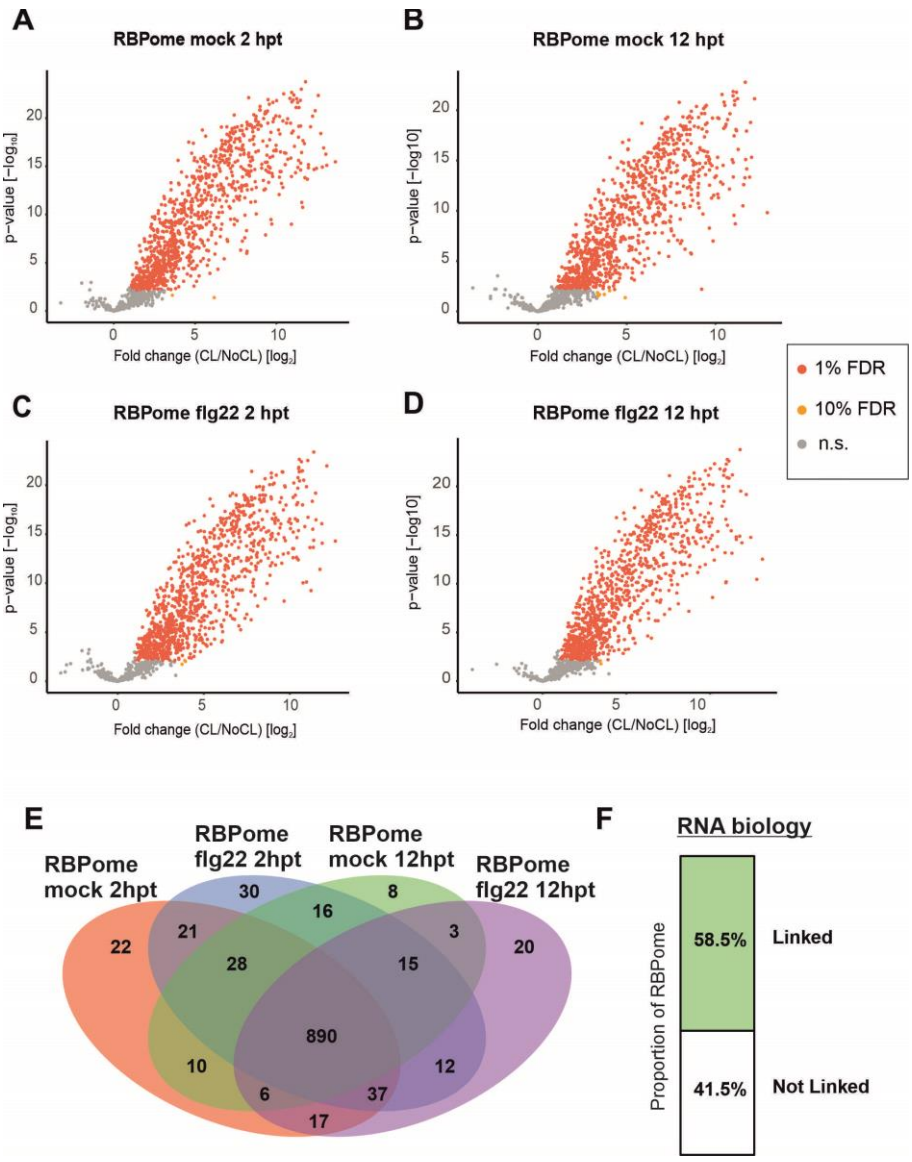


Fig.2. Arabidopsis high confidence leaf RBPomes. A)-D) Volcano plots depicting the log_2 fold change and the significance (p-value) of each protein (dots) between UV-crosslinking (CL) and non-crosslinked (NoCL) treatment using data from four biological replicates. Proteins are coloured in red when FDR ≤ 0.01 and log_2FC [CL/NoCL] ≥ 1.5 and proteins are coloured in yellow when FDR ≤ 0.1 and log_2FC [CL/NoCL] ≥ 3.3. Non-significant proteins are coloured in grey. Red and yellow proteins represent the high confidence leaf RBPome. E) Venn diagram showing the overlap between the four datasets. F) Proportion of RBPs linked to RNA biology or not based on GO annotations as described previously (Beckmann et al., 2015).

the ptRIC-determined leaf RBPome. A number of RBPs were exclusively identified in flg22 or mock-treated plants and could represent proteins that bind to RNA exclusively under a certain cellular state, as has been previously described (Garcia-Moreno et al., 2019). We used stringent criteria to build a high-quality leaf RBPome while minimising the incidence of false positives. Importantly, although the criteria to classify proteins as RBPs here is more stringent than in **Chapter 2** and **Chapter 3**, we identified here a larger number of RBPs. This may be in part because the RBPome described here combines the RBPs found in mock and flg22 treated plants at two different time points, thus uncovering not only housekeeping RBPs, but also condition-specific RNA binders.

2.3. Uncovering the flg22-responsive RBPome

To profile the changes in the leaf RBPome induced by flg22 immune activation, we analysed the RBPs isolated from Arabidopsis plants upon treatment with either H_2O (mock) or the immune elicitor flg22. Flg22 treatment did not induce drastic changes in the set of abundant RBPs detected using silver staining analysis (**Fig. 4A**), in accordance to what was previously observed in other comparative RIC studies (Garcia-Moreno et al., 2019; Sysoev et al., 2016). However, quantitative proteomic analyses revealed dynamic responses of the RBPome to flg22 (**Fig. 3**).

The requisites for an RBP to be classified as 'responsive' was first, to be assigned RNA-binding activity in the comparison to the NoCL control (**Fig. 2**) and second, to display $\log_2$ fold change between flg22-treated and mock-treated leaves ($\log_2$FC [flg22/mock]) higher than 1 or lower than -1 (2-fold up- or downregulated) and a FDR ≤ 0.1. Based on these criteria, a total of 107 RBPs (∼10% of the RBPome) were dynamically altered upon immune activation and we refer to them here as 'high-confidence flg22-responsive RBPs' (**Supplemental table 1**). Moreover, we identified additional 181 RBPs regulated by flg22 with FDR ≤ 0.2 and a $\log_2$FC [flg22/mock] ≥ 0.58

or ≤ -0.58 (1.5-fold up- or downregulated) (**Supplemental table 1**). These proteins were here classified as 'candidate flg22-responsive RBPs'. Importantly, this set contains a number of RBPs with known links to immunity suggesting the presence of true immunoregulatory RBPs.

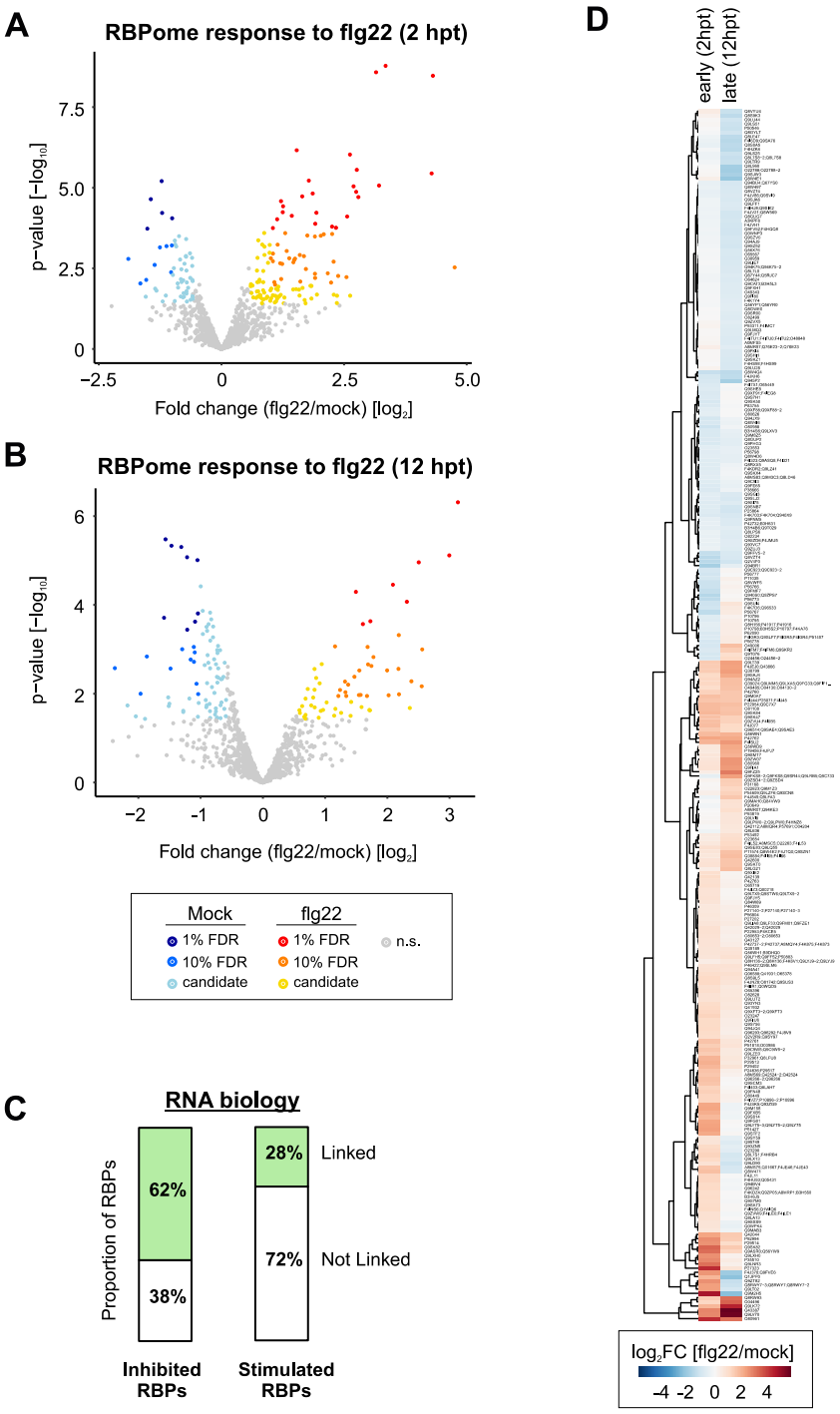


Fig.3. Dynamic responses of the leaf RBPome to flg22 perception. A-B) Volcano plots depicting the $\log_2$ fold change and the significance (p-value) of each protein (dots) between flg22 and mock treatment (2 hpt and 12 hpt) using data from four biological replicates. The colours of the dots indicate the significance

Fig.3 (continued) of the protein as follows: proteins are coloured in red when FDR ≤ 0.01 and log_2FC [flg22/mock] ≥ 1, in orange when FDR ≤ 0.1 and log_2FC [flg22/mock] ≥ 1, in yellow when FDR ≤ 0.2 and log_2FC [flg22/mock] ≥ 0.58, in dark blue when FDR ≤ 0.01 and log_2FC [flg22/mock] ≤ -1, in blue when FDR ≤ 0.1 and log_2FC [flg22/mock] ≤ -1 and in light blue when FDR ≤ 0.2 and log_2FC [flg22/mock] ≤ -0.58. Red, orange and yellow proteins represent the leaf RBPs stimulated upon flg22 perception, whereas dark blue, blue and light blue represent leaf RBPs inhibited upon flg22 perception. Non-significant proteins are coloured in grey. C) Proportion of flg22-inhibited or flg22-stimulated RBPs linked to RNA biology or not based on GO annotations as described previously (Beckmann et al., 2015). D) Heat map representing the differential association with RNA of leaf RBPs at early (2 hpt) and late (12 hpt) time points after flg22 treatment. For each of the proteins the colours represent log_2FC [flg22/mock].

In total, the flg22-responsive RBPome (including high-confidence and candidate flg22-responsive RBPs) included 288 proteins (25% of the leaf RBPome), of which 144 were altered at the early response (2 hpt), 108 were altered in the late response (12 hpt) and 36 were altered in both (**Fig. 3A-B, D**). It has traditionally been described that the responses to flg22, including transcriptional changes, are a rapid phenomenon (Denoux et al., 2008). For instance, MAPK phosphorylation occurs within 1 minute following flg22 perception (Tena et al., 2011). We have identified that RBPs probably regulating transcriptome reprograming were altered both at the early and late stages of immune responses. However, interestingly, just 36 RBPs were common between the early and late time points, suggesting that the plant RBPome is dynamically and temporally modulated to adjust the transitory reprogramming that follows immune elicitation (**Fig. 3D**).

To validate the proteomic results, we performed ptRIC followed by western blot (WB) analyses. We selected two stimulated RBPs at 2 hpt and one inhibited RBP at 12 hpt for which commercial antibodies were available. WB analysis confirmed that both HSC70-3 and METK4 have increased association with RNA at 2 h after flg22 treatment, whereas RH3 has decreased RNA-binding affinity at 12 hpt (**Fig. 4B** 'eluates'). Therefore, results obtained by WB strongly support our proteomics data, validating the accuracy of our approach. The antibody used to detect METK4 recognises 4 members of Arabidopsis METK family (METK1-4). Although we only detected METK4 to bind RNA and to be stimulated upon flg22 perception, we cannot exclude that some other members of the family contributed to the signal. Moreover, although the antibody used

to detect HSC70-3 can also recognise other Arabidopsis HSP70/HSC70 members, the band observed coincides with the molecular size expected for HSC70-3. However, we cannot exclude that other HSP70/HSC70 members contributed to the signal.

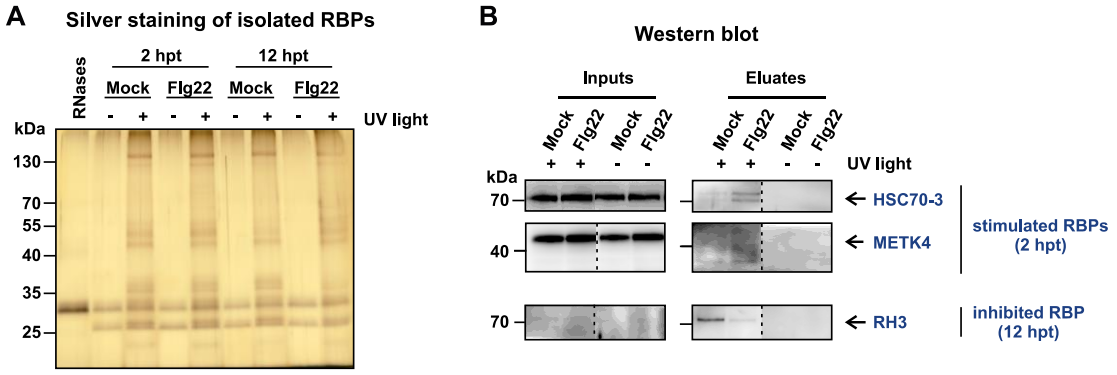


Fig.4. A) Silver staining analyses of the eluates (leaf RBPome) isolated by ptRIC. Arabidopsis mature plants (5-6 weeks old) were infiltrated with either H_2O (mock) or the immune elicitor flg22 (1 µM) and leaf tissue was harvested at 2 hpt or 12 hpt. The RBPs were isolated using ptRIC and analysed by silver staining. The RNases used to release the RBPs from RNAs are loaded as control. **B) Validation of the MS results by WB analyses.** Arabidopsis mature plants (5-6 weeks old) were infiltrated with either H_2O (mock) or the immune elicitor flg22 (1 µM) and leaf tissue was harvested at 2 hpt or 12 hpt. The RBPs isolated by ptRIC (eluates) or the whole cell lysates (inputs) were analysed by WB using specific antibodies against HSC70, METK4 and RH3. The changes in association with RNA detected by WB (eluates) match the changes detected by MS analyses, thus validating the experimental approach.

2.4. RBP pathways dynamically regulated during plant immunity

To identify RBP-mediated processes or pathways altered upon immunity induction by flg22 we analysed the flg22-responsive RBPs with decreased (referred to here as 'inhibited RBPs') or increased (referred to here as 'stimulated RBPs') association with RNA upon flg22 perception and searched for pathways enriched using GO annotations. Below we discuss the identity of the inhibited and stimulated flg22-responsive RBPs.

2.4.1. flg22-inhibited RBPs are enriched in chloroplastic RBPs, editing and translation

The majority of inhibited RBPs (62%) had links to RNA biology (**Fig. 3C, Supplemental table 1**) and many of them had functions related to RNA editing, translation and were known or predicted to be located in the chloroplast.

Chloroplastic RBPs

It is well known that some chloroplastic activities such as photosynthesis are inhibited during immune responses (Serrano et al., 2016; Su et al., 2018); hence, it is possible that some chloroplastic RBPs are involved in regulating chloroplastic processes that are inhibited during immune responses. Remarkably, 25 pentatricopeptide repeat (PPR) and tetratricopeptide repeats (TPR) domain containing proteins were inhibited upon flg22 perception (**Supplemental table 1**), one of them at 2 hpt, twenty-two at 12 hpt and two at both time points. Most of these inhibited proteins are chloroplastic and have roles related to chloroplastic RNA metabolism such as RNA editing, rRNA processing, etc. For instance, SVR7 (suppressor of variegation 7) is a PPR-containing protein involved in chloroplast rRNA processing that is required for normal photosynthetic function and photo-oxidative stress responses (Liu et al., 2010; Lv et al., 2014). The PPR-containing protein family is large in plants (450 members in Arabidopsis), and its members localise mainly in organelles (Silverman et al., 2013; Colcombet et al., 2013). Two recent publications highlight the importance of the PPR gene family in response to biotic stress in poplar and rice (Xing et al., 2018). Hence, the extensive downregulation of PPR-containing RBPs may be involved in controlling organellar RNA metabolism during immune responses.

RBPs involved in RNA editing

In plants, plastid and mitochondrial transcripts can undergo post-transcriptional RNA editing (including C-to-U editing), which often results in amino acid changes (Ichinose and Sugita, 2017). Multiple RBPs involved in chloroplastic editing were inhibited at 12 h post treatment with flg22 (**Supplemental table 1**). One example is VAR3 (variegated 3), a zinc-finger motif containing protein that controls a large number of chloroplastic RNA editing sites and is essential for normal chloroplast development (Næsted et al., 2004; Sun et al., 2015). We also identified early stimulation of the mitochondrial RNA editing protein ORRM2 (Shi et al., 2015; **Supplemental table 1**).

Ribosomal proteins and RBPs involved in translation

Plants can respond to changes in the environment by regulating mRNA translation, which allows rapid regulation of protein abundance (Xu et al., 2017). In addition, the most limiting step in the production of proteins is the initiation of mRNA translation, which is regulated by different eukaryotic initiation factors (eIFs; Richter and Sonenberg, 2005). Five eIFs and three elongation factors as well as other proteins linked to translation were inhibited at late time points after flg22 treatment (**Supplemental table 1**). Most of them were chloroplastic, probably reflecting the decrease in chloroplastic translation upon flg22 perception (Serrano et al., 2016).

Several other inhibited RBPs function in translation (**Supplemental table 1**). These include two cytosolic MA3 domain-containing translation regulatory factor (MRF) proteins, MRF1 and MRF3, that were inhibited at 2 hpt (**Supplemental table 1**). The Arabidopsis genome encodes four MRF proteins that are translation regulators (Lee et al., 2017; Cheng et al., 2013). MRF proteins are under control of the TOR (Target of rapamycin) complex that regulates the translation machinery through phosphorylation in response to signals such as nutrient availability and growing factors (Lee et al., 2017). Thus, the TOR signalling pathway controls protein synthesis according to the cellular conditions. Both MRF1 and MRF3 are induced by dark and starvation, suggesting a translational repressive role (Lee et al., 2017). In our data, both MRF1 and MRF3 were inhibited upon immune activation, suggesting differential regulation of MRF1 and MRF3 according to multiple cellular contexts. To our knowledge, these 2 RBPs have not previously been linked to immunity in Arabidopsis. We also found that 2 additional TOR substrates, ML5 and LARP1a, are inhibited upon flg22 treatment (**Supplemental table 1**).

A large proportion of flg22-responsive RBPs were ribosomal proteins (RPs) or proteins linked with the ribosome (**Supplemental table 1**). With only a few exceptions, most of these proteins were inhibited by flg22 (**Supplemental table 1**). The existence of 'specialised ribosomes' has

been proposed, with varying protein/RNA composition or unique PTM signatures (Xue and Barna, 2012; Mauro and Edelman, 2002; Shi et al., 2017). This has been supported by the discovery of differential RBP sets associated with ribosomes in mice and humans under difference cellular states (Simsek et al., 2017; Garcia-Moreno et al., 2019). In agreement with our data, some RPs were previously found to be differentially accumulated in plants treated with the elicitor chitosan (Fakih et al., 2016). One of these proteins is RPL19-2, which in this study was inhibited at 2 hpt (**Supplemental table 1**). We also identified that RPL12C is stimulated at 12 hpt (**Supplemental table 1**), and this protein has been shown to play an important role in disease resistance in both *N. benthamiana* and Arabidopsis (Nagaraj et al., 2016). Therefore, our results support previous findings suggesting that ribosomes and their associated proteins are dynamically regulated in response to biological cues. However, the effects these changes have on translation control remain unknown. Taken together, flg22-driven regulation of RBPs involved in translation argue for the tight translational regulation of the plant defence responses upon flg22 treatment.

Interestingly, some enzymes related to nascent peptide modification were also inhibited during immune responses, especially at 12 hpt (**Supplemental table 1**). Co-translationally, the nascent peptides can be modified at the *N*-terminal affecting for example the protein half-life and folding characteristics. In eukaryotes two major protein modifications are *N*-terminal Met excision (NME) carried out by Met aminopeptidases (MAP) and *N*- α-terminal acetylation (Nt-acetylation) catalysed by *N*-terminal acetyltransferases (Nats; Ross et al., 2005; Ree et al., 2018). Some of these enzymes involved in *N*-terminal peptide modification were inhibited upon flg22 treatment (**Supplemental table 1**). Two examples are the chloroplastic Met aminopeptidases MAP1B (Ross et al., 2005), and the cytoplasmic NAA10, which is the catalytic subunit of *N*-terminal acetyltransferase A (NATA). NAA10 is involved in defence against pathogens by regulating acetylation of the first Met of the disease resistance (R) proteins SNC1 and RPM1, thus targeting

the proteins to degradation (Xu et al., 2015). In agreement, *naa10* mutants are more resistant to *Hyaloperonospora arabidopsidis* (Xu et al., 2015). Hence, plants may actively inhibit the association of NAA10 with RNA as part of the flg22-triggered immunity. To our knowledge, NAA10 has no RBD but is known to interact with the ribosome, and proteins that interact with the ribosome could potentially bind mRNA. Another common protein modification is *N*-myristoylation (MYR) by *N*-myristoyltransferases (NMTs), which generally grants proteins the ability to target membranes. We identified flg22-driven inhibition of the cytoplasmic NMT1 (**Supplemental table 1**), which *N*-myristoylates different immune receptors (NLRs; Boisson et al., 2003). It has been shown that NMT1 interacts with ribosomes and this interaction is mediated by RNA (Ohta et al., 2015).

2.4.2. Flg22-stimulated RBPs are enriched in RNA-binding moonlighting enzymes and proteins involved in stress responses

Contrarily to the inhibited RBPs, only 28% of the stimulated flg22-responsive RBPs had annotations with obvious links to RNA biology (**Fig. 3C**; **Supplemental table 1**). However, many of these stimulated RBPs with no obvious link to RNA biology have either been shown to bind RNA in other organisms (such as HSPs or PPIs), or mono- or di-nucleotides (ATP/ADP, NADP/NADH), which is a feature found in several RNA-binding moonlighting enzymes (Castello et al., 2015, 2016). Moreover, most of the stimulated RBPs have been previously linked to stress responses such as defence against bacteria or fungus (**Supplemental table 1**), which supports the conclusion that our dataset contains *bona fide* immunoregulators.

Peptidyl-prolyl cis-trans isomerase enzymes (PPIs) moonlighting as RBPs

Interestingly, several peptidyl-prolyl cis-trans isomerase enzymes (PPIs) were stimulated at 2h after flg22 treatment (**Supplemental table 1**) including the PPIs FKBP53, TIG, CYP63 and CYP95. PPIs are known to bind RNA in humans and yeast (Hentze et al., 2018) and we have

demonstrated that multiple PPIs can bind RNA in plants (**Chapter 3**). For example, the human PPIA has been described to be an important regulator of different virus infection (Garcia-Moreno et al., 2019; Rupp and Bartenschlager, 2014). Interestingly, an effector from the cyst nematode *Heterodera schachtii* targets the PPI FKBP53 and the histone deacetylase HDT1 (Vijayapalani et al., 2018), both of which were stimulated upon flg22 perception (**Supplemental table 1**). Moreover, Fan and colleagues identified that *Phytophthora capsici* secretes an RXLP effector that targets a host a PPI to disrupt plant immunity (Fan et al., 2018). These observations highlight the importance of PPI for plant immunity, possibly by regulating RNA metabolism.

Heat shock proteins (HSP) moonlighting as RBPs

A number of heat shock proteins (HSPs) were stimulated upon flg22 treatment including HSP90-1, HSP90-3/4 (protein group), HSP70-3, HSP70-18 (protein group), HSC70-1 and HSC70-2/HSP70-6/HSP70-7 (protein group; **Supplemental table 1**). Notably, HSP90-3/4 and Hsc70-2/HSP70-18/HSP70T-1 were stimulated in both early and late responses to flg22. Although HSPs have been traditionally associated with folding of proteins during stress responses, we have also provided evidence that multiple members of HSP70 and HSP90 families can bind RNA in different plant species (**Chapter 3**). Moreover, it has been described that the human HSP90AB1 has an increased association with RNA upon viral infection and plays an important pro-viral function (Garcia-Moreno et al., 2019). Moreover, HSP70 and HS90 have been associated with plant immunity (Park and Seo, 2015). For example, the *N. benthamiana* HSP90 interacts with the viral RNA of Bamboo mosaic virus (Huang et al., 2012), uncovering a direct link to immunity.

Other RNA-binding moonlighting enzymes

Several glutathione S-transferases (GSTs) were stimulated at both 2 and 12 hpt (**Supplemental table 1**). Glutathione (GSH) plays an important role in immune signalling and some studies have established links between GSH and RNA regulation (Datta et al., 2015; Baena-González et al.,

2001). Moreover, a number of ATPases were stimulated upon flg22 treatment, some of which contribute to acidifying the apoplast in response to pathogen attack (**Supplemental table 1**). Several ATPases have been reported to bind poly(U) (Xu et al., 2007), although their roles in RNA metabolism remain unknown. We have previously identified that many ATPases interact with RNA in plants, and their RNA-binding function is conserved across multiple plant species (**Chapter 3**). In addition, a group of proteins related to glutamine metabolism were stimulated upon flg22 treatment, including the cytosolic glutamine synthetases GLN1-1 and GLN1-3, and the chloroplastic GLN2 (**Supplemental table 1**). However, to our knowledge, it is unknown though if GLNs can bind RNA by themselves, although they are annotated as nucleotide-binding (TAIR, Uniprot). Many enzymes moonlighting as RBPs have been reported to date such as GAPDH or aconitase/IRP1 (Castello et al., 2015). Hence, we hypothesise that the enzymes identified as flg22-responsive RBPs in this study may be involved in RNA binding during immune responses or alternatively, may be allosterically regulated by RNA.

RBPs associated with processing bodies (P-bodies)

Several proteins associated with processing bodies (P-bodies) were stimulated upon immune activation. These include PAT1, PAT1H1, LSM6A/B (protein group), EIN2, LARP1a and RH6 (**Supplemental table 1**). P-bodies are dynamic cytoplasmic assemblies that comprise translationally inactive mRNAs and proteins involved in mRNA turnover and translational repression. Typically, within P-bodies mRNAs can undergo deadenylation, decapping, miRNA-induced gene silencing and nonsense-mediated decay (NMD), but mRNAs can also be stored in P-bodies without degradation (Li et al., 2015; Merchante et al., 2015; Horvathova et al., 2017; Hubstenberger et al., 2017). It is known that during stress conditions some mRNAs can be re-localised to P-bodies (Weber et al., 2008; Chantarachot and Bailey-Serres, 2017) and recent studies have demonstrated that P-bodies greatly contribute to immune responses by rapidly degrading or storing some target mRNAs (Yu et al., 2019; Maldonado-Bonilla et al., 2014; Li et

al., 2015; Merchante et al., 2015). These may include, for example, mRNAs of negative immune regulators (Yu et al., 2019). Moreover, proteins associated to P-bodies have been demonstrated to be targeted by pathogens (Petre et al., 2016).

PAT1 is one of the three Arabidopsis homologs of the yeast PAT1 and functions as an enhancer of decapping, establishing a link between deadenylation and decapping (Roux et al., 2015). Together with LSM1-7 it binds the 3' of deadenylated mRNAs and promotes decapping of specific transcripts (Tharun, 2009; Roux et al., 2015). In addition, PAT1 plays a role in translational inhibition and P-body formation (Roux et al., 2015). Recently, PAT1 has been demonstrated to function in plant immunity since *pat1* loss-of-function mutants exhibit constitutively high expression of *PR* genes that results in autoimmunity and increased resistance to *Pst*DC3000 (Roux et al., 2015; see **Chapter 1** for review). Upon flg22 treatment PAT1 interacts and is phosphorylated by the immune kinase MPK4 (and MPK6), and this leads to re-localisation to P-bodies (Roux et al., 2015). We identified that PAT1 is stimulated upon flg22 treatment. Moreover, we also identified that MPK4 is stimulated at both time points (**Supplemental table 1**). To our knowledge, MPK4 has not previously been described to bind RNA. However, multiple kinases have been proposed to moonlight as RNA binders such as the adenylate kinase, hexokinase-2 and nucleoside diphosphate kinase A, and some kinases such as the human PKR can be allosterically regulated by RNA (Popow et al., 2015; Castello et al., 2015, 2016; Dabo and Meurs, 2012). Thus, it is feasible that MPK4 could also moonlight as an RBP. Even though the mRNA targets of PAT1 remain unknown, the regulation of PAT1 by MPK4 represents an elegant mechanism whereby MPK4 regulates mRNA decay machinery in P-bodies.

PAT1H1, a second Arabidopsis PAT1 homolog, was stimulated at both 2 and 12 hpt (**Supplemental table 1**). Although the role of PAT1H1 in plant immunity remains unknown, PAT1H1 has been linked to jasmonic acid signalling (Yu et al., 2016). The stimulation of PAT1H1

in response to flg22 suggests that PAT1H1 may be involved in immunity, possibly through a similar mechanism to PAT1. Interestingly, the human PATL1 is stimulated after viral infection correlating with a profound remodelling of the host cell transcriptome (Garcia-Moreno et al., 2019). We also identified that LARP1a is inhibited at 12 hpt (**Supplemental table 1**). LARP1 has been shown to re-localise to P-bodies after heat shock and is involved in co-translational decay of specific transcripts together with the exoribonuclease XRN4 (Merret et al., 2013; Deragon and Bousquet-Antonelli, 2015; Merret et al., 2015). Taken together, our results argue for an important role of P-bodies in RNA regulation during immune responses.

2.4.3. RBP families with both stimulated and inhibited RBPs include known and unconventional RBPs

A number of RBP families had members that were stimulated or inhibited upon flg22 treatment (**Supplemental table 1**) and are discussed below.

Photosynthesis-related components moonlighting as RBPs

Many of the RBPs altered early upon flg22 treatment were photosynthesis-related components, 16 of which were inhibited and 10 were stimulated (**Supplemental table 1**). Flg22-inhibited RBPs included 6 Photosystem I (PSI)-related proteins (PSI-C, the apoproteins PSAA and PSAB, and the reaction centre subunits PSI-D1, PSI-F and PSI-L) and 4 Photosystem II (PSII)-related proteins (PSII-D1, CP22 and the reaction centre proteins CP43 and CP47) (**Supplemental table 1**). Flg22 also inhibited the association with RNA of 3 rubisco small chain 1A, 1B and 2B/3B proteins and cytochrome b6, a member of the cytochrome b6f complex (**Supplemental table 1**). Moreover, some photosynthesis-related proteins were stimulated at 2 hpt, and the set of photosynthesis-related proteins stimulated seems to be functionally distinct from that inhibited at the same time point (**Supplemental table 1**). For example, the flg22-stimulated RBPs include 2 carbonic anhydrases (bCA1, bCA2), 3 light-harvesting complex proteins (LHCA3-1, LHCA1/6 and LHCB3-

1), 3 oxygen-evolving complex proteins (PSBP-1, PSBQ1, PSBQ2) and the rubisco activase, amongst others (**Supplemental table 1**).

Remarkably, after 12 h of flg22 treatment, there was no evident alteration of proteins linked to photosynthesis (**Supplemental table 1**). Therefore, some photosynthetic enzymes may moonlight as RNA binders specifically during the initial response to flg22, but not at later stages. Conversely, a general inhibition of chloroplastic RBPs was observed at 12 hpt and some of these regulated proteins are known to control photosynthesis-related transcripts (**Supplemental table 1**). For example, PGR3, which is known to regulate the cytb6f complex (Yamazaki et al., 2004), and LPE1, which is known to regulate photosynthesis biogenesis (Jin et al., 2018; Williams-Carrier et al., 2019), were inhibited at 12 hpt.

RNA helicases

We identified 9 RNA helicases as differentially regulated upon immune elicitation, five of which were stimulated and four inhibited (**Supplemental table 1**). RNA helicases mirror the general trend, and most of the inhibited RNA helicases had a chloroplastic localisation, whereas the stimulated RNA helicases were nuclear/cytosolic. In agreement with our results, the association with RNA of 16 RNA helicases is altered in response to viral infection in human cells (Garcia-Moreno et al., 2019). Even though the genome of Arabidopsis encodes 58 RNA helicases (Mingam et al., 2004), this family of RBPs has been largely uncharacterised and only the role of a few RNA helicases has been described, particularly in response to abiotic stress (Gong et al., 2005; Kant et al., 2007). Our data shows that RH50 and RH3 were inhibited at 12 hpt (**Supplemental table 1**). RH3 is a chloroplastic DEAD-box RNA helicase that plays important roles during salt and cold-shock stresses by regulating splicing and ribosome biogenesis in chloroplasts (Gu et al., 2014). RH50 is also a chloroplastic DEAD-box RNA helicase involved in rRNA maturation and associates with ribosomes (Paieri et al., 2018). Interestingly, the rice

ortholog of RH50, OsBIRH1, is known to play important roles in biotic and oxidative stresses (Li et al., 2008; see **Chapter 1** for review). However, to our knowledge this is the first time that Arabidopsis RH3 and RH50 have been linked to plant immunity. The fact that we identify that around 15% of Arabidopsis RNA helicases are dynamically regulated upon immune activation supports their involvement in plant immunity.

RBPs involved in splicing

A number of flg22-responsive RBPs are involved in splicing (**Supplemental table 1**). At 12 hpt, nine inhibited RBPs were linked to splicing, and most of them are chloroplastic. Of note, three of these contain chloroplast RNA splicing and ribosome maturation (CRM) domains. CRM-containing proteins are known to be involved in splicing and many of them are targeted to organelles (Barkan et al., 2007; Asakura and Barkan, 2007). Many of the flg22-responsive RBPs were involved in splicing or processing of photosynthesis-related transcripts. Hence, the inhibition of splicing proteins regulating transcripts of photosynthetic components in the chloroplast could be a mechanism to control photosynthesis during plant immune responses. Moreover, some of these flg22-responsive RBPs involved in splicing have been previously linked to abiotic stress. For example, the chloroplastic CFM4 is involved in cold and salt stress tolerance (Lee et al., 2014). We also identified the nuclear protein SF1, which has been shown to have a role in splicing of HSPs, amongst others (Jang et al., 2014).

Contrarily, multiple nuclear RBPs linked to splicing were stimulated, particularly at 2 hpt (**Supplemental table 1**). One example is the evolutionarily conserved SKI-interacting protein (SKIP). SKIP functions as both transcription factor and splicing factor and regulates alternative splicing of genes related to tolerance to abiotic stresses (Li et al., 2019; Feng et al., 2015). We also discovered that the pre-mRNA-splicing factor (AT3G49601) was flg22-responsive, and this protein has been previously identified to be differentially accumulated in Arabidopsis plants

challenged with the immune elicitor chitosan (Fakih et al., 2016). Differential regulation of splicing factors has also been observed in human cells infected with Sindbis virus (Garcia-Moreno et al., 2019). Thus, splicing emerges as a tightly regulated process in response to environmental cues and this regulation seems to be driven by the differential association with RNA of splicing factors.

2.5. Does protein abundance change in response to flg22?

Many studies have examined transcriptional changes of plants to a wide range of elicitors, effectors and multiple pathogen species (Denoux et al., 2008; Bektas et al., 2016; Sohn et al., 2014; Yang et al., 2017; Hacquard et al., 2016). These studies have led to the identification of the transcriptional changes that characterise the immune responses in plants. However, to date only a few studies have focused on the effects that these cues have on the plant proteome (Jayaraman et al., 2012; Afroz et al., 2013). In order to characterize the changes in protein abundance that occur during immune responses in leaves we performed mass spectrometry analyses of the whole cell lysates (WCL; inputs of the ptRIC; total proteome) of plants treated with mock (H_2O) or flg22. We identified a total of 3782 proteins, of which only 89 (~2.35%) had significantly altered abundance upon flg22 recognition (FDR ≤ 0.1 and $\log_2$FC [flg22/mock] ≥ 1 or ≤ -1) (**Fig. 5B-D**). Only 21 of these were identified at 2 hpt, whereas 63 were identified at 12 hpt, and 5 were identified at both time points. Notably, most of these proteins increased in abundance upon flg22 treatment, regardless of the time point (**Fig. 5C-D**). In particular, 23 of the altered proteins (~88%) had increased accumulation upon flg22 treatment at 2 hpt, whereas just 3 (~12%) had reduced levels. Accordingly, 51 (~75%) displayed higher abundance at 12 hpt, while 17 (~25%) had reduced levels at this time point. Moreover, just 5 proteins were identified to have increased accumulation at both time points, highlighting the dynamic nature of the cellular proteome (**Fig. 5B**).

Interestingly, a high proportion of the differentially accumulated proteins are known to respond to different type of stresses, including biotic stress (**Supplemental table 2**). For example, we identified significant increased accumulation of WRKY40, which is a transcription factor known to be induced upon flg22 treatment (Birkenbihl et al., 2017). WRKY40 plays an important role in immunity by regulating more than 1000 genes, many of which are involved in PTI signal perception and transduction, production of hormones and secondary metabolites (Birkenbihl et al., 2017). Eight peroxidases and 11 proteins involved in oxidation-reduction processes also accumulated to higher levels upon flg22 treatment (**Supplemental table 2**). These may contribute to the PTI characteristic ROS burst that follows flg22 perception or to the downstream detoxification of ROS. We also identified increased accumulation of six immune receptors such as FRK1 (flg22-induced receptor-like kinase 1), IOS1 (impaired oomycete susceptibility 1) or SIF2 (stress induced factor 2), all of which are known to be important for PTI responses (Asai et al., 2013; Yeh et al., 2016; Yuan et al., 2018; **Supplemental table 2**). Other proteins that differentially accumulated upon flg22 treatment were hydrolases, protease inhibitors, proteins involved in calcium signalling and enzymes involved in cell wall remodelling, all of which contribute to plant defence responses.

It is surprising that only 2.35 % of the total proteome was differentially expressed in response to flg22. In agreement, silver staining analyses did not reveal noticeable changes of protein levels, although this method only detects the most abundant proteins (**Fig. 5A**). This contrasts with the massive transcriptional changes that occur upon immune activation as reported by many RNA-seq studies. It has been described that transcriptional changes do not correlate with the translational changes, and that translation is more tightly regulated that transcription (Xu et al., 2017). Alternatively, differentially expressed proteins could be expressed in low to medium levels and may not be detectable by our proteomic analysis.

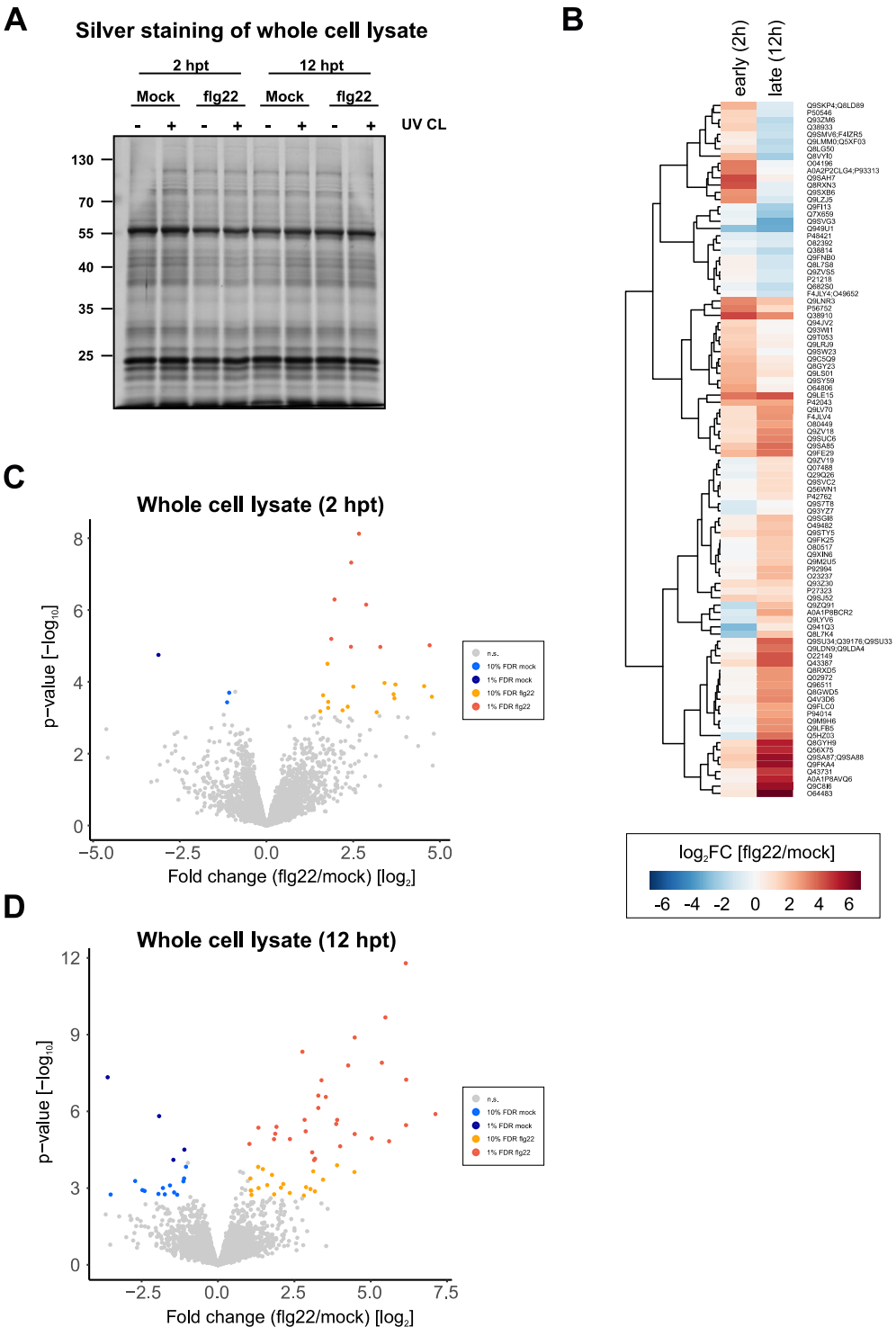


Fig.5. Response of the total proteome (whole cell lysate) to flg22 perception. A) Silver staining analyses of the total proteomes (whole cell lysates). The total proteomes of Arabidopsis plants treated with either mock (H_2O) or flg22 were separated in a 12% acrylamide and analysed by silver staining. B) Heat map representing the differential accumulation of proteins from the total proteome at early (2 hpt) and late (12 hpt) time points after flg22 treatment. For each of the proteins the colours represent log_2FC [flg22/mock]. C-D) Volcano plots depicting the log_2 fold change and the significance (p-value) of each protein (dots) between flg22 and mock treatment using data from four biological replicates. Proteins are coloured in red when FDR ≤ 0.01 and log_2FC [flg22/mock] ≥ 1, in yellow when FDR ≤ 0.1 and log_2FC [flg22/mock] ≥ 1, in dark blue when FDR ≤ 0.01 and log_2FC [flg22/mock] ≤ -1 and in light blue when FDR ≤ 0.1 and log_2FC [flg22/mock] ≤ -1. Red and yellow proteins represent proteins with higher accumulation

Fig. 5 (continued) upon flg22 perception, whereas dark blue and light blue represent proteins with lower accumulation upon flg22 perception. Non-significant proteins are coloured in grey.

2.6. Changes in association with RNA do not correlate with changes in protein abundance

Changes in RBP association with RNA can be due to changes in RNA-binding affinity or to matching changes in protein abundance (Sysoev et al., 2016; Garcia-Moreno et al., 2019). In other words, a given protein can be detected in higher extent by ptRIC analyses only because this protein is expressed in higher levels. To address whether changes in the detection of RBPs can be explained by changes in protein abundance upon flg22 treatment, we cross-analysed the ptRIC data with the whole cell proteome (**Fig. 6**). To be able to compare both datasets, we applied the same criteria (FDR ≤ 0.2 and $\log_2$FC [flg22/mock] ≥ 0.58 or ≤ -0.58) to unify changes in association with RNA (RBPome; ptRIC) and in protein abundance (proteome; WCL). The overlap between the two datasets was 542 proteins, thus we could identify the protein abundance for about half of the RBPs isolated by ptRIC. Among the 542 proteins, 142 were flg22-responsive RBPs (i.e. significantly altered in affinity) detected in both ptRIC and the WCL (**Fig. 6A**).

Next, we categorised the flg22 responsive RBPs (ptRIC) into 3 groups depending on their abundance: i) **Responsive**: RBPs that changed in RNA association (ptRIC) but did not change in protein abundance (WCL); ii) **Abundance-driven**: RBPs that changed in RNA association (ptRIC) but also in protein abundance (WCL); iii) **Unknown**: RBPs that changed in RNA association (ptRIC) but were not detected in the WCL. Overall, most of the flg22-responsive RBPs were classified as 'responsive', indicating that they changed in association with RNA and remained unaltered at the protein level (**Fig. 6B**). This is also in concordance to what has been observed in human cells infected with Sindbis virus (Garcia-Moreno et al., 2019). Only 8 flg22-responsive RBPs showed changes in association with RNA that were abundance-driven (**Fig. 6B**).

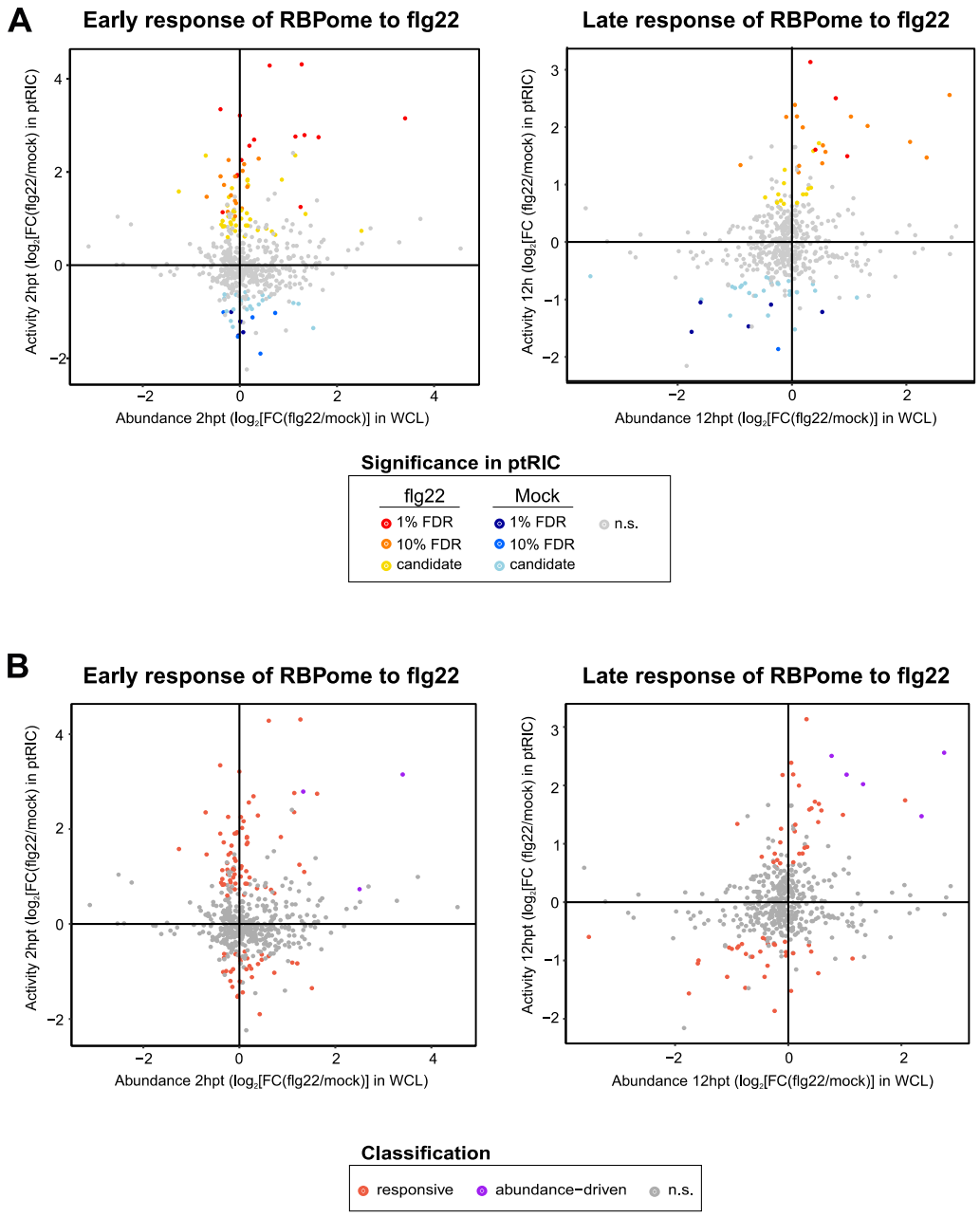


Fig.6. Changes in association with RNA of the RBPome do not correlate with changes in protein abundance. Scatter plots depicting the log_2 fold change in WCL (x-axis) and log_2 fold change in ptRIC (y-axis) of each protein (dots) between flg22 and mock treatment using data from four biological replicates. A) Proteins are coloured according to the significance (FDR) in ptRIC as following: proteins are coloured in red when FDR ≤ 0.01 and log_2FC[flg22/mock] ≥ 1, in orange when FDR ≤ 0.1 and log_2FC [flg22/mock] ≥ 1, in yellow when FDR ≤ 0.2 and log_2FC [flg22/mock] ≥ 0.58, in dark blue when FDR ≤ 0.01 and log_2FC [flg22/mock] ≤ -1, in blue when FDR ≤ 0.1 and log_2FC [flg22/mock] ≤ -1 and in light blue when FDR ≤ 0.2 and log_2FC [flg22/mock] ≤ -0.58. Red, orange and yellow proteins represent the leaf RBPs stimulated upon flg22 perception (ptRIC), whereas dark blue, blue and light blue represent leaf RBPs inhibited upon flg22 perception (ptRIC). Non-significant proteins in ptRIC are coloured in grey. B) Proteins are coloured according to the classification as following: proteins are coloured in red when changes in ptRIC are statistically significant but non-significant in WCL (responsive), in purple when changes in both ptRIC and WCL are statistically significant (abundance-driven), and in grey when both changes are not statistically significant. Non-significant proteins in ptRIC are coloured in grey.

2.7. Functional screening confirms RBPs involved in plant immunity

To validate that the flg22-responsive RBPs indeed play important roles in plant immunity, we performed an immunity-based screen of Arabidopsis mutant lines. We selected 19 RBPs from the flg22-responsive RBPs (including both high-confidence and candidate flg22-responsive RBPs) based on the availability of mutants and information in the literature (**Supplemental table 3**) and obtained T-DNA insertional lines from Nottingham Arabidopsis Stock Centre (NASC; http://arabidopsis.info). The mutants were screened for their susceptibility to *Pseudomonas syringae* pv. tomato DC3000 (*Pst*DC3000) and its mutant hrpA$^-$. We used these two bacterial strains because *Pst*DC3000 causes disease in Arabidopsis whereas the hrpA$^-$ mutant elicits PTI. Since the flg22-responsive RBPs were identified in a context of PTI (i.e. flg22 elicits PTI), we considered that infection with hrpA$^-$ was a required treatment. We reasoned that if a given RBP is involved in plant immunity, knocking out that particular RBP could result in increased susceptibility or resistance to bacterial pathogens. However, it should be noted that if a given mutant line doesn't yield an infection phenotype it doesn't necessarily indicate that the RBP is not involved in plant immunity, since its function could be redundant with other proteins.

Out of the 19 screened mutants, 11 had a clear phenotype of resistance or susceptibility to at least one of the tested pathogens (**Supplemental table 3; Supplemental figure 1**). These included 3 RBPs whose mutants were more resistant to bacterial pathogens, such as RH11, METK4 and TSS, and 8 RBPs whose mutants were more susceptible to bacterial pathogens, such as EIN2, EFG and ClpD (**Supplemental table 3; Supplemental figure 1**). Two of the selected flg22-responsive RBPs were identified as protein groups rather that unique proteins (HSP90-3/4 and HSP70-2/18), and therefore, mutant lines of all the RBPs within the protein group were screened (**Supplemental table 3**). Four out of the 7 high-confidence flg22-responsive RBPs yielded a disease phenotype, whereas 7 out of 12 candidate flg22-responsive RBPs had a disease

phenotype. This indicates that although identified using less stringent statistical criteria, the candidate flg22-responsive RBPome includes *bona fide* immune players. We selected 3 mutants with disease phenotypes for further characterisation.

RNA helicase RH11 (RH11)

We identified the DEAD-box ATP-dependent RNA helicase RH11 (RH11) to be involved in plant immunity by ptRIC and by immunity-based mutant screen. Although RH11 has been largely uncharacterized, some recent studies have linked RH11 to different processes of RNA biology. For example, RH11 has been shown to interact and co-localize with the NMD factor UPF1 in the cytosol and P-bodies, indicating that it might function in mRNA decay (Chicois et al., 2018). RH11 has also been linked to the spliceosome based on information of homologs for different organisms (Koncz et al., 2012). In addition, it has been speculated that RH11 might function as a ribosome biogenesis factor since its human (DDX3X) and yeast (Dbp1p) homologues are involved in different steps of RNA regulation, including ribosome biogenesis (Liu and Imai, 2018). Therefore, RH11 may be functionally important in many different cellular processes, however, further studies are required to dissect the specific role of this RNA helicase.

We determined that RH11 had increased association with RNA at early time points after flg22 perception (2 hpt) with no changes in protein abundance (**Fig. 6A, B**). Although RH11 is known to be phosphorylated (Kanno et al., 2018), whether this is a mechanism by which its association with RNA is regulated needs further validation. Three independent *rh11* mutant lines were obtained (**Fig. 7**) and challenged with the wild type *Pst*DC3000 and its hrpA$^-$ mutant. All three independent *rh11* mutants were more resistant to *Pst*DC3000, as determined by bacterial count and by visual phenotype (**Fig. 7C, D; Supplemental figure 1**). No change in resistance to hrpA$^-$ was observed (data not shown). Moreover, *rh11* mutants had an increased ROS burst (**Fig. 7E**), indicating that RH11 may be involved in plant immunity by controlling the RNA metabolism of

genes associated with oxidative burst. However, the homozygous insertion of the T-DNA could only be confirmed for the *rh11-2* mutant, hence, *rh11-1* and *rh11-3* still need be confirmed.

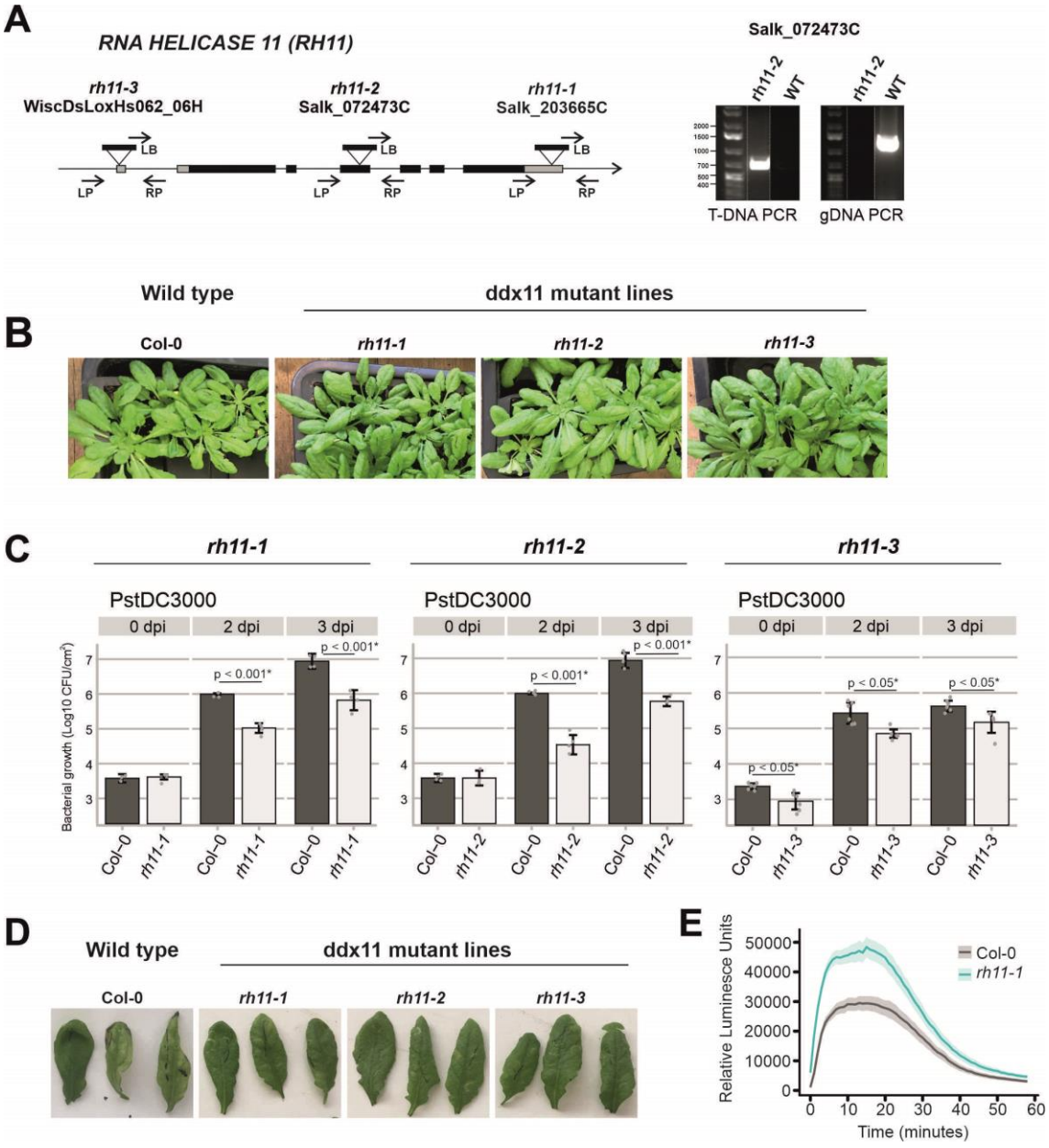


Fig. 7. RH11 is involved in immunity in Arabidopsis. A) Schematic representation of the insertion of T-DNA for each of the mutant lines. Dual genotyping PCRs of the mutant *rh11-2*: T-DNA PCR was performed using the T-DNA LB (Salk) and RH11-specific RP, whereas the gDNA was performed with RH11-specific LP and RP. B) Representative pictures of 6-week old plants of each of the *rh11* mutant lines and Col-0 (WT). C) *rh11* mutants are more resistant to the disease-causing strain *Pst*DC3000. *rh11* and Col-0 plants were infiltrated with *Pst*DC3000 at an inoculum density of $5x10^5$ cells/ml and the bacterial density analysed at 0, 2 and 3 days after inoculation (dpi). Error bars indicate mean ± sd of n = 4. ANOVA with Tukey HSD was used to calculate the p-value. The experiment was repeated multiple times with similar results. D) Infection phenotype of *rh11* mutant lines at 3 dpi. E) Measurement of ROS burst in Col-0 and *rh11-1* mutant after elicitation with flg22 (500 nM). Error bars indicate mean ± SEM of n = 12. The experiment was repeated multiple times with similar results.

S-ADENOSYLMETHIONINE SYNTHASE 4 (METK4)

We identified S-adenosylmethionine synthase 4 (METK4; MAT4; SAMS3; MTO3) to play a role in plant immunity by both ptRIC and immunity-based mutant screen. S-adenosylmethionine synthases are enzymes that catalyse the synthesis of S-adenosylmethionine (SAM) from methionine and ATP (Binet et al., 2011). SAM is important in methionine metabolism and is a widely used molecule for a broad range of biological reactions because is a universal methyl group donor. In addition, SAM-dependent methyltransferases can transfer the methyl group to DNA, RNA, proteins or other metabolites (Sauter et al., 2013). The genome of Arabidopsis encodes for four S-adenosylmethionine synthase genes (METK1-4).

METK4 has been shown to be an important player in DNA and histone methylation (Meng et al., 2018). Additionally, METK4 can interact with TAD3, a tRNA adenosine deaminase involved in tRNA editing (Zhou et al., 2014). Although METK4 does not possess any known RBD, we experimentally uncovered that METK4 can bind RNA in Arabidopsis leaves by ptRIC followed by MS and WB (**Fig. 4B**). In agreement, Reichel and colleagues also identified that METK4 (and METK3) can bind RNA in etiolated seedlings (Reichel et al., 2016). We determined that METK4 was stimulated at 2 h after flg22 treatment with no changes in protein abundance (**Fig. 4B**). Moreover, 2 members of the S-adenosyl-L-methionine-dependent methyltransferases superfamily (AT1G16445 and AT1G55450) were also altered in association with RNA upon flg22 perception.

To validate the role of METK4 in immunity, we obtained one *metk4* mutant line and one line overexpressing METK4-GFP (METK4 OX; **Fig. 8**) and challenged them with *Pst*DC3000 and its hrpA$^-$ mutant. Contrarily to what was previously reported for CRISPR/Cag9 generated *metk4* mutants (Meng et al., 2018) our *metk4* mutant line (Salk_052289C) was viable. *metk4* mutant

plants were more resistant to both *Pst*DC3000 and hrpA$^-$, whereas METK4 OX were more susceptible to *Pst*DC3000 (**Fig. 8C**). In agreement, *metk4* plants were visually more resistant to *Pst*DC3000 whereas METK4 OX lines appeared more susceptible (**Fig. 8D**). However, bacterial resistance was not linked to the ROS burst, since the ROS levels were comparable to those of the wild type (WT) plants (**Fig. 8E**). Our results indicate that METK4 is involved in plant defence responses and moonlights between its enzymatic activity and RNA metabolism.

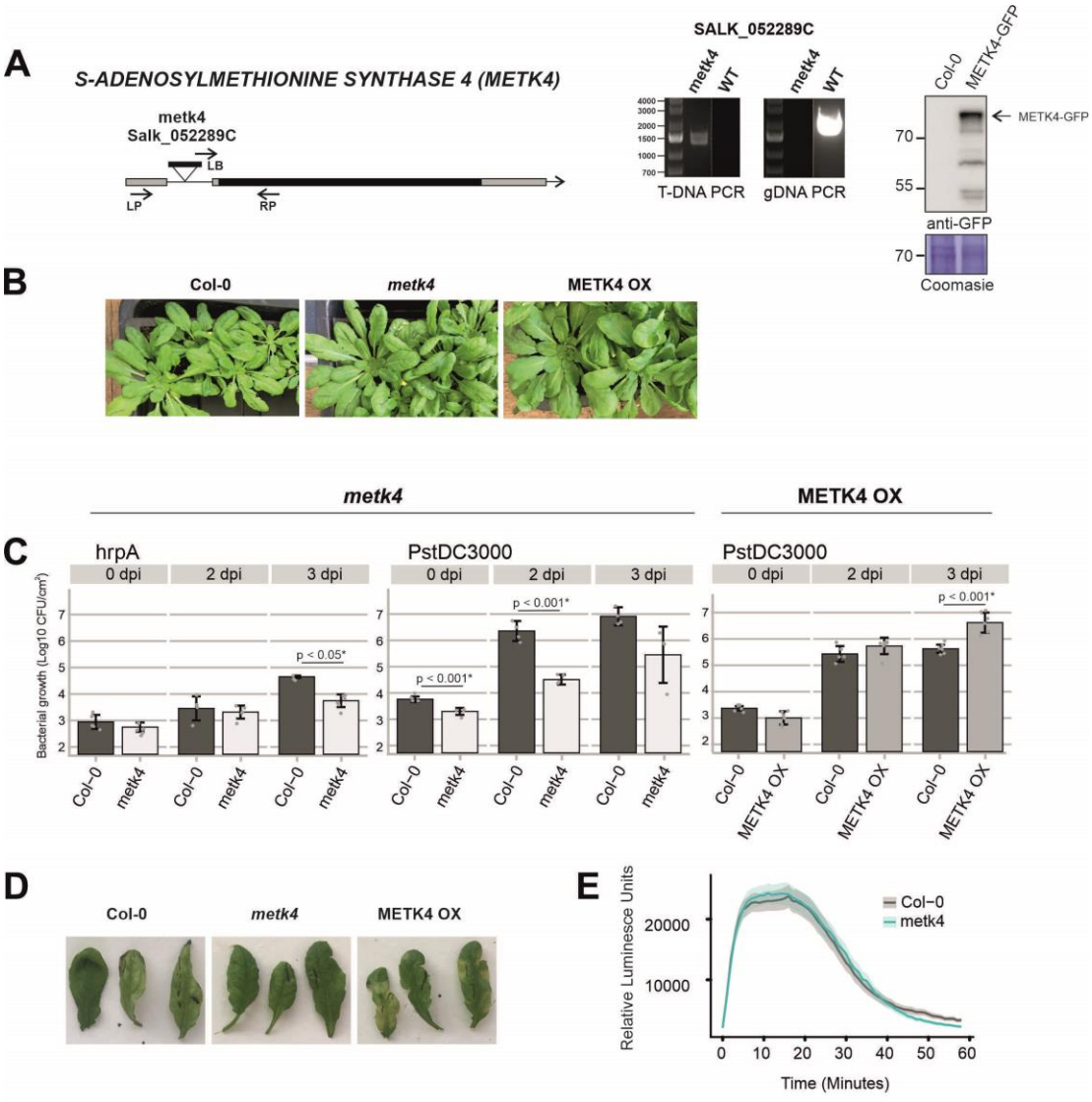


Fig. 8. METK is involved in immunity in Arabidopsis. A) Schematic representation of the insertion of T-DNA for the *metk4* mutant line. Dual genotyping PCRs of the mutant line *metk4*: T-DNA PCR was performed using the T-DNA LB (Salk) and METK4-specific RP, whereas the gDNA was performed with METK4-specific LP and RP. Western blot analyses of Col-0 (WT) and METK4 OX (METK4-GFP) plants using an anti-GPF antibody (Abcam ab6663). Coomassie staining of the blot is shown below as a loading control. B) Representative pictures of 6-week old plants of the mutant and overexpressing lines and Col-0 (WT).

Fig. 8 (continued) C) *metk4* mutant is more resistant to the disease-causing strain *Pst*DC3000 and the *hrpA*- mutant, while the METK4 OX is more susceptible to *Pst*DC3000. *metk4*, METK4 OX and Col-0 (WT) plants were infiltrated with *Pst*DC3000 at an inoculum density of $5x10^5$ cells/ml and the bacterial density analysed at 0, 2 and 3 days after inoculation (dpi). *metk4* and Col-0 (WT) plants were infiltrated with *hrpA*- at an inoculum density of $5x10^5$ and the bacterial density analysed at 0, 2 and 3 days after inoculation (dpi). Error bars indicate mean ± sd of n = 4. ANOVA with Tukey HSD was used to calculate the p-value. The experiment was repeated multiple times with similar results. D) Infection phenotype of *metk4* mutant line and METK4 OX line at 3 dpi. E) Measurement of ROS burst in Col-0 and *metk4* mutant after elicitation with flg22 (500 nM). Error bars indicate mean ± SEM of n = 12. The experiment was repeated multiple times with similar results.

ETHYLENE INSENSITIVE 2 (EIN2)

Using our approach we also identified ethylene insensitive 2 (EIN2), a protein known to be involved in plant immunity (reviewed in **Chapter 1**; Alonso et al., 1999; Mersmann et al., 2010), thus highlighting the quality of our dataset. EIN2 is stimulated at early time points after flg22 perception. Unfortunately, EIN2 was not detected in the total proteome, so we lack information on EIN2 protein abundance. However, interestingly, we also identified that ACCO, an enzyme involved in ethylene biosynthesis, was an RNA binder and was stimulated at 12 hpt.

EIN2 is a component of the ethylene signalling pathway that contains an ER-localised transmembrane domain and a cytoplasmic domain (**Fig. 3** of **Chapter 1**). Two modes of action have been described for EIN2. Firstly, upon ethylene perception a portion of the cytoplasmic domain, the CEND, is cleaved and re-localises to the nucleus to act in the stabilization of two transcription factors (EIN3/EIL1) that positively regulate ethylene responses (Ju et al., 2012; Qiao et al., 2012; Wen et al., 2012). Recently, a second mode of action of EIN2 was identified. Upon ethylene perception, the CEND can also remain cytoplasmic and bind the 3' UTR of mRNAs that negatively regulate ethylene responses (EBF1/2). Together with the NMD machinery, EIN2 re-localises to the P-bodies and promotes translational repression of those mRNAs (Li et al., 2015; Merchante et al., 2015). EBF1/2 act to promote degradation of EIN3/EIL1 transcription factors. Therefore, by promoting translational repression of the negative regulators EBF1/2, EIN2 promotes the expression of ethylene responsive genes. In this study we have confirmed that EIN2 can bind RNA in Arabidopsis leaves.

EIN2 is conserved through evolution (Ju et al., 2015) and has been extensively linked to responses to many environmental cues in different species, including responses to biotic stress (Gazzarrini and Mccourt, 2003; Salvador-Guirao et al., 2018; Rin et al., 2017). EIN2 regulates FLS2, the receptor responsible of sensing flg22. Hence, *ein2* mutants accumulate lower levels of FLS2 mRNA and protein (Tintor et al., 2013; Mersmann et al., 2010; Boutrot et al., 2010). Moreover, *ein2* mutants have been described to have reduced immune responses and to be more susceptible to *Pst*DC3000, especially at early time points (reviewed in **Chapter 1;** Tintor et al., 2013; Mersmann et al., 2010; Clay et al., 2009; Washington et al., 2016), suggesting that EIN2 is important in the early stages of plant immunity. However, there is controversy with respect to *ein2* mutants since other studies have reported opposite phenotypes in terms of defence outputs and diseases phenotypes (Chen et al., 2009; Boutrot et al., 2010). In our experimental set up, e*in2-1* mutants were more susceptible to infection by *Pst*DC3000, especially at early time points, and by hrpA$^-$ at early and late time points (**Fig. 9C, D**). Although not statistically significant, overexpression of EIN2 (EIN2OX) resulted in a slight increase in resistance to *Pst*DC3000 at early time points, and values of bacterial growth similar to the WT in the hrpA$^-$ bacterial strain (**Fig. 9C, D**). This supports the findings by Mersmann and colleagues and Tintor and colleagues (Mersmann et al., 2010; Tintor et al., 2013). Moreover, *ein2-1* mutants had a consistent reduced ROS burst in response to flg22 (**Fig. 9E**).

<u>Is EIN2 cleaved at an alternative site upon flg22 treatment?</u>

Flg22 induces ethylene production (Felix 1999; Boller 1995), and in the presence of ethylene the CEND is cleaved and functions in the nucleus and cytoplasm (Ju et al., 2012; Qiao et al., 2012; Wen et al., 2012; Li et al., 2015; Merchante et al., 2015). However, to our knowledge it has not been directly proven that flg22 induces the cleavage of the CEND. We found that EIN2 has

increased association with RNAs upon flg22 treatment (**Fig. 10A**), suggesting a direct link between flg22 perception and the RNA-binding ability of EIN2.

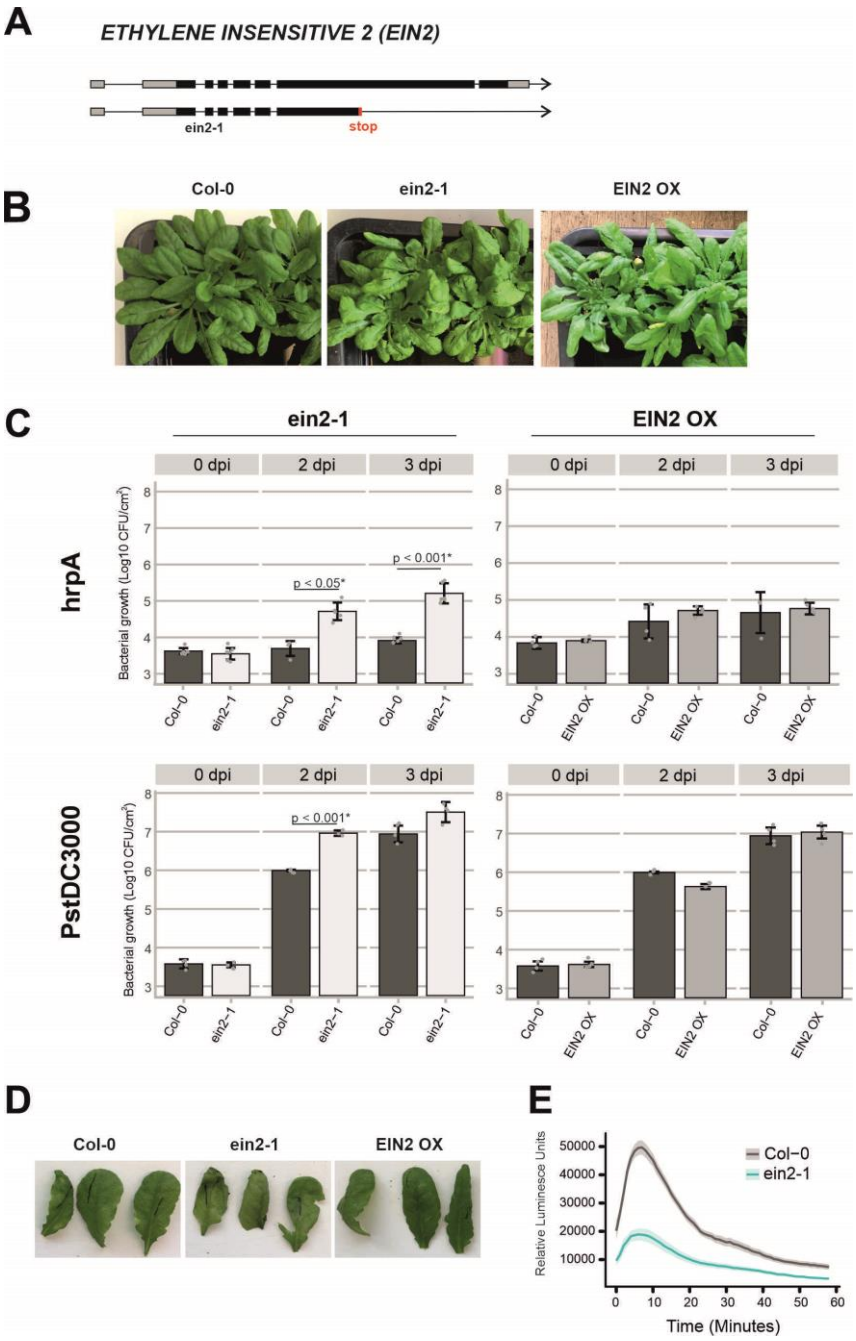


Fig. 9. EIN2 is involved in immunity in Arabidopsis. A) Schematic representation of the *ein2-1* mutant. B) Representative pictures of 6-week old plants of each of *ein2-1*, EIN2 overexpressing line (EIN2 OX) and Col-0 (WT). C) *ein2-1* mutant is more susceptible to the disease-causing strain *Pst*DC3000 and the *hrpA*- mutant, while the EIN2 OX is slightly resistant to *Pst*DC3000. Error bars indicate mean ± sd of n = 4. ein2-1, EIN2 OX and Col-0 (WT) plants were infiltrated with *Pst*DC3000 or *hrpA*- at an inoculum density of $5x10^5$ cells/ml and the bacterial density analysed at 0, 2 and 3 days after inoculation (dpi). ANOVA with Tukey HSD was used to calculate the p-value. The experiment was repeated multiple times with similar results. D) Infection phenotype of *ein2-1* mutant line and EIN2 OX line at 3 dpi. E) Measurement of ROS burst in Col-0 and *ein2-1* mutant after elicitation with flg22 (500 nM). Error bars indicate mean ± SEM of n = 12. The experiment was repeated multiple times with similar results.

To profile the responses of EIN2 upon flg22 perception, we mapped the peptides identified by proteomics and their intensities to EIN2 in the 4 different conditions (mock 2 hpt, flg22 2 hpt, mock 12 hpt and flg22 12 hpt). We identified more peptides and, generally, with higher intensity in plants treated with flg22 than with H_2O at 2hpt (**Fig. 10A**). Moreover, this intensity increment was lost upon 12 hpt, suggesting that the EIN2 response to flg22 is rapid and dynamically regulated. These results align well with previous results suggesting that EIN2 is involved in the early preinvasive immunity response (Mersmann et al., 2010).

Interestingly, the peptides detected were not restricted to the CEND, but mapped to the whole cytoplasmic domain (**Fig. 10A**). This suggests that either the cleavage site is not at position 645 as described by Qiao and colleagues (Qiao et al., 2012), or that an alternative cleavage site that includes all of the cytoplasmic domain exists and that this extended cleavage product is able to interact with RNA. Moreover, this longer product seems to be the predominant form. We would expect that if CEND and the longer form were both equally abundant, the intensity of peptides identified from the CEND region would be twice as high, and this is not the case (**Fig. 10A**). This suggests the existence of an alternative cleaved form that possesses increased RNA binding in response to flg22. Additional experiments should be performed to confirm the existence of this extended cleavage product. Using an algorithm trained with RBDmap data (Shuai *et al. in prep*), we predicted that the cytoplasmic domain of EIN2 harbours 3 putative RNA-binding sites (**Fig. 10B**), and all these regions fall within the CEND.

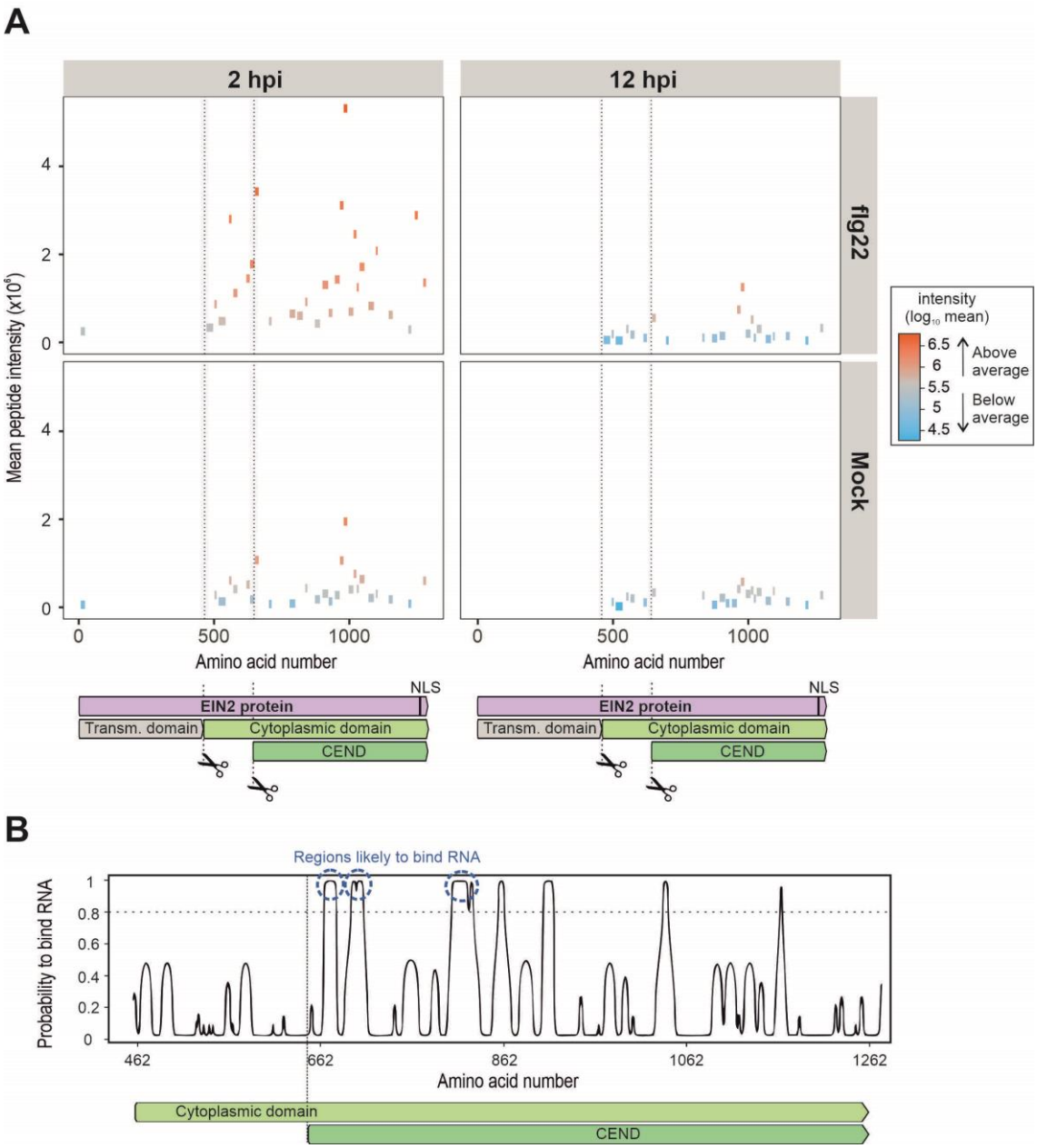


Fig.10. EIN2 is stimulated at early time points following flg22 perception and is cleaved at the C-terminal domain. A) Mapping of the peptides identified by MS to the EIN2 sequence and mean peptide intensity using data from four biological replicates. Colours indicate the $\log_2$ mean intensity of the peptides. Dotted lines indicate the cleavage sites of EIN2: a known cleavage site that yields the CEND, and a newly identified putative alternative cleavage site at the N-terminal part of the cytoplasmic domain. B) Prediction of the RNA-binding regions of the EIN2 cytoplasmic domain. Abbreviations: NLS, nuclear localization signal

3. DISCUSSION

3.1. ptRIC as a method to study the dynamic immune response of the plant RBPome

Recently, the RBPomes of different species have been described to respond to different cellular contexts and environmental cues (Garcia-Moreno et al., 2019; Sysoev et al., 2016; Perez-Perri et al., 2018; Trendel et al., 2019; Shchepachev et al., 2019; Marondedze et al., 2019). Hence, we hypothesised that the Arabidopsis RBPome could be modulated during immune responses, and that we would capture those changes using RIC. We have improved RIC to be applied efficiently to plant leaves (**Chapter 2**). This enhanced method, referred to as 'ptRIC', allowed the establishment of most complete and extensive RBPome of Arabidopsis leaves to date, comprising 1135 RBPs (**Fig. 2**, **Supplemental digital table 1**). This has led to the identification of many potential novel RBPs. Importantly, we have designed a strategy to profile RBP dynamics comprehensively in response to flg22 (**Fig. 1**). This approach can now be applied to study the RBP responses to virtually any physiological or pathological cue.

Our study revealed that about 25% of the leaf RBPome is modulated in response to flg22 (**Fig. 3**). This proportion is in agreement with a recent report showing that 30% of the human RBPs respond to the infection with Sindbis virus (Garcia-Moreno et al., 2019). Similar numbers apply to treatment with inhibitors of the RNA demethylases in human cells (Perez-Perri et al., 2018) or *Drosophila* embryo during maternal-to-zygotic transition (Sysoev et al., 2016), which suggest that the RBPome is highly dynamic. In our study, many of the flg22-responsive RBPs discovered are known regulators of immunity, thus confirming our hypothesis that RBPome dynamics would reveal functionally relevant RBPs (**Supplemental table 1**). However, dozens of flg22-regulated RBPs have not yet been linked to immunity and deserve further in detail characterisation. Moreover the 'candidate responsive RBP' were identified with less stringent criteria and we recommend researchers to experimentally validate them prior to initiate any research, as we

expect a higher incidence of false positives. However, this set contains a number of RBPs with known links to immunity suggesting the presence of true immunoregulatory RBPs.

Many defence-related proteins are dynamically activated in response to pathogen perception, since constitutive expression or/and activity may lead to tissue damage (Chakraborty et al., 2018; van Wersch et al., 2016). Furthermore, during immune response plants downregulate unnecessary processes or pathways that are counterproductive for the immune response, such as photosynthesis (Serrano et al., 2016). It has been speculated that photosynthetic inhibition would be aimed to reduce carbon availability for pathogens or to prioritise defence-related processes over other processes (Serrano et al., 2016). In this context, the pervasive remodelling of the Arabidopsis leaf RBPome possibly reflects the activation of critical defence pathways and the inhibition of host factors potentially exploitable by the pathogen. Although plants rapidly respond to flg22 perception (Denoux et al., 2008), RBP responses are detected both at early and late time points after elicitation (**Fig. 3**). We hypothesise that RBPome temporal remodelling is critical for reprogramming of the plant transcriptome.

3.2. Canonical and non-canonical RBPs regulate immune responses

Flg22-responsive RBPs include RBPs involved in virtually every step of RNA metabolism, from synthesis to decay. However, a number of RBPs that are flg22-regulated had no previous links to RNA and represent RNA-binding moonlighting proteins or enzymes (**Supplemental table 1**).

Canonical RBPs regulating plant responses to flg22

Amongst the flg22-responsive RBPs involved in known steps of RNA metabolism we have identified multiple RBPs involved in splicing (**Supplemental table 1**). It is known that splicing and alternative splicing are of paramount importance for the adaptive response to different challenges (Gulledge et al., 2012; Tanabe et al., 2007; Palusa et al., 2007), and we have been

able to capture some of those responses during plant immune activation. Interestingly, at early time points all of the splicing related RBPs were stimulated and nuclear, whereas at late time points they were inhibited and mostly chloroplastic. We have also identified that multiple helicases are altered in association with RNA upon flg22 treatment (**Supplemental table 1**). RNA helicases are extremely important at many step of RNA metabolism since they catalyse the unwinding of RNA secondary structure, and changes in their activity are expected to affect the capacity of RNPs to remodel (Chen and Shyu, 2014). Hence, it is not unexpected that they have important roles in immunity regulation. Indeed, an ortholog of one of the ptRIC-identified RNA helicases is known to play important roles in immunity (Li et al., 2008).

It is well known that to quickly respond to the changing environment, plants need a plastic regulation of their proteome. This can be achieved by transcriptional regulation, or through regulation of mRNA translation, which has a quicker impact on protein abundance (Xu et al., 2017). Moreover, translation is tightly regulated, especially at the immune response level (Xu et al., 2017). We have found that multiple proteins involved in translation were inhibited at late time points after flg22 treatment including initiation factors and elongation factors (**Supplemental table 1**). We have also suggested the existence of specialized ribosomes functioning in plant immune responses. Differential regulation of ribosomal proteins has been previously observed in other organisms (Simsek et al., 2017; Garcia-Moreno et al., 2019). Moreover, the composition of the ribosomes was demonstrated to change upon treatment of *N. benthamiana* with *Agrobacterium tumefaciens* either carrying infectious cDNA of Potato virus A (PVA) by affinity-tag purification (Eskelin et al., 2019). Hence, the diverse RNA and protein composition of the ribosomes could be important for fine-tuning of translational control in the cell's response to pathogens.

Moreover, several proteins involved in P-bodies were stimulated by flg22 (**Supplemental table 1**), some of which were previously described to play important roles in plant immunity (Roux et al., 2015). This discovery stresses the importance of this membrane-less cellular compartment for storage or RNA decay of defence-related transcripts, allowing rapid responses to pathogen attack. In addition, many chloroplastic RBPs involved in RNA editing, translation and peptide modification are inhibited during flg22-induced immune response (**Supplemental table 1**), which indicates the importance of chloroplasts for plant defences (Serrano et al., 2016).

Non-canonical RBPs regulating plant responses to flg22

However, a large number of RBPs that are regulated by flg22 had no previous links to RNA biology and may represent RNA-binding moonlighting proteins or enzymes (**Supplemental table 1**). Some flg22-responsive RBPs have mono- or di-nucleotide binding domains, which is characteristic of some enzymes moonlighting as RNA binders (Castello et al., 2015, 2016). Other RBPs have orthologs in other organisms with known RNA-binding activity (e.g. PPIs and HSPs). For example, PPIs have been classified as RBPs from yeast to human. We show here (and in **Chapter 3**) that PPIs also bind to RNA in plants. Importantly, a set of PPIs are stimulated upon immune activation (**Supplemental table 1**). This aligns well with observations in human cells infected with Sindbis virus, where PPIA emerged as a key regulator of infection (Garcia-Moreno et al., 2019). Moreover, some of the responsive PPIs discovered here are targeted by pathogens (Vijayapalani et al., 2018; Fan et al., 2018), indicating that they are critical for immune responses or represent susceptibility factors. Similarly, we have identified that multiple HSPs are stimulated upon immune activation (**Supplemental table 1**), some of which are known to play roles in immunity (Park and Seo, 2015; Huang et al., 2012). Our results, together with previous studies, strongly support the potential role of HSPs as regulators of RNA metabolism during immune responses (Castello et al., 2016; Iwasaki et al., 2010; Willmund et al., 2013; Reichel et al., 2016; Garcia-Moreno et al., 2019).

Previous results suggest that many photosynthesis-related proteins interact with RNA *in vivo* in a broad range of plant species (**Chapter 3**). Interestingly, some of the photosynthesis components are inhibited by flg22, including proteins within the PSI (apoproteins and reaction centre proteins), PSII (reaction centre proteins), rubisco subunits and cytochrome b6 (**Supplemental table 1**). By contrast, functionally different proteins such as carbonic anhydrases, components of the light-harvesting complex and components of the oxygen-evolving complex are stimulated (**Supplemental table 1**). Thus, taken together these results support the presence of moonlighting RNA-binding activities in the photosynthetic pathway that can be regulated in response to immune cues. What this responsive behaviour means biologically deserves further investigation.

Because photosynthesis is such a vital process for almost all plants, it is expected to be regulated in multiple ways. It is tempting to speculate that one of these may involve RNA as rapid-response regulator of photosynthetic activity either controlling allosterically protein activity or the fate of mRNAs encoding photosynthetic factors. This hypothesis is supported by the fact that Rubisco large subunit (LSU) and cytochrome f can regulate their own RNAs upon specific conditions (Yosef et al., 2004; Cohen et al., 2006, 2005; Choquet et al., 2003) and that most changes in photosynthetic proteins are triggered at 2 h after flg22 treatment and almost no changes are observed after 12 hpt (**Supplemental table 1**). We hypothesise that the mechanism of gene expression regulation by LSU and cytochrome f may be extended to other photosynthetic components. Allosteric regulation of proteins by RNAs has been described in other organisms. For example, the vault1-1 RNA can bind to autophagy receptor p62 and interfere with p62 function by inhibiting its required oligomerisation (Horos et al., 2019). This emphasizes the importance and tight regulation of the photosynthesis in plants and indicates that enzymes moonlighting as RNA binders under certain cellular contexts may contribute both to normal cellular homeostasis and stress responses.

We previously discovered that many plant metabolic enzymes can bind RNA and that for some of these the RNA-binding function is conserved in multiple species across the plant kingdom (**Chapter 3**). In the last decades, it has become clear that certain metabolic enzymes can moonlight as RNA-binding proteins, often under particular cellular states such as stress (Castello et al., 2015). For example, during iron deficiency or oxidative stress the animal aconitase/IRP1 abandons its role as cytosolic aconitase and becomes an RBP that regulates mRNAs related to iron metabolism genes (Volz, 2008). We hypothesise that some plant proteins or enzymes could function in a similar way and moonlight as RNA binders only upon immune activation. These flg22-responsive enzymes identified in this study may be endowed with moonlighting RNA-binding activity; however, further studies are needed to confirm the role of these enzymes in RNA metabolism.

3.3. Why is important to understand the dynamics of the RBPome?

The total proteome does not undergo pervasive changes upon flg22 treatment, at least at the time points analysed. In this study only 89 proteins were found to have altered levels after exposure to flg22, many of which are hallmarks of plant immunity such as transcription factors, receptor like kinases and hydrolases, while others are uncharacterised and might represent novel immunoregulators (**Fig. 5**, **Supplemental table 2**). *A priori*, it is striking that only 2.35 % of the total proteome is differentially expressed upon flg22 perception, since previous studies reported massive changes occurring at RNA level (Denoux et al., 2008; Bektas et al., 2016; Sohn et al., 2014; Yang et al., 2017; Hacquard et al., 2016). However, protein synthesis is tightly regulated during immune responses and the proteome poorly correlates with transcription (Xu et al., 2017). Our results reinforce this observation as we did not detect pervasive remodelling of the cell proteome at 12 h post flg22 treatment.

Indeed, there is recent evidence that indicates that changes at the proteome level occur at a different scale to those that occur at the transcriptional level. For example, Liu and colleagues identified changes in the levels of 8 redox-sensitive proteins upon 15 min treatment with flg22 treatment (Liu et al., 2015)**.** Moreover, 64 membrane-associated proteins were found to be differentially accumulated at 15 min after treatment with flg22 (Keinath et al., 2010). Similar results have also been observed in other organisms. For example, 202 proteins are altered in abundance upon lipopolysaccharide (LPS) and interferon-gamma (IFN-γ) treatment of mouse microglia (Woo et al., 2017). Taken together, all these findings support the notion that changes at protein level are less dramatic than changes at the RNA level. It is possible that the changes at transcript and protein levels occur at different temporal scales. In other words, it may be necessary to allow longer incubations with flg22 (or other stimuli) to observe substantial changes at the protein level. An alternative explanation is the lower depth of proteomic studies when compared to transcriptomics (Sidoli et al., 2017). It is challenging to detect and quantify low-to-mid abundant proteins; however, these are the most likely to be altered by physiological and pathological cues. Conversely, transcriptome analysis is deeper in coverage that whole proteome analysis, which allows the investigation of RNAs with a broad range of overall abundance.

Most of the flg22-driven changes in association with RNA of RBPs occurred without detectable changes in protein level (**Fig. 6**), suggesting the existence of mechanisms that regulate the RNA-binding affinity of these RBPs. These mechanisms are likely to involve post-translational modifications which can, in principle, stimulate or inhibit the association with RNA of the RBP (Arif et al., 2009, 2011; Castello et al., 2016). Other mechanisms that regulate RBPs activity and that may govern the flg22-driven changes in association with RNA could be cofactor binding or miss-assembly (Arif et al., 2018; Clingman et al., 2014), binding to metabolites (Guiducci et al., 2019), PPIs activity (Haghighat and Sonenberg, 1997), allosteric regulation by RNA (Dabo and

Meurs, 2012; Horos et al., 2019; Choudhury et al., 2017), changes in protein localisation (Arif et al., 2009) or changes in the RNA availability (Garcia-Moreno et al., 2019). For example, it has been seen that an important driving force in the remodelling of the human RBPome upon viral infection is the degradation of cellular RNAs and synthesis of viral RNA (Garcia-Moreno et al., 2019). This stresses the need for activity-based techniques, which can ultimately yield information about the biologically relevant activity of a given protein or enzyme. One of these techniques is ptRIC which allows identification of the RBPs actively engaged in RNA binding in a variety of cellular contexts, for instance, in immune responses.

3.4. RBPs as novel candidates of immunity regulation

Most proteins identified as immunoregulators, including RBPs, have been discovered on an individual basis, by functional screening or disease phenotypes. However, proteins functioning in particular cellular states, or proteins with redundant functions, are difficult to identify by these assays. Previous studies have applied RNA-seq or proteomic approaches to plants treated with elicitors, effectors or pathogens. Although very informative, these studies cannot profile protein activity and regulation. To circumvent these limitations, we have used ptRIC, which profiles RBP dynamics in a proteome-wide scale. ptRIC has identified 288 proteins as 'flg22-responsive' and we have confirmed that some of these proteins play important roles in immunity (**Fig. 7, Fig. 8** and **Fig. 9**). These immunoregulatory RBPs include EIN2, METK4 and RH11, amongst others. Notably, more than half of the tested mutant lines had a phenotype in infection (**Supplemental table 3**), proving the utility of our approach. The remaining proteins with no disease phenotype could represent RBPs with redundant functions, RBPs implicated in defence against other pathogens or simply represent false positives.

Although we have identified almost 300 RBPs with potential roles in immunity, for most of these RBPs their precise role in RNA metabolism and/or immunity is unknown. Because RBPs regulate

RNAs, or are regulated by RNAs, it becomes essential to identify their target RNAs. A number of approaches such as CLIPseq, iCLIP, eCLIP allow identification of target RNAs bound by specific RBPs (Köster and Meyer, 2018). Application of these techniques to flg22-responsive RBPs will identify the target RNAs and the exact position where the binding happens. This data will be critical to decipher the role of a given responsive RBP in immunity.

Additionally, 68 out of the 288 flg22-responsive RBPs were identified in the core RBPome of plants (**Chapter 3**). In other words, they bind to RNA in different plant species, which suggests that their association with RNA is conserved across the plant kingdom. Discovery of RBPs involved in immunity in different species and unveiling their precise role in the defence response will allow researchers to exploit their regulatory network (e.g. a particular target RNA) for breeding programs to achieve durable resistance in crops.

3.5. Conclusion

In conclusion, we have adapted RIC to efficiently identify and profile RBP responses in plant leaves. This improved approach, named 'ptRIC', identified hundreds of RBPs with altered association with RNA upon immune activation by flg22, allowing us to discover RBP networks functioning in plant immunity. Finally, we have uncovered that several of these RBPs either increase resistance or susceptibility to bacterial infection.

4. MATERIALS AND METHODS

4.1. Plant material and treatment

Arabidopsis thaliana plants were grown at neutral day conditions (12 h light, 12 h dark) at 20 °C and light intensity of approximately 100 µmol/m^2/s. The mutant lines used in this study are detailed in **Table 3**. METK-GFP overexpressing lines (METK4-OX) were a kind gift from Gong's lab

(Meng et al., 2018). EIN2 overexpressing lines (EIN2OX) were a kind gift from Qiao's lab (Zhang et al., 2016a, 2017a). For the immune elicitation treatments, Arabidopsis leaves of mature plants (5-6 weeks old) were infiltrated with either flg22 (1µM; Anaspec) or H_2O (mock) using a 1 ml blunt end syringe and leaf tissue was harvested at 2 and 12 h after treatment. For ptRIC experiments, 4 biological replicates per treatment and time point were performed.

4.2. plant RNA-interactome capture (ptRIC)

ptRIC was performed as described in **Chapter 2**.

4.3. SDS-PAGE, silver staining and western blot

Proteins were separated by SDS-PAGE (10-12% acrylamide) and transferred onto a PVDF membrane using the semidry TransBlot system (BioRad). The membrane was blocked with either 5% skimmed milk or 5% BSA (Bovine Serum Albumin) in TBST (Tris-Buffered Saline + 0.1 % tween) and incubated overnight at 4°C with one of the following antibodies: anti-METK1-4 (Agrisera, AS16 3148A), anti-HSC70 (Agrisera, AS08 371), anti-RH3 (AS13 2714), anti-GFP (Abcam, ab6663) or anti-MAPK (Cell Signalling Technology #4370). For the anti-GFP antibody, the blot was washed 3 times with TBST and visualized using LAS4000 (GE healthcare) and SuperSignal West Femto substrate (ThermoScientific). For the rest of the antibodies, the membrane was washed 3 times with TBST, incubated with secondary antibody coupled to HRP for 1 h at room temperature, washed with TBST and visualised as described for the anti-GFP.

Protein extraction for the MAPK phosphorylation assay was performed as described by Flury and colleagues (Flury et al., 2013), and the WB was performed as detailed in above.

4.4. Sample preparation and MS analyses

The isolated RBPs (eluates ptRIC) and total proteome samples (inputs ptRIC) were prepared for MS using the standard FASP (filter aided sample preparation) method (Wiśniewski et al., 2009) as described in **Chapter 2**. The samples from the total proteomes were previously incubated with benzonase (Merk) for 45 minutes at 4 °C to digest nucleic acids. MS analyses were performed as described in **Chapter 2** but using the Uniprot *A. thaliana* reference database (UP000006548_3702.fasta, downloaded 02/01/2018).

4.5. Statistical analysis of the leaf RBPome

The MS data was analysed as described in **Chapter 2** with minor modifications. Protein groups with at least one missing value within a group were filtered out to ensure we identified *bona fide RBPs* with consistent changes in association with RNA. After filtering, raw intensities were normalized using the variance stabilization normalization method implemented in the R/Bioconductor package 'vsn' (Huber et al., 2002). Then the pipeline of analysis continued as described in **Chapter 2**. For the RBPs identified by ptRIC (eluates) we incorporated the effect of 'Batch' into our linear model since plant treatments and sample preparation were performed in different days. However, for the total proteome (WCL) experiment the sample preparation was performed all together and the batch effect was minor. Thus, we decided not to incorporate 'Batch' into our linear model for the WCL. One of the replicates corresponding to the eluate of mock 12 hpt CL was lost during the processing. Therefore, for that condition, the statistical analyses were performed using three biological samples. For both datasets (ptRIC and WCL), some of the identified peptides were not unique and could not be assigned to a single protein. For these, we performed the analysis using the protein groups, although we cannot confidently assign the peptide to one unique protein. The protein groups are specified in the main text. The statistical criteria to define proteins as RBPs or as flg22-responsive RBPs are specified in the main text.

The links to RNA biology of the identified RBPs were determined based on GO annotation as described by Beckmann and colleagues (Beckmann et al., 2015).

4.6. Data visualization

Graphs were generated using the ggplot2 package within R (Wickham, 2009). Heatmaps were generated using the pheatmap package and venn diagrams using the limma package within R. To visualise the RBPomes in each of the conditions the missing values were imputed as described in **Chapter 2** and all the proteins classified as RBPs (by both quantitative and semiquantitative) were included in the Volcano plot.

4.7. Prediction of the RNA-binding regions

The sequence propensity to bind RNA was calculated using an unpublished software (Shuai *et al. in prep*) as described previously (Hobor et al., 2018).

4.8. Genotyping PCRs

To confirm the homozygous presence of the T-DNA insertion in both alleles, a dual PCR was performed as detailed by O'Malley and colleagues (O'Malley et al., 2015). Briefly, gDNA was extracted from the mutants using previously available protocols (Edwards et al., 1991; Kasajima et al., 2004) and two independent PCRs were performed using different pair of specific primers The first PCR (T-DNA PCR) uses one primer located in the left border of the T-DNA (LB) and another primer in the genomic region at the 3' end of the predicted insertion point (RP; **Table 1**). Thus, this PCR selectively amplifies a region that expands from the T-DNA to the specific genomic location within the gene. The second PCR (genomic PCR) uses a pair of primers designed in the genomic region around the T-DNA insertion point (LP and RP; **Table 1**).

Mutant line	Accession code	Primer name	Sequence (5' -> 3')
ddx11-2	Salk_072473C	SALK_072473C_LP	ATACACCAAAAACAGCCAGGG
		SALK_072473C_RP	CAGATTGCGACTGGCTTTAAG
metk4	Salk_052289C	SALK_052289_LP_NEW2	AGAGCGGATCGGAATCCTCTA
		SALK_052289_RP_NEW2	CCACGAACACAGACAATGGC
LB primer for Salk lines		LBb1.3_Salk	ATTTTGCCGATTTCGGAAC
LB primer for Wisc lines		L4_WiscDsLoxHs	TGATCCATGTAGATTTCCCGGACATGAAG

Table 1. Genotyping primers
List of primers used for genotyping the Arabidopsis mutant lines.

4.9. Bacterial growth assays

Leaves of mature Arabidopsis plants (5-6-week old) were syringe-infiltrated with *Pst*DC3000 or *hrpA*- at an inoculum density of $5x10^5$ cells/ml. At 0, 2 and 3 days after inoculation the bacterial density was analysed by homogenising leaf discs in water and plating serial dilutions onto LB (Lysogeny Broth) plates with the appropriate antibiotics.

4.10. Measurement of reactive oxygen species (ROS)

Measurements of ROS production were performed as described by Bach-Pages and Preston (Bach-Pages and Preston, 2018) with minor modification. Briefly, leaf discs of 4 mm diameter were punched from fully expanded leaves of mature plants using a cork borer and floated in water overnight at room temperature. Next day, leaf discs were transferred to a 96 well plate and presented with 200 µl of an eliciting solution containing HRP (20µg/ml; ThermoFIsher), LO-12 (10µM; WAKO) and flg22 (500nM; Anaspec). Luminescence was measured every minute using a M1000Pro microplate reader (TECAN Group Ltd.) for 1h and integration time 200ms.

ACKNOWLEDGEMENTS

We want to thank all the members of Preston, Castello and van der Hoorn lab for their fruitful discussions and inputs on the manuscript. We also want to specially thank Nattapong Sanguankiattichai for his valuable advice and help in the bioinformatics pipeline. We thank Urszula Pyzio, Sarah Rodgers and Caroline O'Brien for excellent technical support.

FUNDING

Marcel Bach-Pages is supported by Biotechnology and Biological Sciences Research Council (BBSRC, grant BB/M011224/1) and by the Lorna Casselton Memorial Scholarship at St. Cross College, Oxford. Alfredo Castello is supported by MRC Career Development Award MR/L019434/1 and MRC grant MR/R021562/1.

AUTHOR CONTRIBUTIONS

MBP designed the experiments with feedback from RvdH, AC and GMP. MBP performed all the experiments except for the MS run, which was performed by FK. HC assisted with the bioinformatics analyses. MBP analysed all the data and wrote the manuscript with feedback from AC and GMP.

SUPPLEMENTAL DATA

APPENDIX IV – Supplemental data Chapter 4

Note: supplemental tables 1 and 2 are a summary of the results; the full data can be found in supplemental digital tables 2 and 3, respectively.

Printed supplemental files

Supplemental table 1. Summary of flg22-responsive RBPome

Supplemental table 2. Summary of total proteome (WCL) flg22-responsive proteins

Supplemental table 3. RBP mutant lines screened

Supplemental figure 1. Bacterial counts of RBP mutant lines

Digital supplemental files

Supplemental digital table 1. Arabidopsis leaf RBPomes

Supplemental digital table 2. Flg22-responsive RBPome

Supplemental digital table 3. Total proteome (WCL) responses to flg22

CHAPTER 5

General discussion

General discussion

Overview

RNA-binding proteins (RBPs) regulate the fate of RNAs and thus are critical factors controlling post-transcriptional gene expression (Glisovic et al., 2008). Because of their sessile nature, plants require a tight regulation of the transcriptome to adjust to the ever-changing environment and RBPs have been shown to be crucial in controlling these processes (**Chapter 1**; Prall et al., 2019). However, plant RBPs have been historically understudied as compared to other organisms such as mammals. Hence, determining the composition and activity of the plant RNA-binding proteome (RBPome) is critical to understand normal cell homeostasis and responses to environmental stresses in plants.

RNA Interactome capture (RIC) has revolutionised the field of RNA biology by allowing systematic identification of RBPs in multiple species, ranging from yeast to humans (Hentze et al., 2018). Recently, RIC has been applied to plants, uncovering, for the first time, the RBPome of the model species *Arabidopsis thaliana* (Reichel et al., 2016; Zhang et al., 2016b; Marondedze et al., 2016). However, RIC was applied to Arabidopsis tissues with no relevance for plant physiology and, to our knowledge, it has not been applied to any other plant species beyond Arabidopsis.

ptRIC is a valuable tool to study plant RBPomes

We have developed a variant of RIC termed 'plant RNA-interactome capture' (ptRIC) that is optimised to efficiently and selectively isolate RBPs actively bound to RNA from leaves of mature plants (**Chapter 2**). This ptRIC protocol allows deeper determination of the RBPome than any other RIC protocol described for plants to date (Reichel et al., 2016; Zhang et al., 2016b;

Marondedze et al., 2016). We have also shown that ptRIC can be used in multiple plant species beyond the model plant *Arabidopsis thaliana* (**Chapter 3**) and envisage that ptRIC could be applied to a variety of plant tissues and organs. However, we have observed that the efficiency of ptRIC varies across different plant species (**Chapter 3**), so we anticipate that minor modifications to the protocol might be required to unveil the differences between RBPomes. An improved version of RIC, called enhanced RIC (eRIC), employs oligo(dT) LNA (locked nucleic acids)-modified probes instead of DNA probes, increasing the specificity and decreasing background noise (Perez-Perri et al., 2018). These advances can be applied to ptRIC to improve the coverage of plant RBPomes, especially for those species where the current version of ptRIC underperformed.

Because ptRIC uses oligo(dT) to capture the polyadenylated RNA, it cannot be used to detect RBPs bound to non-poly(A) RNAs. Recently, this limitation has been circumvented in multiple protocols that do not rely on oligo(dT) to purify RNA. For instance, TRAPP (total RNA-associated protein purification) uses silica beads to isolate total RNA-RBP complexes (Shchepachev et al., 2019) while CARIC (click-chemistry-assisted RNA interactome capture) and RICK (RNA-interactome capture using click chemistry) employ click-chemistry approaches (Huang et al., 2018b, 2018a; Bao et al., 2018). Other recent techniques exploit the physicochemical properties of RNA-protein complexes such as OOPS (orthogonal organic phase separation), PTex (phenol-toluol extraction) and XRNAX (protein-crosslinked RNA extraction) (Queiroz et al., 2019; Urdaneta et al., 2019; Trendel et al., 2019). To our knowledge, none of these techniques has been applied to plants yet.

These later techniques based on the physicochemical properties of RNP complexes have been demonstrated to be easy and of broad applicability. Moreover, PTex has been shown to yield larger numbers of RBPs than OOPS or XRNAX and allows identification of RBPs interacting with RNAs as short as 30 nt. Therefore, it would be of value to optimise PTex for plants to be able to capture plant RBPs binding to non-poly(A) RNAs. The problem of these approaches is that rRNA,

which is the most abundant RNA in the cell and forms part of the ribosomes, accounts for the largest proportion of the RNAs purified, requiring peptide fractionation to obtain enough depth in proteomic analysis. Despite this limitation, PTex can complement ptRIC to provide a more comprehensive census of RBPs. Moreover, R-DeeP (RNA-dependent proteins) uses RNase treatment followed by gradient ultracentrifugation to identify RNA dependent proteins (not necessarily RBPs), allowing reconstruction of protein complexes that contain RNA (Caudron-Herger et al., 2019). Hence, R-DeeP would be a valuable technique for global exploration of the protein complexes that require RNA in plants.

Many studies have performed transcriptomic analyses (RNA-seq) as a proxy to understand the cellular responses to environmental, pathological or physiological cues. However, transcriptomic and proteomic changes often poorly correlate, and strong evidence suggests that translation is as fine-tune regulated as transcription (Xu et al., 2017). Unlike the transcriptome, the proteome does not undergo massive changes in response to elicitors, as illustrated in **Chapter 4**. Moreover, the most effective and rapid way of regulation is by activating or inhibiting pre-existent resources, since the path from transcription to translation to reach the protein levels required for function takes hours. Hence, the cell has developed post-translational mechanisms to control protein activity. The activity of RBPs is known to be regulated by these mechanisms, including post-translational modifications (Castello et al., 2016; Arif et al., 2009, 2011), protein-protein interactions (Haghighat and Sonenberg, 1997), cofactor binding or miss-assembly (Arif et al., 2018; Clingman et al., 2014), binding to metabolites (Guiducci et al., 2019), PPIs activity (Haghighat and Sonenberg, 1997), allosteric modulation (Horos et al., 2019; Choudhury et al., 2017), changes in protein localisation (Arif et al., 2009) or substrate RNA availability (Garcia-Moreno et al., 2019). This calls for the development of methods that profile the activity of RNA-binding proteins, instead their protein levels. The optimal technique to do so is RIC as it has been benchmarked in multiple biological systems and its experimental output

has been largely validated by orthogonal methods (Hentze et al., 2018). Moreover, the capacity of RIC to uncover the changes in activity and composition of the RBPome under different cellular contexts have been successfully tested (Garcia-Moreno et al., 2019; Sysoev et al., 2016; Perez-Perri et al., 2018; Marondedze et al., 2019). Ultimately, the activity of a given protein is what determines its relevance in a particular context.

Riboregulation, the emergent universe in cellular control

In the past years, much of the studies have focused on characterising the role of RBPs binding and regulating target RNAs. However, recent transformational research has uncovered that RBPs can be regulated by RNA (Hentze et al., 2018; Horos et al., 2019). One example is the human vault1-1 RNA that can allosterically regulate the p62 autophagy receptor by binding and inhibiting its oligomerisation, which is required for its function (Horos et al., 2019). In agreement, a number of long non-coding RNAs have been described to regulate transcription by recruiting RBPs such as PRC2, which is member of the transcriptional repressor, polycom complexes (Achour and Aguilo, 2018; Rinn and Chang, 2012; Vance and Ponting, 2014). Moreover, RNAs can also play other important roles such as serving as a scaffold to recruit, organise or sequester proteins within complexes (Hentze et al., 2018). For example, the long non-coding RNA NEAT1 is critical for paraspeckle formation, thus contributing to storing RBPs in these nuclear bodies (Yamazaki et al., 2018). The important role of RNAs as regulators is also evidenced by recent findings from Caudron-Herger and colleagues, which identified that a large proportion of human protein complexes contain RNA (Caudron-Herger et al., 2019). This demands further investigation on the role of RNAs as regulators in protein complexes. To our knowledge, riboregulation is a largely unexplored area of research in plants.

Towards the discovery of the core plant RBPome

We have unveiled the RBPomes of seven different species that span across the plant kingdom. Analysis of these datasets revealed RBPs conserved across these species (core plant RBPome; **Chapter 3**). The core plant RBPome does not only comprise RBPs involved in known RNA-related processes such as translation and splicing, but also included multiple metabolic enzymes and proteins of the photosynthetic machinery. These newly identified plant RNA-binding moonlighting enzymes could represent unexplored layers of cellular regulation (**Chapter 3**). Moreover, by comparing the identified RBPs across all the species we have uncovered multiple potential novel RBDs and RBPs in plants that deserve further characterisation (**Chapter 3**). A number of molecular biology and biochemical assays can be used to test the capacity of these putative RBDs to interact with RNA and identify their binding modes, affinities and specificities. For example, the use of mutated or truncated variants of the domains coupled to techniques such as electrophoretic mobility shift assays (EMSA) or crosslinking and immunoprecipitation assays (CLIP) could help to characterise these RBDs *in vitro* and in cells, respectively. Alternatively, Castello and colleagues developed a technique termed RBDmap that can be applied to determine the regions within proteins that bind RNA at a proteome-wide scale (Castello et al., 2016, 2017). RBDmap is an extension of the RIC that adds an extra step of moderate proteolysis combined with a second oligo (dT) capture to allow global identification of protein regions engaged in RNA binding in cells (Castello et al., 2016, 2017). We envisage that this technique could be key at identifying novel RBDs in plants and confirm the putative RBDs discovered here. In addition, it would be interesting to investigate the conservation of these novel proteins and domains across eukaryotes and explore if any of these is specific to the plant kingdom.

Uncovering the plant immune RBPome

Using ptRIC we have examined the dynamic behaviour of the plant RBPome upon immune activation with flg22. This has led to the identification of nearly 300 RBPs with altered association with RNA during immune responses (**Chapter 4**). Analyses of the immune-responsive RBPome has uncovered RBP networks that are activated or inhibited during plant immunity including in processes such as RNA editing and translation (**Chapter 4**). Interestingly, we have also identified that many components of photosynthesis pathway can bind RNA, and that their association with RNA is altered during immunity in a manner that indicates that different parts of the photosynthetic apparatus have opposite behaviours to immune activation (**Chapter 4**). For example, some proteins of photosystem I and II reaction centre are inhibited upon immunity, whereas members of the light-harvesting complex and the oxygen-evolving complex are stimulated.

To determine if flg22-responsive RBPs play a role in immunity we tested RBP mutant lines and showed that they either increase or decrease resistance to bacterial pathogens (**Chapter 4**). However, their precise function in RNA metabolism and how they confer immunity or susceptibility in plants deserve further characterisation. Hence, we have contributed to support the existence of 'RBP-mediated plant immunity', whereby RBPs orchestrate the reprogramming of RNA metabolism that occur during immune responses. Moreover, we have provided a valuable foundation for the plant community to further dissect the specific roles of RBPs in plant immunity.

However, in nature, plants are constantly challenged by pathogens of diverse nature. Accordingly, the plant immune system is comprehensive and responds to wide range different elicitors, effectors and pathogens. It is known that some of these responses integrate into the same pathways and regulate common sets of genes, while others are specific. Therefore, the RBPs functioning during the flg22-induced defence responses (the flg22-responsive RBPome;

Chapter 4) are just the tip of the iceberg (the global immune RBPome). Moreover, pathogens are known to target host RBPs to disrupt immunity (**Chapter 1**). The most well studied example is AtGRP7, which is targeted by the *P. syringae* effector HopU1 (Fu et al., 2007; Nicaise et al., 2013; reviewed in **Chapter 1**). Therefore, RBPs can also be considered as susceptibility factors, which adds another layer of complexity and yet excitement. Hence, it will be exciting to apply ptRIC to plants treated with a variety of elicitors, effectors, pathogens and hormones, to disentangle the networks of RBPs that work together during different immunity responses or that are targeted by pathogens. Studies using a range of mutants and transgenic lines with alterations in immune signalling will help to dissect the exact function of RBPs within the immune response. One particular area of plant pathology that would greatly benefit from ptRIC is plant-viral interactions. Because viruses rely on the cellular machinery for replication and translation, pervasive changes of the plant RBPome are expected to be identified. For example, Garcia-Moreno and colleagues determined more than 200 RBPs modulated by the RNA virus Sindbis, many of which are hijacked by the virus to promote viral replication and spread (Garcia-Moreno et al., 2019).

In a broader context, the RBPomes established in this thesis represent only small subset of all the RBPs functioning in plants, since we have focused on leaves and a limited range of plant species and conditions (flg22). Hence, application of ptRIC to other plant tissues, species and conditions will greatly expand our knowledge of plant RBPs.

The need for target RNA identification

Because RBPs regulate RNAs and RNAs regulate RBPs, it becomes critical to identify the RNAs bound by each functional RBP. Recently, a number of approaches have been developed that allow the determination of target RNAs of RBPs as well as their binding sites *in vivo*. Multiple techniques and variants of them differ in their resolution, sensitivity and ease of use, but they

are all based on UV crosslinking, RNase treatment and immunoprecipitation (CLIP) of the RBP of interest, followed by sequencing of the covalently linked RNA fragments. In other words, this method reveals the footprints of RBPs in the cellular transcriptome. These techniques include CLIP-seq, PAR-CLIP, iCLIP and eCLIP (reviewed in Lee and Ule, 2018) and although fully stablished in mammals, they have proven to be challenging in plants. This is probably due to the low crosslinking efficiency of plant tissues (Köster and Meyer, 2018). However, our recently established UV crosslinking approach could now be applied to CLIP-based methods. These future studies of the RNA networks bound by these newly identified plant RBPs will serve to determine their specificity and precise function. One particularly interesting flg22-responsive RBP is EIN2. It has been described that a fraction of the cytoplasmic domain (CEND) of EIN2 is cleaved upon ethylene perception and binds RNA. Our proteomic data suggest that upon flg22 treatment, a longer isoform comprising the whole cytoplasmic domain is produced and it binds to RNA. It will be exciting to use CLIP approaches to identify the scope of RNAs bound by the full length or shorter forms of EIN2.

RBPs as targets for breeding programs

The results provided in this thesis, together with recent work, highlight the important role of RBPs as a hub regulating cellular homeostasis both in steady-state and in response to environmental, physiological and pathological stimuli. Since RBPs are critical for plant growth, development and survival, they represent excellent targets for breeding programs to improve plant traits, including pathogen resistance. Hence, knowledge on the regulation, specificity, structure and function of RBPs can be used to engineer RBPs to have improved features. For example, these could include modification of the RNA-binding affinity or specificity or to avoid being targeted by pathogens. Finally, by gaining information on RBPs from different species, we could also generate chimeric RBPs with improved features to achieve durable plant resistance to pathogens.

Conclusions

- We have improved RIC to be efficiently applied to plant leaves. This improved method, called plant RNA interactome capture (ptRIC), has allowed us to uncover the RBPome of relevant plant tissues and how it responds to biological cues.
- We have determined the most extensive Arabidopsis RBPome to date.
- We have identified the complement of RBPs functioning in multiple species across plant kingdom, many of which are novel RBPs.
- Our core RBPome reveals a large set of protein domains that frequently occur in RBPs and may harbour RNA-binding activity.
- We have discovered nearly 300 RBPs that respond to flg22, providing evidence of the existence of an RBP-mediated immune response.
- The contents of this thesis advance our knowledge of RNA-binding proteins in plants and represent an important resource for the plant and RNA community.

APPENDICES

Please note that the digital supplemental files are provided separately in a digital format.

APPENDIX I – Supplemental data Chapter 1

Supplemental table 1. RNA-binding proteins involved in plant immunity

APPENDIX II – Supplemental data Chapter 2

Supplemental table 1. Buffers used in ptRIC

APPENDIX III – Supplemental data Chapter 3

Note: supplemental table 4 is a summary of the results; the full data can be found in supplemental digital table 3.

Printed supplemental files

Supplemental Table 1. Classical RBDs

Supplemental table 2. Non-classical RBDs

Supplemental table 3. Selected putative novel RBDs

Supplemental table 4. Summary table of the core plant RBPome

Supplemental table 5. Putative lineage-specific RBPs

Digital supplemental files

Supplemental digital table 1. Superset of plant RBPs for all the species

Supplemental digital table 2 Domains not classified as RBDs and associated with RBPs

Supplemental digital table 3. Core RBPome

APPENDIX IV – Supplemental data Chapter 4

Note: supplemental tables 1 and 2 are a summary of the results; the full data can be found in supplemental digital tables 2 and 3, respectively.

Printed supplemental files

Supplemental table 1. Summary of flg22-responsive RBPome

Supplemental table 2. Summary of total proteome (WCL) flg22-responsive proteins

Supplemental table 3. RBP mutant lines screened

Digital supplemental files

Supplemental digital table 1. Arabidopsis leaf RBPomes

Supplemental digital table 2. Flg22-responsive RBPome

Supplemental digital table 3. Total proteome (WCL) responses to flg22

APPENDIX V – List of publications

List of publications

APPENDIX I – Supplemental data Chapter 1

Supplemental table 1. RNA-binding proteins involved in plant immunity

RNA metabolic Pathway	Protein name	Complete protein name	Alternative name(s)	Gene ID	Species	Specific pathway(s)	Possible function(s)	Pathogen / Elicitor / Effector	Reference(s)
Capping	**CPL1**	RNA POLYMERASE II C-TERMINAL DOMAIN (CTD) PHOSPHATASE-LIKE1	*FRY2; SHI4; JOE1; RCF2*	AT4G21670	*Arabidopsis thaliana*	RNA capping / RNA decay	RNA capping and decay by dephosphorylating the CTD of RNA Pol II	*Fusarium oxysporum; Alternaria brassicicola*	Jiang et al., 2013; Thatcher et al., 2018
Splicing machinery	**PRL1**	PLEIOTROPIC REULATORY LOCUS 1	MAC2	AT4G15900	*Arabidopsis thaliana*	Pre-mRNA splicing	Component of MAC (splicing) and sRNAs biogenesis	*Pseudomonas syringae; Hyaloperonospora parasitica*	Palma et al., 2007; Zhang et al., 2014
	MAC5A	MOS4-ASSOCIATED COMPLEX SUBUNIT 5A	AtC3H4	AT1G07360	*Arabidopsis thaliana*	Pre-mRNA splicing	Component of MAC (splicing)	-	Monaghan et al., 2010
	MAC5B	MOS4-ASSOCIATED COMPLEX SUBUNIT 5B	AtC3H25	AT2G29580	*Arabidopsis thaliana*	Pre-mRNA splicing	Component of MAC (splicing)	-	Monaghan et al., 2010
	MOS2	MODIFIER OF SNC1 2	-	AT1G33520	*Arabidopsis thaliana*	Pre-mRNA splicing	Linked to spliceosome and interacts with miRNA biogenesis machinery; basal defence as well as *R*-gene-mediated resistance	*Pseudomonas syringae; Phytophthora parasitica*; oomycetes	Zhang et al., 2005; Copeland et al., 2013; Wu et al., 2013
	RSZ21	SERINE/ ARGININE-RICH PROTEIN 21	-	-	*Arachis diogoi*	Pre-mRNA splicing	HR and induction of defence-related genes	*Phaeoisariopsis personata*	Kumar and Kirti, 2012
	SR45	SERINE/ARGININE-RICH 45	RNPS1	AT1G16610	*Arabidopsis thaliana*	Alternative splicing	Alternative splicing; genome stability; RNA-directed DNA methylation	*Pseudomonas syringae; Hyaloperonospora parasitica*	Zhang et al., 2017
	TCI14	TOBACCO CRYPTOGEIN-INDUCED 14	-	-	*Nicotiana tabacum* (tobacco)	Alternative splicing	Splicing	Cryptogein (elicitor of *Phytophthora)*	Petitot et al., 1997
	NSRA	NUCLEAR SPECKLE RNA BINDING PROTEIN A	-	AT1G76940	*Arabidopsis thaliana*	Alternative splicing	Alternative splicing	-	Bazin et al., 2018
	NSRB	NUCLEAR SPECKLE RNA	-	AT1G21320	*Arabidopsis thaliana*	Alternative splicing	Alternative splicing	-	Bazin et al., 2018

		BINDING PROTEIN B							
	GRP7	GLYCINE RICH PROTEIN 7	CCR2; GR-RBP7; RBGA3	AT2G21660	*Arabidopsis thaliana*	Alternative splicing	Alternative splicing; Polyadenylation; mRNA stability; targeted by *P. syringae effector*; regulate FLS2 and EFR	*Pseudomonas syringae; Pectobacterium carotovorum; Botrytis cinerea*	Fu et al., 2007; Jeong et al., 2011a; Lee et al., 2012b; Streitner et al., 2012; Nicaise et al., 2013
	GRP8	GLYCINE RICH PROTEIN 8	CCR1; GR-RBP8; RBGA6	AT4G39260	*Arabidopsis thaliana*	Alternative splicing	Alternative splicing; mRNA stability; targeted by *P. syringae effector*; regulate FLS2 and EFR	*Pseudomonas syringae*	Fu et al., 2007; Streitner et al., 2012
Polyadenylation	**FPA**	FLOWERING TIME CONTROL PROTEIN FPA	-	AT2G43410	*Arabidopsis thaliana*	Polyadenylation	Regulates the poly(A) site choice of ERF4 and possibly other defence related genes	*Pseudomonas syringae*	Lyons et al., 2013
	CPSF30	CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR30	OXT6; AtC3H11	AT1G30460	*Arabidopsis thaliana*	Polyadenylation	Controls of mRNA 3' end processing (poly(A) site selection) of defence related genes	*Pseudomonas syringae*	Bruggeman et al., 2014; Zhang et al., 2008; Shimberg et al., 2016
	PAPS1	POLY(A) POLYMERASE 1	-	AT1G17980	*Arabidopsis thaliana*	Polyadenylation	Poly(A) polymerase of specific mRNAs	*Hyaloperonospora arabidopsidis*	Kappel et al., 2015; Trost et al., 2014
	EDM3	ENHANCED DOWNY MILDEW 3	AIPP1	AT1G05970	*Arabidopsis thaliana*	Polyadenylation	Suppresses proximal transcript polyadeylation/termination of the immune receptor RPP7 mRNA; suppressor of H3K9me2 and cytosine methylation.	*Hyaloperonospora arabidopsidis*	Lai et al., 2018
RNA export	**SDE5**	SILENCING DEFECTIVE 5	-	AT3G15390	*Arabidopsis thaliana*	Nuclear mRNA export	mRNA export; transgene silencing and the production of trans-acting siRNAs	*Pseudomonas syringae; Erwinia caratovora*	Uddin et al., 2017; Hernandez-Pinzon et al., 2007
	MOS11	MODIFIER OF SNC1 11	-	AT5G02770	*Arabidopsis thaliana*	Nuclear mRNA export	mRNA export and/or RNA processing	*Pseudomonas syringae*; *Hyaloperonospora arabidopsidis*	Germain et al., 2010
	MOS3	*MODIFIER OF SNC1 3*	NUP96; PRE; SAR3	AT1G80680	*Arabidopsis thaliana*	Nuclear mRNA export	mRNA export	*Pseudomonas syringae; Phytophthora parasitica*	Zhang, 2005
	NUP160	NUCLEAR PORE COMPLEX PROTEIN NUP160	SAR1	AT1G33410	*Arabidopsis thaliana*	Nuclear mRNA export	mRNA export	*Pseudomonas syringae*	Wiermer et al., 2012

	SEH1	-	SEH1H	AT1G64350	*Arabidopsis thaliana*	Nuclear mRNA export	mRNA export	*Pseudomonas syringae*	Wiermer et al., 2012
	HPR1	HYPER RECOMBINATION1	THO1	AT5G09860	*Arabidopsis thaliana*	Nuclear mRNA export	mRNA export; may be involved in splicing	*Pseudomonas syringae; Hyaloperonospora arabidopsidis*	Pan et al., 2012
RNA decay / Deadenylation	**CAF1a; CAF1b**	CCR4-ASSOCIATED FACTOR 1	-	AT3G44260 ; AT5G22250	*Arabidopsis thaliana*	RNA stability/turnover (deadenylation)	3'-5' exoribonuclease; Regulates deadenylation of stress-related mRNAs.	*Pseudomonas syringae*	Liang et al., 2009; Walley et al., 2010
	CAF1a	CCR4-ASSOCIATED FACTOR 1	-	-	*Capsicum annuum* (Hot pepper)	RNA stability/turnover (deadenylation)	-	*Xanthomonas axonopodis; P. infestans*	Sarowar et al., 2007a
	PARN	POLY(A) RIBONUCLEASE	AHG2	AT1G55870	*Arabidopsis thaliana*	RNA stability/turnover (deadenylation)	Degrades poly(A) tails of specific RNAs	Cellotriose (elicitor from Piriformospora indica)	Johnson et al., 2018
RNA decay / Decapping	**DCP1**	DECAPPING 1	-	AT1G08370	*Arabidopsis thaliana*	RNA stability/turnover (decapping/PBs)	Decapping enhancer; involved in PBs; interacts with XRN4 upon PTI induction	*Pseudomonas syringae*	Yu et al., 2019
	DCP2	DECAPPING 2	ITS1, TDT, TRIDENT	AT5G13570	*Arabidopsis thaliana*	RNA stability/turnover (decapping/PBs)	Decapping enzyme	*Pseudomonas syringae*	Yu et al., 2019
	PAT1	PROTEIN-ASSOCIATED WITH TOPOISOMERASE 1	-	AT1G79090	*Arabidopsis thaliana*	RNA stability/turnover (decapping/deadenylation/PBs)	Decapping enhancer; translation repressor; involved in PBs	*Pseudomonas syringae*	Roux et al., 2015
	EDC4	ENHANCER OF MRNA DECAPPING 4	VCS	Traes_6DL_3FBA5B70E	*Nicotiana benthamiana/ Triticum aestivum* (wheat)	RNA stability/turnover (decapping/PBs)	Core decapping complex scaffold	Effector of *Puccinia striiformis*	Petre et al., 2016
mRNA decay	**BSR-K1**	BROAD-SPECTRUM RESISTANCE KITAAKE-1	-	Os10G0548200	*Oryza sativa* (Rice)	RNA stability/turnover (mRNA decay)	Promotes turnover of the defense-related *OsPAL* genes	*Magraporthe oryzae, Xanthomonas oryzae*	Zhou et al., 2018
	PRP-BP	PVPRP7 MRNA BINDING PROTEIN	-	-	Phaseolus vulgaris (Bean)	RNA stability/turnover	Binds to U-rich region of PvPRP1 mRNA and promotes its sdegradation	Elicitors of Colletotrichum lindemuthianum	Zhang and Mehdy, 1994
	RBP-DR1	RNA-BINDING PROTEIN-DEFENSE RELATED 1	BRN1	AT4G03110	*Arabidopsis thaliana*	RNA stability/turnover (mRNA decay)	Positive regulation of salicylic acid-mediated immunity and HR	*Pseudomonas syringae*	Qi et al., 2010; Kim et al., 2013
	RBP45b	POLYADENYLATE-BINDING PROTEIN RBP45B	-	AT1G11650	*Arabidopsis thaliana*	RNA stability/translation (suggested)	Potential role in mRNA stability and translation initiation	-	Peal et al., 2011; Muthuramalingam et al., 2017

PRs	**PR-10**	PATHOGENESIS-RELATED PROTEIN 10	-	-	Multiple species	RNA stability/turnover	Ribonuclease	Multiple pathogens	Jain and Kumar, 2015
	PR4	PATHOGENESIS-RELATED PROTEIN 4	-	-	Multiple species	RNA stability/turnover	Ribonuclease	Multiple pathogens	Filipenko et al., 2013
NMD factors	**UPF1**	REGULATOR OF NONSENSE TRANSCRIPTS UPF1	LBA1	AT5G47010	*Arabidopsis thaliana*	NMD	Nonsense-mediated decay	*Pseudomonas syringae*	Jeong et al., 2011b; Riehs-Kearnan et al., 2012; Shi et al., 2012; Rayson et al., 2012
	UPF3	REGULATOR OF NONSENSE TRANSCRIPTS UPF3	-	AT1G33980	*Arabidopsis thaliana*	NMD	Nonsense-mediated decay	*Pseudomonas syringae*	Jeong et al., 2011b
	SMG7	-	-	AT5G19400	*Arabidopsis thaliana*	NMD/PBs	Nonsense-mediated decay, recruitment of NMD complex to the PB	*Pseudomonas syringae*	Gloggnitzer et al., 2014; Merai et al., 2013
Stress granules and P body	**EIN2**	ETHYLENE-INSENSITIVE PROTEIN 2	CKR1; ORE 3	AT5G03280	*Arabidopsis thaliana*	PBs	CEND is cleaved, binds RNA and promotes translational repression in PB; Regulates FLS2 RNA and protein levels	-	Merchante et al., 2015; Tintor et al., 2013; Mersmann et al., 2010; Boutrot et al., 2010; Chen et al., 2009; Li et al., 2015
	AtTZF9	TANDEM ZINC FINGER PROTEIN 9	AtC3H66	AT5G58620	*Arabidopsis thaliana*	PBs and SGs	Associates with PBs and SGs	*Pseudomonas syringae*	Maldonado-Bonilla et al., 2014
	ZFP1	*ZINC FINGER PROTEIN 1*	-	-	*Gossypium hirsutum* (Cotton)	PBs and SGs	Involved in SA-dependednt plant defence repsonse	*Rhizoctonia solani*	Guo et al., 2009
	AN	ANGUSTIFOLIA	DOQ	AT1G01510	*Arabidopsis thaliana*	SGs	Interacts with SG-associated proteins and localizes in SG.	*Pseudomonas syringae*	Bhasin and Hülskamp, 2017
	TAF15b	TBP-ASSOCIATED FACTOR 1B	-	AT5G58470	*Arabidopsis thaliana*	PBs	TNL-mediated plant immunity	-	Dong et al., 2016
RNA editing	OCP3	OVEREXPRESSOR OF CATIONIC PEROXIDASE 3	-	AT5G11270	*Arabidopsis thaliana*	RNA editing (chloroplast)	Disease resistance regulator; controls editing efficiency of plastid ndhB transcripts	*Plectosphaerella cucumerina; Pseudomonas syringae; Botrytis cinerea; Hyaloperonospora arabidopsidis*	Ramírez et al., 2010

	SLO2	SLOW GROWTH 2	PCMP-E76	AT2G13600	*Arabidopsis thaliana*	RNA editing (mitochondria)	Required for response to ABA, ethylene, biotic, and abiotic stress.	*Botrytis cinerea*	Zhu et al., 2014
Translation	**RPL12C**	60S RIBOSOMAL PROTEIN L12	-	AT5G60670	*Arabidopsis thaliana, Nicotiana benthamiana*	Translation/Ribosomal proteins	Nonhost bacterial disease resistance	*Pseudomonas syringae; Xanthomonas campestris*	Nagaraj et al., 2016
	RPL19B	60S RIBOSOMAL PROTEIN L19	-	AT3G16780	*Arabidopsis thaliana, Nicotiana benthamiana*	Translation/Ribosomal proteins	Nonhost bacterial disease resistance	*Pseudomonas syringae; Xanthomonas campestris*	Nagaraj et al., 2016
	JIP60	JASMONATE-INDUCED PROTEIN 60	-	X66376.1	*Hordeum vulgare* (barley)	Translation	Recruits a subset of mRNA for translation	-	Rustgi et al., 2014
RNA helicases	**BIRH1**	BTH-INDUCED RNA HELICASE 1	-	Os03G0183 0	*Oryza sativa* (Rice)	RNA helicase	RNA helicase	*Alternaria brassicicola;Pseudomonas syringae*	Li et al., 2008
	PINP1	PSR1-INTERACTING PROTEIN 1	CUV; EMB3011; PRP16	AT5G13010	*Arabidopsis thaliana*	RNA helicase/Splicing/small RNA regulation	Regulates small RNAs; targeted by *Phytophthora* effector	*Phytophthora capsici*	Qiao et al., 2015; Tsugeki et al., 2015
General / Unknown pathways	**KRBP1**	K-HOMOLOGY (KH) RNA-BINDING PROTEIN 1	-	PGSC0003 DMT400066 837	*Solanum tuberosum (potato)*	Unknown	Susceptibility factor targeted by *P. infestans* effector	*Phytophthora infestans*	Wang et al., 2015
	RCF3	REGULATOR OF CBF GENE EXPRESSION 3	ESR1; SHI1;HOS5 -1	AT5G53060	*Arabidopsis thaliana*	Unknown	Negative regulator of biotic and abiotic stress resistance	*F*usarium *oxysporum*	Jiang et al., 2013a; Thatcher et al., 2015
	LIF2	LHP1-INTERACTING FACTOR 2	hnRNP-Q	AT4G00830	*Arabidopsis thaliana*	Unknown	Regulates transcription of SA and JA-genes	*Pseudomonas syringae; Botrytis cinerea*	Le Roux et al., 2014; Molitor et al., 2016
	UBA2a/b; UBA2c	UBP1-ASSOCIATED PROTEIN 2	-	-	*Solanum tuberosum (potato)*	Unknown	HR and induction of defence-related genes	-	Na et al., 2015
	G3BP1	RAS-GTPASE–ACTIVATING PROTEIN SH3-DOMAIN–BINDING PROTEINS		AT5G48650	*Arabidopsis thaliana*	Unknown	Negative regulator of stomatal and apoplastic immunity via SA-mediated defence	*Pseudomonas syringae*	Abulfaraj et al., 2018; Krapp et al., 2017
	GRP2; GRP3	GLYCINE-RICH PROTEIN 2 AND 3	-	-	*Hordeum vulgare* (barley)	Unknown	Upregulated upon infection with fungal pathogens	*Erysiphe graminis; Rhynchosporium secalis*	Molina et al., 1997
	PGN	PENTATRICOPEPTIDE REPEAT PROTEIN FOR GERMINATION ON NACL	PCMP-E64	AT1G56570	*Arabidopsis thaliana*	Unknown	Regulation of reactive oxygen species homeostasis	*Botrytis cinerea; Alternaria brassicicola*	Laluk et al., 2011

	RBP1	RNA-BINDING PROTEIN1 GENE	-	-	*Capsicum annuum* (Hot pepper)	Unknown	Required for HR and defence responses	*Xanthomonas campestris; Hyaloperonospora arabidopsidis*	Lee et al., 2012a
	RAP	RNA-BINDING DOMAIN ABUNDANT IN APICOMPLEXANS	-	AT2G31890	*Arabidopsis thaliana*	Unknown	Negative regulator of plant immunity by suppressing LSU2 and GLK1; targeted by bacterial siRNA	*Pseudomonas syringae; Hyaloperonospora arabidopsidis*	Wang et al., 2017
	GRS1	GLYCINE-RICH SHEPERIN 1	-	-	*Capsella bursa-pastoris*	Unknown	Activity against several bacteria and fungi	Multiple bacteria and fungi	Park et al., 2000

Supplemental table 1. RNA-binding proteins involved in plant immunity.
List of RBPs that have been linked to plant immunity in the literature. RBPs explained in the main text are marked in bold. Abbreviations: NMD, nonsense-mediated decay; PBs, Protein bodies; SGs, Stress granules

APPENDIX II – Supplemental data Chapter 2

Supplemental table 1. Buffers used in ptRIC

RIC Buffers	Tris-HCl, pH 7.5	LiCL	EDTA	LiDS	IGEPAL	PVP40	B-ME	DTT	Protease Inh.	RNase inh.
Lysis buffer	20 mM	500 mM	1 mM	0.50%	0.02%	2.50%	1%	5 mM	+	**100 ul**
Harsh Buffer	20 mM	2 M	1 mM	1%	0.02%	-	-	5 mM	-	-
Buffer I	20 mM	500 mM	1 mM	0.10%	0.02%	-	-	5 mM	-	-
Buffer II	20 mM	500 mM	1 mM	-	0.02%	-	-	5 mM	-	-
Buffer III + IGEPAL	20 mM	200 mM	1 mM	-	0.02%	-	-	5 mM	-	-
Buffer III NO IGEPAL	20 mM	200 mM	1 mM	-	-	-	-	5 mM	-	-
Elution buffer MS	20 mM	-	1 mM	-	-	-	-	-	-	-

Supplemental Table 1. Buffers used in ptRIC

The buffers containing Tris-HCl, LiCL, EDTA and LiDS are prepared, filtered and can be stored up to 3 months. IGEPAL, PVP40, B-ME, DTT, protease inhibitor and RNase inhibitor are added immediately before use. The catalogue numbers for each the reagents are the following: Tris (Sigma-aldrich cat.no. T1503), LiCL (Sigma-aldrich cat.no. 62476), EDTA (Sigma-aldrich cat.no. E5134), LiDS (Sigma-aldrich cat.no. L4632), IGEPAL (Sigma-aldrich cat.no. I3021), PVP40 (Sigma-aldrich cat.no. PVP40), B-ME (Sigma-aldrich cat.no M6250), DTT (Thermo Scientific cat.no. R0862), Protease inhibitor (Roche cat.no. 11873580001) and RNase inhibitor (Promega cat.no. N2611).

APPENDIX III – Supplemental data Chapter 3

Note: supplemental table 4 is a summary of the results; the full data can be found in supplemental digital table 3.

Printed supplemental files

Supplemental table 1. Classical RBDs

Domain name	Mp	Pp	Os	Pv	At	Nb	Sl	TOTAL
RRM	75	126	111	87	110	186	124	*819*
DEAD	18	32	9	6	24	22	17	*128*
zf-CCCH	13	15	18	16	12	31	20	*125*
KH	10	10	13	10	10	19	14	*86*
PUF	2	6	3	3	6	7	4	*31*
CSD	1	3	2	4	4	6	5	*25*
Piwi	1	2	0	0	2	4	6	*15*
dsrm	1	1	0	0	1	0	1	*4*

Supplemental table 1. Classical RBDs
Number of proteins containing each of the classical RBDs in each of the plant species
Abbreviations: Mp, *Marchantia polymorpha*; Pp, *Physcomitrella patens*; Os, *Oryza sativa*; Pv, *Phaseolus vulgaris*; At, *Arabidopsis thaliana*; Nb, *Nicotiana benthamiana*; Sl, *Solanum lycopersicum*.

Supplemental table 2. Non-classical RBDs

Domain name	Mp	Pp	Os	Pv	At	Nb	Sl	TOTAL
Ribosomal	124	268	102	152	194	242	171	1253
PPR	15	28	6	30	55	61	42	237
Helicase_C	19	31	9	6	23	21	16	125
zf-CCHC	8	24	11	9	13	20	16	101
YTH	1	1	5	5	10	9	4	35
GTP_EFTU	6	13	4	7	8	10	7	55
MMR_HSR1	8	16	0	3	8	8	8	51
GTP_EFTU_D2	5	12	4	6	7	8	6	48
S1	8	14	7	7	6	12	7	61
WD40	3	21	0	0	4	4	2	34
zf-RanBP	4	8	4	2	4	3	2	27
HABP4_PAI-RBP1	1	4	4	3	3	6	3	24
Glycolytic	1	1	1	1	2	2	3	11
HSP70	1	5	0	2	2	2	1	13
HTH_3	2	2	1	1	2	1	1	10
SNase	1	1	1	1	2	4	2	12
SSB	2	3	1	3	2	3	2	16
Cpn60_TCP1	0	1	0	1	1	0	1	4
FDF	1	3	2	1	1	2	1	11
HSP90	0	4	0	1	1	1	1	8
Ldh_1_C	1	2	0	0	1	0	0	4
Ldh_1_N	1	2	0	0	1	0	0	4
LSM	3	11	2	2	1	10	0	29
Pro_isomerase	2	3	1	0	1	1	1	9
PseudoU_synth_2	0	1	0	0	1	1	1	4
Thioredoxin	0	4	1	1	1	3	1	11
TMA7	0	0	0	0	1	0	0	1
14/03/2003	0	1	0	1	0	0	0	2
ACBP	0	1	0	0	0	1	0	2
Aldedh	1	0	0	0	0	0	0	1
FKBP_C	0	1	0	0	0	0	0	1
fn3	0	1	0	0	0	0	0	1
HMG_box	0	1	0	0	0	0	0	1
LRR_6	0	1	0	0	0	0	0	1
PAP_assoc	1	1	0	0	0	0	0	2
R3H	2	1	0	0	0	1	1	5
SAM_1	0	2	0	0	0	0	0	2
SAP	1	0	0	0	0	0	0	1

Supplemental table 2. Non-classical RBDs
Number of proteins containing each of the non-classical RBDs in the different plant species
Abbreviations: Mp, *Marchantia polymorpha*; Pp, *Physcomitrella patens*; Os, *Oryza sativa*; Pv, *Phaseolus vulgaris*; At, *Arabidopsis thaliana*; Nb, *Nicotiana benthamiana*; Sl, *Solanum lycopersicum*.

Supplemental table 3. Selected putative novel RBDs

RBD(s)	Annotation
Chloroa_b-bind	Part of the Light-harvesting complex (LHC)
Rubisco_large	Part of the large catalytic RuBisCO subunit
Rubisco_small (RbcS)	Part of the small RuBisCO subunit
PsaA_PsaB, PsaD, PsaL, PSI_PsaE, PSI_PsaF, PSI_PSAK,	Part of Photosystem I (PSI)
Psb28, PsbH, PsbP, PsbQ, PsbR, PSII, PSII_BNR, Cytochrom_B559	Part of Photosystem II (PSII)
MIP	Major intrinsic protein family of transmembrane channels.
HSP	Heat Shock Proteins functioning as chaperones
Whirly	Transcription factors
Alba	Transcriptional repression; novel chromosomal protein that coats archaeal DNA
DUF1296/GBF-interacting protein 1	Domain of unknown function
AAA	ATPases Associated with diverse cellular Activities
FoP_duplication	Friend of Prmt1; nuclear proteins associated with heterochromatin
MBF1	Multiprotein bridging factor 1
NTF2	Nuclear transport factor 2; nuclear import of cargo protein
PCI	Proteasome, COP9, Initiation factor 3; found in different proteasome subunits, the CSN and some eIF3
Peptidase_M24	Metallopeptidases
PGK	Phosphoglycerate kinase (glucose metabolism)
S10_plectin	Found in ribosomal S10 protein and cytoskeletal muscle protein plectin
Stm1_N	Suppressor of Tom1
Ubiquitin	Ubiquitination
TUDOR	Found in proteins that bind DNA and RNA
PAM2	PABP-interacting motif; important binding site for the PABC domain
Myb_DNA-bind_4	Family expanded in plants and present in several transposon proteins
FAD_binding	Mono o dinucleotide binding
DUF3355	Domain of unknown function
Epimerase	NADH dehydrogenase (ubiquinone)

Supplemental table 3. Selected putative novel RBDs
List of selected putative novel RBDs. The annotated function was obtained from Pfam.
Abbreviations: Mp, *Marchantia polymorpha*; Pp, *Physcomitrella patens*; Os, *Oryza sativa*; Pv, *Phaseolus vulgaris*; At, *Arabidopsis thaliana*; Nb, *Nicotiana benthamiana*; Sl, *Solanum lycopersicum*.

Supplemental table 4. Summary table of the core plant RBPome

Orthogroup	Generic protein name
OG0000000	Polyadenylate-binding protein RBP45
OG0000001	YTH domain-containing protein ECT
OG0000002	Serine/arginine-rich-splicing factor SR34
OG0000003	Polyubiquitin
OG0000004	Heterogeneous nuclear ribonucleoprotein (hnRNP)
OG0000005	Chlorophyll a-b binding protein, chloroplastic
OG0000006	Ribulose bisphosphate carboxylase small chain, chloroplastic (RuBisCO small subunit)
OG0000008	Alba DNA/RNA-binding protein
OG0000009	Cold, circadian rhythm; Glycine-rich RNA-binding protein
OG0000010	Eukaryotic initiation factor 4A (DEAD-box ATP-dependent RNA helicase)
OG0000011	Protein MEI2-like (AML)
OG0000012	U2 snRNP auxiliary factor large subunit
OG0000013	Heterogeneous nuclear ribonucleoprotein Q (hnRNP Q)
OG0000014	Heat shock cognate protein 70
OG0000015	ADP, ATP carrier protein, mitochondrial (ADP/ATP translocase)
OG0000016	Nuclear transport factor 2 (AtNTF2)
OG0000017	RNA-binding protein CP29, chloroplastic (Ribonucleoprotein)
OG0000019	Tetratricopeptide repeat (TPR)-like superfamily protein
OG0000020	Serine/arginine-rich splicing factor SC35
OG0000021	RNA-binding KH domain-containing protein
OG0000022	Elongation factor 1-alpha (EF-1-alpha) (eEF-1A2)
OG0000023	DEAD-box ATP-dependent RNA helicase
OG0000024	CRS1 / YhbY (CRM) domain-containing protein
OG0000025	Serine/arginine-rich splicing factor
OG0000026	RNA-binding (RRM/RBD/RNP motifs) family protein (Ribonucleoprotein-like)
OG0000027	Ran BP2/NZF zinc finger-like superfamily protein
OG0000029	40S ribosomal protein S4
OG0000030	Plasma membrane ATPase
OG0000031	Serine/arginine-rich splicing factor (RS-containing zinc finger protein)
OG0000032	Oligouridylate binding protein
OG0000035	CTC-interacting domain; Polyadenylate-binding protein-interacting protein (PABP-interacting protein)
OG0000036	60S ribosomal protein L7
OG0000037	RGG repeats nuclear RNA binding protein
OG0000039	(S)-2-hydroxy-acid oxidase; Aldolase-type TIM barrel family protein
OG0000040	Nucleic acid-binding, OB-fold-like protein
OG0000041	Fructose-bisphosphate aldolase
OG0000042	Zinc finger CCCH domain-containing protein
OG0000043	40S ribosomal protein S17
OG0000044	Ribosomal protein L14/L23 family protein
OG0000045	RNA-binding (RRM/RBD/RNP motifs) family protein; THO complex subunit
OG0000046	40S ribosomal protein S7
OG0000047	Pumilio homolog
OG0000048	Zinc finger CCCH domain-containing protein
OG0000049	Polyadenylate-binding protein
OG0000050	Polypyrimidine tract-binding protein
OG0000051	60S ribosomal protein L22-2
OG0000052	RNA-binding (RRM/RBD/RNP motifs) family protein
OG0000053	Cold shock protein; Glycine-rich protein
OG0000054	60S ribosomal protein L19
OG0000055	DExH-box ATP-dependent RNA helicase
OG0000056	Zinc finger CCCH domain-containing protein
OG0000057	Polyadenylate-binding protein
OG0000058	60S ribosomal protein L35

OG0000059	40S ribosomal protein S11
OG0000060	60S ribosomal protein L24
OG0000061	40S ribosomal protein S8
OG0000062	Elongation factor Ts, mitochondrial (EF-Ts)
OG0000063	RNA-binding (RRM/RBD/RNP motifs) family protein
OG0000064	Alba DNA/RNA-binding protein
OG0000065	40S ribosomal protein S24
OG0000066	30S ribosomal protein S7, chloroplastic
OG0000067	60S ribosomal protein L34
OG0000068	60S ribosomal protein L35a
OG0000069	40S ribosomal protein S13-2
OG0000070	60S ribosomal protein L13a
OG0000072	40S ribosomal protein S18;40S ribosomal protein S18;40S ribosomal protein S18
OG0000073	Nucleolin 1
OG0000074	UBP1-associated protein 2C
OG0000075	60S ribosomal protein L11
OG0000076	40S ribosomal protein S9
OG0000077	Serine/arginine-rich splicing factor
OG0000079	Glycine-rich RNA-binding protein
OG0000080	40S ribosomal protein S20
OG0000081	60S ribosomal protein L28
OG0000082	Ubiquitin-60S ribosomal protein L40
OG0000083	60S ribosomal protein L8
OG0000084	60S ribosomal protein L18a
OG0000085	60S ribosomal protein L7a-
OG0000086	40S ribosomal protein S3
OG0000087	UBP1-associated protein 2
OG0000088	60S ribosomal protein L13
OG0000089	40S ribosomal protein S25
OG0000090	60S ribosomal protein L4-1
OG0000091	Elongation factor 2 (EF-2)
OG0000092	Protein MLN51 homolog (Protein barentsz)
OG0000093	60S ribosomal protein L26
OG0000094	60S ribosomal protein L15
OG0000095	60S ribosomal protein L10
OG0000096	40S ribosomal protein SA
OG0000097	60S ribosomal protein L5
OG0000098	40S ribosomal protein S2
OG0000099	40S ribosomal protein S10
OG0000100	Nascent polypeptide-associated complex subunit beta
OG0000101	Zinc finger CCCH domain-containing protein 41
OG0000102	40S ribosomal protein S5
OG0000103	60S ribosomal protein L27
OG0000104	40S ribosomal protein S16
OG0000105	Ubiquitin-40S ribosomal protein S27a
OG0000106	Zinc finger CCCH domain-containing protein
OG0000107	Beta carbonic anhydrase 1, chloroplastic
OG0000108	Pentatricopeptide repeat-containing protein
OG0000110	60S ribosomal protein L17
OG0000111	60S ribosomal protein L37
OG0000112	40S ribosomal protein S3a
OG0000113	60S ribosomal protein L14
OG0000114	Heat shock protein 90
OG0000115	40S ribosomal protein S28
OG0000116	Protein decapping 5
OG0000117	Splicing factor U2AF
OG0000119	Aquaporin PIP2
OG0000121	60S ribosomal protein L39
OG0000122	40S ribosomal protein S14

OG0000123	40S ribosomal protein S6
OG0000124	60S ribosomal protein L30
OG0000125	Flowering time control protein FCA
OG0000126	Putative RNA-binding protein
OG0000127	RNA-binding protein BRN
OG0000128	60S ribosomal protein L27a
OG0000129	60S ribosomal protein L31
OG0000130	60S ribosomal protein L32
OG0000131	40S ribosomal protein S30
OG0000132	Translation initiation factor SUI1 family protein
OG0000133	DEAD-box ATP-dependent RNA helicase
OG0000134	60S ribosomal protein L3
OG0000135	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase; Cobalamin-independent methionine synthase
OG0000136	Protein argonaute 1
OG0000137	DEAD-box ATP-dependent RNA helicase
OG0000138	40S ribosomal protein S23
OG0000139	Ribulose bisphosphate carboxylase/oxygenase activase, chloroplastic (RA)
OG0000140	60S ribosomal protein L2, mitochondrial
OG0000141	40S ribosomal protein S19
OG0000143	Eukaryotic translation initiation factor 3 subunit D
OG0000144	50S ribosomal protein L2, chloroplastic
OG0000145	30S ribosomal protein S5, chloroplastic
OG0000146	40S ribosomal protein S27
OG0000147	Ribosomal protein L18e/L15 superfamily protein
OG0000148	Glycine-rich RNA-binding protein, mitochondrial
OG0000149	RNA-binding (RRM/RBD/RNP motifs) family protein)
OG0000150	Nucleic acid-binding proteins superfamily
OG0000152	ERBB-3 BINDING PROTEIN 1; Proliferation-associated protein G2p
OG0000154	Pentatricopeptide repeat-containing protein, mitochondrial
OG0000155	DNA topoisomerase, type IA, core
OG0000156	60S ribosomal protein L6
OG0000157	Serine/arginine-rich SC35-like splicing factor SCL30A
OG0000158	Oxygen-evolving enhancer protein 1, chloroplastic
OG0000159	Multiprotein-bridging factor 1
OG0000160	Pentatricopeptide repeat-containing protein, chloroplastic
OG0000162	Nucleic acid-binding, OB-fold-like protein; Putative translation initiation factor eIF-1A
OG0000163	THO complex subunit 4
OG0000165	Eukaryotic translation initiation factor 4B
OG0000166	Eukaryotic translation initiation factor 5A
OG0000167	Phosphoglycerate kinase 1, chloroplastic
OG0000168	Glyceraldehyde-3-phosphate dehydrogenase GAPA, chloroplastic
OG0000169	RNA-binding (RRM/RBD/RNP motifs) family protein
OG0000172	Ribonuclease TUDOR
OG0000173	Ubiquitin-NEDD8-like protein RUB
OG0000174	Eukaryotic translation initiation factor 3 subunit
OG0000175	Pumilio 5
OG0000176	DEAD-box ATP-dependent RNA helicase, mitochondrial
OG0000178	Elongation factor Tu, chloroplastic (EF-Tu)
OG0000180	30S ribosomal protein S18, chloroplastic
OG0000181	40S ribosomal protein S26
OG0000182	50S ribosomal protein L15, chloroplastic
OG0000183	DNA-binding protein (Putative helicase); P-loop containing nucleoside triphosphate hydrolases superfamily protein
OG0000184	DEAD-box ATP-dependent RNA helicase 3, chloroplastic
OG0000185	Chlorophyll a-b binding protein, chloroplastic
OG0000187	Polypyrimidine tract-binding protein
OG0000189	60S acidic ribosomal protein P0
OG0000190	Eukaryotic translation initiation factor 3 subunit

OG0000195	40S ribosomal protein S15
OG0000196	Uncharacterized protein
OG0000198	Photosystem I reaction center subunit III, chloroplastic (PSI-F)
OG0000199	50S ribosomal protein L9, chloroplastic
OG0000200	Eukaryotic translation initiation factor 3 subunit
OG0000202	Ribonuclease J
OG0000203	Probable mediator of RNA polymerase II transcription subunit 36 (Histone-glutamine methyltransferase) (SKP1-interacting partner) (rRNA 2'-O-methyltransferase fibrillarin)
OG0000204	Probable nucleolar protein 5 (MAR-binding NOP56/58 homolog)
OG0000206	Thioredoxin H3
OG0000210	50S ribosomal protein L4, chloroplastic
OG0000211	Eukaryotic translation initiation factor 3 subunit
OG0000212	La-related protein 1A
OG0000213	Photosystem I subunit l; Photosystem I reaction center subunit XI, chloroplastic (PSI-L) (PSI subunit V)
OG0000215	Glycine dehydrogenase (decarboxylating), mitochondrial
OG0000216	Protein BTR1 (Binding to ToMV RNA 1)
OG0000218	GTP-binding protein OBGC, chloroplastic
OG0000220	Photosystem I reaction center subunit IV, chloroplastic (PSI-E A)
OG0000221	Pentatricopeptide repeat-containing protein
OG0000222	30S ribosomal protein S1, chloroplastic
OG0000223	Glutamate--glyoxylate aminotransferase; Alanine aminotransferase; Alanine--glyoxylate aminotransferase; Alanine-2-oxoglutarate aminotransferase
OG0000224	Sm-like protein LSM36B; U6 snRNA-associated Sm-like protein LSM6B
OG0000225	Ribosome maturation factor
OG0000226	Zinc finger CCCH domain-containing protein 17
OG0000229	Transcription factor Pur-alpha 1 (Purine-rich single-stranded DNA-binding protein alpha 1)
OG0000234	60S ribosomal protein L23a
OG0000235	Protein HHL1, chloroplastic (Hypersensitive to high light 1)
OG0000236	Apoptosis inhibitory protein 5
OG0000238	50S ribosomal protein L20, chloroplastic
OG0000239	50S ribosomal protein L6, chloroplastic
OG0000240	Putative heat shock factor protein hsf8; Nucleic acid-binding, OB-fold-like protein
OG0000241	High chlorophyll fluorescence phenotype 173
OG0000243	50S ribosomal protein L21, chloroplastic
OG0000244	DExH-box ATP-dependent RNA helicase DExH15 chloroplastic
OG0000245	Pentatricopeptide repeat-containing protein, mitochondrial
OG0000246	50S ribosomal protein L13, chloroplastic
OG0000247	Translation initiation factor IF3
OG0000248	50S ribosomal protein L1, chloroplastic
OG0000249	Zinc finger CCCH domain-containing protein
OG0000250	Nucleic acid-binding proteins superfamily
OG0000251	Chaperone protein ClpC1, chloroplastic
OG0000252	50S ribosomal protein L24
OG0000257	30S ribosomal protein S2, chloroplastic
OG0000258	KH domain-containing protein
OG0000259	ATPase E1 (AtE1)
OG0000260	Putative H/ACA ribonucleoprotein complex subunit 1
OG0000261	Ribosomal protein S21 family protein
OG0000270	Single-stranded DNA-binding protein WHY, chloroplastic
OG0000271	Chlorophyll a-b binding protein, chloroplastic
OG0000272	Pentatricopeptide repeat-containing protein, chloroplastic
OG0000273	DEAD-box ATP-dependent RNA helicase 50
OG0000274	Pentatricopeptide repeat-containing protein
OG0000275	Proton pump interactor 1
OG0000276	Pentatricopeptide repeat-containing protein, chloroplastic
OG0000277	DEAD-box ATP-dependent RNA helicase 26
OG0000278	Carboxylate clamp-tetratricopeptide repeat protein; Heat shock protein 70 (HSP70)-interacting protein; Putative myosin adapter B2

OG0000279	Elongation factor G, chloroplastic
OG0000280	CRS2-associated factor 2, chloroplastic
OG0000282	Ribosomal protein S13/S18 family
OG0000287	Glycine-rich protein
OG0000288	Splicing factor
OG0000289	Elongation factor 4; Translation factor GUF1 homolog, chloroplastic
OG0000291	Pentatricopeptide repeat-containing protein, mitochondrial
OG0000302	La protein 1
OG0000303	RNA-binding (RRM/RBD/RNP motifs) family protein
OG0000304	Photosystem II 22 kDa protein, chloroplastic; Chlorophyll A-B binding family protein
OG0000305	30S ribosomal protein S12, chloroplastic;30S ribosomal protein S12, chloroplastic
OG0000306	Ribulose bisphosphate carboxylase large chain (RuBisCO large subunit)
OG0000307	50S ribosomal protein L16, chloroplastic
OG0000308	Photosystem I reaction center subunit II-2, chloroplastic (PSI-D2)
OG0000309	Oxygen-evolving enhancer protein 2, chloroplastic
OG0000311	Single-stranded DNA-binding protein WHY2, mitochondrial
OG0000314	50S ribosomal protein L5, chloroplastic
OG0000315	50S ribosomal protein L27, chloroplastic
OG0000316	Pentatricopeptide repeat-containing protein, chloroplastic
OG0000317	La-related protein
OG0000318	Antitermination NusB domain-containing protein
OG0000319	30S ribosomal protein
OG0000321	DNA repair ATPase-like protein; Myosin heavy chain-like protein
OG0000324	Nuclear/nucleolar GTPase 2
OG0000325	Pescadillo homolog
OG0000336	Transketolase, chloroplastic
OG0000338	Chloroplast RNA binding protein
OG0000340	GTPase ERA-like, chloroplastic
OG0000341	30S ribosomal protein S4, chloroplastic
OG0000342	30S ribosomal protein S3, chloroplastic
OG0000343	50S ribosomal protein L22, chloroplastic
OG0000344	Translation initiation factor IF-2, chloroplastic
OG0000345	30S ribosomal protein S9, chloroplastic
OG0000346	Photosystem II 10 kDa polypeptide, chloroplastic
OG0000349	Polyribonucleotide nucleotidyltransferase 1, chloroplastic
OG0000350	Chlorophyll a-b binding protein CP29
OG0000351	RNA methyltransferase family protein (RNA methyltransferase-like protein)
OG0000353	DEAD-box ATP-dependent RNA helicase 39
OG0000354	50S ribosomal protein L35, chloroplastic
OG0000355	GTP-binding family protein
OG0000356	Protein WHAT'S THIS FACTOR 1 homolog, chloroplastic
OG0000359	GTP binding protein-like
OG0000361	Guanine nucleotide-binding protein-like NSN; Nucleolar GTP-binding protein NSN
OG0000362	Zinc finger CCCH domain-containing protein 37
OG0000363	Aquaporin TIP1
OG0000364	Ribonuclease III domain-containing protein RNC1, chloroplastic
OG0000367	Organelle RRM domain-containing protein 1, chloroplastic
OG0000368	Cryptochrome DASH, chloroplastic/mitochondrial
OG0000377	Oxygen-evolving enhancer protein 3-1
OG0000378	Pentatricopeptide repeat-containing protein, chloroplastic
OG0000380	ATP synthase subunit alpha, chloroplastic
OG0000381	Photosystem II CP43 reaction center protein
OG0000385	T1N15.19 (Zinc finger (Ran-binding) family protein)
OG0000386	50S ribosomal protein L3-1, chloroplastic
OG0000387	30S ribosomal protein S20, chloroplastic
OG0000389	Chloroplast stem-loop binding protein of 41 kDa, chloroplastic
OG0000390	Chlorophyll a-b binding protein CP26, chloroplastic
OG0000392	RNA-binding CRS1 / YhbY (CRM) domain protein
OG0000395	Pentatricopeptide repeat-containing protein, chloroplastic

OG0000398	Pentatricopeptide repeat-containing protein
OG0000402	60S ribosomal export protein NMD3
OG0000404	Nucleic acid-binding, OB-fold-like protein
OG0000406	ABC transporter F family member 5 (ABC transporter ABCF.5) (AtABCF5) (GCN20-type ATP-binding cassette protein GCN5); ABC transporter F family member 2 (ABC transporter ABCF.2) (AtABCF2) (GCN20-type ATP-binding cassette protein GCN2)
OG0000407	Ribosomal protein L25/Gln-tRNA synthetase, anti-codon-binding domain-containing protein
OG0000408	RNA-binding protein
OG0000412	Monocarboxylic acid transporters
OG0000413	Uncharacterized protein
OG0000423	Photosystem II protein D1
OG0000424	50S ribosomal protein L14, chloroplastic
OG0000426	Nucleic acid-binding protein, putative
OG0000429	30S ribosomal protein S17, chloroplastic
OG0000430	Putative chloroplast RNA binding protein (RNA-binding (RRM/RBD/RNP motifs) family protein)
OG0000434	RNA-binding protein CP33, chloroplastic
OG0000435	Translation initiation factor IF-1, chloroplastic
OG0000447	Tumor necrosis factor receptor family protein
OG0000453	Uncharacterized protein
OG0000473	Photosystem II D2 protein
OG0000474	Cytochrome b6
OG0000475	30S ribosomal protein S11, chloroplastic
OG0000487	GTP-binding protein BRASSINAZOLE INSENSITIVE PALE GREEN 2, chloroplastic

Supplemental table 4. Summary table of the core plant RBPome

List of orthogroups comprising the core plant RBPome. Orthogroups are classified as core plant RBPome if found in at least five out of the seven plant RBPomes. For each of the orthogroups a generic protein name is included based on the names of the individual RBPs within the orthogroup.

Supplemental table 5. Putative lineage-specific RBPs

Lineage-specific	Orthogroup	Generic protein name
Bryophyte-specific	OG0000118	chalcone synthase-like
	OG0000151	pumilio homolog-like;
	OG0000191	hornerin-like isoform X
	OG0000192	epithelial splicing regulatory protein 1-like isoform X; heterogeneous nuclear ribonucleoprotein F-like isoform X
	OG0000207	uncharacterized protein
	OG0000208	uncharacterized protein
	OG0000230	protein EXORDIUM-like
	OG0000293	pre-mRNA-processing factor 39-like isoform X
	OG0000294	uncharacterized protein; protein terminal ear1 homolog isoform X2
	OG0000296	CAP-Gly domain-containing linker protein -like; major antigen-like isoform X;myosin-1-like isoform X;
	OG0000369	uncharacterized protein
	OG0000410	probable ADP-ribosylation factor GTPase-activating protein AGD8
	OG0000449	tRNA 2-phosphotransferase 1-like isoform X
	OG0000450	40S ribosomal protein S15a
	OG0000452	uncharacterized protein
	OG0000454	pentatricopeptide repeat-containing protein, mitochondrial-like
	OG0000457	uncharacterized protein
	OG0000458	RNA-binding protein cabeza-like
	OG0000491	germin-like protein 9-3
	OG0000492	uncharacterized protein
	OG0000494	xyloglucan endotransglucosylase/hydrolase protein 9-like
	OG0000495	protein SUPPRESSOR OF GENE SILENCING 3 homolog isoform X
	OG0000496	glucomannan 4-beta-mannosyltransferase 9-like
	OG0000497	probable sugar phosphate/phosphate translocator
	OG0000498	multiprotein-bridging factor 1a-like
	OG0000536	uncharacterized protein
	OG0000537	translocon-associated protein subunit alpha-like
	OG0000538	uncharacterized protein
	OG0000539	nuclear pore complex protein GP210-like isoform X
	OG0000540	pentatricopeptide repeat-containing protein, mitochondrial-like
	OG0000542	uncharacterized protein
	OG0000591	protein transport protein Sec61 subunit beta-like isoform X
	OG0000592	peroxidase 55-like
	OG0000594	zinc finger CCCH domain-containing protein 11-like
	OG0000595	acetyl-coenzyme A synthetase, chloroplastic/glyoxysomal-like
	OG0000597	golgin subfamily A member 6-like protein 22
	OG0000598	tudor domain-containing protein 3-like isoform X
	OG0000600	guanosine nucleotide diphosphate dissociation inhibitor 1-like
	OG0000601	keratin, type I cytoskeletal 9-like; eukaryotic translation initiation factor 4B2-like isoform X2
	OG0000602	methionine aminopeptidase 2B-like
	OG0000604	U1 small nuclear ribonucleoprotein 70 kDa-like
	OG0000605	V-type proton ATPase subunit a3-like
	OG0000606	mitochondrial outer membrane protein porin of 36 kDa-like
	OG0000607	RNA-binding protein CP31B, chloroplastic-like
	OG0000608	RNA-binding protein FUS-like
	OG0000609	ABC transporter C family member 14-like
	OG0000610	multiple RNA-binding domain-containing protein 1-like
	OG0000611	pentatricopeptide repeat-containing protein -like
	OG0000612	uncharacterized protein
	OG0000656	NADH dehydrogenase subunit 5 (mitochondrion)
	OG0000657	obg-like ATPase 1

	OG0000658	splicing factor 3B subunit 1-like
	OG0000659	dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B-like
	OG0000660	hornerin-like
	OG0000661	dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A-like
	OG0000662	dnaJ protein ERDJ2A-like
	OG0000663	trans-cinnamate 4-monooxygenase-like
	OG0000665	DExH-box ATP-dependent RNA helicase DExH14-like isoform X
	OG0000666	probable pyridoxal 5-phosphate synthase subunit PDX1
	OG0000667	photosystem I reaction center subunit V, chloroplastic-like
	OG0000668	transducin beta-like protein 2
	OG0000669	pentatricopeptide repeat-containing protein OTP51, chloroplastic-like
	OG0000670	UDP-xylose transporter 2-like
	OG0000671	DEAD-box ATP-dependent RNA helicase 7-like
	OG0000672	rRNA biogenesis protein RRP5-like
	OG0000673	phospholipid:diacylglycerol acyltransferase 1-like isoform X
	OG0000674	signal recognition particle subunit SRP72-like
	OG0000675	eukaryotic translation initiation factor 3 subunit E-like
	OG0000676	transmembrane 9 superfamily member-like
	OG0000677	ATP-dependent RNA helicase HAS1-like
	OG0000678	uncharacterized protein
	OG0000679	U2 small nuclear ribonucleoprotein B 2-like
	OG0000680	uncharacterized protein
	OG0000681	ribosomal L1 domain-containing protein 1-like
	OG0000682	sulfate transporter 1.1-like
Angiosperm-specific	OG0000047	Pumilio homolog (APUM)
	OG0000056	Zinc finger C-x8-C-x5-C-x3-H type family protein; Zinc finger CCCH domain-containing protein
	OG0000079	Glycine-rich RNA-binding protein
	OG0000087	UBP1-associated protein 2A (UBP1-interacting protein 2a)
	OG0000106	Zinc finger C-x8-C-x5-C-x3-H type family protein; Zinc finger CCCH domain-containing protein
	OG0000163	ALWAYS EARLY 4; THO complex subunit 4 (ALYREF homolog 3)
	OG0000169	RNA-binding (RRM/RBD/RNP motifs) family protein
	OG0000175	Pumilio 5
	OG0000251	Chaperone protein ClpC1, chloroplastic; Casein lytic proteinase C1; Protein DE-REGULATED CAO ACCUMULATION 1; Protein IRON-RESCUED MUTANT 1
	OG0000270	Single-stranded DNA-binding protein WHY, chloroplastic; Protein PLASTID TRANSCRIPTIONALLY ACTIVE 1
	OG0000275	Proton pump interactor 1
	OG0000303	RNA-binding (RRM/RBD/RNP motifs) family protein
	OG0000311	Single-stranded DNA-binding protein WHY2, mitochondrial (Protein WHIRLY 2)
	OG0000367	Organelle RRM domain-containing protein 1, chloroplastic
	OG0000426	Putative nucleic acid-binding protein; RNA-binding (RRM/RBD/RNP motifs) family protein)
	OG0000430	Putative chloroplast RNA binding protein (RNA-binding (RRM/RBD/RNP motifs) family protein)
	OG0000434	RNA-binding protein CP33, chloroplastic; Protein PIGMENT DEFECTIVE 322
Dicot-specific	OG0000007	Pentatricopeptide repeat-containing protein, chloroplastic
	OG0000188	Pentatricopeptide repeat-containing protein, mitochondrial
	OG0000313	Pentatricopeptide repeat-containing protein, chloroplastic; Protein MATERNAL EFFECT EMBRYO ARREST 40
	OG0000366	RNA-binding (RRM/RBD/RNP motifs) family protein
	OG0000379	Pentatricopeptide repeat (PPR-like) superfamily protein
	OG0000384	Glyceraldehyde-3-phosphate dehydrogenase GAPB, chloroplastic
	OG0000405	Eukaryotic translation initiation factor 4B3 (eIF4B3)
	OG0000421	Pentatricopeptide repeat-containing protein, chloroplastic
	OG0000427	Organelle RRM domain-containing protein 6, chloroplastic

	OG0000428	Pentatricopeptide repeat-containing protein, chloroplastic
	OG0000436	50S ribosomal protein L28, chloroplastic (CL28)
	OG0000446	Pentatricopeptide repeat-containing protein, mitochondrial
	OG0000484	Multiple organellar RNA editing factor 6, mitochondrial (RNA editing-interacting protein 6)
	OG0000485	Pentatricopeptide repeat-containing protein, mitochondrial
	OG0000528	Pentatricopeptide repeat-containing protein, chloroplastic; Protein EMBRYO DEFECTIVE 2750
Solanaceae-specific	OG0000300	Ataxin-2, C-terminal; Polyadenylate-binding protein-interacting protein 9
	OG0000416	Multiple organellar RNA editing factor 1, mitochondrial-like isoform X1; multiple organellar RNA editing factor 1, mitochondrial-like isoform X2
	OG0000464	Heterogeneous nuclear ribonucleoprotein Q
	OG0000502	Vasa; Cold shock protein 2;Cold shock domain protein 1
	OG0000505	Far upstream element-binding protein 2; KH domain-containing protein;F14M2.18 protein
	OG0000547	Putative G3BP-like protein; BnaC09g17770D protein
	OG0000550	Lupus la ribonucleoprotein, putative; JHL20J20.1 protein;LA-related protein 6 LA RNA-binding domain protein
	OG0000553	Heterogeneous nuclear ribonucleoprotein Q; Water-stress protein
	OG0000555	Uncharacterised protein
	OG0000613	RNA-binding protein 1
	OG0000615	Histone H2A
	OG0000616	Histone H2A
	OG0000618	Pentatricopeptide repeat 5
	OG0000621	Pentatricopeptide repeat-containing protein, chloroplastic
	OG0000623	DNA/RNA-binding protein Alba-like protein
	OG0000626	Protein argonaute 8
	OG0000627	Glycine-rich RNA-binding protein 3, mitochondrial
	OG0000685	Histone H2A
	OG0000686	Histone H2A
	OG0000688	RAP domain-containing protein, chloroplastic
	OG0000690	Ataxin-2, C-terminal
	OG0000691	Pentatricopeptide repeat-containing protein, mitochondrial
	OG0000692	Protein argonaute 4
	OG0000693	Putative uncharacterized protein Sb01g003150
	OG0000694	Heterogeneous nuclear ribonucleoprotein D0
	OG0000695	Glycine-rich RNA-binding protein 5, mitochondrial

Supplemental table 5. Putative lineage-specific RBPs
Proteins identified as putative lineage specific RBPs. For each of the proteins we include a generic protein name based on the names of the RBPs from all the species.

APPENDIX IV – Supplemental data Chapter 4

Note: supplemental tables 1 and 2 are a summary of the results; the full data can be found in supplemental digital tables 2 and 3, respectively.

Printed supplemental files

Supplemental table 1. Summary of flg22-responsive RBPome

Gene ID(s)	Protein name(s)	Summary activity	Statistical criteria	RBP Function / Family
AT2G33380	PXG3 (Probable peroxygenase 3); RESPONSIVE TO DESICCATION 20	Inhibited (12h)	candidate	Calcium binding
AT2G39200	MLO12 (MLO-like protein 12)	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Calcium binding
AT2G41100	TCH3; CML12 (Calmodulin-like protein 12); CAL4 (Calmodulin-like protein 4)	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Calcium binding
AT4G37640	ACA2 (Calcium-transporting ATPase 2)	Stimulated (2h)	candidate	Calcium binding
AT5G23060	Calcium sensing receptor, chloroplastic; Sulfurtransferase 3	Stimulated (2h)	high-confidence	Calcium binding
AT5G61790	Calnexin homolog 1	Stimulated (2h)	high-confidence	Calcium binding
AT1G10290, AT1G59610	Dynamin-2A/B	Stimulated (2h); Stimulated (12h)	candidate (early); high-confidence (late)	Cytoskeleton
AT1G20010	TUB5 (Tubulin beta-5 chain)	Stimulated (2h)	high-confidence	Cytoskeleton
AT1G49240, AT3G18780, AT3G18780	Actin-8; Actin-2; Actin 2	Stimulated (2h)	high-confidence	Cytoskeleton
AT2G03680	Protein SPIRAL1; Protein NAP16kDa	Inhibited (12h)	high-confidence	Cytoskeleton
AT5G09810	Actin-7 (Actin-2)	Stimulated (12h)	candidate	Cytoskeleton
AT5G12250	TUB6 (Tubulin beta-6 chain)	Stimulated (2h)	high-confidence	Cytoskeleton
AT5G19770, AT5G19780	TUA3/5 (Tubulin alpha-3 chain)	Stimulated (2h); Stimulated (12h)	candidate (both)	Cytoskeleton
AT5G44340, AT4G20890	TUB4/9 (Tubulin beta-4/9 chain)	Stimulated (2h)	high-confidence	Cytoskeleton
AT5G62700, AT5G62690	TUB3/2 (Tubulin beta-3/2 chain)	Stimulated (2h); Stimulated (12h)	high-confidence (early); candidate (late)	Cytoskeleton
AT5G17790	VAR3 (VARIEGATED 3), chloroplastic; Organelle Zinc finger 1	Inhibited (12h)	high-confidence	Editing
AT5G54580	ORRM2 (Organelle RRM domain-containing protein 2), mitochondrial	Stimulated (2h)	candidate	Editing
AT3G22150	AEF1; PCMP-E95; Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	candidate	Editing/PPR

AT5G48910	LPA66; Pentatricopeptide repeat-containing protein	Inhibited (12h)	candidate	Editing/PPR
AT5G66520	CREF7; Pentatricopeptide repeat-containing protein	Inhibited (12h)	candidate	Editing/PPR
AT1G02930	GSTF6 (Glutathione S-transferase F6); EARLY RESPONSE TO DEHYDRATION 11	Stimulated (2h)	high-confidence	GST
AT2G30870	GSTF10 (Glutathione S-transferase F10); EARLY RESPONSE TO DEHYDRATION 13	Stimulated (2h)	candidate	GST
AT2G47730	GSTF10; Isoform 2 of Glutathione S-transferase F8, chloroplastic	Stimulated (2h)	candidate	GST
AT4G02520, AT2G02930	GSTF2/3 (Glutathione S-transferase F2/3)	Stimulated (2h); Stimulated (12h)	candidate (both)	GST
AT4G23100	GCS (Glutamate--cysteine ligase), chloroplastic; ROOT MERISTEMLESS 1	Stimulated (2h)	candidate	GST
AT5G43940	GSH-FDH; S-(hydroxymethyl)glutathione dehydrogenase;Alcohol dehydrogenase class-3	Inhibited (2h)	candidate	GST
AT1G12770	RH47 (DEAD-box ATP-dependent RNA helicase 47), mitochondrial	Inhibited (12h)	candidate	Helicases
AT2G42520	RH37 (DEAD-box ATP-dependent RNA helicase 37)	Stimulated (2h)	candidate	Helicases
AT3G01540, AT5G14610	RH14/46 (DEAD-box ATP-dependent RNA helicase 14/46)	Stimulated (12h)	candidate	Helicases
AT3G06980	RH50 (DEAD-box ATP-dependent RNA helicase 50)	Inhibited (12h)	high-confidence	Helicases
AT3G58510	RH11 (DEAD-box ATP-dependent RNA helicase 11)	Stimulated (2h)	candidate	Helicases
AT5G62190	RH7 (DEAD-box ATP-dependent RNA helicase 7)	Stimulated (2h)	candidate	Helicases
AT3G09440	HSC70-3; Heat shock cognate protein 70-3); HSP70-3 (Heat shock protein 70-3)	Stimulated (2h)	candidate	HSP
AT5G49910, AT4G24280	HSC70-2 (Heat shock 70 kDa protein 2); HSP70-6 (Heat shock 70 kDa protein 6); HSP70-7 (Heat shock 70 kDa protein 7) , chloroplastic	Stimulated (2h)	candidate	HSP
AT5G52640	HSP90-1 (Heat shock protein 90-1)	Stimulated (2h)	high-confidence	HSP
AT5G56010, AT5G56000	HSP90-3 (Heat shock protein 90-3); HSP90-4 (Heat shock protein 90-4); MUTANT SNC1-ENHANCING 10	Stimulated (2h); Stimulated (12h)	high-confidence (early); candidate (late)	HSP
AT1G56410; AT5G02490	Heat shock 70 kDa protein 18; Heat shock cognate 70 kDa protein 2	Stimulated (2h); Stimulated (12h)	candidate (both)	HSP/Transcription
AT5G02500	Probable mediator of RNA polymerase II transcription subunit 37e; Heat shock cognate protein 70-1; Protein EARLY-RESPONSIVE TO DEHYDRATION 2	Stimulated (2h)	candidate	HSP/Transcription
AT1G02360	Putative chitinase	Stimulated (12h)	high-confidence	Hydrolase
AT1G69830	Alpha-amylase 3, chloroplastic	Stimulated (2h)	candidate	Hydrolase
AT3G13790	Beta-fructofuranosidase; CWINV1 (Cell wall beta-fructosidase 1)	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Hydrolase
AT1G17420	LOX3 (Lipoxygenase 3) chloroplastic	Stimulated (2h)	high-confidence	Lipid metabolism
AT1G45196, AT1G45201	TLL1 (Triacylglycerol lipase-like 1)	Stimulated (2h)	candidate	Lipid metabolism
AT1G65352	Defensin-like protein 266	Inhibited (2h)	high-confidence	Lipid metabolism

AT2G33150	3-ketoacyl-CoA thiolase 2, peroxisomal	Stimulated (12h)	high-confidence	Lipid metabolism
AT2G45180	T14P1.1; Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	Stimulated (2h)	high-confidence	Lipid metabolism
AT4G00165	Putative lipid-binding protein	Stimulated (2h)	high-confidence	Lipid metabolism
AT4G14070	AEE15 (Long-chain-fatty-acid--[acyl-carrier-protein] ligase AEE15), chloroplastic	Stimulated (2h)	candidate	Lipid metabolism
AT5G55050	GDSL esterase/lipase	Stimulated (2h)	candidate	Lipid metabolism
AT1G04410	MDH1 (Malate dehydrogenase 1), cytoplasmic	Stimulated (12h)	candidate	Metabolism
AT1G05010, AT1G62380, AT1G12010	ACC oxidase 4/2/3 (1-aminocyclopropane-1-carboxylate oxidase 4/2/3)	Stimulated (2h); Stimulated (12h)	high-confidence (early); candidate (late)	Metabolism
AT1G11860	Aminomethyltransferase, mitochondrial; GCVT (Glycine cleavage T-protein family)	Stimulated (2h)	candidate	Metabolism
AT1G12900	GAPA2 (Isoform 2 of Glyceraldehyde-3-phosphate dehydrogenase GAPA2)	Stimulated (12h)	candidate	Metabolism
AT1G22400	UDP-glycosyltransferase 85A1; AtZOG2 (Zeatin O-glucosyltransferase 2)	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Metabolism
AT1G22410	Phospho-2-dehydro-3-deoxyheptonate aldolase; Class-II DAHP synthetase family protein	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Metabolism
AT1G51680	4CL 1 (4-coumarate--CoA ligase 1)	Stimulated (2h)	high-confidence	Metabolism
AT1G53310, AT3G14940	PEPC1/3 (Phosphoenolpyruvate carboxylase 1/3)	Stimulated (12h)	candidate	Metabolism
AT1G68010	GDH; AtHPR1 (Isoform 2 of Glycerate dehydrogenase HPR), peroxisomal	Stimulated (2h)	high-confidence	Metabolism
AT1G76680	OPR1 (12-oxophytodienoate reductase 1)	Stimulated (2h)	high-confidence	Metabolism
AT1G79550	PGK3 (Phosphoglycerate kinase), cytosolic	Stimulated (12h)	high-confidence	Metabolism
AT2G24270	ALDH11A3; NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	Stimulated (2h)	candidate	Metabolism
AT2G26080	Glycine dehydrogenase (decarboxylating) 2, mitochondrial	Inhibited (2h)	candidate	Metabolism
AT2G30490	C4H (Trans-cinnamate 4-monooxygenase)	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Metabolism
AT2G37040	Phenylalanine ammonia-lyase 1	Stimulated (2h)	high-confidence	Metabolism
AT2G38240	Probable 2-oxoglutarate-dependent dioxygenase ANS (Anthocyanidin synthase)	Stimulated (2h)	high-confidence	Metabolism
AT2G47470	PDIL2-1 (Protein disulfide-isomerase like 2-1); Protein MATERNAL EFFECT EMBRYO ARREST 30	Stimulated (12h)	candidate	Metabolism
AT3G29360, AT5G15490, AT5G39320, AT1G26570	UGD2/3/4/1; UDP-glucose 6-dehydrogenase 2/3/4/1	Stimulated (2h)	candidate	Metabolism
AT3G48990	AAE3 (Oxalate--CoA ligase); 4CL8 (4-coumarate--CoA ligase isoform 8 soform 8)	Stimulated (12h)	high-confidence	Metabolism
AT3G52880	MDAR1 (Monodehydroascorbate reductase 1), peroxisomal	Stimulated (12h)	high-confidence	Metabolism

AT3G52990	Pyruvate kinase; Pyruvate kinase	Stimulated (12h)	candidate	Metabolism
AT4G15640	Adenylyl cyclase	Stimulated (2h); Inhibited (12h)	high-confidence (both)	Metabolism
AT4G17090	Beta-amylase 3, chloroplastic	Inhibited (2h)	candidate	Metabolism
AT4G33510	3-deoxy-d-arabino-heptulosonate 7-phosphate synthase; Phospho-2-dehydro-3-deoxyheptonate aldolase 2, chloroplastic	Stimulated (2h)	candidate	Metabolism
AT4G34200, AT1G17745	PGDH1/PGDH2 (D-3-phosphoglycerate dehydrogenase 1/2), chloroplastic; EMBRYO SAC DEVELOPMENT ARREST 9	Stimulated (2h); Stimulated (12h)	candidate (early); high-confidence (late)	Metabolism
AT5G09660	PMDH2 (Malate dehydrogenase 2), glyoxysomal	Stimulated (2h)	candidate	Metabolism
AT5G10160	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase-like protein	Stimulated (2h)	candidate	Metabolism
AT5G38670	Isoform 2 of F-box/kelch-repeat protein	Inhibited (2h)	high-confidence	Metabolism
AT5G50850	PDHE1-B (Pyruvate dehydrogenase E1 component subunit beta-1), mitochondrial	Stimulated (12h)	high-confidence	Metabolism
AT5G11350	CCR4 homolog 6 (Carbon catabolite repressor protein 4 homolog 6)	Inhibited (12h)	high-confidence	mRNA decay
AT1G13270	MAP 1B (Methionine aminopeptidase 1B), chloroplastic	Inhibited (12h)	high-confidence	Nascent peptide modification
AT5G13780	NAA10 (N-terminal acetyltransferase A complex catalytic subunit NAA10)	Inhibited (12h)	candidate	Nascent peptide modification
AT5G57020	NMT1 (Glycylpeptide N-tetradecanoyltransferase 1)	Inhibited (12h)	candidate	Nascent peptide modification
AT1G37130	NR2 (Nitrate reductase [NADH] 2)	Inhibited (2h)	high-confidence	Nitrogen metabolim
AT3G17820	GLN1;3 (Glutamine synthetase cytosolic isozyme 1-3)	Stimulated (12h)	candidate	Nitrogen metabolim
AT3G45060	NRT2:6 (High affinity nitrate transporter 2.6)	Stimulated (2h)	high-confidence	Nitrogen metabolim
AT5G35630	GLN2 (Glutamine synthetase), chloroplastic/mitochondrial	Stimulated (2h)	candidate	Nitrogen metabolim
AT5G37600	GLN1-1 (Glutamine synthetase cytosolic isozyme 1-1)	Stimulated (2h); Stimulated (12h)	candidate (early); high-confidence (late)	Nitrogen metabolim
AT5G64290	DCT1 (Dicarboxylate transporter 2.1), chloroplastic	Inhibited (2h)	candidate	Nitrogen metabolim
AT1G07890	APX1 (L-ascorbate peroxidase 1), cytosolic	Stimulated (2h)	high-confidence	Other
AT1G13110, AT1G13100, AT1G13090	Cytochrome P450 71B7/29/28	Stimulated (2h)	candidate	Other
AT1G14670, AT2G01970	Transmembrane 9 superfamily member 2/3	Inhibited (2h)	high-confidence	Other
AT1G20440	Dehydrin COR47 (Cold-induced COR47 protein)	Stimulated (12h)	high-confidence	Other
AT1G22910	RNA recognition motif-containing protein	Inhibited (2h)	candidate	Other
AT1G29370	Kinase-related protein; RNA polymerase II degradation factor-like protein (DUF1296)	Inhibited (2h)	candidate	Other
AT1G29400	AML5 (Protein MEI2-like 5)	Inhibited (2h)	high-confidence	Other
AT1G33470	RNA-binding (RRM/RBD/RNP motifs) family protein	Inhibited (12h)	candidate	Other
AT1G54410	Dehydrin HIRD11	Inhibited (12h)	candidate	Other

AT1G56660	F25P12.91 protein (MAEBL domain protein)	Stimulated (2h)	high-confidence	Other
AT1G60000	RNA-binding (RRM/RBD/RNP motifs) family protein	Inhibited (12h)	candidate	Other
AT1G76180	Dehydrin ERD14	Stimulated (2h)	candidate	Other
AT1G78900	VHA-A; V-type proton ATPase catalytic subunit A	Stimulated (12h)	candidate	Other
AT2G14095	Expressed protein	Inhibited (12h)	candidate	Other
AT2G20890	Protein THYLAKOID FORMATION 1, chloroplastic	Stimulated (12h)	candidate	Other
AT2G24850	TAT3 (Tyrosine aminotransferase 3)	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Other
AT2G28540, AT2G28540	RNA recognition motif-containing protein	Stimulated (2h)	candidate	Other
AT2G32240	PICC; Early endosome antigen	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Other
AT2G32690	GRP23 (Glycine-rich protein 23)	Inhibited (12h)	candidate	Other
AT2G40300	Ferritin-4, chloroplastic	Stimulated (2h)	high-confidence	Other
AT3G01920	F28J7.25 (DHBP synthase RibB-like alpha/beta domain-containing protein)	Inhibited (2h); Inhibited (12h)	candidate (both)	Other
AT3G09840, AT5G03340, AT3G53230	CDC48A/D/E (Cell division control protein 48 homolog A/D/E)	Stimulated (12h)	high-confidence	Other
AT3G11820	Isoform 2 of Syntaxin-121; Syntaxin-121	Stimulated (12h)	candidate	Other
AT3G14310	PME27 (Pectin methylesterase 27); Pectinesterase/pectinesterase inhibitor 3	Inhibited (2h); Stimulated (12h)	candidate (early); high-confidence (late)	Other
AT3G16530	Lectin-like protein At3g16530	Stimulated (2h)	high-confidence	Other
AT3G20390	Reactive Intermediate Deaminase A, chloroplastic	Stimulated (2h)	candidate	Other
AT3G20820	Leucine-rich repeat (LRR) family protein (Polygalacturonase inhibitor-like protein)	Stimulated (2h)	high-confidence	Other
AT3G21300	TRM2A (TRNA METHYLTRANSFERASE 2A); RNA methyltransferase family protein	Inhibited (12h)	candidate	Other
AT3G44310	Nitrilase 1	Stimulated (2h)	candidate	Other
AT3G46780	PTAC16 (PLASTID TRANSCRIPTIONALLY ACTIVE 16), chloroplastic	Stimulated (2h)	high-confidence	Other
AT3G58410	MATH domain and coiled-coil domain-containing protein At3g58410 (RTM3-like protein)	Stimulated (2h)	high-confidence	Other
AT4G00090	F6N15_8 (Transducin/WD40 repeat-like superfamily protein)	Stimulated (2h)	candidate	Other
AT4G01050	Rhodanese-like domain-containing protein 4, chloroplastic; Protein THYLAKOID RHODANESE-LIKE; Sulfurtransferase 4	Stimulated (2h)	high-confidence	Other
AT4G01370, AT1G01560, AT5G10530, AT5G06740, AT5G59260, AT4G02410,	MPK 4/11 (Mitogen-activated protein kinase 4/11); L-type lectin-domain containing receptor kinase IX.1 (LecRK-IX.1); ,Probable L-type lectin-domain containing receptor kinase S.5 (LecRK-S.5); Probable L-type lectin-domain containing receptor kinase II.1 (LecRK-II.1); L-type lectin-domain containing receptor kinase IV.3 (LecRK-IV.3); L-type lectin-	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Other

AT5G65600, AT2G37710, AT3G53810, AT5G55830, AT5G01540, AT5G01550, AT5G01560, AT3G08870, AT5G59270, AT5G03140, AT3G53380, AT2G32800	domain containing receptor kinase IX.2 (LecRK-IX.2); L-type lectin-domain containing receptor kinase IV.1 (AthlecRK-e) (LecRK-IV.1); (Lectin Receptor Kinase 1),L-type lectin-domain containing receptor kinase IV.2 (Arabidopsis thaliana lectin-receptor kinase a4) (AthlecRK-a4) (LecRK-IV.2) (EC 2.7.11.1) (Protein SMALL, GLUED-TOGETHER, AND COLLAPSED POLLEN),Probable L-type lectin-domain containing receptor kinase S.7 (LecRK-S.7); L-type lectin-domain containing receptor kinase VI.2 (LecRK-VI.2);(Lectin receptor kinase A4.1),Lectin-domain containing receptor kinase VI.3 (LecRK-VI.3); (Lectin receptor kinase A4.2),Lectin-domain containing receptor kinase VI.4 (LecRK-VI.4); (Lectin receptor kinase A4.3),Probable L-type lectin-domain containing receptor kinase VI.1 (LecRK-VI.1); Putative L-type lectin-domain containing receptor kinase II.2 (LecRK-II.2);,L-type lectin-domain containing receptor kinase VIII.2 (LecRK-VIII.2); L-type lectin-domain containing receptor kinase VIII.1 (LecRK-VIII.1); Receptor like protein kinase S.2 (LecRK-S.2)			
AT4G03550	Callose synthase 12; Protein POWDERY MILDEW RESISTANT 4	Stimulated (2h)	candidate	Other
AT4G09650	ATPD (ATP synthase subunit delta), chloroplastic	Stimulated (2h)	candidate	Other
AT4G26650	RNA recognition motif-containing protein	Inhibited (12h)	candidate	Other
AT4G27520	Early nodulin-like protein 2 (Phytocyanin-like protein)	Inhibited (2h); Stimulated (12h)	high-confidence (early); candidate (late)	Other
AT4G32260	PDE334; ATPase, F0 complex, subunit B/B', bacterial/chloroplast	Stimulated (2h)	candidate	Other
AT4G36090	Oxidoreductase, 2OG-Fe(II) oxygenase family protein	Inhibited (12h)	candidate	Other
AT4G36220	Cytochrome P450 84A1	Stimulated (2h); Stimulated (12h)	candidate (early); high-confidence (late)	Other
AT5G03460	Transmembrane protein	Stimulated (2h)	candidate	Other
AT5G05100	MUG13.4; Single-stranded nucleic acid binding R3H protein	Stimulated (2h)	high-confidence	Other
AT5G12860	pOMT1 (Dicarboxylate transporter 1), chloroplastic	Inhibited (2h)	candidate	Other
AT5G42650	Allene oxide synthase, chloroplastic	Stimulated (2h)	candidate	Other
AT5G49210	Stress response NST1-like protein	Inhibited (12h)	candidate	Other
AT5G51070	Chaperone protein ClpD, chloroplastic	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Other
AT5G64120	Peroxidase 71	Stimulated (2h)	high-confidence	Other
AT5G03280	EIN2 (Ethylene-insensitive protein 2)	Stimulated (2h)	high-confidence	P-bodies
AT2G45810	RH6 (DEAD-box ATP-dependent RNA helicase 6)	Stimulated (2h)	candidate	P-bodies/Helicases
AT1G79090	PAT1 (Protein PAT1 homolog)	Stimulated (2h)	candidate	P-bodies/mRNA decay
AT3G22270	PAT1H1 (Topoisomerase II-associated protein PAT1)	Stimulated (2h); Stimulated (12h)	high-confidence (both)	P-bodies/mRNA decay
AT5G21160	LARP1a; LA RNA-binding protein; La-related protein 1A	Inhibited (12h)	candidate	P-bodies/mRNA decay

AT1G06680	PSBP-1; OEE2; Isoform 2 of Oxygen-evolving enhancer protein 2-1, chloroplastic	Stimulated (2h)	candidate	Photosynthesis
AT1G31330	PSI-F (Photosystem I reaction center subunit III), chloroplastic	Inhibited (2h)	candidate	Photosynthesis
AT1G44575	CP22 (Photosystem II 22 kDa protein), chloroplastic	Inhibited (2h)	candidate	Photosynthesis
AT1G61520	LHCA3-1; PSI type III chlorophyll a/b-binding protein; PSI type III chlorophyll a/b-binding protein	Stimulated (2h)	candidate	Photosynthesis
AT1G67090	RuBisCO small subunit 1A (Ribulose bisphosphate carboxylase small chain 1A), chloroplastic	Inhibited (2h)	candidate	Photosynthesis
AT1G79040	PSBR; Photosystem II 10 kDa polypeptide, chloroplastic	Stimulated (2h)	candidate	Photosynthesis
AT2G39730	RuBisCO activase (Ribulose bisphosphate carboxylase/oxygenase activase), chloroplastic	Stimulated (2h)	high-confidence	Photosynthesis
AT3G01500	AtbCA1 (Beta carbonic anhydrase 1), chloroplastic; SABP3 (Protein SALICYLIC ACID-BINDING PROTEIN 3)	Stimulated (2h)	candidate	Photosynthesis
AT3G08940	LHCB4.2; Chlorophyll a-b binding protein CP29.2, chloroplastic	Inhibited (2h)	candidate	Photosynthesis
AT3G54890	LHCA1/6 (Light-harvesting complex I chlorophyll a/b binding protein 1); Chlorophyll a-b binding protein 6, chloroplastic; LHCI-730	Stimulated (2h)	candidate	photosynthesis
AT4G02770	PSI-D1 (Photosystem I reaction center subunit II-1, chloroplastic)	Inhibited (2h)	candidate	Photosynthesis
AT4G05180	PSBQ2; OEE3; Oxygen-evolving enhancer protein 3-2, chloroplastic	Stimulated (2h); Stimulated (12h)	high-confidence (early); candidate (late)	Photosynthesis
AT4G12800	PSI-L (Photosystem I reaction center subunit XI), chloroplastic	Inhibited (2h); Stimulated (12h)	high-confidence (early); candidate (late)	Photosynthesis
AT4G21280	PSBQ1; OEE3; Isoform 2 of Oxygen-evolving enhancer protein 3-1, chloroplastic	Stimulated (2h)	high-confidence	Photosynthesis
AT5G14740	AtbCA2;(Beta carbonic anhydrase 2), chloroplastic	Stimulated (2h)	candidate	Photosynthesis
AT5G38410, AT5G38410, AT5G38420, AT5G38410	RuBisCO small subunit 3B/2B (Ribulose bisphosphate carboxylase small chain 3B/2B), chloroplastic	Inhibited (2h)	candidate	Photosynthesis
AT5G38430	RuBisCO small subunit 1B (Ribulose bisphosphate carboxylase small chain 1B), chloroplastic	Inhibited (2h)	candidate	Photosynthesis
AT5G54270	LHCB3-1; MDK4_9	Stimulated (2h)	candidate	Photosynthesis
ATCG00020	Photosystem II protein D1	Inhibited (2h)	candidate	Photosynthesis
ATCG00280	Photosystem II CP43 reaction center protein	Inhibited (2h)	candidate	Photosynthesis
ATCG00340	PsaB (Photosystem I P700 chlorophyll a apoprotein A2)	Inhibited (2h)	high-confidence	Photosynthesis
ATCG00350	PsaA (Photosystem I P700 chlorophyll a apoprotein A1)	Inhibited (2h)	high-confidence	Photosynthesis
ATCG00680	Photosystem II CP47 reaction center protein	Inhibited (2h)	high-confidence	Photosynthesis
ATCG00720	Cytochrome b6	Inhibited (2h)	high-confidence	Photosynthesis
ATCG01060	PsaC (Photosystem I iron-sulfur center)	Inhibited (2h)	candidate	Photosynthesis
AT1G16720	HCF173; High chlorophyll fluorescence phenotype 173 protein	Inhibited (2h)	candidate	Photosynthesis/Translation

AT4G25340	PPIase FKBP53 (Peptidyl-prolyl cis-trans isomerase FKBP53)	Stimulated (2h)	high-confidence	PPIs
AT4G32420	PPIase CYP95 (Peptidyl-prolyl cis-trans isomerase CYP95)	Stimulated (2h)	high-confidence	PPIs
AT1G02150	Pentatricopeptide repeat-containing protein	Inhibited (12h)	candidate	PPR
AT1G19720	Pentatricopeptide repeat-containing protein	Inhibited (12h)	candidate	PPR
AT1G30610	EMB2279 (Protein EMBRYO DEFECTIVE 2279); Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	high-confidence	PPR
AT1G71460	Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (2h)	candidate	PPR
AT1G79080	Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	candidate	PPR
AT2G16880	Pentatricopeptide repeat-containing protein	Inhibited (12h)	candidate	PPR
AT2G18940	Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	candidate	PPR
AT2G30780	Pentatricopeptide repeat-containing protein	Inhibited (12h)	candidate	PPR
AT3G04760	Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	candidate	PPR
AT3G26630	PCMP-A6; Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	candidate	PPR
AT3G46610	Pentatricopeptide repeat-containing protein	Inhibited (2h); Inhibited (12h)	candidate (early); high-confidence (late)	PPR
AT3G53700	MEE40 (MATERNAL EFFECT EMBRYO ARREST 40); Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	candidate	PPR
AT4G30825	BFA2 (BIOGENESIS FACTOR REQUIRED FOR ATP SYNTHASE 2); Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	candidate	PPR
AT4G31850	PGR3 (PROTON GRADIENT REGULATION 3); Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	high-confidence	PPR
AT5G02830	Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	high-confidence	PPR
AT5G46580	SOT1; Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	high-confidence	PPR
AT1G09750	AED3 (Apoplastic EDS1-dependent protein 3); Aspartyl protease-like protein	Stimulated (2h)	candidate	Protease
AT4G30530	GGP1 (Gamma-glutamyl peptidase 1)	Stimulated (2h); Stimulated (12h)	high-confidence (both)	Protease
AT5G48430	Aspartyl protease family protein	Stimulated (2h)	high-confidence	Protease
AT1G06190	RHON1 (Rho-N domain-containing protein 1), chloroplastic	Inhibited (12h)	candidate	Ribosomal
AT1G12800	SDP; Nucleic acid-binding, OB-fold-like protein	Inhibited (12h)	candidate	Ribosomal
AT1G15810	S15/NS1 RNA-binding protein	Inhibited (12h)	candidate	Ribosomal
AT1G18080	RACK1A (Receptor for activated C kinase 1A); ATARCA	Stimulated (12h)	candidate	Ribosomal
AT1G30580	Obg-like ATPase 1; Ribosome-binding ATPase YchF	Stimulated (2h)	high-confidence	Ribosomal
AT1G52980	NUG2; Nuclear/nucleolar GTPase 2	Inhibited (2h)	candidate	Ribosomal
AT1G69620	60S ribosomal protein L34-2	Inhibited (2h)	candidate	Ribosomal
AT2G21580	RPS25-2 (40S ribosomal protein S25-2)	Inhibited (12h)	candidate	Ribosomal
AT2G27710, AT2G27720	60S acidic ribosomal protein P2-2/1	Inhibited (2h)	candidate	Ribosomal

AT2G39140	SVR1 (SUPPRESSOR OF VARIEGATION 1); Putative ribosomal large subunit pseudouridine synthase family protein	Inhibited (12h)	high-confidence	Ribosomal
AT2G39390	60S ribosomal protein L35-2	Inhibited (2h)	candidate	Ribosomal
AT2G43030	50S ribosomal protein L3-1, chloroplastic	Inhibited (2h)	candidate	Ribosomal
AT3G09200, AT3G09200, AT3G11250, AT2G40010	60S acidic ribosomal protein P0-2/3/1	Stimulated (12h)	candidate	Ribosomal
AT3G11940	RPS5A (40S ribosomal protein S5-2)	Stimulated (2h)	candidate	Ribosomal
AT3G23390, AT4G14320	RPL36AA/RPL36AB; 60S ribosomal protein L36a	Stimulated (2h)	candidate	Ribosomal
AT3G44750	HDT1 (Histone deacetylase HDT1)	Stimulated (2h); Inhibited (12h)	high-confidence (early); candidate (late)	Ribosomal
AT3G44890	RPL9 (50S ribosomal protein L9), chloroplastic	Inhibited (2h); Inhibited (12h)	candidate (early); high-confidence (late)	Ribosomal
AT3G53020	60S ribosomal protein L24-2; SHORT VALVE 1	Inhibited (2h)	candidate	Ribosomal
AT3G53430, AT5G60670, AT2G37190	60S ribosomal protein L12-2/3/1	Stimulated (12h)	candidate	Ribosomal
AT3G55280, AT3G55280, AT2G39460	RPL23AB (60S ribosomal protein L23a-2/1)	Inhibited (2h)	candidate	Ribosomal
AT3G57490	RPS2D; 40S ribosomal protein S2-4	Stimulated (2h)	candidate	Ribosomal
AT4G39200	RPS25-4 (40S ribosomal protein S25-4)	Inhibited (12h)	candidate	Ribosomal
AT5G02610	60S ribosomal protein L35-4; Ribosomal L29 family protein	Inhibited (2h)	candidate	Ribosomal
AT5G14320	30S ribosomal protein S13, chloroplastic	Inhibited (12h)	candidate	Ribosomal
AT5G18570	OBGC (GTP-binding protein OBGC), chloroplastic	Inhibited (12h)	candidate	Ribosomal
AT5G30510	PRPS1 (30S ribosomal protein S1), chloroplastic	Inhibited (12h)	candidate	Ribosomal
AT5G46420	RimM-like 16S rRNA processing protein	Inhibited (2h)	candidate	Ribosomal
AT5G47190	50S ribosomal protein L19-2, chloroplastic	Inhibited (2h)	candidate	Ribosomal
AT5G66470	ERA-1; GTP-binding protein Era	Inhibited (12h)	candidate	Ribosomal
ATCG00330	RPS14 (30S ribosomal protein S14), chloroplastic	Stimulated (2h)	candidate	Ribosomal
ATCG00800	30S ribosomal protein S3, chloroplastic	Inhibited (2h)	candidate	Ribosomal
ATMG00090	RPS3 (Ribosomal protein S3), mitochondrial	Stimulated (2h); Inhibited (12h)	candidate (both)	Ribosomal
ATMG00560 ,AT2G07715	60S ribosomal protein L2, mitochondrial	Inhibited (12h)	candidate	Ribosomal
AT5G66540	U3 small nucleolar ribonucleoprotein protein MPP10	Stimulated (2h)	candidate	Ribosomal (processing pre-r18S rRNA)
AT3G07050	NSN1 (Guanine nucleotide-binding protein-like NSN1)	Inhibited (2h)	candidate	Ribosomal (modulates ribosome biogenesis)

AT3G57180	BPG2 (GTP binding protein BRASSINAZOLE INSENSITIVE PALE GREEN 2)	Inhibited (2h)	candidate	Ribosomal regulate rRNA)
AT3G57150	NAP57 (Nopp-140-associated protein of 57 kDa homolog); H/ACA ribonucleoprotein complex subunit 4	Inhibited (12h)	candidate	Ribosomal/Editing
AT4G09730	RH39 (DEAD-box ATP-dependent RNA helicase 39)	Inhibited (12h)	candidate	Ribosomal/Helicases
AT5G55220	PPIase TIG; Trigger factor-like protein TIG, Chloroplastic	Stimulated (2h); Stimulated (12h)	high-confidence (early); candidate (late)	Ribosomal/PPIs
AT4G16390	SVR7 (SUPPRESSOR OF VARIEGATION 7); Pentatricopeptide repeat-containing protein, chloroplastic	Inhibited (12h)	candidate	Ribosomal/PPR
AT3G05060	MAR-binding NOP56/58 homolog 2; Probable nucleolar protein 5-2	Stimulated (2h)	candidate	Ribosomal/Splicing
AT5G26742	RH3 (DEAD-box ATP-dependent RNA helicase 3), chloroplastic	Inhibited (12h)	high-confidence	Ribosomal/Splicing/H elicases
AT1G16445	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	Inhibited (2h)	candidate	SAM-related
AT1G55450	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	Stimulated (2h); Stimulated (12h)	high-confidence (both)	SAM-related
AT3G17390	METK4 (S-adenosylmethionine synthase 4)	Stimulated (2h)	candidate	SAM-related
AT1G77180	SKIP (SNW/SKI-interacting protein)	Stimulated (2h)	candidate	Splicing
AT2G16940	Splicing factor, CC1-like protein	Stimulated (2h)	candidate	Splicing
AT2G24590	RSZ22A (Serine/arginine-rich splicing factor RSZ22A)	Inhibited (12h)	candidate	Splicing
AT3G18390	CRS1 / YhbY (CRM) domain-containing protein	Inhibited (12h)	high-confidence	Splicing
AT3G49601	Pre-mRNA-splicing factor	Stimulated (2h)	high-confidence	Splicing
AT4G01037	WTF1 (WHAT'S THIS FACTOR 1 homolog), chloroplastic	Inhibited (12h)	candidate	Splicing
AT4G14510	CFM3B (CRM family member 3B)	Inhibited (12h)	high-confidence	Splicing
AT4G37510	RNC1 (Ribonuclease III domain-containing protein RNC1), chloroplastic	Inhibited (12h)	candidate	Splicing
AT4G39040	CFM4; Putative RNA-binding protein containing KH domain	Inhibited (12h)	candidate	Splicing
AT5G06160	SF3a60 (Splicing factor 3a); ATROPOS	Stimulated (2h)	candidate	Splicing
AT5G51300	SF1 (Splicing factor-like protein 1)	Inhibited (12h)	high-confidence	Splicing
AT4G18520	PCMP-A2; Pentatricopeptide repeat-containing protein	Inhibited (12h)	high-confidence	Splicing/Editing/PPR
AT2G43810, AT3G59810	LSM6A/B (U6 snRNA-associated Sm-like protein 6A/B); LSM36B (U6 snRNA-associated Sm-like protein 36B)	Stimulated (12h)	high-confidence	Splicing/P-bodies/mRNA decay
AT3G63400	PPIase CYP63 (Peptidyl-prolyl cis-trans isomerase CYP63)	Stimulated (2h)	high-confidence	Splicing/PPIs
AT1G01320	REC1; Tetratricopeptide repeat (TPR)-like superfamily protein	Inhibited (12h)	candidate	TPR
AT1G15290	Tetratricopeptide repeat (TPR)-like superfamily protein	Inhibited (12h)	high-confidence	TPR
AT4G28080	Protein TSS (TPR-domain suppressor of STIMPY)	Inhibited (2h); Inhibited (12h)	high-confidence (both)	TPR
AT1G10170	NFXL1 (NF-X1-type zinc finger protein NFXL1)	Stimulated (2h)	candidate	Transcription
AT2G32080	Transcription factor Pur-alpha 1	Inhibited (12h)	candidate	Transcription
ATCG00190	DNA-directed RNA polymerase subunit beta	Inhibited (12h)	candidate	Transcription

AT4G01290	CBE1 (CONSERVED BINDING OF EIF4E 1); Chorismate synthase	Inhibited (12h)	candidate	Transcription/Translation
AT1G17220	IF-2 (Translation initiation factor IF-2), chloroplastic	Inhibited (12h)	candidate	Translation
AT1G62750	SCO1 (SNOWY COTYLEDON 1); Elongation factor G, chloroplastic	Inhibited (2h); Inhibited (12h)	candidate (both)	Translation
AT1G73180	eIF-2A (Eukaryotic translation initiation factor eIF2A-like protein)	Inhibited (12h)	candidate	Translation
AT2G19870	tRNA/rRNA methyltransferase (SpoU) family protein	Inhibited (12h)	candidate	Translation
AT2G24060	IF3-2 (Translation initiation factor IF-3)	Inhibited (12h)	candidate	Translation
AT2G46280, AT2G46290	eIF3i (Eukaryotic translation initiation factor 3 subunit I)	Stimulated (12h)	high-confidence	Translation
AT3G49470	Nascent polypeptide-associated complex subunit alpha-like protein 2	Inhibited (2h)	candidate	Translation
AT3G60240	eIF4G (Eukaryotic translation initiation factor 4G)	Inhibited (12h)	candidate	Translation
AT4G11175	IF-1 (Translation initiation factor IF-1), chloroplastic	Inhibited (12h)	candidate	Translation
AT4G24800	MRF3 (MA3 DOMAIN-CONTAINING TRANSLATION REGULATORY FACTOR 3); ECIP1	Inhibited (2h); Inhibited (12h)	candidate (both)	Translation
AT4G26300	Arginine--tRNA ligase, chloroplastic/mitochondrial	Stimulated (2h)	candidate	Translation
AT4G29060	EF-TsMt (Elongation factor Ts), mitochondrial	Inhibited (12h)	candidate	Translation
AT5G08650	EF-4; Translation factor GUF1 homolog, chloroplastic	Inhibited (12h)	candidate	Translation
AT5G63190	MRF1 (MA3 DOMAIN-CONTAINING TRANSLATION REGULATORY FACTOR 1)	Inhibited (2h)	high-confidence	Translation
AT1G61870	PPR336; Pentatricopeptide repeat-containing protein, mitochondrial	Inhibited (12h)	candidate	Translation/PPR
AT2G29200	PUM1 (Pumilio homolog 1)	Stimulated (12h)	high-confidence	Translation/RNA stability
AT1G59870	ABC transporter G family member 36; PENETRATION 3	Stimulated (2h)	candidate	Transporter
AT1G76030, AT1G20260, AT4G38510	VACB1/2/3 (V-type proton ATPase subunit B1/2/3)	Stimulated (12h)	high-confidence	Transporter
AT2G18960	Proton pump 1; AHA1 (ATPase 1, plasma membrane-type)	Stimulated (12h)	candidate	Transporter
AT2G20990	Synaptotagmin A; Synaptotagmin-1	Stimulated (12h)	candidate	Transporter
AT2G24820	TIC55 (Translocon at the inner envelope membrane of chloroplasts 55), chloroplastic	Inhibited (2h)	candidate	Transporter
AT2G37920	COPT4; Copper ion transmembrane transporter	Stimulated (12h)	candidate	Transporter
AT2G39210	T16B24.1; Major facilitator superfamily protein; Nodulin-like protein	Stimulated (12h)	high-confidence	Transporter
AT3G01390	VMA10; V-type proton ATPase subunit G1	Stimulated (2h)	candidate	Transporter
AT4G11380, AT4G23460	Beta-adaptin-like protein B	Stimulated (2h)	candidate	Transporter
AT4G16143	Importin subunit alpha-2	Stimulated (2h)	high-confidence	Transporter
AT4G30190	AHA2 (ATPase 2, plasma membrane-type)	Stimulated (12h)	high-confidence	Transporter

AT5G23630	Probable manganese-transporting ATPase PDR2; PHOSPHATE DEFICIENCY RESPONSE 2	Stimulated (2h)	candidate	Transporter
AT5G26340	Sugar transport protein 13; Multicopy suppressor of snf4 deficiency protein 1	Stimulated (12h)	high-confidence	Transporter
AT5G40780, AT1G67640, AT1G24400, AT1G48640	Lysine histidine transporter 1/2	Stimulated (12h)	candidate	Transporter
AT5G55190, AT5G20020	RAN3/2/1; GTP-binding nuclear protein Ran-3/2/1	Inhibited (2h)	candidate	Transporter
AT5G60980	Nuclear transport factor 2 and RNA recognition motif domain-containing protein	Inhibited (12h)	candidate	Transporter
AT1G53210	AtNCL (Sodium/calcium exchanger NCL)	Stimulated (12h)	candidate	Transporter/Calcium-binding
AT1G41830, AT4G22010	Dicarboxylate transporter 1; SKU5 similar 6/4; Pectinesterase-like protein	Inhibited (2h)	candidate	Transporter/Metabolism

Supplemental table 1. Summary of flg22-responsive RBPome

Proteins identified as flg22-responsive RBPome, including high-confidence and candidate flg22-responsive RBPs. For each of the proteins a summary of the significant changes in association with RNA after flg22 treatment are included. The classification of 'RBP function/family' is based on predicted GO annotations.

Supplemental table 2. Summary of total proteome (WCL) flg22-responsive proteins

Gene ID(s)	Protein name(s)	Summary abundance	Protein Pathway(s) / Families
AT5G55530	Calcium-dependent lipid-binding (CaLB domain) family protein	Upregulated (12h)	Calcium signalling
AT3G01830	Probable calcium-binding protein CML40	Upregulated (12h)	Calcium signalling
AT2G22950	Putative calcium-transporting ATPase 7, plasma membrane-type	Upregulated (12h)	Calcium signalling
AT4G12490, AT4G12480, AT4G12500	pEARLI1-like lipid transfer protein 2/1/3	Upregulated (12h)	Defence
AT1G09560	Germin-like protein subfamily 2 member 1	Upregulated (12h)	Defence
AT2G30490	CA4H (Trans-cinnamate 4-monooxygenase)	Upregulated (12h)	Defence
AT3G52400	SYP122 (Syntaxin-122)	Upregulated (12h)	Defence
AT2G35980	NDR1/HIN1-like protein 10; ELLOW-LEAF-SPECIFIC GENE 9	Upregulated (12h)	Defence
AT3G25780	Allene oxide cyclase 3, chloroplastic	Upregulated (12h)	Defence
AT1G17420	LOX3 (Lipoxygenase 3), chloroplastic	Upregulated (12h)	Defence
AT1G65390	PP2-A5 (PHLOEM PROTEIN 2-LIKE A5)	Upregulated (12h)	Defence
AT1G80840	Probable WRKY transcription factor 40	Upregulated (12h)	Defence
AT4G11850	Phospholipase D gamma 1	Upregulated (12h)	Defence
AT1G10170	NFXL1 (NF-X1-type zinc finger protein NFXL1)	Upregulated (12h)	Defence
AT4G25030	NHR2B (NON HOST RESISTANCE 2B); Serine/Threonine-kinase	Upregulated (12h)	Defence
AT5G26742	RH3 (DEAD-box ATP-dependent RNA helicase 3), chloroplastic	Upregulated (12h)	Defence/RNA regulation
AT3G54420	Endochitinase EP3	Upregulated (12h)	Hydrolase
AT4G25810	XTH23 (Probable xyloglucan endotransglucosylase/hydrolase protein 23)	Upregulated (12h)	Hydrolase
AT2G39930	ISA1 (Isoamylase 1), chloroplastic	Upregulated (12h)	Hydrolase
AT2G19190	FLG22-induced receptor-like kinase 1; Senescence-induced receptor-like serine/threonine-protein kinase	Upregulated (12h)	Immune receptor
AT1G51800	IMPAIRED OOMYCETE SUSCEPTIBILITY 1; LRR receptor-like serine/threonine-protein kinase IOS1	Upregulated (12h)	Immune receptor
AT3G46280	Kinase-like protein	Upregulated (12h)	Immune receptor
AT1G51850	SIF2 (STRESS INDUCED FACTOR 2); Leucine-rich repeat protein kinase family protein	Upregulated (12h)	Immune receptor
AT3G22060	Cysteine-rich repeat secretory protein 38	Upregulated (12h)	Immune receptor
AT1G11310	Mlo2 (MLO-like protein 2)	Upregulated (12h)	Immune receptor
AT2G45220	PME17 (Probable pectinesterase/pectinesterase inhibitor 17)	Upregulated (12h)	Inhibitor
AT2G38870	Putative protease inhibitor (Serine protease inhibitor, potato inhibitor I-type family protein)	Upregulated (12h)	Inhibitor
AT4G22470	Protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	Upregulated (12h)	Inhibitor
AT1G73260	KTI1 (Kunitz trypsin inhibitor 1)	Upregulated (12h)	Inhibitor
AT3G47380	PMEI11 (Pectinesterase inhibitor 11)	Upregulated (12h)	Inhibitor
AT1G47710	Serpin-ZX; Serpin-1	Upregulated (12h)	Inhibitor
AT2G39518	CASPL4D2 (CASP-like protein 4D2)	Upregulated (12h)	Lignin-related
AT2G39530	CASPL4D1 (CASP-like protein 4D1)	Upregulated (12h)	Lignin-related
AT4G37990	CAD8 (Cinnamyl alcohol dehydrogenase 8)	Upregulated (12h)	Lignin-related
AT4G34230	CAD5 (Cinnamyl alcohol dehydrogenase 5)	Upregulated (12h)	Lignin-related
AT5G54160	OMT1 (Flavone 3-O-methyltransferase 1)	Upregulated (12h)	Lignin-related
AT4G15610	CASPL1D1 (CASP-like protein 1D1)	Upregulated (12h)	Lignin-related
AT4G27440	PCR B (Protochlorophyllide reductase B), chloroplastic	Downregulated (12h)	Other
AT5G13630	GENOMES UNCOUPLED 5; Magnesium-chelatase subunit ChlH, chloroplastic	Downregulated (12h)	Other
AT2G30520, AT2G30510	Root phototropism protein 2	Downregulated (12h)	Other
AT3G57650	1-acyl-sn-glycerol-3-phosphate acyltransferase 2	Downregulated (12h)	Other
AT5G54770	Thiamine thiazole synthase, chloroplastic	Downregulated (12h)	Other
AT1G01610, AT4G00400, AT4G00410	Glycerol-3-phosphate 2-O-acyltransferase 4/ 8	Downregulated (12h)	Other

AT5G44520	Probable ribose-5-phosphate isomerase 4, chloroplastic	Downregulated (12h)	Other
AT2G17790	Vacuolar protein sorting-associated protein 35A	Downregulated (12h)	Other
AT1G16410	Dihomomethionine N-hydroxylase; BUSHY 1; Protein SUPERSHOOT 1	Downregulated (12h)	Other
AT5G26030	FC1 (Ferrochelatase-1), chloroplastic/mitochondrial	Downregulated (12h)	Other
AT1G55860	E3 ubiquitin-protein ligase UPL1	Downregulated (12h)	Other
AT2G31810	AHAS (Acetolactate synthase small subunit 2), chloroplastic	Downregulated (12h)	Other
AT3G27050	MOJ10_14	Downregulated (12h)	Other
AT1G32920	F9L11_25	Downregulated (12h)	Other
AT5G33320	PPT1 (Phosphoenolpyruvate/phosphate translocator 1), chloroplastic	Downregulated (12h)	Other
AT5G25260	Flotillin-like protein 2 (Nodulin-like protein 2)	Upregulated (12h)	Other
AT5G01210	F7J8_190; Anthranilate N-benzoyltransferase-like protein	Upregulated (12h)	Other
AT1G71697	CK1 (Probable choline kinase 1)	Upregulated (12h)	Other
AT2G44790	Uclacyanin-2	Upregulated (12h)	Other
AT2G27290	FAM210B-like protein, putative (DUF1279)	Upregulated (12h)	Other
AT4G18950	BLUE LIGHT-DEPENDENT H+-ATPASE PHOSPHORYLATION	Upregulated (12h)	Other
AT2G17390	Ankyrin repeat domain-containing protein 2B	Upregulated (12h)	Other
AT5G20230	BCB (Blue copper protein)	Upregulated (12h)	Other
AT2G29630	Phosphomethylpyrimidine synthase, chloroplastic	Upregulated (12h)	Other
AT5G37600	GLN1;1 (Glutamine synthetase cytosolic isozyme 1-1)	Downregulated (12h)	Other
ATMG00580	NADH-ubiquinone oxidoreductase chain 4	Downregulated (2h)	Oxidoreduction
AT1G30700	Berberine bridge enzyme-like 8	Downregulated (12h)	Oxidoreduction
AT4G22690, AT4G22710	Cytochrome P450, family 706, subfamily A, polypeptide 1/2	Upregulated (2h)	Oxidoreduction
AT4G20840	BBE-like 21 (Berberine bridge enzyme-like 21)	Upregulated (2h)	Oxidoreduction
AT1G30720, AT1G30730	Berberine bridge enzyme-like 10/11	Upregulated (2h)	Oxidoreduction
ATCG01010	NAD(P)H-quinone oxidoreductase subunit 5, chloroplastic	Upregulated (2h)	Oxidoreduction
AT4G20860	Berberine bridge enzyme-like 22	Upregulated (both)	Oxidoreduction
AT2G38240	Probable 2-oxoglutarate-dependent dioxygenase; ANS (Anthocyanidin synthase)	Upregulated (both)	Oxidoreduction
AT5G38900	Thioredoxin superfamily protein	Upregulated (both)	Oxidoreduction
AT2G38860	DJ-1 protein homolog E; YELLOW-LEAF-SPECIFIC GENE 5	Upregulated (both)	Oxidoreduction
AT4G13770	Cytochrome P450 83A1; REDUCED EPIDERMAL FLUORESCENCE 2	Upregulated (both)	Oxidoreduction
AT5G64100	Peroxidase 69	Downregulated (2h)	Peroxidase
AT5G39580	Peroxidase 62	Upregulated (2h)	Peroxidase
AT4G37520	Peroxidase 50	Upregulated (2h)	Peroxidase
AT5G64120	Peroxidase 71	Upregulated (2h)	Peroxidase
AT1G14540	Peroxidase 4	Upregulated (2h)	Peroxidase
AT4G08770, AT4G08780	Peroxidase 37/ 38	Upregulated (2h)	Peroxidase
AT5G05340	Peroxidase 52	Upregulated (2h)	Peroxidase
AT4G36430	Peroxidase 49	Upregulated (2h)	Peroxidase
AT5G48430	Eukaryotic aspartyl protease family protein	Upregulated (2h)	Protease
AT1G03230	Eukaryotic aspartyl protease family protein	Upregulated (2h)	Protease
AT5G51070	Chaperone protein ClpD, chloroplastic	Downregulated (2h)	RNA metabolism/Defence
AT5G52640	HSP90.1 (Heat shock protein 90-1); Hsp81-1 (Heat shock protein 81-1)	Upregulated (2h)	RNA regulation/Defence
AT1G65700	Sm-like protein LSM8 (AtLSM8) (U6 snRNA-associated Sm-like protein LSM8)	Upregulated (2h)	RNA regulation/Defence
AT3G12680	HUA1; C3H37 (Zinc finger CCCH domain-containing protein 37)	Upregulated (2h)	RNA regulation/Other
AT5G17020, AT3G03110	Protein EXPORTIN 1A/ 1B; HEAT-INTOLERANT 2	Upregulated (2h)	Transport
AT3G62700	ABC transporter C family member 14	Upregulated (2h)	Transport

Supplemental table 2. Summary of total proteome (WCL) flg22-responsive proteins

Proteins significantly altered in abundance in the total proteome (whole cell lysate). For each of the proteins a summary of the significant changes in abundance after flg22 treatment are included. The classification 'Protein pathways(s)/functions' is based on predicted GO annotations

Supplemental table 3. RBP mutant lines screened

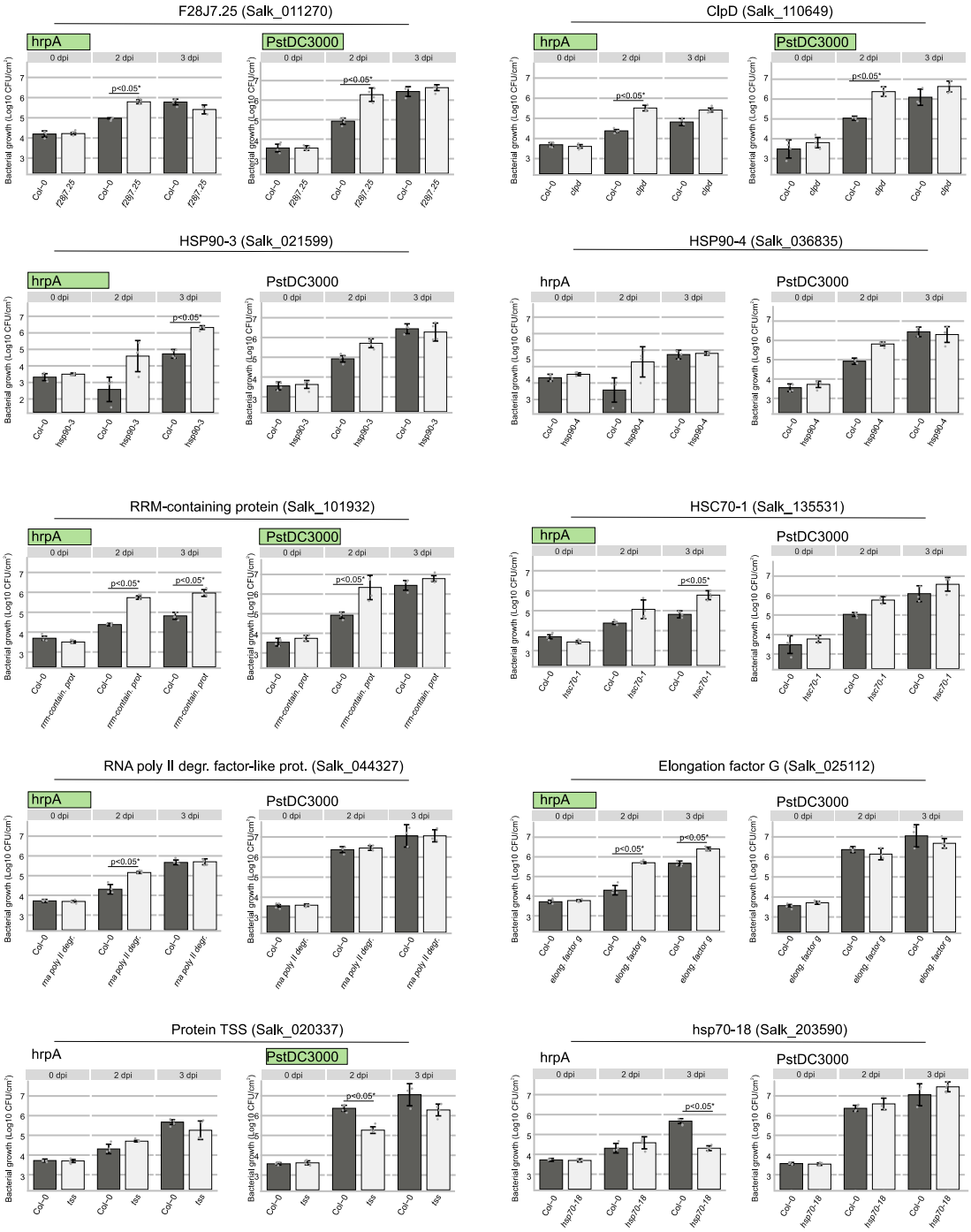
Gene ID(s)	Protein ID(s)	Protein name	Mutant lines	Phenotype mutants	log2FC 2hpt	FDR 2hpt	log2FC 12hpt	FDR 12hpt	flg22-responsive RBPome
AT3G01920	Q9SGI3	F28J7.25 (DHBP synthase RibB-like alpha/beta domain-containing protein)	Salk_011270	Susceptible to PstDC3000 & hrpA	-0.599	0.030	-0.830	0.007	candidate
AT5G51070	P42762	Chaperone protein ClpD, chloroplastic	Salk_110649	Susceptible to PstDC3000 & hrpA	2.255	0.008	2.184	0.011	high-confidence
AT5G03280	Q9S814	EIN2 (Ethylene-insensitive protein 2)	ein2-1	Susceptible to PstDC3000 & hrpA	1.859	0.002	-0.231	0.705	high-confidence
AT5G56010; AT5G56000	P51818; O03986	Heat shock protein 90-3; Heat shock protein 90-4	Salk_036835C (*hsp90-4*); Salk_021599 *(hsp90-3*)	Susceptible to hrpA (*hsp90-3*); no (*hsp90-4*)	1.434	0.004	0.809	0.102	high-confidence
AT1G22910	F4I323; Q9ASQ8; F4I321	RNA recognition motif-containing protein	Salk_101932	Susceptible to hrpA	-0.707	0.091	-0.522	0.254	candidate
AT5G02500	P22953; F4KCE5	Probable mediator of RNA polymerase II transcription subunit 37e; Heat shock cognate protein 70-1; Protein EARLY-RESPONSIVE TO DEHYDRATION 2	Salk_135531	Susceptible to hrpA	0.872	0.010	0.364	0.312	candidate
AT1G29370	Q8VZT4	RNA polymerase II degradation factor-like protein (DUF1296); Kinase-related protein	Salk_044327C	Susceptible to hrpA	-1.543	0.156	-0.940	0.476	candidate
AT1G62750	Q9SI75	Elongation factor G, chloroplastic; SCO1 (SNOWY COTYLEDON 1)	Salk_025112	Susceptible to hrpA	-0.720	0.013	-0.720	0.018	candidate
AT3G17390	Q9LUT2	METK4 (S-adenosylmethionine synthase 4)	Salk_052289	Resistant to PstDC3000 & hrpA	0.850	0.123	0.641	0.302	candidate
AT3G58510	Q8LA13	RH11 (DEAD-box ATP-dependent RNA helicase 11)	ddx11-1 (Salk_203665); ddx11-2 (Salk_072473); ddx11-3 (WiscDsLoxHs062_06H)	Resistant to PstDC3000	0.876	0.017	0.005	0.994	candidate
AT4G28080	F4JKH6	Protein TSS (TPR-domain suppressor of STIMPY)	Salk_020337	Resistant to PstDC3000	-1.005	0.005	-1.469	0.000	high-confidence
AT1G56410; AT5G02490	P22954; Q9C7X7	Heat shock 70 kDa protein 18; Heat shock cognate 70 kDa protein 2	Salk_203590 (*hsp70-18*); Salk_085076 (*hsc70-2*)	no (*hsp70-18*); no (hsp70-2)	1.815	0.105	1.720	0.154	candidate
AT3G09440	O65719	HSC70-3; Heat shock cognate protein 70-3); HSP70-3 (Heat shock protein 70-3)	Salk_013280	no	0.983	0.017	0.279	0.572	candidate

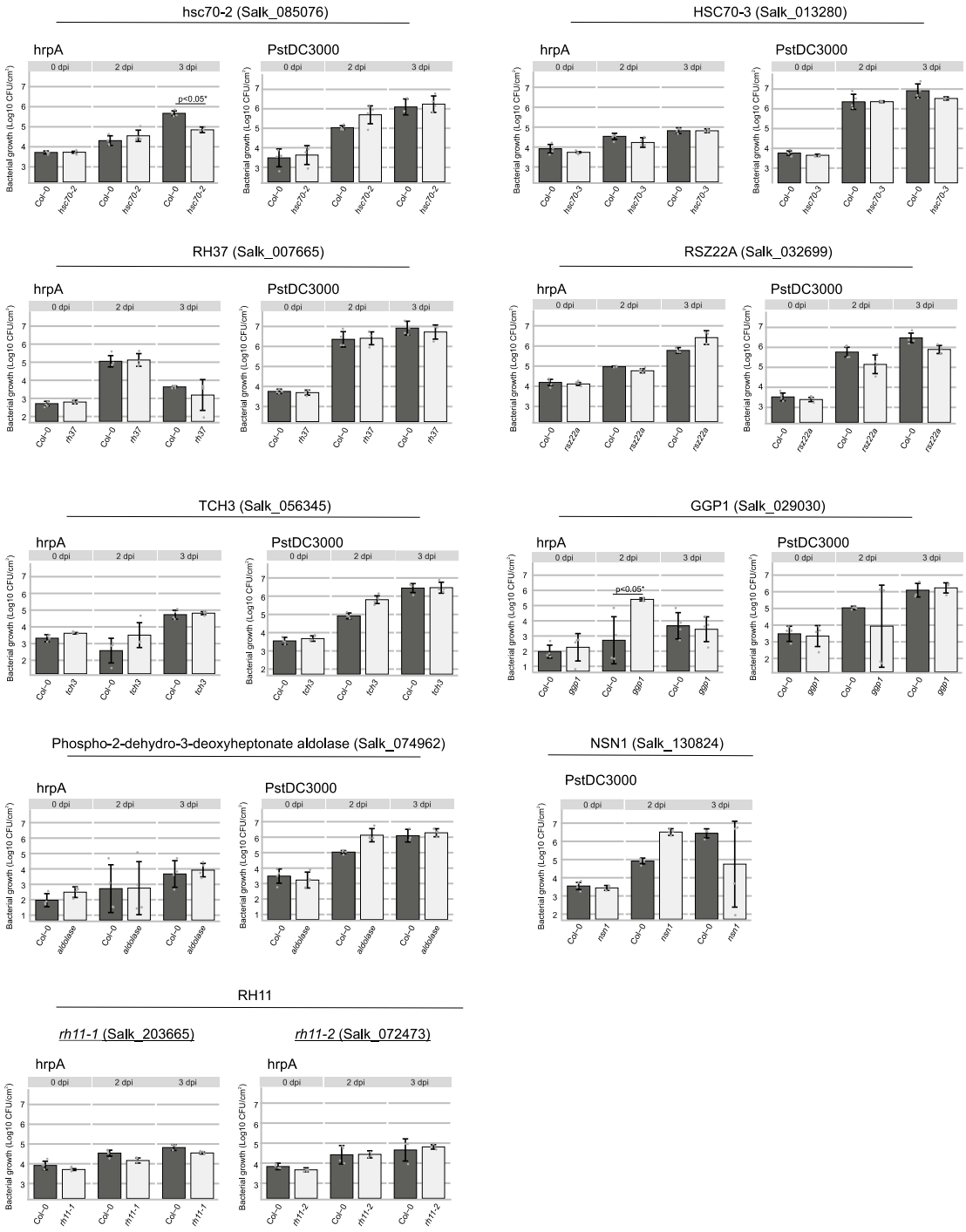
AT2G42520	Q84W89	RH37 (DEAD-box ATP-dependent RNA helicase 37)	Salk_007665	no	0.752	0.013	0.245	0.491	candidate
AT3G07050	Q9M8Z5	NSN1 (Guanine nucleotide-binding protein-like NSN1)	Salk_130824	no	-0.865	0.011	-0.509	0.152	candidate
AT2G24590	Q9SJA6	RSZ22A (Serine/arginine-rich splicing factor RSZ22A)	Salk_032699	no	-0.250	0.665	-0.856	0.129	candidate
AT2G41100	F4IJ44; P25071; F4IJ45	TCH3; CML12 (Calmodulin-like protein 12); CAL4 (Calmodulin-like protein 4)	Salk_056345	no	1.598	0.027	1.550	0.038	high-confidence
AT4G30530	Q9M0A7	GGP1 (Gamma-glutamyl peptidase 1)	Salk_029030	no	1.722	0.047	1.570	0.094	high-confidence
AT1G22410	Q9SK84	Phospho-2-dehydro-3-deoxyheptonate aldolase; Class-II DAHP synthetase family protein	Salk_074962	no	1.913	0.007	1.682	0.018	high-confidence

Supplemental table 3. RBP mutant lines screened

List of mutant lines screened and their phenotype of resistance/susceptibility to *Pst*DC3000 and hrpA$^-$. For each of the proteins, the $\log_2$FC[Flg22/Mock] and the FDR at each time point are included. RBPs are classified as 'high-confidence flg22-responsive RBPs' when FDR ≤ 0.1 and $\log_2$FC [flg22/mock] ≥ 1 or ≤ -1 or 'candidate flg22-responsive RBPs' when FDR ≤ 0.2 and a $\log_2$FC [flg22/mock] ≥ 0.58 or ≤ -0.58. A summary column ('flg22-responsive RBPome') includes the statistical changes in association with RNA at either 2 or 12 hours post treatment with flg22.

Supplemental figure 1. Bacterial counts of RBP mutant lines





Supplemental figure 1. Bacteial count of RBP mutant lines. RBP mutant lines and Col-0 plants were infiltrated with *Pst*DC3000 or with its mutant *hrpA* at an inoculum density of $5x10^5$ cells/ml and the bacterial density analysed at 0, 2 and 3 days after inoculation (dpi). Error bars indicate mean ± sd of n = 4. ANOVA with Tukey HSD was used to calculate the p-value. The experiment was repeated multiple times with similar results. Green boxes indicate consistent significant disease phenotypes of the RBP mutant lines.

APPENDIX V – List of publications

Garcia-Moreno M, Noerenberg M, Ni Shuai, Jarvelin AI, Gonzalez-Almela E, Lenz C, **Bach-Pages M**, Cox V, Avolio R, Davis T, Hester S, Sohier TJM, Li B, Sanz MA, Carrasco L, Ricci EP, Pelechano V, Fisher B, Mohammed S, Castello A (2019). Understanding RNP remodelling uncovers RBPs functionally required for viral replication. *Molecular Cell*. doi: https://doi.org/10.1101/350686

Bach-Pages M, and Preston GM (2018). Methods to Quantify Biotic-Induced Stress in Plants. *Methods Mol Biol*., 1734:241-255. doi: 10.1007/978-1-4939-7604-1_19

Bach-Pages M, Castello A, and Preston GM (2017). Plant RNA Interactome Capture: Revealing the Plant RBPome. *Trends in Plant Science*, 22: 449-451. doi: 10.1016/j.tplants.2017.04.006

Misas-Villamil JC, van der Burgh AM, Grosse-Holz F, **Bach-Pages M**, Kovács J, Kaschani F, Schilasky S, Emon AE, Ruben M, Kaiser M, Overkleeft HS, van der Hoorn RA (2017). Subunit-selective proteasome activity profiling uncovers uncoupled proteasome subunit activities during bacterial infections. *The Plant Journal*, 90:418-430. doi: 10.1111/tpj.13494

REFERENCES

Abulfaraj, A.A., Mariappan, K., Bigeard, J., Manickam, P., Blilou, I., Guo, X., Al-Babili, S., Pflieger, D., Hirt, H., and Rayapuram, N. (2018). The *Arabidopsis* homolog of human G3BP1 is a key regulator of stomatal and apoplastic immunity. Life Sci. Alliance **1**: e201800046.

Achour, C. and Aguilo, F. (2018). Long non-coding RNA and Polycomb: an intricate partnership in cancer biology. Front. Biosci. - Landmark **23**: 2106–2132.

Afroz, A., Zahur, M., Zeeshan, N., and Komatsu, S. (2013). Plant-bacterium interactions analyzed by proteomics. Front. Plant Sci. **4**: 1–18.

Ahmad, S. and Hur, S. (2015). Helicases in antiviral immunity: dual properties as sensors and effectors. Trends Biochem Sci. **40**: 576–585.

Aksoy, E., Jeong, I.S., and Koiwa, H. (2013). Loss of function of Arabidopsis C-terminal domain phosphatase-like1 activates iron deficiency responses at the transcriptional level. Plant Physiol. **161**: 330–345.

Alonso, J.M., Hirayama, T., Roman, G., Nourizadeh, S., and Ecker, J.R. (1999). EIN2, a bifunctional transducer of ethylene and stress responses in Arabidopsis. Science (80-.). **284**: 2148–2152.

Ambrosone, A., Batelli, G., Nurcato, R., Aurilia, V., Punzo, P., Bangarusamy, D.K., Ruberti, I., Sassi, M., Leone, A., Costa, A., and Grillo, S. (2015). The Arabidopsis RNA-binding protein AtRGGA regulates tolerance to salt and drought stress. Plant Physiol. **168**: 292–306.

Anantharaman, V., Koonin, E. V, and Aravind, L. (2002). Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acids Res. **30**: 1427–1464.

Anzi, C., Pelucchi, P., Vazzola, V., Murgia, I., Gomarasca, S., Beretta Piccoli, M., and Morandini, P. (2008). The proton pump interactor (Ppi) gene family of Arabidopsis thaliana: Expression pattern of Ppi1 and characterisation of knockout mutants for Ppi1

and 2. Plant Biol. **10**: 237–249.

Arif, A., Jia, J., Moodt, R.A., DiCorleto, P.E., and Fox, P.L. (2011). Phosphorylation of glutamyl-prolyl tRNA synthetase by cyclin-dependent kinase 5 dictates transcriptselective translational control. Proc. Natl. Acad. Sci. U. S. A. **108**: 1415–1420.

Arif, A., Jia, J., Mukhopadhyay, R., Willard, B., Kinter, M., and Fox, P.L. (2009). Two-site phosphorylation of EPRS coordinates multimodal regulation of noncanonical translational control activity. Mol. Cell **35**: 164–180.

Arif, A., Yao, P., Terenzi, F., Jia, J., Ray, P.S., and Fox, P.L. (2018). The GAIT translational control. Wiley Interdiscip Rev RNA **9**: 1441.

Ariumi, Y., Kuroki, M., Kushima, Y., Osugi, K., Hijikata, M., Maki, M., Ikeda, M., and Kato, N. (2011). Hepatitis C virus hijacks P-body and stress granule components around lipid droplets. J. Virol. **85**: 6882–6892.

Asai, T., Tena, G., Plotnikova, J., Willmann, M.R., Chiu, W., Gomez-gomez, L., Boller, T., Ausubel, F.M., and Sheen, J. (2013). MAP kinase signalling cascade in Arabidopsis innate immunity. **415**: 1–7.

Asakura, Y. and Barkan, A. (2007). A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. Plant Cell **19**: 3864–3875.

Bach-Pages, M., Castello, A., and Preston, G.M. (2017). Plant RNA interactome capture: revealing the plant RBPome. Trends Plant Sci. **22**: 449–451.

Bach-Pages, M. and Preston, G.M. (2018). Methods to quantify biotic-induced stress in plants. In Host-Pathogen Interactions, pp. 241–255.

Baena-González, E., Baginsky, S., Mulo, P., Summer, H., Aro, E.M., and Link, G. (2001). Chloroplast transcription at different light intensities. Glutathione-mediated phosphorylation of the major RNA polymerase involved in redox-regulated organellar gene expression. Plant Physiol. **127**: 1044–1052.

Baer, B. and Millar, A.H. (2016). Proteomics in evolutionary ecology. J. Proteomics **135**: 4–11.

Baldrich, P., Rutter, B.D., Karimi, H.Z., Podicheti, R., Meyers, B.C., and Innes, R.W. b. (2019). Plant extracellular vesicles contain diverse small RNA species and are enriched in 10- to 17-nucleotide "Tiny" RNAs. Plant Cell **31**: 315–324.

Baltz, A.G. et al. (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell **46**: 674–690.

Bang, W.Y., Kim, S.W., Jeong, I.S., Koiwa, H., and Bahk, J.D. (2008). The C-terminal region (640-967) of Arabidopsis CPL1 interacts with the abiotic stress- and ABA-responsive transcription factors. Biochem. Biophys. Res. Commun. **372**: 907–912.

Bao, X. et al. (2018). Capturing the interactome of newly transcribed RNA. Nat. Methods **15**: 213–220.

Barakat, A., Szick-miranda, K., Chang, I., Guyot, R., Blanc, G., Cooke, R., Delseny, M., and Bailey-serres, J. (2001). The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. Plant Phy **127**: 398–415.

Barkan, A., Kawamura, T., Asakura, Y., Watkins, K.P., Klipcan, L., and Ostersetzer, O. (2007). The CRM domain: An RNA binding module derived from an ancient ribosome-associated protein. Rna **13**: 55–64.

Barkan, A. and Small, I. (2014). Pentatricopeptide repeat proteins in plants. Annu. Rev. Plant Biol. **65**: 415–442.

Bass, J., Wilkinson, D., Rankin, D., Phillips, B., Szewczyk, N., Smith, K., and Atherton, P. (2017). An overview of technical considerations for Western blotting applications to physiological research. Scand J Med Sci Sport. **27**: 4–25.

Bazin, J., Romero, N., Rigo, R., Charon, C., Blein, T., Ariel, F., and Crespi, M. (2018). Nuclear speckle RNA binding proteins remodel alternative splicing and the non-coding Arabidopsis transcriptome to regulate a cross-talk between auxin and immune responses. Front. Plant Sci. **9**: 1–13.

Beckmann, B.M. (2017). RNA interactome capture in yeast. Methods **118**–**119**: 82–92.

Beckmann, B.M., Horos, R., Fischer, B., Castello, A., Eichelbaum, K., Alleaume, A.-M., Schwarzl, T., Curk, T., Foehr, S., Huber, W., Krijgsveld, J., and Hentze, M.W. (2015). The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. Nat. Commun. **6**: 10127.

Bektas, Y., Rodriguez-Salus, M., Schroeder, M., Gomez, A., Kaloshian, I., and Eulgem, T. (2016). The synthetic elicitor DPMP (2,4-dichloro-6-{(E)-[(3-methoxyphenyl)imino]methyl}phenol) triggers strong immunity in Arabidopsis thaliana and tomato. Sci. Rep. **6**: 1–16.

Bell, S.D., Botting, C.H., Wardleworth, B.N., Jackson, S.P., and White, M.F. (2002). The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. Science (80-.). **296**: 148–151.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. Science (80-.). **301**: 1090–1093.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate : a practical and powerful approach to multiple testing. J. R. Stat. Soc. **57**: 289–300.

Bhasin, H. and Hülskamp, M. (2017). ANGUSTIFOLIA, a plant homolog of CtBP/BARS localizes to stress granules and regulates their formation. Front. Plant Sci. **8**.

Binet, R., Fernandez, R.E., Fisher, D.J., and Maurelli, A.T. (2011). Identification and characterization of the Chlamydia trachomatis L2 S-adenosylmethionine transporter. MBio **2**: e00051-11.

Birkenbihl, R.P., Kracher, B., and Somssich, I.E. (2017). Induced genome-wide binding of three arabidopsis WRKY transcription factors during early MAMP-triggered immunity. Plant Cell **29**: 20–38.

Biswas, D., Yu, Y., Prall, M., Formosa, T., and Stillman, D.J. (2005). The Yeast FACT complex has a role in transcriptional initiation. Mol. Cell. Biol. **25**: 5812–5822.

Bogamuwa, S.P. and Jang, J.C. (2014). Tandem CCCH zinc finger proteins in plant growth, development and stress response. Plant Cell Physiol. **55**: 1367–1375.

Boisson, B., Giglione, C., and Meinnel, T. (2003). Unexpected protein families including cell defense components feature in the N-myristoylome of a higher eukaryote. J. Biol. Chem. **278**: 43418–43429.

Boller, T. and Felix, G. (2009). A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. Annu. Rev. Plant Biol. **60**: 379–406.

Bolton, M.D. (2009). Primary metabolism and plant defense - fuel for the fire. Mol. Plant-Microbe Interact. **22**: 487–497.

Boucas, J., Fritz, C., Schmitt, A., Riabinska, A., Thelen, L., Peifer, M., Leeser, U., Nuernberg, P., Altmueller, J., Gaestel, M., Dieterich, C., and Reinhardt, H.C. (2015). Label-free protein-RNA interactome analysis identifies Khsrp signaling downstream of the p38/Mk2 kinase complex as a critical modulator of cell cycle progression. PLoS One **10**: 1–19.

Boussardon, C., Avon, A., Kindgren, P., Bond, C.S., Challenor, M., Lurin, C., and Small, I. (2014). The cytidine deaminase signature HxE(x)nCxxC of DYW1 binds zinc and is necessary for RNA editing of ndhD-1. New Phytol. **203**: 1090–1095.

Boutrot, F., Segonzac, C., Chang, K.N., Qiao, H., Ecker, J.R., Zipfel, C., and Rathjen, J.P. (2010). Direct transcriptional control of the Arabidopsis immune receptor FLS2 by the ethylene-dependent transcription factors EIN3 and EIL1. Proc. Natl. Acad. Sci. U. S. A. **107**: 14502–14507.

Brant, E.J. and Budak, H. (2018). Plant small non-coding RNAs and their roles in biotic stresses. Front. Plant Sci. **9**: 1–9.

Brown, I., Trethowan, J., Kerry, M., Mansfield, J., and Bolwell, G.P. (1998). Localization of components of the oxidative cross-linking of glycoproteins and of callose synthesis in papillae formed during the interaction between non-pathogenic strains of Xanthomonas

campestris and French bean mesophyll cells. Plant J. **15**: 333–343.

Bruggeman, Q., Garmier, M., de Bont, L., Soubigou-Taconnat, L., Mazubert, C., Benhamed, M., Raynaud, C., Bergounioux, C., and Delarue, M. (2014). The polyadenylation factor subunit CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR30: A key factor of programmed cell death and a regulator of immunity in arabidopsis. Plant Physiol. **165**: 732–746.

Bunnik, E.M., Batugedara, G., Saraf, A., Prudhomme, J., Florens, L., and Le Roch, K.G. (2016). The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum. Genome Biol. **17**: 147.

van der Burgh, A.M. and Joosten, M.H.A.J. (2019). Plant immunity: thinking outside and inside the box. Trends Plant Sci. **24**: 587–601.

Buscaill, P., Chandrasekar, B., Sanguankiattichai, N., Kourelis, J., Kaschani, F., Thomas, E.L., Morimoto, K., Kaiser, M., Preston, G.M., Ichinose, Y., and Van Der Hoorn, R.A.L. (2019). Glycosidase and glycan polymorphism control hydrolytic release of immunogenic flagellin peptides. Science (80-.). **364**.

Buscaill, P. and Rivas, S. (2014). Transcriptional control of plant defence responses. Curr. Opin. Plant Biol. **20**: 35–46.

del Campo, E.M. (2009). Post-transcriptional control of chloroplast gene expression. Gene Regul. Syst. Bio. **3**: 31–47.

Cappadocia, L., Parent, J.S., Sygusch, J., and Brisson, N. (2013). A family portrait: structural comparison of the Whirly proteins from Arabidopsis thaliana and Solanum tuberosum. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. **69**: 1207–1211.

Carballo, E. (1998). Feedback inhibition of macrophage tumor necrosis factor- production by tristetraprolin. Science (80-.). **281**: 1001–1005.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2012).

Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell **149**: 1393–1406.

Castello, A., Fischer, B., Frese, C.K., Horos, R., Alleaume, A.M., Foehr, S., Curk, T., Krijgsveld, J., and Hentze, M.W. (2016). Comprehensive identification of RNA-binding domains in human cells. Mol. Cell **63**: 696–710.

Castello, A., Fischer, B., Hentze, M.W., and Preiss, T. (2013a). RNA-binding proteins in Mendelian disease. Trends Genet. **29**: 318–327.

Castello, A., Frese, C.K., Fischer, B., Järvelin, A.I., Horos, R., Alleaume, A.-M., Foehr, S., Curk, T., Krijgsveld, J., and Hentze, M.W. (2017). Identification of RNA-binding domains of RNA-binding proteins in cultured cells on a system-wide scale with RBDmap. Nat. Protoc. **12**: 2447–2464.

Castello, A., Hentze, M.W., and Preiss, T. (2015). Metabolic enzymes enjoying new partnerships as RNA-binding proteins. Trends Endocrinol. Metab. **26**: 746–757.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2013b). System-wide identification of RNA-binding proteins by interactome capture. Nat. Protoc. **8**: 491–500.

Caudron-Herger, M., Rusin, S.F., Adamo, M.E., Seiler, J., Schmid, V.K., Barreau, E., Kettenbach, A.N., and Diederichs, S. (2019). R-DeeP: proteome-wide and quantitative identification of RNA-dependent proteins by density gradient ultracentrifugation. Mol. Cell **75**: 184-199.e10.

Chadha, P. and Das, R.H. (2006). A pathogenesis related protein, AhPR10 from peanut: An insight of its mode of antifungal activity. Planta **225**: 213–222.

Chakraborty, J., Ghosh, P., and Das, S. (2018). Autoimmunity in plants. Planta **248**: 751–767.

Chakraburtty, K. (2001). Translational regulation by ABC systems. Res. Microbiol. **152**: 391–399.

Chang, C.H. et al. (2013). Posttranscriptional control of T cell effector function by aerobic

glycolysis. Cell **153**: 1239.

Chantarachot, T. and Bailey-Serres, J. (2017). Polysomes, stress granules and processing bodies: a dynamic triumvirate controlling cytoplasmic mRNA fate and function. Plant Physiol. **176**: pp.01468.2017.

Chaudhary, S., Khokhar, W., Jabre, I., Reddy, A.S.N., Byrne, L.J., Wilson, C.M., and Syed, N.H. (2019). Alternative splicing and protein diversity: plants versus animals. Front. Plant Sci. **10**: 1–14.

Chen, C.Y.A. and Shyu, A. Bin (2014). Emerging mechanisms of mRNP remodeling regulation. Wiley Interdiscip. Rev. RNA **5**: 713–722.

Chen, H. and Boutros, P.C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics: 35.

Chen, H., Xue, L., Chintamanani, S., Germain, H., Lin, H., Cui, H., Cai, R., Zuo, J., Tang, X., Li, X., Guo, H., and Zhou, J.M. (2009). ETHYLENE INSENSITIVE3 and ETHYLENE INSENSITIVE3-LIKE1 repress SALICYLIC ACID INDUCTION DEFICIENT2 expression to negatively regulate plant innate immunity in Arabidopsis. Plant Cell **21**: 2527–2540.

Chen, T., Cui, P., Chen, H., Ali, S., Zhang, S., and Xiong, L. (2013). A KH-domain RNA-binding protein interacts with FIERY2/CTD phosphatase-Like 1 and splicing factors and is important for pre-mRNA splicing in Arabidopsis. PLoS Genet. **9**: 1–14.

Cheng, S., Liu, R., and Gallie, D.R. (2013). The unique evolution of the programmed cell death 4 protein in plants. BMC Evol. Biol. **13**.

Chiang, Y.-H. and Coaker, G. (2015). Effector triggered immunity: NLR immune perception and downstream defense responses. Arab. B. **13**: e0183.

Chicois, C., Scheer, H., Garcia, S., Zuber, H., Mutterer, J., Chicher, J., Hammann, P., Gagliardi, D., and Garcia, D. (2018). The UPF1 interactome reveals interaction networks between RNA degradation and translation repression factors in Arabidopsis. Plant J. **96**: 119–132.

Chinchilla, D., Bauer, Z., Regenass, M., Boller, T., and Felix, G. (2006). The Arabidopsis

receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. Plant Cell **18**: 465–476.

Cho, H., Cho, H.S., and Hwang, I. (2019). Emerging roles of RNA-binding proteins in plant development. Curr. Opin. Plant Biol. **51**: 51–57.

Choi, D.S., Hwang, I.S., and Hwang, B.K. (2012). Requirement of the cytosolic interaction between PATHOGENESIS-RELATED PROTEIN10 and LEUCINE-RICH REPEAT PROTEIN1 for cell death and defense signaling in pepper. Plant Cell **24**: 1675–1690.

Choquet, Y., Zito, F., Wostrikoff, K., and Wollman, F.A. (2003). Cytochrome f translation in Chlamydomonas chloroplast is autoregulated by its carboxyl-terminal domain. Plant Cell **15**: 1443–1454.

Choudhury, N.R., Heikel, G., Trubitsyna, M., Kubik, P., Nowak, J.S., Webb, S., Granneman, S., Spanos, C., Rappsilber, J., Castello, A., and Michlewski, G. (2017). RNA-binding activity of TRIM25 is mediated by its PRY/SPRY domain and is required for ubiquitination. BMC Biol. **15**: 1–20.

Clay, N.K., Adio, A.M., Denoux, C., Jander, G., and Ausubel, F.M. (2009). Glucosinolate metabolites required for an Arabidopsis innate immune response. Science (80-.). **323**: 95–102.

Clerch, L.B., Wright, A., and Massaro, D. (1996). Dinucleotide-binding site of bovine liver catalase mimics a catalase mRNA-binding protein domain. Am. J. Physiol. **270**: 790–794.

Clingman, C.C., Deveau, L.M., Hay, S.A., Genga, R.M., Shandilya, S.M.D., Massi, F., and Ryder, S.P. (2014). Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite. Elife **2014**: 1–26.

Cohen, I., Knopf, J.A., Irihimovitch, V., and Shapira, M. (2005). A proposed mechanism for the inhibitory effects of oxidative stress on Rubisco assembly and its subunit expression. Plant Physiol. **137**: 738–746.

Cohen, I., Sapir, Y., and Shapira, M. (2006). A conserved mechanism controls translation of

Rubisco large subunit in different photosynthetic organisms. **141**: 1089–1097.

Colcombet, J., Lopez-Obando, M., Heurtevin, L., Bernard, C., Martin, K., Berthomé, R., and Lurin, C. (2013). Systematic study of subcellular localization of Arabidopsis PPR proteins confirms a massive targeting to organelles. RNA Biol. **10**: 1557–1575.

Conrad, T., Albrecht, A., Rodrigues, V., Costa, D.M., Sauer, S., Meierhofer, D., and Ørom, U.A. (2016). Serial interactome capture of the human cell nucleus. Nat. Commun.: 1–11.

Copeland, C., Xu, S., Qi, Y., and Li, X. (2013). MOS2 has redundant function with its homolog MOS2H and is required for proper splicing of SNC1. Plant Signal. Behav. **8**: 8–10.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics **13**: 2513–2526.

Cox, J. and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. **26**: 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J. V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. **10**: 1794–1805.

Cui, H., Tsuda, K., and Parker, J.E. (2015). Effector-triggered immunity: from pathogen perception to robust defense. Annu. Rev. Plant Biol. **66**: 487–511.

Cui, P., Chen, T., Qin, T., Ding, F., Wang, Z., Chen, H., and Xiong, L. (2016). The RNA Polymerase II C-terminal domain phosphatase-like protein FIERY2/CPL1 interacts with eIF4AIII and is essential for nonsense-mediated mRNA decay in Arabidopsis. Plant Cell **28**: 770–785.

Dabo, S. and Meurs, E.F. (2012). dsRNA-dependent protein kinase PKR and its role in stress, signaling and HCV infection. Viruses **4**: 2598–2635.

Datta, R., Kumar, D., Sultana, A., Hazra, S., Bhattacharyya, D., and Chattopadhyay, S. (2015).

Glutathione regulates ACC synthase transcription via WRKY33 and ACC oxidase by modulating mRNA stability to induce ethylene synthesis during stress. Plant Physiol. **169**: pp.01543.2015.

Davidson, A.L., Dassa, E., Orelle, C., and Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol. Mol. Biol. Rev. **72**: 317–364.

Deng, X., Lu, T., Wang, L., Gu, L., Sun, J., Kong, X., Liu, C., and Cao, X. (2016). Recruitment of the NineTeen Complex to the activated spliceosome requires AtPRMT5. Proc. Natl. Acad. Sci. **113**: 5447–5452.

Denoux, C., Galletti, R., Mammarella, N., Gopalan, S., Werck, D., De Lorenzo, G., Ferrari, S., Ausubel, F.M., and Dewdney, J. (2008). Activation of defense response pathways by OGs and Flg22 elicitors in Arabidopsis seedlings. Mol. Plant **1**: 423–445.

Deragon, J.M. and Bousquet-Antonelli, C. (2015). The role of LARP1 in translation and beyond. Wiley Interdiscip. Rev. RNA **6**: 399–417.

Despic, V., Dejung, M., Gu, M., Krishnan, J., Zhang, J., Herzel, L., Straube, K., Gerstein, M.B., Butter, F., and Neugebauer, K.M. (2017). Dynamic RNA-protein interactions underlie the zebrafish maternal-to-zygotic transition. Genome Res. **1**: gr.215954.116.

Dessau, M., Halimi, Y., Erez, T., Chomsky-Hecht, O., Chamovitz, D.A., and Hirsch, J.A. (2008). The Arabidopsis COP9 signalosome subunit 7 is a model PCI domain protein with subdomains involved in COP9 signalosome assembly. Plant Cell **20**: 2815–2834.

Desveaux, D., Maréchal, A., and Brisson, N. (2005). Whirly transcription factors: Defense gene regulation and beyond. Trends Plant Sci. **10**: 95–102.

van Dijk, T.B., Gillemans, N., Stein, C., Fanis, P., Demmers, J., van de Corput, M., Essers, J., Grosveld, F., Bauer, U.-M., and Philipsen, S. (2010). Friend of Prmt1, a Novel Chromatin Target of Protein Arginine Methyltransferases. Mol. Cell. Biol. **30**: 260–272.

Doma, M.K. and Parker, R. (2007). RNA quality control in eukaryotes. Cell **131**: 660–668.

Dong, O.X., Meteignier, L.-V., Plourde, M.B., Ahmed, B., Wang, M., Jensen, C., Jin, H.,

Moffett, P., Li, X., and Germain, H. (2016). Arabidopsis TAF15b localizes to RNA processing bodies and contributes to snc1 -mediated autoimmunity. Mol. Plant-Microbe Interact. **29**: 247–257.

Van Dyke, M.W., Nelson, L.D., Weilbaecher, R.G., and Mehta, D. V. (2004). Stm1p, a G4 quadruplex and purine motif triplex nucleic acid-binding protein, interacts with ribosomes and subtelomeric Y' DNA in Saccharomyces cerevisiae. J. Biol. Chem. **279**: 24323–24333.

Eckmann, C.R., Rammelt, C., and Wahle, E. (2011). Control of poly(A) tail length. Wiley Interdiscip. Rev. RNA **2**: 348–361.

Edwards, K., Johnstone, C., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res. **19**: 1349.

El-Gebali, S. et al. (2019). The Pfam protein families database in 2019. Nucleic Acids Res. **47**: D427–D432.

Emms, D.M. and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. **16**: 1–14.

Eskelin, K., Varjosalo, M., Ravantti, J., and Mäkinen, K. (2019). Ribosome profiles and riboproteomes of healthy and Potato virus A- and Agrobacterium-infected Nicotiana benthamiana plants. Mol. Plant Pathol. **20**: 392–409.

Fakih, Z., Ahmed, M.B., Letanneur, C., and Germain, H. (2016). An unbiased nuclear proteomics approach reveals novel nuclear protein components that participates in MAMP-triggered immunity. Plant Signal. Behav. **11**: 1–6.

Fan, G., Yang, Y., Li, T., Lu, W., Du, Y., Qiang, X., Wen, Q., and Shan, W. (2018). A Phytophthora capsici RXLR effector targets and inhibits a plant PPIase to suppress endoplasmic reticulum-iediated Immunity. Mol. Plant **11**: 1067–1083.

Feng, J., Li, J., Gao, Z., Lu, Y., Yu, J., Zheng, Q., Yan, S., Zhang, W., He, H., Ma, L., and Zhu, Z.

(2015). SKIP confers osmotic tolerance during salt stress by controlling alternative gene splicing in Arabidopsis. Mol. Plant **8**: 1038–1052.

Filipenko, E.A., Kochetov, A. V., Kanayama, Y., Malinovsky, V.I., and Shumny, V.K. (2013). PR-proteins with ribonuclease activity and plant resistance against pathogenic fungi. Russ. J. Genet. Appl. Res. **3**: 474–480.

Flury, P., Klauser, D., Boller, T., and Bartels, S. (2013). MAPK phosphorylation assay with leaf disks of Arabidopsis. Bio-protocol **3**: 3–6.

Forterre, P., Confalonieri, F., and Knapp, S. (1999). Identification of the gene encoding archeal-specific DNA-binding proteins of the Sac10b family. Mol. Microbiol. **32**: 669–670.

Francischini, C.W. and Quaggio, R.B. (2009). Molecular characterization of Arabidopsis thaliana PUF proteins - Binding specificity and target candidates. FEBS J. **276**: 5456–5470.

Fribourg, S., Braun, I.C., Izaurralde, E., and Conti, E. (2001). Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor. Mol. Cell **8**: 645–656.

Fu, Z.Q., Guo, M., Jeong, B., Tian, F., Elthon, T.E., Cerny, R.L., Staiger, D., and Alfano, J.R. (2007). A type III effector ADP-ribosylates RNA-binding proteins and quells plant immunity. Nature **447**: 284–288.

Galán, J.E., Lara-Tejero, M., Marlovits, T.C., and Wagner, S. (2014). Bacterial type III secretion systems: specialized nanomachines for protein delivery into target cells. Annu. Rev. Microbiol. **68**: 415–438.

García-Andrade, J., Ramírez, V., López, A., and Vera, P. (2013). Mediated plastid RNA editing in plant immunity. PLoS Pathog. **9**: 1–13.

Garcia-Moreno, M. et al. (2019). System-wide profiling of RNA-binding proteins uncovers key regulators of virus infection. Mol. Cell **74**: 196-211.e11.

Garcia, D., Garcia, S., and Voinnet, O. (2014). Nonsense-mediated decay serves as a general viral restriction mechanism in plants. Cell Host Microbe **16**: 391–402.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. **8**: 113–126.

Gaspar, M., Bousser, A., Sissoëff, I., Roche, O., Hoarau, J., and Mahé, A. (2003). Cloning and characterization of ZmPIP1-5b, an aquaporin transporting water and urea. Plant Sci. **165**: 21–31.

Gazzarrini, S. and Mccourt, P. (2003). Cross-talk in plant hormone signalling: What arabidopsis mutants are telling us. Ann. Bot. **91**: 605–612.

Germain, H. et al. (2010). MOS11: A new component in the mRNA export pathway. PLoS Genet. **6**: 1–9.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nat. Rev. Genet. **15**: 829–845.

Gimenez-Ibanez, S., Boter, M., Fernández-Barbero, G., Chini, A., Rathjen, J.P., and Solano, R. (2014). The bacterial effector HopX1 targets JAZ transcriptional repressors to activate jasmonate signaling and promote infection in Arabidopsis. PLoS Biol. **12**.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. **582**: 1977–1986.

Gloggnitzer, J., Akimcheva, S., Srinivasan, A., Kusenda, B., Riehs, N., Stampfl, H., Bautor, J., Dekrout, B., Jonak, C., Jiménez-Gómez, J.M., Parker, J.E., and Riha, K. (2014). Nonsense-mediated mRNA decay modulates immune receptor levels to regulate plant antibacterial defense. Cell Host Microbe **16**: 376–390.

Gong, Z., Dong, C.-H., Lee, H., Zhu, J., Xiong, L., Gong, D., Stevenson, B., and Zhu, J.-K. (2005). A DEAD box RNA helicase is essential for mRNA export and important for development and stress responses in Arabidopsis. Plant Cell **17**: 256–67.

Gosai, S.J., Foley, S.W., Wang, D., Silverman, I.M., Selamoglu, N., Nelson, A.D.L., Beilstein, M.A., Daldal, F., Deal, R.B., and Gregory, B.D. (2015). Global analysis of the RNA-protein interaction and RNA secondary structure landscapes of the arabidopsis nucleus. Mol. Cell

57: 376–388.

Goverse, A. and Smant, G. (2014). The activation and suppression of plant innate immunity by parasitic nematodes. Annu. Rev. Phytopathol. **52**: 243–265.

Goyal, M., Banerjee, C., Nag, S., and Bandyopadhyay, U. (2016). The Alba protein family: Structure and function. Biochim. Biophys. Acta **1864**: 570–583.

Grainger, J.L. and Winkler, M.M. (1987). Fertilization triggers unmasking of maternal mRNAs in sea urchin eggs. Mol. Cell. Biol. **7**: 3947–3954.

Gu, L., Xu, T., Lee, K., Lee, K.H., and Kang, H. (2014). A chloroplast-localized DEAD-box RNA helicaseAtRH3 is essential for intron splicing and plays an important role in the growth and stress response in Arabidopsis thaliana. Plant Physiol. Biochem. **82**: 309–318.

Gu, Y., Zebell, S.G., Liang, Z., Wang, S., Kang, B.H., and Dong, X. (2016). Nuclear pore permeabilization is a convergent signaling event in effector-triggered immunity. Cell **166**: 1526-1538.e11.

Guiducci, G. et al. (2019). The moonlighting RNA-binding activity of cytosolic serine hydroxymethyltransferase contributes to control compartmentalization of serine metabolism. Nucleic Acids Res. **47**: 4240–4254.

Gulledge, A.A., Roberts, A.D., Vora, H., Patel, K., and Loraine, A.E. (2012). Mining Arabidopsis thaliana RNA-seq data with integrated genome browser reveals stress-induced alternative splicing of the putative splicing regulator SR45A. Am. J. Bot. **99**: 219–231.

Guo, H. (2018). Specialized ribosomes and the control of translation. Biochem. Soc. Trans. **46**: 855–869.

Guo, R., Xue, H., and Huang, L. (2003). Ssh10b, a conserved thermophilic archaeal protein, binds RNA in vivo. Mol. Microbiol. **50**: 1605–1615.

Guo, Y.H., Yu, Y.P., Wang, D., Wu, C.A., Yang, G.D., Huang, J.G., and Zheng, C.C. (2009). GhZFP1, a novel CCCH-type zinc finger protein from cotton, enhances salt stress tolerance and fungal disease resistance in transgenic tobacco by interacting with

GZIRD21A and GZIPR5. New Phytol. **183**: 62–75.

Hacquard, S. et al. (2016). Survival trade-offs in plant roots during colonization by closely related beneficial and pathogenic fungi. Nat. Commun. **7**: 11362.

Haghighat, A. and Sonenberg, N. (1997). eIF4g dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. J. Biol. Chem. **272**: 21677–21680.

Hammani, K. and Giegé, P. (2014). RNA metabolism in plant mitochondria. Trends Plant Sci. **19**: 380–389.

Hemetsberger, C., Herrberger, C., Zechmann, B., Hillmer, M., and Doehlemann, G. (2012). The Ustilago maydis effector Pep1 suppresses plant immunity by inhibition of host peroxidase activity. PLoS Pathog. **8**.

Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. Nat. Rev. Mol. Cell Biol. **19**: 327–341.

Hernandez-Pinzon, I., Yelina, N.E., Schwach, F., Studholme, D.J., Baulcombe, D., and Dalmay, T. (2007). SDE5, the putative homologue of a human mRNA export factor, is required for transgene silencing and accumulation of trans-acting endogenous siRNA. Plant J. **50**: 140–148.

Hobor, F., Dallmann, A., Ball, N.J., Cicchini, C., Battistelli, C., Ogrodowicz, R.W., Christodoulou, E., Martin, S.R., Castello, A., Tripodi, M., Taylor, I.A., and Ramos, A. (2018). A cryptic RNA-binding domain mediates Syncrip recognition and exosomal partitioning of miRNA targets. Nat. Commun. **9**.

Hofmann, K. and Bucher, P. (1998). The PCI domain: a common theme in three multiprotein complexes. Trends Biochem. Sci. **23**: 204–205.

Holm, L.M., Jahn, T.P., Møller, A.L.B., Schjoerring, J.K., Ferri, D., Klaerke, D.A., and Zeuthen, T. (2005). NH3 and NH4+ permeability in aquaporin-expressing Xenopus oocytes. Pflugers Arch. Eur. J. Physiol. **450**: 415–428.

Horos, R. et al. (2019). The small non-coding vault RNA1-1 acts as a riboregulator of

autophagy. Cell **176**: 1054-1067.e12.

Horvathova, I., Voigt, F., Kotrys, A. V., Zhan, Y., Artus-Revel, C.G., Eglinger, J., Stadler, M.B., Giorgetti, L., and Chao, J.A. (2017). The dynamics of mRNA turnover revealed by single-molecule imaging in single cells. Mol. Cell **68**: 615-625.e9.

Hsin, J.P. and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. **26**: 2119–2137.

Hua, C., Zhao, J.H., and Guo, H.S. (2018). Trans-kingdom RNA silencing in plant–fungal pathogen interactions. Mol. Plant **11**: 235–244.

Huang, J. et al. (2017). An oomycete plant pathogen reprograms host pre-mRNA splicing to subvert immunity. Nat. Commun. **8**: 2051.

Huang, R., Han, M., Meng, L., and Chen, X. (2018a). Capture and identification of RNA-binding proteins by using click chemistry-assisted RNA-interactome capture (CARIC) strategy. J. Vis. Exp.

Huang, R., Han, M., Meng, L., and Chen, X. (2018b). Transcriptome-wide discovery of coding and noncoding RNA-binding proteins. Proc. Natl. Acad. Sci. U. S. A. **115**: E3879–E3887.

Huang, Y.W., Hu, C.C., Liou, M.R., Chang, B.Y., Tsai, C.H., Meng, M., Lin, N.S., and Hsu, Y.H. (2012). Hsp90 interacts specifically with viral RNA and differentially regulates replication initiation of Bamboo mosaic virus and associated satellite RNA. PLoS Pathog. **8**.

Huber, W., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. Bioinformatics **18**: S96–S104.

Hubstenberger, A. et al. (2017). P-body purification reveals the condensation of repressed mRNA regulons. Mol. Cell **68**: 144-157.e5.

Huh, S.U., Kim, M.J., and Paek, K.-H. (2013). Arabidopsis pumilio protein APUM5 suppresses cucumber mosaic virus infection via direct binding of viral RNAs. Proc. Natl. Acad. Sci. **110**: 779–784.

Huh, S.U. and Paek, K.-H. (2013). Plant RNA binding proteins for control of RNA virus infection. Front. Physiol. **4**: 1–5.

Huh, S.U. and Paek, K.H. (2014). APUM5, encoding a pumilio RNA binding protein, negatively regulates abiotic stress responsive gene expression. BMC Plant Biol. **14**: 1–17.

Ichinose, M. and Sugita, M. (2017). RNA editing and its molecular mechanism in plant organelles. Genes (Basel). **8**: 1–15.

Ichinose, M. and Sugita, M. (2018). The DYW domains of pentatricopeptide repeat RNA editing factors contribute to discriminate target and non-target editing sites. Plant Cell Physiol. **59**: 1652–1659.

Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T., and Tomari, Y. (2010). Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. Mol. Cell **39**: 292–299.

Jain, S. and Kumar, A. (2015). The pathogenesis related class 10 proteins in plant defense against biotic and abiotic stresses. Adv. Plants Agric. Res. **2**: 305–314.

Jang, Y.H., Park, H.Y., Lee, K.C., Thu, M.P., Kim, S.K., Suh, M.C., Kang, H., and Kim, J.K. (2014). A homolog of splicing factor SF1 is essential for development and is involved in the alternative splicing of pre-mRNA in Arabidopsis thaliana. Plant J. **78**: 591–603.

Järvelin, A.I., Noerenberg, M., Davis, I., and Castello, A. (2016). The new (dis) order in RNA regulation. Cell Commun. Signal.

Jayaraman, D., Forshey, K.L., Grimsrud, P.A., and Ané, J.M. (2012). Leveraging proteomics to understand plant-microbe interactions. Front. Plant Sci. **3**: 1–6.

Jelinska, C., Conroy, M.J., Craven, C.J., Hounslow, A.M., Bullough, P.A., Waltho, J.P., Taylor, G.L., and White, M.F. (2005). Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. Structure **13**: 963–971.

Jenkins, N.A., Kaumeyer, J.F., Young, E.M., and Raff, R.A. (1978). A test for masked message: The template activity of messenger ribonucleoprotein particles isolated from sea urchin

eggs. Dev. Biol. **63**: 279–298.

Jeong, B.R., Lin, Y., Joe, A., Guo, M., Korneli, C., Yang, H., Wang, P., Yu, M., Cerny, R.L., Staiger, D., Alfano, J.R., and Xu, Y. (2011a). Structure function analysis of an ADP-ribosyltransferase type III effector and its RNA-binding target in plant immunity. J. Biol. Chem. **286**: 43272–43281.

Jeong, H.J., Kim, Y.J., Kim, S.H., Kim, Y.H., Lee, I.J., Kim, Y.K., and Shin, J.S. (2011b). Nonsense-mediated mRNA decay factors, UPF1 and UPF3, contribute to plant defense. Plant Cell Physiol. **52**: 2147–2156.

Jeong, I.S., Aksoy, E., Fukudome, A., Akhter, S., Hiraguri, A., Fukuhara, T., Bahk, J.D., and Koiwa, H. (2013). Arabidopsis C-terminal domain phosphatase-like 1 functions in miRNA accumulation and DNA methylation. PLoS One **8**: 1–7.

Jiang, J., Wang, B., Shen, Y., Wang, H., Feng, Q., and Shi, H. (2013a). The Arabidopsis RNA Binding Protein with K Homology Motifs, SHINY1, Interacts with the C-terminal Domain Phosphatase-like 1 (CPL1) to Repress Stress-Inducible Gene Expression. PLoS Genet. **9**.

Jiang, S., Yao, J., Ma, K.W., Zhou, H., Song, J., He, S.Y., and Ma, W. (2013b). Bacterial effector activates jasmonate signaling by directly targeting JAZ transcriptional repressors. PLoS Pathog. **9**.

Jiao, Y., Riechmann, J.L., and Meyerowitz, E.M. (2008). Transcriptome-wide analysis of uncapped mRNAs in Arabidopsis reveals regulation of mRNA degradation. Plant Cell **20**: 2571–2585.

Jin, H., Fu, M., Duan, Z., Duan, S., Li, M., Dong, X., Liu, B., Feng, D., Wang, J., Peng, L., and Wang, H. Bin (2018). Low photosynthetic efficiency 1 is required for light-regulated photosystem II biogenesis in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. **115**: E6075–E6084.

Johnson, J.M. et al. (2018). A poly(A) ribonuclease controls the cellotriose-based interaction between piriformospora indica and its host arabidopsis. Plant Physiol. **176**: 2496–2514.

Johnson, K.C.M., Dong, O.X., and Li, X. (2011). The evolutionarily conserved MOS4-associated

complex. Cent. Eur. J. Biol. **6**: 776–784.

Ju, C. et al. (2012). CTR1 phosphorylates the central regulator EIN2 to control ethylene hormone signaling from the ER membrane to the nucleus in Arabidopsis. PNAS **109**: 19486–19491.

Ju, C., Van De Poel, B., Cooper, E.D., Thierer, J.H., Gibbons, T.R., Delwiche, C.F., and Chang, C. (2015). Conservation of ethylene as a plant hormone over 450 million years of evolution. Nat. Plants **1**: 1–7.

Kanno, T., Venhuizen, P., Wen, T.N., Lin, W.D., Chiou, P., Kalyna, M., Matzke, A.J.M., and Matzke, M. (2018). PRP4KA, a putative spliceosomal protein kinase, is important for alternative splicing and development in arabidopsis Thaliana. Genetics **210**: 1267–1285.

Kant, P., Kant, S., Gordon, M., Shaked, R., and Barak, S. (2007). STRESS RESPONSE SUPPRESSOR1 and STRESS RESPONSE SUPPRESSOR2, two DEAD-box RNA helicases that attenuate Arabidopsis responses to multiple abiotic stresses. Plant Physiol. **145**: 814–830.

Kappel, C., Trost, G., Czesnick, H., Ramming, A., Kolbe, B., Vi, S.L., Bispo, C., Becker, J.D., de Moor, C., and Lenhard, M. (2015). Genome-Wide Analysis of PAPS1-Dependent Polyadenylation Identifies Novel Roles for Functionally Specialized Poly(A) Polymerases in Arabidopsis thaliana. PLoS Genet. **11**: 1–30.

Kasajima, I., Ide, Y., Ohkama-Ohtsu, N., Hayashi, H., Yoneyama, T., and Fujiwara, T. (2004). A protocol for rapid DNA extraction from Arabidopsis thaliana for PCR analysis. Plant Mol. Biol. Report. **22**: 49–52.

Katahira, J. (2015). Nuclear export of messenger RNA. Genes (Basel). **6**: 163–184.

Katahira, J., Dimitrova, L., Imai, Y., and Hurt, E. (2015). NTF2-like domain of Tap plays a critical role in cargo mRNA recognition and export. Nucleic Acids Res. **43**: 1894–1904.

Keinath, N.F., Kierszniowska, S., Lorek, J., Bourdais, G., Kessler, S.A., Shimosato-Asano, H., Grossniklaus, U., Schulze, W.X., Robatzek, S., and Panstruga, R. (2010). PAMP (Pathogen-associated Molecular Pattern)-induced changes in plasma membrane

compartmentalization reveal novel components of plant immunity. J. Biol. Chem. **285**: 39140–39149.

Kerk, D., Templeton, G., and Moorhead, G.B.G. (2008). Evolutionary radiation pattern of novel protein phosphatases revealed by analysis of protein data from the completely sequenced genomes of humans, green algae, and higher plants. Plant Physiol. **146**: 351–367.

Kim, H.S., Abbasi, N., and Choi, S.B. (2013). Bruno-like proteins modulate flowering time via 3' UTR-dependent decay of SOC1 mRNA. New Phytol. **198**: 747–756.

Klopfenstein, D. V. et al. (2018). GOATOOLS: A Python library for Gene Ontology analyses. Sci. Rep. **8**: 1–17.

Koiwa, H. et al. (2002). C-terminal domain phosphatase-like family members (AtCPLs) differentially regulate Arabidopsis thaliana abiotic stress signaling, growth, and development. Proc. Natl. Acad. Sci. U. S. A. **99**: 10893–10898.

Koncz, C., DeJong, F., Villacorta, N., Szakonyi, D., and Koncz, Z. (2012). The spliceosome-activating complex: molecular mechanisms underlying the function of a pleiotropic regulator. Front. Plant Sci. **3**: 1–12.

Kondrashov, A., Meijer, hedda a., Barthet-barateig, A., Parker, hannah n., Khurshid, A., Tessier, S., Sicard, M., Knox, alan j., Pang, L., and de moor, C. h. (2009). Inhibition of polyadenylation reduces inflammatory gene induction. RNA **18**: 2236–2250.

Köster, T. and Meyer, K. (2018). Plant ribonomics: proteins in search of RNA partners. Trends Plant Sci. **23**: 352–365.

Köster, T., Meyer, K., Weinholdt, C., Smith, L.M., Lummer, M., Speth, C., Grosse, I., Weigel, D., and Staiger, D. (2014). Regulation of pri-miRNA processing by the hnRNP-like protein AtGRP7 in Arabidopsis. Nucleic Acids Res. **42**: 9925–9936.

Kouba, T., Rutkai, E., Karásková, M., and Valášek, L.S. (2012). The eIF3c/NIP1 PCI domain interacts with RNA and RACK1/ASC1 and promotes assembly of translation preinitiation

complexes. Nucleic Acids Res. **40**: 2683–2699.

Kourelis, J., Kaschani, F., Grosse-Holz, F.M., Homma, F., Kaiser, M., and van der Hoorn, R.A.L. (2019). Homology-guided re-annotation improves the gene models of the alloploid Nicotiana benthamiana. bioRxiv: 373506.

Kramer, K., Sachsenberg, T., Beckmann, B.M., Qamar, S., Boon, K.-L., Hentze, M.W., Kohlbacher, O., and Urlaub, H. (2014). Photo-cross-linking and high-resolution mass spectrometry for assignment of RNA-binding sites in RNA-binding proteins. Nat. Methods **11**: 1064–1070.

Krapp, S., Greiner, E., Amin, B., Sonnewald, U., and Krenz, B. (2017). The stress granule component G3BP is a novel interaction partner for the nuclear shuttle proteins of the nanovirus pea necrotic yellow dwarf virus and geminivirus abutilon mosaic virus. Virus Res. **227**: 6–14.

Krause, K., Herrmann, U., Fuss, J., Miao, Y., and Krupinska, K. (2009). Whirly proteins as communicators between plant organelles and the nucleus?: 51–62.

Kumar, K.R.R. and Kirti, P.B. (2012). Novel role for a serine/arginine-rich splicing factor, AdRSZ21 in plant defense and HR-like cell death. Plant Mol. Biol. **80**: 461–476.

Kwon, S.C., Yi, H., Eichelbaum, K., Fohr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nat. Struct. Mol. Biol. **20**: 1122–1130.

Labadorf, A., Link, A., Rogers, M.F., Thomas, J., Reddy, A.S., and Ben-Hur, A. (2010). Genome-wide analysis of alternative splicing in Chlamydomonas reinhardtii. BMC Genomics **11**: 114.

Lai, Y., Cuzick, A., Lu, X.M., Wang, J., Katiyar, N., Tsuchiya, T., Le Roch, K., McDowell, J.M., Holub, E., and Eulgem, T. (2018). The Arabidopsis RRM domain protein EDM3 mediates race-specific disease resistance by controlling H3K9me2-dependent alternative polyadenylation of RPP7 immune receptor transcripts. Plant J.: 10.1111/tpj.14148.

Laloum, T., Martín, G., and Duque, P. (2018). Alternative splicing control of abiotic stress Responses. Trends Plant Sci. **23**: 140–150.

Laluk, K., Abuqamar, S., and Mengiste, T. (2011). The Arabidopsis mitochondria-localized pentatricopeptide repeat protein PGN functions in defense against necrotrophic fungi and abiotic stress tolerance. Plant Physiol. **156**: 2053–68.

Lee, D.H., Kim, D.S., and Hwang, B.K. (2012a). The pepper RNA-binding protein CaRBP1 functions in hypersensitive cell death and defense signaling in the cytoplasm. Plant J. **72**: 235–248.

Lee, D.H., Park, S.J., Ahn, C.S., and Pai, H.S. (2017). MRF family genes are involved in translation control, especially under energy-deficient conditions, and their expression and functions are modulated by the TOR signaling pathway. Plant Cell **29**: 2895–2920.

Lee, F.C.Y. and Ule, J. (2018). Advances in CLIP technologies for studies of protein-RNA interactions. Mol. Cell **69**: 354–369.

Lee, H.J., Kim, J.S., Yoo, S.J., Kang, E.Y., Han, S.H., Yang, K.Y., Kim, Y.C., McSpadden Gardener, B., and Kang, H. (2012b). Different roles of glycine-rich RNA-binding protein7 in plant defense against Pectobacterium carotovorum, Botrytis cinerea, and tobacco mosaic viruses. Plant Physiol. Biochem. **60**: 46–52.

Lee, K., Lee, H.J., Kim, D.H., Jeon, Y., Pai, H.-S., and Kang, H. (2014). A nuclear-encoded chloroplast protein harboring a single CRM domain plays an important role in the Arabidopsis growth and stress response. BMC Plant Biol. **14**: 98.

Li, B., Meng, X., Shan, L., and He, P. (2016). Transcriptional regulation of pattern-triggered immunity in plants. Cell Host Microbe **19**: 641–650.

Li, D., Liu, H., Zhang, H., Wang, X., and Song, F. (2008). OsBIRH1, a DEAD-box RNA helicase with functions in modulating defence responses against pathogen infection and oxidative stress. J. Exp. Bot. **59**: 2133–2146.

Li, D., Zhang, H., Hong, Y., Huang, L., Li, X., Zhang, Y., Ouyang, Z., and Song, F. (2014).

Genome-wide identification, biochemical characterization, and expression analyses of the YTH domain-containing RNA-binding protein family in Arabidopsis and rice. Plant Mol. Biol. Rep **32**: 1169–1186.

Li, W., Ma, M., Feng, Y., Li, H., Wang, Y., Ma, Y., Li, M., An, F., and Guo, H. (2015). EIN2-directed translational regulation of ethylene signaling in Arabidopsis. Cell **163**.

Li, Y., Yang, J., Shang, X., Lv, W., Xia, C., Wang, C., Feng, J., Cao, Y., He, H., Li, L., and Ma, L. (2019). SKIP regulates environmental fitness and floral transition by forming two distinct complexes in Arabidopsis. New Phytol. **1**.

Liang, W., Li, C., Liu, F., Jiang, H., Li, S., Sun, J., Wu, X., and Li, C. (2009). The Arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. Cell Res. **19**: 307–316.

Liao, S., Sun, H., and Xu, C. (2018). YTH domain: a family of N6-methyladenosine (m6A) readers. Genomics, Proteomics Bioinforma. **16**: 99–107.

Liao, Y. et al. (2016). The cardiomyocyte RNA-binding proteome: links to intermediary metabolism and heart disease. Cell Rep. **16**: 1456–1469.

Liepelt, A. et al. (2016). Identification of RNA-binding proteins in macrophages by interactome capture. Mol. Cell. Proteomics **15**: 2699–2714.

Liu, P., Zhang, H., Yu, B., Xiong, L., and Xia, Y. (2015). Proteomic identification of early salicylate-and flg22-responsive redox-sensitive proteins in Arabidopsis. Sci. Rep. **5**: 1–7.

Liu, X., Huang, B., Lin, J., Fei, J., Chen, Z., Pang, Y., Sun, X., and Tang, K. (2006). A novel pathogenesis-related protein (SsPR10) from Solanum surattense with ribonucleolytic and antimicrobial activity is stress- and pathogen-inducible. J. Plant Physiol. **163**: 546–556.

Liu, X., Yu, F., and Rodermel, S. (2010). An arabidopsis pentatricopeptide repeat protein, SUPPRESSOR OF VARIEGATION7, is required for FtsH-mediated chloroplast biogenesis. Plant Physiol. **154**: 1588–1601.

Liu, Y. and Imai, R. (2018). Function of plant DExD/H-Box RNA helicases associated with

ribosomal RNA biogenesis. Front. Plant Sci. **9**: 1–7.

de Longevialle, A.F., Meyer, E.H., Andrés, C., Taylor, N.L., Lurin, C., Millar, A.H., and Small, I.D. (2007). The pentatricopeptide repeat gene OTP43 is required for trans -splicing of the mitochondrial nad1 intron 1 in Arabidopsis thaliana. Plant Cell **19**: 3256–3265.

Lorković, Z.J. and 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**: 623–35.

Lozano-Durán, R., Bourdais, G., He, S.Y., and Robatzek, S. (2014). The bacterial effector HopM1 suppresses PAMP-triggered oxidative burst and stomatal immunity. New Phytol. **202**: 259–269.

Lu, Y. (2018). RNA editing of plastid-encoded genes. Photosynthetica **56**: 48–61.

Lueong, S., Merce, C., Fischer, B., Hoheisel, J.D., and Erben, E.D. (2016). Gene expression regulatory networks in Trypanosoma brucei: Insights into the role of the mRNA-binding proteome. Mol. Microbiol. **100**: 457–471.

Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nat. Rev. Mol. Cell Biol. **8**: 479–490.

Lurin, C. et al. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell **16**: 2089–2103.

Lv, H.X., Huang, C., Guo, G.Q., and Yang, Z.N. (2014). Roles of the nuclear-encoded chloroplast SMR domain-containing PPR protein SVR7 in photosynthesis and oxidative stress tolerance in Arabidopsis. J. Plant Biol. **57**: 291–301.

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**: 1–10.

Macho, A.P. and Zipfel, C. (2015). Targeting of plant pattern recognition receptor-triggered immunity by bacterial type-III secretion system effectors. Curr. Opin. Microbiol. **23**: 14–

22.

Maldonado-Bonilla, L.D., Eschen-Lippold, L., Gago-Zachert, S., Tabassum, N., Bauer, N., Scheel, D., and Lee, J. (2014). The arabidopsis tandem zinc finger 9 protein binds RNA and mediates pathogen-associated molecular pattern-triggered immune responses. Plant Cell Physiol. **55**: 412–425.

Manavella, P.A., Hagmann, J., Ott, F., Laubinger, S., Franz, M., MacEk, B., and Weigel, D. (2012). Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. Cell **151**: 859–870.

Marondedze, C., Thomas, L., Gehring, C., and Lilley, K.S. (2019). Changes in the Arabidopsis RNA-binding proteome reveal novel stress response mechanisms. BMC Plant Biol. **19**: 1–11.

Marondedze, C., Thomas, L., Serrano, N.L., Lilley, K.S., and Gehring, C. (2016). The RNA-binding protein repertoire of Arabidopsis thaliana. Sci. Rep. **6**: 29766.

Marquez, Y., Brown, J.W.S., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. Genome Res. **22**: 1184–1195.

Matia-González, A.M., Laing, E.E., and Gerber, A.P. (2015). Conserved mRNA-binding proteomes in eukaryotic organisms. Nat. Struct. Mol. Biol. **22**: 1027–1033.

Matsuda, O., Sakamoto, H., Nakao, Y., Oda, K., and Iba, K. (2009). CTD phosphatases in the attenuation of wound-induced transcription of jasmonic acid biosynthetic genes in Arabidopsis. Plant J. **57**: 96–108.

Mauro, V.P. and Edelman, G.M. (2002). The ribosome filter hypothesis. PNAS **99**: 12031–12036.

McCarty, D.R. and Chory, J. (2000). Conservation and innovation in plant signaling pathways. Cell **103**: 201–209.

McGlincy, N.J. and Smith, C.W.J. (2008). Alternative splicing resulting in nonsense-mediated

mRNA decay: what is the meaning of nonsense? Trends Biochem. Sci. **33**: 385–393.

Melotto, M., Underwood, W., Koczan, J., Nomura, K., and He, S.Y. (2006). Plant stomata function in innate immunity against bacterial invasion. Cell **126**: 969–980.

Meng, J., Wang, L., Wang, J., Zhao, X., Cheng, J., Yu, W., Jin, D., Li, Q., and Gong, Z. (2018). METHIONINE ADENOSYLTRANSFERASE4 mediates DNA and histone methylation. Plant Physiol. **177**: 652–670.

Meng, X. and Zhang, S. (2013). MAPK cascades in plant disease resistance Signaling. Annu. Rev. Phytopathol. **51**: 245–266.

Merai, Z., Benkovics, A.H., Nyiko, T., Debreczeny, M., Hiripi, L., Kerenyi, Z., Kondorosi, E., and Silhavy, D. (2013). The late steps of plant nonsense-mediated mRNA decay. Plant J. **73**: 50–62.

Merchante, C., Brumos, J., Yun, J., Hu, Q., Spencer, K.R., Enríquez, P., Binder, B.M., Heber, S., Stepanova, A.N., and Alonso, J.M. (2015). Gene-specific translation regulation mediated by the hormone-signaling molecule EIN2. Cell **163**: 684–697.

Merret, R., Descombin, J., Juan, Y. ting, Favory, J.J., Carpentier, M.C., Chaparro, C., Charng, Y. yung, Deragon, J.M., and Bousquet-Antonelli, C. (2013). XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. Cell Rep. **5**: 1279–1293.

Merret, R., Nagarajan, V.K., Carpentier, M.C., Park, S., Favory, J.J., Descombin, J., Picart, C., Charng, Y.Y., Green, P.J., Deragon, J.M., and Bousquet-Antonelli, C. (2015). Heat-induced ribosome pausing triggers mRNA co-translational decay in Arabidopsis thaliana. Nucleic Acids Res. **43**: 4121–4132.

Mersmann, S., Bourdais, G., Rietz, S., and Robatzek, S. (2010). Ethylene signaling regulates accumulation of the FLS2 receptor and is required for the oxidative burst contributing to plant immunity. Plant Physiol. **154**: 391–400.

Meyer, K., Köster, T., Nolte, C., Weinholdt, C., Lewinski, M., Grosse, I., and Staiger, D. (2017).

Adaptation of iCLIP to plants determines the binding landscape of the clock-regulated RNA-binding protein AtGRP7. Genome Biol. **18**: 1–22.

Michalski, A. et al. (2012). Ultra high resolution linear ion trap orbitrap mass spectrometer (orbitrap elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes. Mol. Cell. Proteomics **11**: 1–11.

Mingam, A., Toffano-Nioche, C., Brunaud, V., Boudet, N., Kreis, M., and Lecharny, A. (2004). DEAD-box RNA helicases in Arabidopsis thaliana: establishing a link between quantitative expression, gene structure and evolution of a family of genes. Plant Biotechnol. J. **2**: 401–415.

Mitchell, A.L. et al. (2019). InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res. **47**: D351–D360.

Mitchell, S.F., Jain, S., She, M., and Parker, R. (2013). Global analysis of yeast mRNPs. Nat. Struct. Mol. Biol. **20**: 127–133.

Molina, A., Mena, M., Carbonero, P., and García-Olmedo, F. (1997). Differential expression of pathogen-responsive genes encoding two types of glycine-rich proteins in barley. Plant Mol. Biol. **33**: 803–10.

Molitor, A.M., Latrasse, D., Zytnicki, M., Andrey, P., Houba-Hérin, N., Hachet, M., Battail, C., Del Prete, S., Alberti, A., Quesneville, H., and Gaudin, V. (2016). The Arabidopsis hnRNP-Q protein LIF2 and the PRC1 subunit LHP1 function in concert to regulate the transcription of stress-responsive genes. Plant Cell **28**: 2197–2211.

Monaghan, J., Xu, F., Gao, M., Zhao, Q., Palma, K., Long, C., Chen, S., Zhang, Y., and Li, X. (2009). Two Prp19-like U-box proteins in the MOS4-associated complex play redundant roles in plant innate immunity. PLoS Pathog. **5**.

Monaghan, J., Xu, F., Xu, S., Zhang, Y., and Li, X. (2010). Two putative RNA-binding proteins function with unequal genetic redundancy in the MOS4-associated complex. Plant Physiol. **154**: 1783–1793.

Morandini, P., Valera, M., Albumi, C., Bonza, M.C., Giacometti, S., Ravera, G., Murgia, I., Soave, C., and De Michelis, M.I. (2002). A novel interaction partner for the C-terminus of Arabidopsis thaliana plasma membrane H+-ATPase (AHA1 isoform): Site and mechanism of action on H+-ATPase activity differ from those of 14-3-3 proteins. Plant J. **31**: 487–497.

Moriyama, T. and Sato, N. (2014). Enzymes involved in organellar DNA replication in photosynthetic eukaryotes. Front. Plant Sci. **5**: 1–12.

Musidlak, O., Nawrot, R., and Goździcka-Józefiak, A. (2017). Which plant proteins are involved in antiviral defense? Review on in vivo and in vitro activities of selected plant proteins against viruses. Int. J. Mol. Sci. **18**.

Muthuramalingam, M., Wang, Y., Li, Y., and Mahalingam, R. (2017). Interacting protein partners of Arabidopsis RNA-binding protein AtRBP45b. Plant Biol. **19**: 327–334.

Na, J.K., Kim, J.K., Kim, D.Y., and Assmann, S.M. (2015). Expression of potato RNA-binding proteins StUBA2a/b and StUBA2c induces hypersensitive-like cell death and early leaf senescence in Arabidopsis. J. Exp. Bot. **66**: 4023–4033.

Næsted, H., Holm, A., Jenkins, T., Nielsen, H.B., Harris, C.A., Beale, M.H., Andersen, M., Mant, A., Scheller, H., Camara, B., Mattsson, O., and Mundy, J. (2004). Arabidopsis VARIEGATED 3 encodes a chloroplast-targeted, zinc-finger protein required for chloroplast and palisade cell development. J. Cell Sci. **117**: 4807–4818.

Nagai, K., Oubridge, C., Ito, N., Avis, J., and Evans, P. (1995). The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. Trends Biochem. Sci. **20**: 235–240.

Nagaraj, S., Senthil-Kumar, M., Ramu, V.S., Wang, K., and Mysore, K.S. (2016). Plant ribosomal proteins, RPL12 and RPL19, play a role in nonhost disease resistance against bacterial pathogens. Front. Plant Sci. **6**: 1–10.

Nandan, D., Thomas, S.A., Nguyen, A., Moon, K.-M., Foster, L.J., and Reiner, N.E. (2017). Comprehensive Identification of mRNA-Binding Proteins of Leishmania donovani by

Interactome Capture. PLoS One **12**: e0170068.

Narsai, R., Howell, K.A., Millar, A.H., O'Toole, N., Small, I., and Whelan, J. (2007). Genome-wide analysis of mRNA decay rates and their determinants in Arabidopsis thaliana. Plant Cell **19**: 3418–3436.

Nicaise, V., Joe, A., Jeong, B.R., Korneli, C., Boutrot, F., Westedt, I., Staiger, D., Alfano, J.R., and Zipfel, C. (2013). Pseudomonas HopU1 modulates plant immune receptor levels by blocking the interaction of their mRNAs with GRP7. EMBO J. **32**: 701–712.

Nicholson, P. and Mühlemann, O. (2010). Cutting the nonsense: The degradation of PTC-containing mRNAs. Biochem. Soc. Trans. **38**: 1615–1620.

Nishioka, K., Wang, X.F., Miyazaki, H., Soejima, H., and Hirose, S. (2018). Mbf1 ensures polycomb silencing by protecting E(Z) mRNA from degradation by Pacman. Dev. **145**.

Ohta, H., Takamune, N., Kishimoto, N., Shoji, S., and Misumi, S. (2015). N-myristoyltransferase 1 enhances human immunodeficiency virus replication through regulation of viral RNA expression level. Biochem. Biophys. Res. Commun. **463**: 988–993.

Ohtani, M. and Wachter, A. (2019). NMD-based gene regulation - a strategy for fitness enhancement in plants? Plant Cell Physiol. **00**: 1–8.

Olsen, J. V., de Godoy, L.M.F., Li, G., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005). Parts per million mass accuracy on an orbitrap mass spectrometer via lock mass injection into a C-trap. Mol. Cell. Proteomics **4**: 2010–2021.

De Pablos, L.M., Ferreira, T.R., Dowle, A., Forrester, S., Parry, E., Newling, K., and Walrad, P.B. (2019). The mRNA-bound proteome of Leishmania mexicana : novel genetic insight into an ancient parasite . Mol. Cell. Proteomics: mcp.RA118.001307.

Paieri, F., Tadini, L., Manavski, N., Kleine, T., Ferrari, R., Morandini, P., Pesaresi, P., Meurer, J., and Leister, D. (2018). The DEAD-box RNA helicase RH50 is a 23S-4.5S rRNA maturation factor that functionally overlaps with the plastid signaling factor GUN1. Plant

Physiol. **176**: 634–648.

Palma, K., Zhao, Q., Yu, T.C., Bi, D., Monaghan, J., Cheng, W., Zhang, Y., and Li, X. (2007). Regulation of plant innate immunity by three proteins in a complex conserved across the plant and animal kingdoms. Genes Dev. **21**: 1484–1493.

Palusa, S.G., Ali, G.S., and Reddy, A.S.N. (2007). Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses. Plant J. **49**: 1091–1107.

Pan, H., Liu, S., and Tang, D. (2012). HPR1, a component of the THO/TREX complex, plays an important role in disease resistance and senescence in Arabidopsis. Plant J. **69**: 831–843.

Parenteau, J., Durand, M., Morin, G., Gagnon, J., Lucier, J.F., Wellinger, R.J., Chabot, B., and Elela, S.A. (2011). Introns within ribosomal protein genes regulate the production and function of yeast ribosomes. Cell **147**: 320–331.

Park, C.J., Kim, K.J., Shin, R., Park, J.M., Shin, Y.C., and Paek, K.H. (2004). Pathogenesis-related protein 10 isolated from hot pepper functions as a ribonuclease in an antiviral pathway. Plant J. **37**: 186–198.

Park, C.J., Park, C.B., Hong, S.S., Lee, H.S., Lee, S.Y., and Kim, S.C. (2000). Characterization and cDNA cloning of two glycine- and histidine-rich antimicrobial peptides from the roots of shepherd's purse, Capsella bursa-pastoris. Plant Mol. Biol. **44**: 187–197.

Park, C.J. and Seo, Y.S. (2015). Heat shock proteins: a review of the molecular chaperones for plant immunity. Plant Pathol. J. **31**: 323–333.

Pashev, I.G., Dimitrov, S.I., and Angelov, D. (1991). Crosslinking proteins to nucleic acids by ultraviolet laser irradiation. Trends Biochem. Sci. **16**: 323–326.

Peal, L., Jambunathan, N., and Mahalingam, R. (2011). Phylogenetic and expression analysis of RNA-binding proteins with triple RNA recognition motifs in plants. Mol. Cells **31**: 55–64.

Perez-Perri, J.I., Rogell, B., Schwarzl, T., Stein, F., Zhou, Y., Rettel, M., Brosig, A., and Hentze,

M.W. (2018). Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. Nat. Commun. **9**.

Pérez-Vilaró, G., Fernández-Carrillo, C., Mensa, L., Miquel, R., Sanjuan, X., Forns, X., Pérez-Del-Pulgar, S., and Díez, J. (2015). Hepatitis C virus infection inhibits P-body granule formation in human livers. J. Hepatol. **62**: 785–790.

Petitot, A.S., Blein, J.P., Pugin, A., and Suty, L. (1997). Cloning of two plant cDNAs encoding a β-type proteasome subunit and a transformer-2-like SR-related protein: early induction of the corresponding genes in tobacco cells treated with cryptogein. Plant Mol. Biol. **35**: 261–269.

Petre, B. and Kamoun, S. (2014). How do filamentous pathogens deliver effector proteins into plant cells? PLoS Biol. **12**.

Petre, B., Saunders, D.G.O., Sklenar, J., Lorrain, C., Krasileva, K. V., Win, J., Duplessis, S., and Kamoun, S. (2016). Heterologous expression screens in nicotiana benthamiana identify a candidate effector of the wheat yellow rust pathogen that associates with processing bodies. PLoS One **11**: 1–16.

Pisarev, A. V., Kolupaeva, V.G., Yusupov, M.M., Hellen, C.U.T., and Pestova, T. V. (2008). Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. EMBO J. **27**: 1609–1621.

Plaschka, C., Lin, P.C., Charenton, C., and Nagai, K. (2018). Prespliceosome structure provides insights into spliceosome assembly and regulation. Nature **559**: 419–422.

Popow, J., Alleaume, A.M., Curk, T., Schwarzl, T., Sauer, S., and Hentze, M.W. (2015). FASTKD2 is an RNA-binding protein required for mitochondrial RNA processing and translation. Rna **21**: 1873–1884.

Prall, W., Sharma, B., and Gregory, B.D. (2019). Transcription is just the beginning of gene expression tegulation : The functional significance of RNA-binding proteins to post-transcriptional processes in plants. **60**: 1939–1952.

Pramanik, S.K. and Bewley, J.D. (1996). Post-transcriptional regulation of protein synthesis during alfalfa embryogenesis: proteins associated with the cytoplasmic polysomal and non-polysomal mRNAs (messenger ribonucleoprotein complex). J. Exp. Bot. **45**: 1871–1879.

Prikryl, J., Watkins, K.P., Friso, G., van Wijk, K.J., and Barkan, A. (2008). A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. Nucleic Acids Res. **36**: 5152–5165.

Pumplin, N. and Voinnet, O. (2013). RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. Nat. Rev. Microbiol. **11**: 745–760.

Pungartnik, C., da Silva, A.C., de Melo, S.A., Gramacho, K.P., de Mattos Cascardo, J.C., Brendel, M., Micheli, F., and da Silva Gesteira, A. (2009). High-affinity copper transport and Snq2 export permease of Saccharomyces cerevisiae modulate cytotoxicity of PR-10 from Theobroma cacao. Mol. Plant-Microbe Interact. **22**: 39–51.

Qi, J., Wang, J., Gong, Z., and Zhou, J.M. (2017). Apoplastic ROS signaling in plant immunity. Curr. Opin. Plant Biol. **38**: 92–100.

Qi, Y., Tsuda, K., Joe, A., Sato, M., Nguyen, L. V, Glazebrook, J., Alfano, J.R., Cohen, J.D., and Katagiri, F. (2010). A putative RNA-binding protein positively regulates salicylic acid-mediated immunity in Arabidopsis. Mol. Plant. Microbe. Interact. **23**: 1573–83.

Qiao, H., Shen, Z., Huang, S.S.C., Schmitz, R.J., Urich, M.A., Briggs, S.P., and Ecker, J.R. (2012). Processing and subcellular trafficking of ER-tethered EIN2 control response to ethylene gas. Science (80-.). **338**: 390–393.

Qiao, Y., Shi, J., Zhai, Y., Hou, Y., and Ma, W. (2015). *Phytophthora* effector targets a novel component of small RNA pathway in plants to promote infection. Proc. Natl. Acad. Sci. **112**: 5850–5855.

Queiroz, R.M.L. et al. (2019). Comprehensive identification of RNA–protein interactions in any organism using orthogonal organic phase separation (OOPS). Nat. Biotechnol. **37**: 169–

178.

Quigley, F., Rosenberg, J.M., Shachar-hill, Y., and Bohnert, H.J. (2001). From genome to function: the Arabidopsis aquaporins. Genome Biol. **3**: 1–17.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Ramanathan, A., Robb, G.B., and Chan, S.H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res. **44**: 7511–7526.

Ramírez, V., Van der Ent, S., García-Andrade, J., Coego, A., Pieterse, C.M.J., and Vera, P. (2010). OCP3 is an important modulator of NPR1-mediated jasmonic acid-dependent induced defenses in Arabidopsis. BMC Plant Biol. **10**: 1–13.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. **2**: 1896–1906.

Rayson, S., Arciga-Reyes, L., Wootton, L., Zabala, M.D.T., Truman, W., Graham, N., Grant, M., and Davies, B. (2012). A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in Arabidopsis thaliana NMD mutants. PLoS One **7**: e31917.

Reddy, A.S.N. (2004). Plant serine/arginine-rich proteins and their role in pre-mRNA splicing. Trends Plant Sci. **9**: 541–547.

Ree, R., Varland, S., and Arnesen, T. (2018). Spotlight on protein N-terminal acetylation. Exp. Mol. Med. **50**.

Reichel, M., Liao, Y., Rettel, M., Ragan, C., Evers, M., Alleaume, A.-M., Horos, R., Hentze, M.W., Preiss, T., and Millar, A.A. (2016). In planta determination of the mRNA-binding proteome of Arabidopsis etiolated seedlings. Plant Cell **28**: 2435–2452.

Ribbeck, K. (1998). NTF2 mediates nuclear import of Ran. EMBO J. **17**: 6587–6598.

Richter, J.D. and Sonenberg, N. (2005). Regulation of cap-dependent translation by eIF4E inhibitory proteins. Nature **433**: 477–480.

Riehs-Kearnan, N., Gloggnitzer, J., Dekrout, B., Jonak, C., and 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.

Rigo, R., Bazin, J., Crespi, M., and Charon, C. (2019). Alternative splicing in the regulation of plant–microbe interactions. Plant Cell Physiol. **0**: 1–11.

Rin, S., Mizuno, Y., Shibata, Y., Fushimi, M., Katou, S., Sato, I., Chiba, S., Kawakita, K., and Takemoto, D. (2017). EIN2-mediated signaling is involved in pre-invasion defense in Nicotiana benthamiana against potato late blight pathogen, Phytophthora infestans. Plant Signal. Behav. **12**.

Rinn, J.L. and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. **81**: 145–166.

Rodrigues, J.P., Rode, M., Gatfield, D., Blencowe, B.J., Carmo-Fonseca, M., and Izaurralde, E. (2001). REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. Proc. Natl. Acad. Sci. U. S. A. **98**: 1030–1035.

Rodrigues, N.F., Christoff, A.P., da Fonseca, G.C., Kulcheski, F.R., and Margis, R. (2017). Unveiling chloroplast RNA editing events using next generation small RNA sequencing data. Front. Plant Sci. **8**.

Ross, S., Giglione, C., Pierre, M., Espagne, C., and Meinnel, T. (2005). Functional and developmental impact of cytosolic protein N-terminal methionine excision in Arabidopsis. Plant Physiol. **137**: 623–637.

Le Roux, C., Del Prete, S., Boutet-Mercey, S., Perreau, F., Balagué, C., Roby, D., Fagard, M., and Gaudin, V. (2014). The hnRNP-Q protein LIF2 participates in the plant immune response. PLoS One **9**.

Roux, M.E., Rasmussen, M.W., Palma, K., Lolle, S., Regué, À.M., Bethke, G., Glazebrook, J., Zhang, W., Sieburth, L., Larsen, M.R., Mundy, J., and Petersen, M. (2015). The mRNA decay factor PAT1 functions in a pathway including MAP kinase 4 and immune receptor

SUMM2. EMBO J. **34**: 593–608.

Rupp, D. and Bartenschlager, R. (2014). Targets for antiviral therapy of hepatitis C. Semin. Liver Dis. **34**: 9–21.

Rustgi, S., Pollmann, S., Buhr, F., Springer, A., Reinbothe, C., von Wettstein, D., and Reinbothe, S. (2014). JIP60-mediated, jasmonate- and senescence-induced molecular switch in translation toward stress and defense protein synthesis. Proc. Natl. Acad. Sci. **111**: 14181–14186.

Rutter, B.D. and Innes, R.W. (2017). Extracellular vesicles isolated from the leaf apoplast carry stress-response proteins. Plant Physiol. **173**: 728–741.

Van Ruyskensvelde, V., Van Breusegem, F., and Van Der Kelen, K. (2018). Post-transcriptional regulation of the oxidative stress response in plants. Free Radic. Biol. Med. **122**: 181–192.

Rybak, K. and Robatzek, S. (2019). Functions of extracellular vesicles in immunity and virulence. Plant Physiol. **179**: 1236–1247.

Ryder, S.P., Recht, M.I., and Williamson, J.R. (2003). Quantitative analysis of protein-RNA Gel Mobility Shift Assay GIVES UPPER AND LOWER LIMITS.: 463.

Salone, V., Rüdinger, M., Polsakiewicz, M., Hoffmann, B., Groth-Malonek, M., Szurek, B., Small, I., Knoop, V., and Lurin, C. (2007). A hypothesis on the identification of the editing enzyme in plant organelles. FEBS Lett. **581**: 4132–4138.

Salvador-Guirao, R., Hsing, Y., and San Segundo, B. (2018). The Polycistronic miR166k-166h Positively Regulates Rice Immunity via Post transcriptional Control of EIN2. Front. Plant Sci. **9**: 337.

Sarowar, S., Oh, H.W., Cho, H.S., Baek, K.H., Seong, E.S., Joung, Y.H., Choi, G.J., Lee, S., and Choi, D. (2007). Capsicum annuum CCR4-associated factor CaCAF1 is necessary for plant development and defence response. Plant J. **51**: 792–802.

Sauter, M., Moffatt, B., Saechao, M.C., Hell, R., and Wirtz, M. (2013). Methionine salvage and S-adenosylmethionine: Essential links between sulfur, ethylene and polyamine

biosynthesis. Biochem. J. **451**: 145–154.

Schad, E., Tompa, P., and Hegyi, H. (2011). The relationship between proteome size, proteome complexity and disorder. Genome Biol. **12**.

Scherrer, T., Mittal, N., Janga, S.C., and Gerber, A.P. (2010). A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. PLoS One **5**: e15499.

Schmitz-Linneweber, C. and Small, I. (2008). Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci. **13**: 663–670.

Serrano, I., Audran, C., and Rivas, S. (2016). Chloroplasts at work during plant innate immunity. J. Exp. Bot. **67**: 3845–3854.

Sharp, P.A. (2005). The discovery of split genes and RNA splicing. Trends Biochem. Sci. **30**: 275–279.

Shaul, O. (2015). Unique aspects of plant nonsense-mediated mRNA decay. Trends Plant Sci. **20**: 767–779.

Shchepachev, V., Bresson, S., Spanos, C., Petfalski, E., Fischer, L., Rappsilber, J., and Tollervey, D. (2019). Defining the RNA interactome by total RNA -associated protein purification . Mol. Syst. Biol. **15**: 1–23.

Sheng, J., D'Ovidio, R., and Mehdy, M.C. (1991). Negative and positive regulation of a novel proline-rich protein mRNA by fungal elicitor and wounding. Plant J. **1**: 345–254.

Shi, C., Baldwin, I.T., and Wu, J. (2012). Arabidopsis plants having defects in nonsense-mediated mRNA decay factors UPF1, UPF2, and UPF3 show photoperiod-dependent phenotypes in development and stress responses. J. Integr. Plant Biol. **54**: 99–114.

Shi, X., Hanson, M.R., and Bentolila, S. (2015). Two RNA recognition motif-containing proteins are plant mitochondrial editing factors. Nucleic Acids Res. **43**: 3814–3825.

Shi, Z., Fujii, K., Kovary, K.M., Genuth, N.R., Röst, H.L., Teruel, M.N., and Barna, M. (2017). Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. Mol. Cell **67**: 71–83.

Shimberg, G.D., Michalek, J.L., Oluyadi, A.A., Rodrigues, A. V., Zucconi, B.E., Neu, H.M., Ghosh, S., Sureschandra, K., Wilson, G.M., Stemmler, T.L., and Michel, S.L.J. (2016). Cleavage and polyadenylation specificity factor 30: An RNA-binding zinc-finger protein with an unexpected 2Fe–2S cluster. Proc. Natl. Acad. Sci. **113**: 4700–4705.

Sidoli, S., Kulej, K., and Garcia, B.A. (2017). Why proteomics is not the new genomics and the future of mass spectrometry in cell biology. J. Cell Biol. **216**: 21–24.

Silverman, I.M., Li, F., and Gregory, B.D. (2013). Genomic era analyses of RNA secondary structure and RNA-binding proteins reveal their significance to post-transcriptional regulation in plants. Plant Sci. **0**: 55–62.

Simsek, D., Tiu, G.C., Flynn, R.A., Byeon, G.W., Leppek, K., Xu, A.F., Chang, H.Y., and Barna, M. (2017). The mammalian ribo-interactome reveals ribosome functional diversity and heterogeneity. Cell **169**: 1051-1065.e18.

Singh, G., Pratt, G., Yeo, G.W., and Moore, M.J. (2015). The clothes make the mRNA: past and present trends in mRNP fashion. Annu. Rev. Biochem. **84**: 583–592.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. **3**: 1–25.

Sohn, K.H. oon, Segonzac, C., Rallapalli, G., Sarris, P.F., Woo, J.Y. ong, Williams, S.J., Newman, T.E., Paek, K.H. ee, Kobe, B., and Jones, J.D.G. (2014). The nuclear immune receptor RPS4 is required for RRS1SLH1-dependent constitutive defense activation in Arabidopsis thaliana. PLoS Genet. **10**: e1004655.

Spirin, A.S., Belitsina, N. V., and Lerman, M.I. (1965). Use of formaldehyde fixation for studies of ribonucleoprotein particles by caesium chloride density-gradient centrifugation. J. Mol. Biol. **14**: 611–615.

Staiger, D., Korneli, C., Lummer, M., and Navarro, L. (2013). Emerging role for RNA-based regulation in plant immunity. New Phytol. **197**: 394–404.

Stern, D.B., Goldschmidt-Clermont, M., and Hanson, M.R. (2010). Chloroplast RNA

Metabolism. Annu. Rev. Plant Biol. **61**: 125–155.

Streitner, C., Köster, T., Simpson, C.G., Shaw, P., Danisman, S., Brown, J.W.S., and Staiger, D. (2012). An hnRNP-like RNA-binding protein affects alternative splicing by in vivo interaction with transcripts in Arabidopsis thaliana. Nucleic Acids Res. **40**: 11240–11255.

Stuwe, T., Hothorn, M., Lejeune, E., Rybin, V., Bortfeld, M., Scheffzek, K., and Ladurner, A.G. (2008). The FACT Spt16 "peptidase" domain is a histone H3-H4 binding module. Proc. Natl. Acad. Sci. U. S. A. **105**: 8884–8889.

Su, J., Yang, L., Zhu, Q., Wu, H., He, Y., Liu, Y., Xu, J., Jiang, D., and Zhang, S. (2018). Active photosynthetic inhibition mediated by MPK3/MPK6 is critical to effector-triggered immunity. PLoS Biol. **16**: 1–29.

Suarez Rodriguez, M.C., Petersen, M., and Mundy, J. (2010). Mitogen-activated protein kinase signaling in plants. Annu. Rev. Plant Biol. **61**: 621–649.

Sun, T., Shi, X., Friso, G., Van Wijk, K., Bentolila, S., and Hanson, M.R. (2015). A zinc finger motif-containing protein Is essential for chloroplast RNA editing. PLoS Genet. **11**: 1–23.

Syed, N.H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J.W.S. (2012). Alternative splicing in plants - coming of age. Trends Plant Sci. **17**: 616–623.

Sysoev, V.O., Fischer, B., Frese, C.K., Gupta, I., Krijgsveld, J., Hentze, M.W., Castello, A., and Ephrussi, A. (2016). Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in Drosophila. Nat. Commun. **7**: 12128.

Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., Jensen, L.J., and Von Mering, C. (2017). The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. **45**: D362–D368.

Tanabe, N., Yoshimura, K., Kimura, A., Yabuta, Y., and Shigeoka, S. (2007). Differential expression of alternatively spliced mRNAs of Arabidopsis SR protein homologs, atSR30 and atSR45a, in response to environmental stress. Plant Cell Physiol. **48**: 1036–1049.

Tena, G., Boudsocq, M., and Sheen, J. (2011). Protein kinase signaling networks in plant innate immunity Guillaume. Curr Opin Plant Biol. **14**: 519–529.

Thapar, R. (2015). Roles of prolyl isomerases in RNA-mediated gene expression. Biomolecules **5**: 974–999.

Tharun, S. (2009). Lsm1-7-Pat1 complex: a link between 3' and 5'-ends in mRNA decay? RNA Biol. **6**: 228–232.

Thatcher, L.F., Foley, R., Casarotto, H.J., Gao, L.L., Kamphuis, L.G., Melser, S., and Singh, K.B. (2018). The Arabidopsis RNA Polymerase II Carboxyl Terminal Domain (CTD) Phosphatase-Like1 (CPL1) is a biotic stress susceptibility gene. Sci. Rep. **8**: 1–14.

Thatcher, L.F., Kamphuis, L.G., Hane, J.K., Oñate-Sánchez, L., and Singh, K.B. (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**: 1–31.

Thomma, B.P.H.J., Nürnberger, T., and Joosten, M.H.A.J. (2011). Of PAMPs and effectors: The blurred PTI-ETI dichotomy. Plant Cell **23**: 4–15.

Tintor, N., Ross, A., Kanehara, K., Yamada, K., Fan, L., Kemmerling, B., Nürnberger, T., Tsuda, K., and Saijo, Y. (2013). Layered pattern receptor signaling via ethylene and endogenous elicitor peptides during Arabidopsis immunity to bacterial infection. Proc Natl Acad Sci U S A **110**: 6211–6216.

Toruño, T.Y., Stergiopoulos, I., and Coaker, G. (2016). Plant-pathogen effectors: cellular probes interfering with plant defenses in spatial and temporal manners. Annu. Rev. Phytopathol. **54**: 419–441.

Trendel, J., Schwarzl, T., Horos, R., Prakash, A., Bateman, A., Hentze, M.W., and Krijgsveld, J. (2019). The human RNA-binding proteome and Its dynamics during translational arrest. Cell **176**: 391-403.e19.

Trost, G., Vi, S.L., Czesnick, H., Lange, P., Holton, N., Giavalisco, P., Zipfel, C., Kappel, C., and Lenhard, M. (2014). Arabidopsis poly(A) polymerase PAPS1 limits founder-cell

recruitment to organ primordia and suppresses the salicylic acid-independent immune response downstream of EDS1/PAD4. Plant J. **77**: 688–699.

Trotman, J.B. and Schoenberg, D.R. (2019). A recap of RNA recapping. Wiley Interdiscip. Rev. RNA **10**: 1–13.

Tsuda, K. and Somssich, I.E. (2015). Transcriptional networks in plant immunity. New Phytol. **206**: 932–947.

Tsuda, K., Tsuji, T., Hirose, S., and Yamazaki, K.I. (2004). Three Arabidopsis MBF1 homologs with distinct expression profiles play roles as transcriptional co-activators. Plant Cell Physiol. **45**: 225–231.

Tsugeki, R., Tanaka-Sato, N., Maruyama, N., Terada, S., Kojima, M., Sakakibara, H., and Okada, K. (2015). CLUMSY VEIN, the Arabidopsis DEAH-box Prp16 ortholog, is required for auxin-mediated development. Plant J. **81**: 183–197.

Tsvetanova, N.G., Klass, D.M., Salzman, J., and Brown, P.O. (2010). Proteome-wide search reveals unexpected RNA-binding proteins in Saccharomyces cerevisiae. PLoS One **5**: 1–12.

Uddin, M.N., Akhter, S., Chakraborty, R., Baek, J.H., Cha, J.Y., Park, S.J., Kang, H., Kim, W.Y., Lee, S.Y., Mackey, D., and Kim, M.G. (2017). SDE5, a putative RNA export protein, participates in plant innate immunity through a flagellin-dependent signaling pathway in Arabidopsis. Sci. Rep. **7**: 1–11.

Urdaneta, E.C., Vieira-Vieira, C.H., Hick, T., Wessels, H.H., Figini, D., Moschall, R., Medenbach, J., Ohler, U., Granneman, S., Selbach, M., and Beckmann, B.M. (2019). Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. Nat. Commun. **10**: 1–17.

Vance, K.W. and Ponting, C.P. (2014). Transcriptional regulatory functions of nuclear long noncoding RNAs. Trends Genet. **30**: 348–355.

Vélez-Bermúdez, I.C. and Schmidt, W. (2014). The conundrum of discordant protein and mRNA expression. Are plants special? Front. Plant Sci. **5**: 1–4.

Venezia, N.D., Vincent, A., Marcel, V., Catez, F., and Diaz, J.J. (2019). Emerging role of eukaryote ribosomes in translational control. Int. J. Mol. Sci. **20**.

Verma, A. et al. (2018). The novel cyst nematode effector protein 30D08 targets host nuclear functions to alter gene expression in feeding sites. New Phytol. **219**: 697–713.

Vijayapalani, P., Hewezi, T., Pontvianne, F., and Baum, T.J. (2018). An effector from the cyst nematode heterodera schachtii derepresses host rRNA genes by altering histone acetylation. Plant Cell **30**: 2795–2812.

Viotti, C., Luoni, L., Morandini, P., and De Michelis, M.I. (2005). Characterization of the interaction between the plasma membrane H +-ATPase of Arabidopsis thaliana and a novel interactor (PPI1). FEBS J. **272**: 5864–5871.

Volz, K. (2008). The functional duality of iron regulatory protein 1. Curr Opin Struct Biol. **18**: 106–111.

Walley, J.W., Coughlan, S., Hudson, M.E., Covington, M.F., Kaspi, R., Banu, G., Harmer, S.L., and Dehesh, K. (2007). Mechanical stress induces biotic and abiotic stress responses via a novel cis-element. PLoS Genet. **3**: 1800–1812.

Walley, J.W., Kelley, D.R., Nestorova, G., Hirschberg, D.L., and Dehesh, K. (2010). Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. Plant Physiol. **152**: 866–875.

Wang, H., Seo, J.-K., Gao, S., Cui, X., and Jin, H. (2017). Silencing of AtRAP, a target gene of a bacteria-induced small RNA, triggers antibacterial defense responses through activation of LSU2 and down-regulation of GLK1. New Phytol. **2017**: 1144–1155.

Wang, M., Ogé, L., Perez-Garcia, M.D., Hamama, L., and Sakr, S. (2018). The PUF protein family: overview on PUF RNA targets, biological functions, and post transcriptional regulation. Int. J. Mol. Sci. **19**: 1–13.

Wang, X., Boevink, P., McLellan, H., Armstrong, M., Bukharova, T., Qin, Z., and Birch, P.R.J. (2015). A host KH RNA-binding protein is a susceptibility factor targeted by an RXLR

effector to promote late blight disease. Mol. Plant **8**: 1385–1395.

Wang, Y., Han, R., Zhang, W., Yuan, Y., Zhang, X., Long, Y., and Mi, H. (2008). Human CyP33 binds specifically to mRNA and binding stimulates PPIase activity of hCyP33. FEBS Lett. **582**: 835–839.

Warner, J.R. and McIntosh, K.B. (2009). How common are extraribosomal functions of ribosomal proteins? Mol. Cell **34**: 3–11.

Washington, E.J., Mukhtar, M.S., Finkel, O.M., Wan, L., Banfield, M.J., Kieber, J.J., and Dangl, J.L. (2016). Pseudomonas syringae type III effector HopAF1 suppresses plant immunity by targeting methionine recycling to block ethylene induction. Proc. Natl. Acad. Sci. U. S. A. **113**: E3577–E3586.

Waskom, M. et al. (2014). seaborn: v0.5.0.

Weber, C., Nover, L., and Fauth, M. (2008). Plant stress granules and mRNA processing bodies are distinct from heat stress granules. Plant J. **56**: 517–530.

Wen, X., Zhang, C., Ji, Y., Zhao, Q., He, W., An, F., Jiang, L., and Guo, H. (2012). Activation of ethylene signaling is mediated by nuclear translocation of the cleaved EIN2 carboxyl terminus. Cell Res. **22**: 1613–1616.

van Wersch, R., Li, X., and Zhang, Y. (2016). Mighty dwarfs: Arabidopsis autoimmune mutants and their usages in genetic dissection of plant immunity. Front. Plant Sci. **7**: 1–8.

Wessels, H.-H., Imami, K., Baltz, A.G., Kolinski, M., Beldovskaya, A., Selbach, M., Small, S., Ohler, U., and Landthaler, M. (2016). The mRNA-bound proteome of the early fly embryo. Genome Res. **26**: 1000–1009.

Wickham, H. (2009). ggplot2 - Elegant Graphics for Data Analysis (2nd Edition) (New York).

Wiermer, M., Cheng, Y.T., Imkampe, J., Li, M., Wang, D., Lipka, V., and Li, X. (2012). Putative members of the Arabidopsis Nup107-160 nuclear pore sub-complex contribute to pathogen defense. Plant J. **70**: 796–808.

Wilhelmsson, P.K.I., Mühlich, C., Ullrich, K.K., and Rensing, S.A. (2017). Comprehensive

genome-wide classification reveals that many plant-specific transcription factors evolved in streptophyte algae. Genome Biol. Evol. **9**: 3384–3397.

Williams-Carrier, R., Brewster, C., Belcher, S.E., Rojas, M., Chotewutmontri, P., Ljungdahl, S., and Barkan, A. (2019). The Arabidopsis pentatricopeptide repeat protein LPE1 and its maize ortholog are required for translation of the chloroplast psbJ RNA. Plant J. **99**: 56–66.

Willmund, F., Del Alamo, M., Pechmann, S., Chen, T., Albanèse, V., Dammer, E.B., Peng, J., and Frydman, J. (2013). The cotranslational function of ribosome-associated Hsp70 in eukaryotic protein homeostasis. Cell **152**: 196–209.

Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat. Methods **6**: 359–62.

Woo, J., Han, D., Wang, J.I., Park, J., Kim, H., and Kim, Y. (2017). Quantitative proteomics reveals temporal proteomic changes in signaling pathways during BV2 mouse microglial cell activation. J. Proteome Res.: 3419−3432.

Wu, C.H., Abd-El-Haliem, A., Bozkurt, T.O., Belhaj, K., Terauchi, R., Vossen, J.H., and Kamoun, S. (2017). NLR network mediates immunity to diverse plant pathogens. Proc. Natl. Acad. Sci. U. S. A. **114**: 8113–8118.

Wu, X., Liu, M., Downie, B., Liang, C., Ji, G., Li, Q.Q., and Hunt, A.G. (2011). Genome-wide landscape of polyadenylation in Arabidopsis provides evidence for extensive alternative polyadenylation. Proc. Natl. Acad. Sci. **108**: 12533–12538.

Wu, X., Shi, Y., Li, J., Xu, L., Fang, Y., Li, X., and Qi, Y. (2013). A role for the RNA-binding protein MOS2 in microRNA maturation in Arabidopsis. Cell Res. **23**: 645–657.

Xiao, R. et al. (2019). Pervasive chromatin-RNA binding protein enteractions Enable RNA-based regulation of transcription. Cell **178**: 107–121.

Xing, H., Fu, X., Yang, C., Tang, X., Guo, L., Li, C., Xu, C., and Luo, K. (2018). Genome-wide investigation of pentatricopeptide repeat gene family in poplar and their expression

analysis in response to biotic and abiotic stresses. Sci. Rep. **8**: 1–9.

Xu, F. et al. (2015). Two N-terminal acetyltransferases antagonistically regulate the stability of a nod-like receptor in arabidopsis. Plant Cell **27**: 1547–1562.

Xu, G., Greene, G.H., Yoo, H., Liu, L., Marqués, J., Motley, J., and Dong, X. (2017). Global translational reprogramming is a fundamental layer of immune regulation in plants. Nature **545**: 487–490.

Xu, Y., Wang, B.C., and Zhu, Y.X. (2007). Identification of proteins expressed at extremely low level in Arabidopsis leaves. Biochem. Biophys. Res. Commun. **358**: 808–812.

Xue, S. and Barna, M. (2012). Specialized ribosomes: A new frontier in gene regulation and organismal biology. Nat. Rev. Mol. Cell Biol. **13**: 355–369.

Yamazaki, H., Tasaka, M., and Shikanai, T. (2004). PPR motifs of the nucleus-encoded factor, PGR3, function in the selective and distinct steps of chloroplast gene expression in Arabidopsis. Plant J. **38**: 152–163.

Yamazaki, T., Souquere, S., Chujo, T., Kobelke, S., Chong, Y.S., Fox, A.H., Bond, C.S., Nakagawa, S., Pierron, G., and Hirose, T. (2018). Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation. Mol. Cell **70**: 1038-1053.e7.

Yang, L., Teixeira, P.J.P.L., Biswas, S., Finkel, O.M., He, Y., Salas-Gonzalez, I., English, M.E., Epple, P., Mieczkowski, P., and Dangl, J.L. (2017). Pseudomonas syringae type III effector HopBB1 promotes host transcriptional repressor degradation to regulate Phytohormone responses and virulence. Cell Host Microbe **21**: 156–168.

Yeh, Y.H., Panzeri, D., Kadot, Y., Huang, Y.C., Huang, P.Y., Tao, C.N., Roux, M., Chien, H.C., Chin, T.C., Chu, P.W., Zipfel, C., and Zimmerli, L. (2016). The arabidopsis malectin-like/LRR-RLK IOS1 is critical for BAK1-dependent and BAK1-independent pattern-triggered immunity. Plant Cell **28**: 1701–1721.

Yosef, I., Irihimovitch, V., Knopf, J.A., Cohen, I., Orr-Dahan, I., Nahum, E., Keasar, C., and

Shapira, M. (2004). RNA binding activity of the bibulose-1,5-bisphosphate carboxylase/oxygenase large subunit from Chlamydomonas reinhardtii. J. Biol. Chem. **279**: 10148–10156.

Yu, Q., Liu, J., Zheng, H., Jia, Y., Tian, H., and Ding, Z. (2016). Topoisomerase II-associated protein PAT1H1 is involved in the root stem cell niche maintenance in Arabidopsis thaliana. Plant Cell Rep. **35**: 1297–1307.

Yu, X., Li, B., Jang, G.-J., Jiang, S., Jiang, D., Jang, J.-C., Wu, S.-H., Shan, L., and He, P. (2019). Orchestration of processing body dynamics and mRNA decay in Arabidopsis immunity. Cell Rep. **28**: 2194-2205.e6.

Yuan, N., Yuan, S., Li, Z., Zhou, M., Wu, P., Hu, Q., Mendu, V., Wang, L., and Luo, H. (2018). STRESS INDUCED FACTOR 2, a leucine-rich repeat kinase regulates basal plant pathogen defense. Plant Physiol. **176**: 3062–3080.

Yuan, W., Zhou, J., Tong, J., Zhuo, W., Wang, L., Li, Y., Sun, Q., and Qian, W. (2019). ALBA protein complex reads genic R-loops to maintain genome stability in Arabidopsis. Sci. Adv. **5**.

Zeng, W. and He, S.Y. (2010). A prominent role of the flagellin receptor FLAGELLIN-SENSING2 in mediating stomatal response to Pseudomonas syringae pv tomato DC3000 in Arabidopsis. Plant Physiol. **153**: 1188–1198.

Zhang, F., Qi, B., Wang, L., Zhao, B., Rode, S., Riggan, N.D., Ecker, J.R., and Qiao, H. (2016a). EIN2-dependent regulation of acetylation of histone H3K14 and non-canonical histone H3K23 in ethylene signalling. Nat. Commun. **7**.

Zhang, F., Wang, L., Qic, B., Zhao, B., Ko, E.E., Riggan, N.D., Chin, K., and Qiao, H. (2017a). EIN2 mediates direct regulation of histone acetylation in ethylene response. Proc Natl Acad Sci U S A **114**: 10274–10279.

Zhang, J., Addepalli, B., Yun, K.Y., Hunt, A.G., Xu, R., Rao, S., Li, Q.Q., and Falcone, D.L. (2008). A polyadenylation factor subunit implicated in regulating oxidative signaling in

Arabidopsis thaliana. PLoS One **3**.

Zhang, S., Liu, Y., and Yu, B. (2014a). PRL1, an RNA-binding protein, positively regulates the accumulation of miRNAs and siRNAs in Arabidopsis. PLoS Genet. **10**.

Zhang, S., Liu, Y., and Yu, B. (2014b). PRL1, an RNA-Binding Protein, Positively Regulates the Accumulation of miRNAs and siRNAs in Arabidopsis. PLoS Genet. **10**.

Zhang, S. and Mehdy, M.C. (1994). Binding of a 50-kD Protein to a U-Rich Sequence in an mRNA Encoding a Proline-Rich Protein That Is Destabilized by Fungal Elicitor. Plant Cell **6**: 135–145.

Zhang, S., Sheng, J., Liu, Y., and Mehdy, M.C. (1993). Fungal Elicitor-Induced Bean Proline-Rich Protein mRNA Down-Regulation Is Due to Destabilization That Is Transcription and Translation Dependent. Plant Cell **5**: 1089–1099.

Zhang, X.N., Shi, Y., Powers, J.J., Gowda, N.B., Zhang, C., Ibrahim, H.M.M., Ball, H.B., Chen, S.L., Lu, H., and Mount, S.M. (2017b). Transcriptome analyses reveal SR45 to be a neutral splicing regulator and a suppressor of innate immunity in Arabidopsis thaliana. BMC Genomics **18**: 1–17.

Zhang, Y., Cheng, Y.T., Bi, D., Palma, K., and Li, X. (2005). MOS2, a protein containing G-patch and KOW motifs, is essential for innate immunity in Arabidopsis thaliana. Curr. Biol. **15**: 1936–1942.

Zhang, Y. and Li, X. (2005). A putative nucleoporin 96 is required for both basal defense and constitutive resistance responses mediated by suppressor of npr1-1,constitutive 1. Plant Cell Online **17**: 1306–1316.

Zhang, Z., Boonen, K., Ferrari, P., Schoofs, L., Janssens, E., van Noort, V., Rolland, F., and Geuten, K. (2016b). UV crosslinked mRNA-binding proteins captured from leaf mesophyll protoplasts. Plant Methods **12**: 42.

Zhang, Z., Boonen, K., Li, M., and Geuten, K. (2017c). mRNA Interactome Capture from Plant Protoplasts. J. Vis. Exp.: e56011.

Zhou, W., Karcher, D., and Bock, R. (2014). Identification of enzymes for adenosine-to-inosine editing and discovery of cytidine-to-uridine editing in nucleus-encoded transfer RNAs of arabidopsis. Plant Physiol. **166**: 1985–1997.

Zhou, X. et al. (2018). Loss of function of a rice TPR-domain RNA-binding protein confers broad-spectrum disease resistance Xiaogang. Proc. Natl. Acad. Sci. **115**: 3174–3179.

Zhou, X.J., Lu, S., Xu, Y.H., Wang, J.W., and Chen, X.Y. (2002). A cotton cDNA (GaPR-10) encoding a pathogenesis-related 10 protein with in vitro ribonuclease activity. Plant Sci. **162**: 629–636.

Zhu, C., Liu, T., Chang, Y.N., and Duan, C.G. (2019). Small RNA functions as a trafficking effector in plant immunity. Int. J. Mol. Sci. **20**.

Zhu, Q., Dugardeyn, J., Zhang, C., Mühlenbock, P., Eastmond, P.J., Valcke, R., De Coninck, B., Öden, S., Karampelias, M., Cammue, B.P.A., Prinsen, E., and Van Der Straeten, D. (2014). The arabidopsis thaliana RNA editing factor SLO2, which affects the mitochondrial electron transport chain, participates in multiple stress and hormone responses. Mol. Plant **7**: 290–310.

Zhuo, Z., Yu, Y., Wang, M., Li, J., Zhang, Z., Liu, J., Wu, X., Lu, A., Zhang, G., and Zhang, B. (2017). Recent advances in SELEX technology and aptamer applications in biomedicine. Int. J. Mol. Sci. **18**: 1–19.

Zipfel, C. (2014). Plant pattern-recognition receptors. Trends Immunol. **35**: 345–351.

Züst, R. et al. (2011). Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. Nat. Immunol. **12**: 137–143.