Trends in Genetics

CellPress

Review

# How intrinsically disordered proteins order plant gene silencing

Baoshuan Shang,[1,2] Changhao Li,[2] and Xiuren Zhang[2,3,*]

**Intrinsically disordered proteins (IDPs) and proteins with intrinsically disordered regions (IDRs) possess low sequence complexity of amino acids and display non-globular tertiary structures. They can act as scaffolds, form regulatory hubs, or trigger biomolecular condensation to control diverse aspects of biology. Emerging evidence has recently implicated critical roles of IDPs and IDR-contained proteins in nuclear transcription and cytoplasmic post-transcriptional processes, among other molecular functions. We here summarize the concepts and organizing principles of IDPs. We then illustrate recent progress in understanding the roles of key IDPs in machineries that regulate transcriptional and post-transcriptional gene silencing (PTGS) in plants, aiming at highlighting new modes of action of IDPs in controlling biological processes.**

## Highlights

The genome of *Arabidopsis thaliana* encodes thousands of intrinsically disordered proteins (IDPs). Several IDPs, exemplified by H2B.8, LHP1, ADCP1, and EMB1579, regulate transcription and epigenetic silencing in various condensation-dependent manners. Diverse IDPs such as FCA, FLL2, FRI, and/or their associated long non-coding RNAs, regulate the transcription of *FLC* via phase separation.

SE promotes Microprocessor activity via its condensation. The protein can be hijacked by SAID1 and 2 to form nuclear condensates, where SAID1/2 sequester primary miRNA precursors (pri-miRNAs) and inhibit Microprocessor activity. SAID1/2 promote phosphorylation of SE, which triggers its destruction by the 20S proteasome and reduces its binding affinity for HYL1, while isomerization of SE promotes its condensation and miRNA production.

SGS3 constitutes a regulatory hub to control the production of 21–22 nt small interfering RNAs (siRNAs) via phase separation. SGS3 condensation is critical for the recruitment of RDR6 and RNA substrates. However, SGS3 association with its RNA substrates is enhanced by FVE, which also bridges SGS3/RDR6 and DRD4/DCL4 to facilitate siRNA production.

## IDPs are pivotal regulators of diverse biological processes

IDPs are a group of proteins that lack a fixed tertiary structure. Frequently, proteins that harbor IDRs while having globular structures in other domains are also classified as IDPs. Unlike well-folded proteins, IDPs possess a biased and heterogeneous amino acid composition, with more hydrophilic and fewer hydrophobic residues [1]. These residues constitute the driving force for IDP interactions via electrostatic interactions (charge–charge, charge–π, cation–π, π–π, and hydrogen bonds) and/or hydrophobic interactions of nonpolar amino acid residues with hydrophobic side chains (Figure 1A) [2]. The amino acid composition and molecular features of IDPs enable them to serve as scaffolds or hubs for multivalent and/or transient interactions with their partners for various functions [3]. IDPs often undergo liquid–liquid phase separation (LLPS), also termed biomolecular condensation, via their IDRs and/or low-complexity sequences. LLPS is a feature of numerous membraneless organelles in eukaryotic cells and allows the compartmentalization and organization of ribonucleoprotein (RNP) macromolecules to fulfill specific cellular activities [4].

Although IDPs have been discovered for decades, their pivotal biological roles and regulatory mechanisms have been only appreciated until very recently. IDPs participate in cell proliferation, polarity, and metabolism in bacteria [5]. In mammals, dysregulation of IDPs can cause brain disorders such as Parkinson's and Alzheimer's diseases, as well as cancers, making IDPs promising targets for therapy [6,7]. In plants, IDPs are engaged in hormone signaling [8], seed germination [9], intra-chloroplast cargo sorting [10], flowering [11,12], plant immune response [13], and temperature and light perception [14,15], among others. In this review we first summarize basic biochemical features of IDPs and then specifically address their roles in gene silencing.

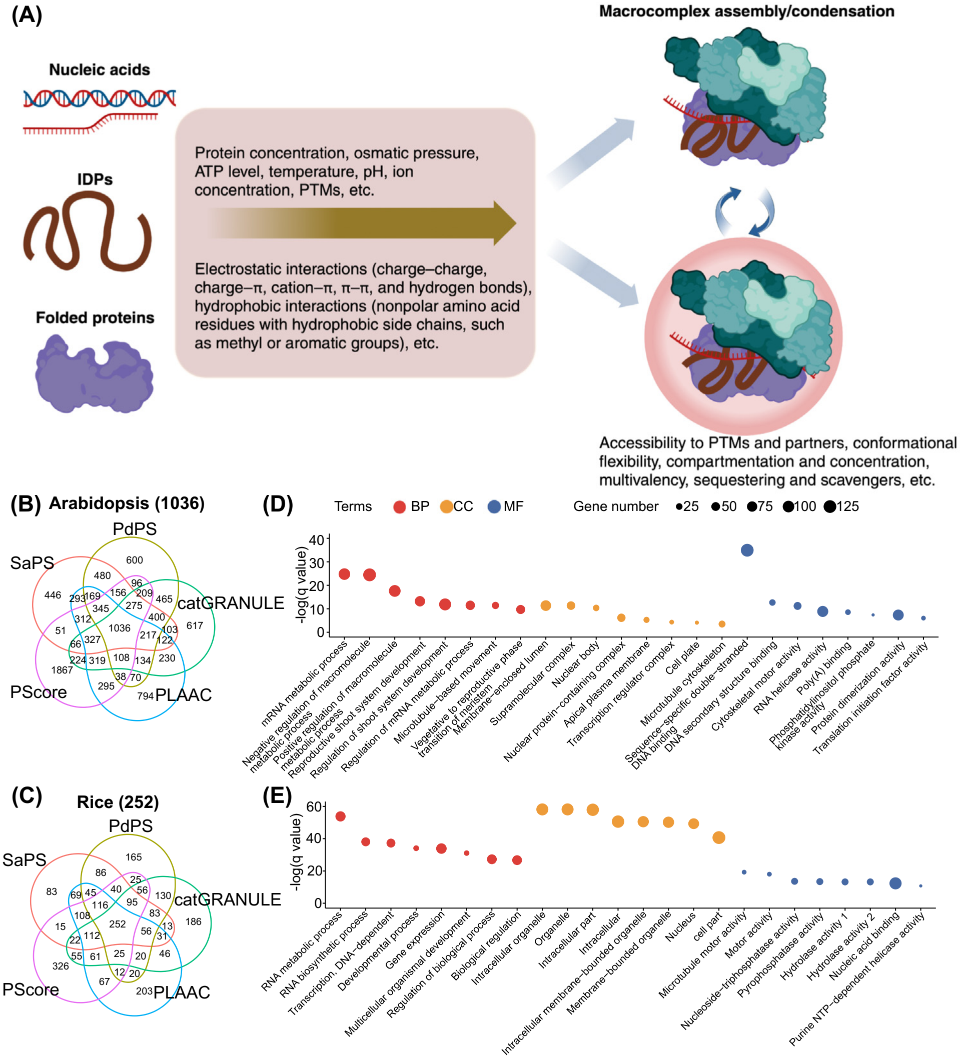
## Biochemical characteristics of IDPs

IDPs are frequently involved in formation of biological condensates via LLPS and can also undergo phase transition to less reversible gel-like and solid aggregates [16]. IDP-triggered condensation can enrich substrates, thus facilitating biochemical reactions. Conversely, the

[1]National Key Laboratory of Cotton Bio-breeding and Integrated Utilization (Henan University), State Key Laboratory of Crop Stress Adaptation and Improvement, School of Life Sciences, Henan University, Kaifeng 475004, China
[2]Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843, USA
[3]Department of Biology, Texas A&M University, College Station, TX 77843, USA

*Correspondence:
xiuren.zhang@tamu.edu (X. Zhang).

https://doi.org/10.1016/j.tig.2023.12.009



**Trends in Genetics**

Figure 1. Interaction principles of intrinsically disordered proteins (IDPs) and the predicted fraction of IDPs with liquid–liquid phase separation (LLPS) properties in *Arabidopsis thaliana* and rice (*Oryza sativa*). (A) IDPs are involved in macromolecular complex assembly or condensation via multivalent or transient interactions with their partners. IDPs recruit their partners, such as nucleic acids and folded proteins, via electrostatic or hydrophobic interactions, to trigger macromolecular complex assembly, disassembly, condensation, and/or aggregation. The interactions are mostly multivalent and transient, and are determined by thermodynamic factors of IDPs or their partners, such as protein concentration, ATP level, temperature, balance of net charge, cellular physiological conditions, and post-translational modifications (PTMs). The features of IDPs and their assembly/condensation underlie their roles in cellular compartmentation, conformational change, accessibility to PTM modifications, sequestering and scavenging, among others, to fine-tune reaction activities, storage, and/or cleavage of proteins. (B–E) Overlap of the top 30% of proteins predicted from five computational analyses to retain LLPS capacity (B and C) and their Gene Ontology (GO) enrichment (D,E). The top 30% of proteins with LLPS capacity were predicted by SaPS, PdPS, catGRANULE, PLAAC, and PScore in *A. thaliana* (B) and rice (C). GO enrichment for the 1036 and 252 overlapping proteins with LLPS capacity in *A. thaliana* (D) and rice (E). Abbreviations: BP, biological process; CC, cellular component; MF, molecular function.

condensates may also sequester biomolecules from other niches to stimulate or suppress the reactions of the sequestered factors. Proteins in the condensates may either be prevented from, or be prone to, degradation depending on biophysiological conditions. Biological condensates or aggregates can also enable cellular compartmentalization, concentration of cellular components, rapid assembly and disassembly, and rapid switchable protein–protein and protein–nucleic acid interactions, and/or facilitate the dynamic formation of local reaction centers with spatiotemporal specificity (Figure 1A) [4]. Overall, the aforementioned features of IDPs are evolutionarily adaptive strategies that enable protective and energy-saving processes for the recovery of cellular proteins without widespread misfolding or degradation.

IDPs have the potential to bind to multiple partners to attain distinct conformations that enable them to interact with different targets and display flexible conformational preferences [17]. Generally, IDRs, short linear motifs (SLiMs), or low-complexity domains (LCDs) of IDPs mediate their multivalent interactions [4]. One type of LCD is a prion-like domain (PrLD) that resembles the yeast prion domain in that it is enriched in polar, uncharged amino acids such as Asn, Gln, Tyr, Ser, and Gly [18]. Upon binding to their partners, many IDPs often undergo conformational changes, depending on their foldability and microenvironment [19,20]. This exposes variable surfaces of different recognition motifs and/or sites of post-translational modifications (PTMs), thus increasing the total repertoire of structures resulting from an IDP sequence and allowing binding promiscuity and plasticity spatial-temporally and transiently [21]. The assembly and disassembly of IDPs depend on the local microenvironment, including ionic concentration, temperature, pH, and other parameters. These processes are also determined by the concentration of the IDPs in aqueous solution and their features of length, hydrophobicity, and charge distribution (Figure 1A) [22].

Multiple datasets including PONDR (http://www.pondr.com) and DisProt (https://www.disprot.org) have been developed for the prediction of IDPs [23]. The propensity of IDPs to undergo phase separation can also be predicted through bioinformatic tools, although the results can vary considerably [24]. Five distinct methods (SaPS and PdPS [25], catGRANULE [26], PLAAC [27], and PScore [28]) have been popular for this purpose. Comparative computational analyses reveal that ~1036 and 252 proteins from the dicotyledon and monocotyledon model organisms *Arabidopsis thaliana* and rice (*Oryza sativa*), reflecting the top 30% overlapping proteins from the five methods, tend to undergo phase-separation events (Figure 1B,C and Tables S1 and S2 in the supplemental information online). Gene Ontology (GO) analysis implicates these proteins in macromolecule metabolism, mRNA metabolism, and developmental regulation in *A. thaliana*. The proteins are over-represented with molecular functions in DNA binding, protein dimerization, kinase activity, RNA helicase, and translation initiation (Figure 1D). Similarly, in rice, the proteins are enriched in RNA metabolic processes, transcription, and developmental regulation, and have molecular functions in nucleic acid binding as well as pyrophosphatase, hydrolase, and purine NTP-dependent helicase activities (Figure 1E). Notably, many of the computationally predicted phase separation-prone proteins in *A. thaliana* could also be recovered experimentally via biotinylated isoxazole microcrystals that tend to precipitate proteins with random coil polypeptides [29], highlighting the power and complementarity of the two strategies to predict plant phase-separating proteins at a proteome level.

## IDPs participate in transcriptional regulation

The transcription apparatus in eukaryotes is dynamically constituted of RNA polymerases, transcription factors (TFs), mediators, chromatin-remodeling complexes, and numerous epigenetic regulators. Recent studies have shown that a large fraction of the transcription apparatus comprises IDPs, and the resultant condensates of distinct compositions may drive consecutive transcription cycles. The C-terminal domain (CTD) of RNA polymerase II (Pol II) has an IDR that retains multiple heptapeptide YSPTSPS repeats ranging from 26 in yeast to 52 repeats in human [30]. Boehning *et al.* showed that the human and yeast RNA Pol II CTDs possess the capacity to form condensates. Moreover, the length of the CTD influences the stability and dynamics of droplets, as an extended CTD can lead to stronger CTD–CTD interactions and more stable droplets *in vitro* [31]. While the RNA Pol II CTD triggers the formation of various transcription-dependent condensates with distinct components, phosphorylation of the protein alters its conformation in subsequent transcription apparatus assemblies [32]. Unphosphorylated RNA Pol II is recruited to promoters via LLPS of the CTD and the transcription initiation machinery [31]. Subsequently, RNA Pol II is released from transcription initiation condensates with RNA accumulation and CTD phosphorylation, and is relayed to elongation or splicing condensates [32]. In addition, IDRs of TFs can determine their binding specificity to target DNA regions through multivalent determinants distributed within the IDR sequences [33].

In *A. thaliana*, the *FLOWERING LOCUS C* (*FLC*) locus encodes a master transcriptional regulator that controls flowering time [34]. Recent progress has shed light on the importance and mechanism of IDPs in regulating *FLC* transcription [11,12]. Briefly, FLC protein is a MADS box TF that represses the expression of genes in the downstream flowering pathway, leading to delayed flowering [34]. The expression of *FLC* itself is enhanced by the transcriptional factor, FRIGIDA (FRI), an IDP with two coiled-coil motifs [12]. FRI can recruit transcriptional coactivators and histone modifiers such as histone H3 lysine 36 trimethylation (H3K36me3) and H3K4me3 methyltransferases to the *FLC* promoter under warm conditions, leading to activation of *FLC* transcription. However, this process can be reverted by vernalization [35]. While FRI is unstable in a warm environment, prolonged cold exposure results in accumulation of the protein and promotes FRI to form nuclear condensates that depend on its C-terminal IDR and two coiled-coil domains. Importantly, the condensation can stabilize FRI but sequester the protein away from *FLC* promoter, leading to repression of *FLC* expression (Figure 2), a prerequisite for flowering transition after vernalization [12].

Transcription of *FLC* is also regulated by the antisense long non-coding RNA *COOLAIR* from the *FLC* locus [34]. Notably, the RNA-binding protein FCA is an IDP with a prion-like domain and can form condensates. The size and number of the FCA condensates can be further stimulated by another coiled-coil IDP protein, FLL2. This can promote 3′ processing of *COOLAIR* RNA and cause chromatin silencing of *FLC* (Figure 2) [11]. In addition, Zhu *et al.* reported that cold exposure can alter FRI interaction with *COOLAIR* and other partners, and can promote FRI condensation, as shown in root cells and tobacco leaf nuclei. Consequently, the condensate sequesters FRI away from *FLC* promoter and triggers the repression of *FLC* transcription. Deletion of the *COOLAIR* promoter significantly alters the overall size distribution and reduces the number of cold-induced FRI condensates, leading to elevated FRI occupancy at the *FLC* promoter and inefficient repression of *FLC* transcription under cold exposure [12,36].

The abovementioned model has recently been challenged because Zhang *et al.* argue that *COOLAIR* is neither required for FRI condensation nor for *FLC* repression under constant cold exposure [37]. They reported that, following deletion of the promoter cold-responsive *cis* elements, there was no significant change in the size or number of FRI condensates, either before or after cold exposure, between functional *pFRI:FRI-GFP* and the cold-insensitive *pFRI:FRI-GFP*; *ΔCOOLAIR* lines in which the *COOLAIR* promoter was deleted. Furthermore, transcriptional repression of *FLC* following cold exposure was not affected in *pFRI:FRI-GFP*; *ΔCOOLAIR* lines which lack *COOLAIR* expression [37]. Interestingly, Zhu *et al.* revisited the experiments later and insisted on the correctness of their model [36]. The conflicting results might be due to differences in growth conditions and seedling densities in the two laboratories. Clearly, more experiments will be necessary to clarify the regulatory mechanism of the FRI transcriptional condensates in controlling *FLC* locus expression. Given that IDP-triggered condensates are context-specific and condition-sensitive, it will be crucial to directly quantitate how cold exposure might alter the

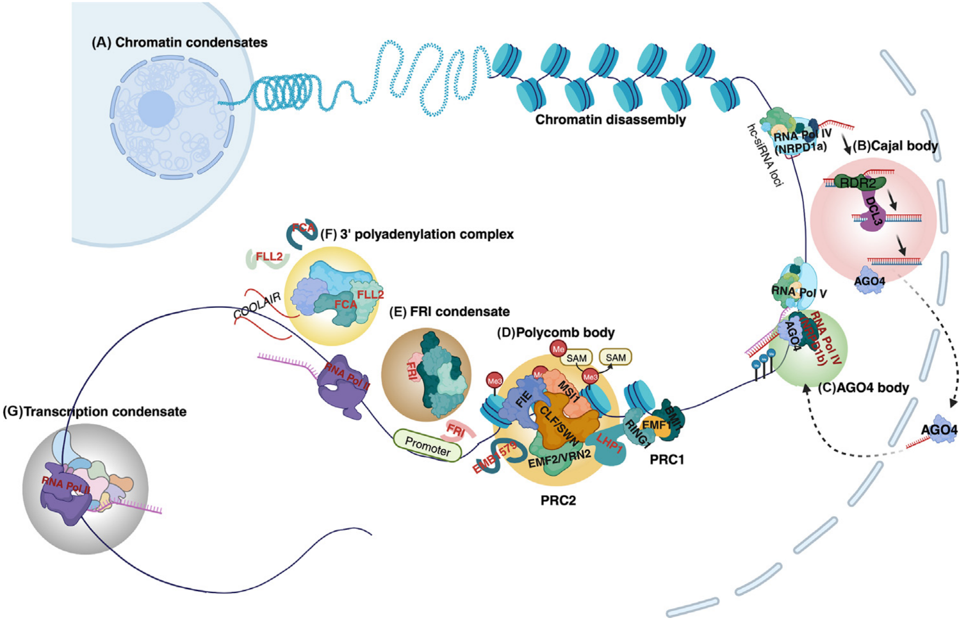


Figure 2. Intrinsically disordered proteins (IDPs) regulate transcription and transcriptional silencing in plants. Chromatin accessibility regulates gene transcription activity. (A) Chromatin compaction is regulated by condensation of histone variant H2B.8 and diverse regulators of histone and DNA modifications. (B) In the 24 nt siRNA-mediated RNA-directed DNA methylation (RdDM) pathway, RDR2 physically interacts with NRPD1 (NUCLEAR RNA POLYMERASE D1) of RNA polymerase IV (Pol IV), and converts RNA Pol IV transcripts into double-stranded (ds)RNAs of ~30 bp. The dsRNAs are further cleaved into 24 nt siRNAs by DCL3, and loaded onto AGO4 to target gene loci via base-pairing with Pol V transcripts and recruiting DNA methyltransferase DRM2 and histone-modifying enzymes to facilitate silencing of the target loci. In the process, RDR2, DCL3, AGO4 colocalize with the Cajal bodies, which are sites for efficient ribonucleoprotein (RNP) complex assembly. (C) The 24 nt siRNA-loaded AGO4 colocalizes with NRPD1b of RNA Pol IV in so-called AGO4–NRPD1b bodies which are proposed to facilitate silencing of the target loci by base-pairing to Pol V-derived non-coding RNA transcripts and by recruiting DRM2 and histone-modifying enzymes. (D) LHP1 and EMB1579 are IDPs that can drive the assembly of PRC1/PRC2 into Polycomb group (PcG) bodies which trigger histone methylation (Me) and chromatin compaction. (E,F) Transcription of *FLC* is regulated not only by histone and DNA modifications but also by FRI condensates (E) and 3′ polyadenylation condensates (F). FRI itself is an IDP and undergoes phase separation, which in turn sequesters FRI from the *FLC* promoter to repress FRI activation of *FLC* transcription. FRI condensation is regulated by the *FLC* antisense transcript *COOLAIR*. FLL2 can act as an IDP to trigger the IDP FCA to undergo liquid–liquid phase separation (LLPS) and assemble into 3′ polyadenylation condensates, thereby regulating the 3′ polyadenylation and stability of *COOLAIR*. Intriguingly, the RNA Pol II C-terminal domain (CTD) has an IDR and its phosphorylation status triggers the dynamic exchange of transcriptional condensates (G), adding to the complexity of transcriptional regulation. Abbreviations: AGO4, ARGONAUTE4; BMI1, also known as DRIP1, DREB2A-INTERACTING PROTEIN 1; CLF, CURLY LEAF; DCL3, DICER-LIKE 3; EMB1579, EMBRYO DEFECTIVE 1579; EMF1, EMBRYONIC FLOWER 1; EMF2, EMBRYONIC FLOWER 2; FCA, FLOWERING CONTROL LOCUS A; FIE, FERTILIZATION INDEPENDENT ENDOSPERM; FLL2, FLX-LIKE 2; FRI, FRIGIDA; hc-siRNA, heterochromatic small interfering RNA; LHP1, LIKE HETEROCHROMATIN PROTEIN 1; MSI1, MULTICOPY SUPRESSOR OF IRA1; PRC1, Polycomb repressive complex 1; PRC2, Polycomb repressive complex 2; RDR2, RNA-DEPENDENT RNA POLYMERASE 2; SAM, *S*-adenosyl methionine; SWN, SWINGER; VRN2, VERNALIZATION 2.

interaction of FRI with *COOLAIR* within FRI condensates and the consequences thereof. In addition, it will be crucial to analyze the underlying mechanism of FRI condensation in aerial parts under physiological conditions, even though this is technically more challenging than in root tip cells.

## IDPs regulate transcriptional silencing

Heterochromatin compartmentalization acts to repress transcription, and IDPs are involved in chromatin condensation in both animals and plants. Chromatin compaction and accessibility are regulated by histone variants and modifications [38], DNA methylation [39], and chromatin remodeling (Figure 2) [40]. In a study to explore the mechanism of sperm chromatin condensation in flowering plants, Buttress *et al.* reported that the N-terminal IDR of the *A. thaliana* histone variant H2B.8 mediates its phase separation which drives chromatin compaction of sperm nuclei. Whereas dysfunction of H2B.8 results in enlarged sperm nuclei with decondensed chromatin, ectopic expression of H2B.8 in somatic cells triggers smaller nuclei with compacted chromatin. Intriguingly, the authors found that H2B.8 condensates mainly associate with and compact transcriptionally inactive AT-rich euchromatin, thus facilitating nuclear compaction while retaining active gene expression [41].

Chromatin compaction is also regulated by histone methylation which can be mediated through the activities of the evolutionally conserved Polycomb group (PcG) complexes, including Polycomb repressive complex 1 (PRC1) and PRC2, in both animals and plants (Figure 2). PRC2 catalyzes mono-, di-, and tri-methylation of histone lysine 27 (H3K27me1–3) that are hallmarks of silent chromatin. Canonical PRC1 facilitates histone H2A ubiquitination (H2AK119/121ub) and chromatin compaction, and this process relies on PRC2-deposited H3K27me3 marks. By contrast, the noncanonical PRC1 promotes these processes independently of PRC2 activity [42,43]. New insights are gained on the activity and recruitment of Polycomb complex, with the identification of IDPs involved in chromatin condensation [44]. Mammalian CHROMOBOX 2 (CBX2) and its homologs may act as IDPs that interact and condense with PRC1, thus facilitating chromatin compaction [45]. LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) also contains an intrinsically disordered hinge region. This IDR can bind to RNA *in vitro* and is required for the formation of subnuclear foci *in vivo* that are reminiscent of Polycomb bodies. Notably, disruption of the IDR abolishes the function of LHP1 in H3K27me3 maintenance [46]. The IDP EMBRYO DEFECTIVE 1579 (EMB1579) undergoes LLPS to recruit DNA DAMAGE-BINDING PROTEIN 1 (DDB1), CULLIN 4 (CUL4), and MULTIPLE SUPPRESSOR OF IRA 4 (MSI4) to promote PRC2-mediated H3K27me3 modification at the *FLC* locus [47]. HETEROCHROMATIN PROTEIN 1 (HP1) recognizes H3K9me3, another hallmark of heterochromatin, and self-oligomerizes to drive chromatin condensation and compaction [48]. In plants, the HP1 functional analog and H3K9me2 reader protein, AGENET DOMAIN-CONTAINING PROTEIN 1 (ADCP1), fulfils a similar function to drive heterochromatin formation [49]. Chromatin remodelers utilize ATP hydrolysis to disrupt electrostatic interactions between the positively charged histones and negatively charged DNA [50]. Several members of the SWI/SNF remolding components, namely BRM (BRAHMA), SWI3C (SWITCH/SUCROSE NONFERMENTING 3C), AN3 (ARATHNICTABA 3), and SWP73B (SWI/SNF-ASSOCIATED PROTEIN 73B), are predicted to contain PrLDs [40], and might assume their roles as IDPs in chromatin remodelling.

Like histone methylation, DNA methylation is also strongly associated with heterochromatin compaction and transcriptional silencing [51]. In *A. thaliana*, whereas heterochromatic DNA methylation is facilitated by the chromatin remodelers DDM1 (DECREASED DNA METHYLATION 1) and CMT2 (CHROMOMETHYLASE 2) [52], interspersed DNA methylation within transposon- or repeat-containing euchromatic regions is mediated by an RNA-directed DNA methylation (RdDM) pathway [53]. The RdDM pathway is initiated by the production of short transcripts by DNA-dependent RNA polymerase IV from transposable elements (TEs) and repetitive elements. The nascent RNAs are converted into double-stranded (ds)RNAs by RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) in *A. thaliana* [54,55], and these are further processed by DICER-LIKE 3 (DCL3) to produce 24 nt siRNAs [56,57]. The 24 nt siRNA-loaded ARGONAUTE 4 (AGO4) then directs RNA polymerase V to the correct loci and recruits DNA methyltransferases and additional epigenetic components to initiate DNA methylation [58,59]. In addition, MICRORCHIDIA (MORC) family proteins are further recruited and interact with the DNA methylation-binding protein

SUVH2/9, as well as with chromatin remodelers DRD1, DMS3, and RDM1, at RdDM loci to facilitate efficient establishment of RdDM and transcriptional silencing [60].

The biogenesis and action of the 24 nt siRNAs are linked to discrete nuclear speckles, called Cajal bodies, whose condensation relies on COILIN protein [61]. COILIN comprises a N-terminal globular domain, a CTD with a Tudor-like structure, and a highly disordered central region that contains a nuclear localization signal (NLS) and an RG (Arg/Gly) box that is subject to Arg dimethylation. Unexpectedly, the N-terminal domain of COILIN mediates its multivalent interaction with partners and condensation of Cajal bodies [62]. The assembly and activity of Cajal bodies are modulated by phosphorylation [63] and symmetrical dimethylation of Arg residues [64] in COILIN protein. Several other proteins, for instance AGO4, in the 24 nt siRNA pathway are predicted to be IDPs and have been reported to colocalize with Cajal bodies [65,66]. In fact, a proportion of AGO4 localizes to AGO4/NRPD1b bodies, although the majority of AGO4 localizes in Cajal bodies (Figure 2). Notably, disruption of COILIN affects the integrity of Cajal bodies and AGO4 colocalization [67]. Both NRPD1b and AGO4 contain PrLDs [68], but how they condense and the role of condensation in RdDM remain to be studied.

## IDPs regulate PTGS

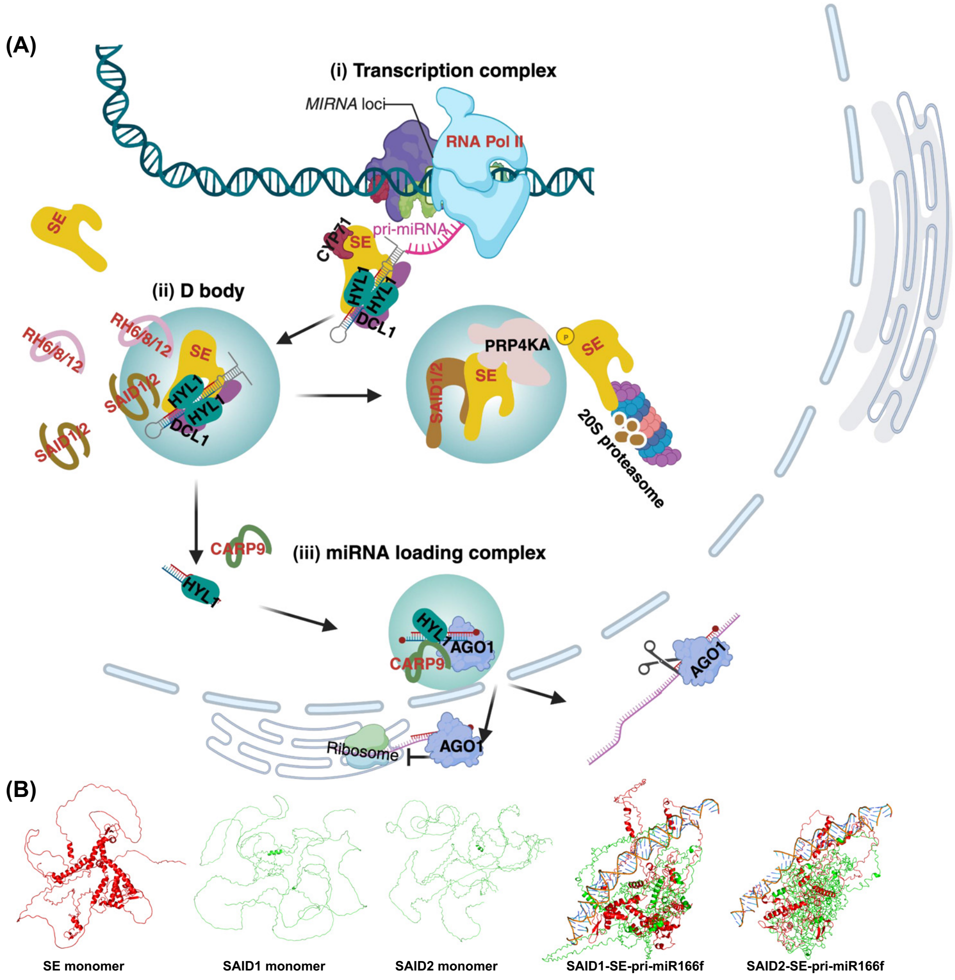
PTGS is primarily guided by small RNAs (sRNAs) exemplified by miRNAs and small interfering RNAs (siRNAs). miRNAs and siRNAs are loaded into AGO proteins to form the RNA-induced silencing complex (RISC) that either cleaves target mRNAs or triggers translational inhibition followed by RNA decay [69]. These small noncoding RNAs are universally present in eukaryotic organisms, and incongruity of sRNAs and dysregulation of the sRNA machinery can perturb normal development and fitness in both plants and animals. Hence, sRNA pathways afford promising targets for the development of therapeutics and for agronomic improvement [70–72].

### IDPs participate in miRNA biogenesis

The biogenesis of miRNAs is initiated by the transcription of the primary substrates of miRNAs (pri-miRNAs) by RNA Pol II. In metazoans, pri-miRNAs are processed by the nuclear Microprocessor complex that primarily consists of DROSHA and two DGCR8 proteins that generate the 60–75 nt miRNA precursors (pre-miRNAs) [73,74]. Pre-miRNAs are further cleaved by the RNase III enzyme DICER1 to yield miRNA/* duplexes [75]. Interestingly, Cui *et al.* recently reported that the animal DGCR8 is an IDP and its elevated accumulation can cause imbalanced DGCR8/DROSHA protein stoichiometry, which in turn can drive irreversible aggregation of Microprocessor and impede pri-miRNA processing [76].

In plants, pri-miRNAs are sequentially processed into miRNA/* by the core Microprocessor complex composed of the RNase III enzyme DCL1, the multifunctional zinc-finger (ZF) protein SERRATE (SE), and the dsRNA-binding protein HYL1 [77]. Once ectopically expressed, these components tend to be enriched in nuclear speckles, often called dicing bodies (D-bodies) (Figure 3A) [78]. Notably, some auxiliary factors that affect miRNA production often colocalize with the core Microprocessor complex when transiently overexpressed, whereas their mutants display compromised miRNA production accompanied by alteration of D-body number or size [79,80]. Thus, one prevailing view is that D-body is the main factory site for miRNA production. However, increasing evidence has documented that initial processing of pri-miRNA occurs cotranscriptionally, and subsequent processing may take place in D-bodies or even in structures in the nucleoplasm that are not detectable at the resolution of classic microscopy [81,82].

The core Microprocessor component SE is an IDP that interacts with multiple partners, and has multiple functions in RNA metabolism including miRNA production [83] and pri-miRNA decay



Trends in Genetics

Figure 3. Intrinsically disordered proteins (IDPs) in the regulation of miRNA pathways in plants. (A) The core of the Microprocessor complex is composed of DCL1, HYL, and SE. As an IDP, SE plays multifaceted roles in the assembly of diverse macrocomplexes, especially in the recruitment of HYL1, DCL1, and other cofactors of the Microprocessor complex. (i) Primary substrates of miRNAs (pri-miRNAs) transcribed by RNA polymerase II (Pol II) are processed cotranscriptionally by the Microprocessor complex. (ii) SE phase separation triggers the recruitment of HYL1, DCL1, and other cofactors into the Microprocessor complex, leading to efficient production of miRNAs. The phase-separation capacity of SE may be regulated by phosphorylation (P) catalyzed by PRP4KA, and phosphorylated SE is subject to 20S proteasome cleavage. Intriguingly, a family of IDPs, SAID1/2, are recruited by SE to undergo phase separation, but they can sequester pri-miRNAs from SE, impede Microprocessor activity, and trigger SE phosphorylation and degradation, thus maintaining SE homeostasis and preventing miRNA overproduction. The RNA helicases RH6/8/12 are also IDPs that act to modulate SE phase separation. The peptidyl-prolyl isomerase CYP71, which catalyzes *cis–trans* isomerization of SE, also acts to promote SE condensation and the production of some miRNAs. (iii) After production, mature miRNAs are associated with HYL1 and subsequently routed into a hypothesized loading complex recruited by the IDP CARP9. As an IDP, CARP9 could recruit and interact with HYL1 and AGO1, thus facilitating miRNA loading into an RNA-induced silencing complex (RISC) for mRNA cleavage or translation inhibition. (B) AlphaFold2 modeling predicts conformational changes in SE and SAID1/2 upon binding to pri-miRNAs. SAID1 and SAID2 monomers are less-structured IDPs, and upon interaction with the IDP SE and pri-miRNAs, all proteins are predicted to change their conformation and become structured.

[84]. Excessive accumulation of SE could lead to the formation of intermediate Microprocessor complexes that are unable to efficiently produce miRNAs [83]. However, these intermediate complexes appear to remain soluble, and their mode of action may differ from that of animal DGCR8/Drosha aggregates [76]. SE has generally been believed to act as a scaffold to recruit the core

processing machinery including DCL1/HYL1 to the proper RNA substrates, or vice versa, to generate miRNAs [77,85]. SE has also been recently reported to mediate LLPS to enhance the processing activity of Microprocessor (Figure 3A) [86]. Whereas SE harbors two major IDRs located in its N-terminal and C-terminal tails, the N-terminal IDR confers its propensity to form condensates [86].

### PTMs of IDPs control the homeostasis and activity of the Microprocessor complex

Eukaryotic proteins are mostly degraded through a ubiquitin-dependent 26S proteasome pathway. However, IDPs are prone to degradation independently of the process. Indeed, SE protein can undergo ATP- and ubiquitin-independent degradation, clearly reflecting to its inherent feature as an IDP (Figure 3A). *In vivo*, excess SE impedes the assembly of macromolecular complexes and/or is not shielded from the complexes. Consequently, the protein is readily scavenged by PAG1 (PROTEASOME A SUBUNIT 1) and channeled to the 20S proteasome for destruction [83]. Subsequent work by Wang *et al.* revealed that the degradome signal for triggering SE degradation is protein phosphorylation via PRP4KA (PRE-MRNA PROCESSING 4 KINASE A) and its homologs. Hypophosphorylated SE variants display elevated binding affinity for HYL1 and are relatively less vulnerable to the activity of 20S proteasome than hyperphosphorylated SE counterparts [87]. Phosphorylation of SE represents a regulatory mechanism at the post-translational level to rapidly clear excess SE and to maintain homeostasis of SE *in vivo* to secure its proper functions (Figure 3A) [87]. Because PTMs of IDPs can alter protein conformations, and phase behaviors [88], it could be hypothesized that phosphorylation might impact on the LLPS behavior of SE and thus compromise the assembly of Microprocessor complexes.

Prolyl isomerization can contribute to the regulation of phase separation. A recent report showed that the recruitment of peptidyl prolyl isomerase (PPIA) into IDP droplets triggers condensate dissolution via *cis/trans* interconversion of prolines in TAU and/or the PRO–ARG DIPEPTIDE REPEAT PROTEIN PR20 [89]. CYP71 is a PPIA that could interact with Microprocessor components and promote droplet formation (Figure 3A). Furthermore, this relies on CYP71 PPIase activity *in vitro*. Consistently, HYL1–YFP foci that serve as a proxy for D-bodies disappear in *cyp71* mutants [90]. Notwithstanding, the precise proline residue(s) of SE that are targeted by CYP71 remain to be pinpointed, and how proline isomerization failure might impact on SE function *in vivo* awaits future clarification.

The phase behavior and function of SE protein can be also regulated by other IDPs. *A. thaliana* encodes 11 hypothetical DENTIN SIALOPHOSPHOPROTEIN-LIKE (DSPPL) proteins, and at least three of the proteins are IDPs, namely SAID1/2 (SE-ASSOCIATED INHIBITORS WITH DWARFISM APPEARANCE IN THE MUTANTS 1/2) and CARP9 (CONSTITUTIVE ALTERATIONS IN THE SMALL RNA PATHWAYS 9). Interestingly, the three IDPs all act in miRNA pathway in plants, but with different mechanisms [91,92]. SAID1 and 2 act as negative regulators of SE stability and miRNA biogenesis, as *said1*; *said2* loss-of-function mutants display increased assembly of Microprocessor and elevated accumulation of miRNAs. Mechanistically, SAID1 and 2 promote PRP4KA-mediated phosphorylation of SE, causing its degradation *in vivo*. In addition, SAID1/2 display strong binding affinity for pri-miRNAs and can sequester them from SE, leading to the direct inhibition of pri-miRNA processing by Microprocessor *in vitro*. Importantly, SAID1/2-mediated dual inhibition of miRNA production appears to take place in SE-centered condensates *in vivo* (Figure 3A) [91]. This notion is highlighted by the observation that SAID1/2 display clear nuclear puncta, suggestive of condensation in wild-type plants, whereas the condensates disappear in *se* mutants. Because SAID1/2 and SE are all IDPs, it is possible that they might mutually affect biochemical features and functions for each other through co-condensation [91]. It is also possible that the IDPs might undergo conformational switches into structured proteins upon their association or binding to pri-miRNAs. The latter hypothesis appears to be supported by artificial intelligence (AI) modeling (Figure 3B) [93].

## RNA helicases regulate condensation of D-bodies

Recently, the DEAD-box RNA helicases RH6, RH8, and RH12 IDPs have been identified as HYL partners by immunoprecipitation and mass spectrometry (IP-MS) of the HYL1 interactome [94]. The purified RH6, RH8, and RH12 proteins can undergo phase separation in an RNA- and ATP-dependent manner *in vitro*. The proteins can also physically interact and colocalize with Microprocessor components in nuclear speckles resembling D-bodies *in planta*. Whereas HYL–YFP foci are eliminated in the knockdown lines of three helicases, the addition of RH12 elevates the size and number of SE droplets *in vitro*, indicative of the critical roles of these helicases in promoting SE phase separation and D-body assembly (Figure 3A) [94].

RH6, RH8, and RH12 also contribute to the assembly and subcellular dynamics of processing bodies (P-bodies) and stress granules (SGs) *in vivo* [95]. Because the homologs of these helicases in yeast have long been known to drive the formation and turnover of P-bodies and SGs through LLPS [96], RH6, RH8, and RH12 might mechanistically behave similarly to regulate the two RNP complexes in *A. thaliana*. In addition, RH6, RH8 and RH12 can be also translocated into another cellular compartmentation niche on the chloroplast periphery, called virus bodies because these are where virus replication actively takes place [94]. Thus, these helicases seem to concurrently regulate the assembly and disassembly of multiple macromolecular complexes in both nuclei and cytoplasm to adapt to physiological changes. Whether all these functions are fulfilled via modulation of the condensation of the complexes remains unclear; however, the regulation is likely to involve RNA substrates given that the RH proteins are RNA helicases.

The SWI2/SNF2 ATPase CHR2/BRM canonically acts to remodel chromatin to regulate gene transcription. The protein can also act as an RNA helicase that interacts with SE to remodel the secondary structures of pri-miRNAs, thus impeding Microprocessor processing efficiency [97]. Notably, CHR2 protein itself contains long stretch of low-complexity regions. It is possible that CHR2 and SE might form cocondensates to orchestrate transcription and cotranscriptional processing of transcripts including pri-miRNAs. In addition, SE has been reported to enhance the enzymatic activities of histone methyltransferases ATXR5 (ARABIDOPSIS TRITHORAX-RELATED PROTEIN 5) and ATXR6 *in vitro* and H3K27me1 accumulation *in vivo* [98]. A retrospective thought of this process is that the promotive function of SE might also be granted through its tendency to form condensates.

## IDPs in the regulation of miRNA loading

miRNAs are loaded onto AGO proteins to form RISCs. This also involves IDPs and phase separation. In metazoans, the scaffold protein TNRC6 (TRINUCLEOTIDE REPEAT-CONTAINING 6) sets up a bridge between AGO2 and the downstream deadenylase complex through multivalent interactions between the Gly/Trp (GW)-rich domain of TNRC6B, which is located in the IDR, and three evenly spaced Trp-binding pockets in the AGO2 PIWI domain, facilitating subsequent translational repression and decapping [99]. Another IDP, YBX1, can form condensates to selectively recruit and sort miR-223 into exosomes to mediate intercellular communication [100].

Tomassi *et al.* identified CARP9, a large IDP, as a new partner of the miRNA pathway in *A. thaliana*. Mutations in *CARP9* produce morphological alterations, a mild reduction in miRNA accumulation, and impaired RNA silencing [92]. *CARP9* mutants also present low accumulation of AGO1 protein and apparent depletion of miRNAs in the cytoplasm. Notably, CARP9 interacts with HYL1, but not with other Microprocessor components. CARP9 is colocalized in SE/HYL1-containing nuclear speckles, but not with DCL1 protein. Intriguingly, CARP9 also associates with AGO1 and HEAT

SHOCK PROTEIN 90 (HSP90), as well as with mature miRNAs. This prompted the authors to propose that CARP9 might utilize its IDP feature to act as a scaffold protein to connect HYL1 to AGO1, thus promoting miRNA loading into the RISC before its export to the cytosol (Figure 3A) [92].

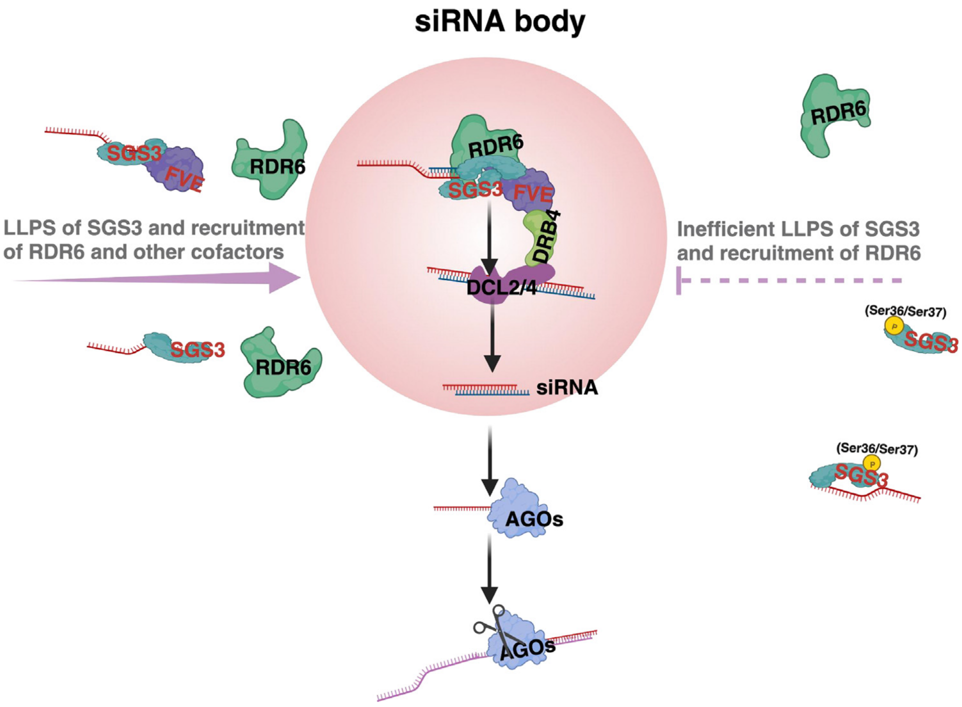
### IDPs in the regulation of 21–22 nt siRNA biogenesis in PTGS

Different from miRNAs, siRNA production in plants starts with the conversion of single-stranded (ss)RNA substrate to dsRNA via a complex composed of RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) and SUPPRESSOR OF GENE SILENCING 3 (SGS3) [101,102]. In turn, dsRNAs are processed by the DCL2/4-DRB4 (DOUBLE-STRANDED RNA-BINDING PROTEIN 4) complex to generate 21–22 nt siRNAs [103,104]. For the production of endogenous *trans*-acting sRNAs or phased siRNAs, SGS3 interacts directly with the 3′ end of miRNAs in AGO1/AGO7 RISCs that also bind to the processed target transcripts, thus enhancing miRNA-mediated ribosome stalling and the amplification of secondary siRNAs [105].

SGS3 is an IDP and has been speculated to undergo LLPS to form siRNA bodies where PTGS-destined siRNAs are produced (Figure 4). SGS3 contains two putative PrLDs located in its N and C termini. The protein also contains ZF, rice gene X and SGS3 (XS), and coiled-coil (CC) domains. The XS domain is involved in RNA binding, whereas the CC domain is involved in protein–protein interactions. Indeed, Kim *et al.* observed that full-length SGS3 protein is distributed into cytoplasmic foci, whereas PrLD-deleted SGS3 protein is in the nucleus. Consistently, recombinant GFP–SGS3 protein forms liquid droplets *in vitro*, but the PrLD-deleted SGS3 does not exhibit such phase-separating behavior [106]. Tan *et al.* carefully assessed the contributions of individual domains to SGS3 condensation. It turns out that the N-terminal PrLD domain is indispensable for SGS3 phase separation, whereas the XS domain is required for the fluid properties of SGS condensates. Notably, the CC domain of SGS3 is responsible for the recruitment of RDR6 into SGS3 condensates [107]. Interestingly, Tan *et al.* also reported that both ssRNA and dsRNA can be incorporated into SGS3 condensates. They showed that SGS3 has a higher affinity for ssRNA than for dsRNA [107]. However, this finding differs from the observations of two other groups [108,109]. The contradiction might be explained by the fact that Tan *et al.* used a dual tagged recombinant SGS3 protein (GST–SGS3–GFP) that might interfere with the conformation of SGS3 in the assay, whereas the other teams used tag-free SGS3 in their electrophoretic mobility shift assays. Under physiological conditions, SGS3 protein might bind to ssRNA, but this binding likely entails additional RNA-binding protein(s) exemplified by FVE (FLOWERING LOCUS VE, discussed in the following section).

SGS3 is predicted to contain multiple phosphorylation sites in its N-terminal LCD [110]. Han *et al.* selected two such sites (Ser36 and Ser37) for mutagenesis assays. They reported that the hypophophorylated SGS3 variant (Ser to Ala) undergoes droplet formation whereas the phosphomimetic (Ser to Asp) variant compromises phase separation in a fluorescence recovery assay (Figure 4). Importantly, SGS3 phase separation is critical for plant antiviral immunity and fertility. For instance, the phosphomimetic (Ser to Asp) variant lacks antiviral defense capability, reminiscent of the *sgs3* mutant, whereas the hypophosphorylated variant can restore the antiviral ability in the *sgs3* mutant to the wild-type level [111]. However, how the phase separation of SGS3 regulates antiviral immunity and fertility in plants awaits future clarification. Furthermore, whether SGS3 is phosphorylated at the two sites *in planta* and the identity of the kinase(s) that catalyze the phosphorylation of SGS3 remain unknown.

The behavior and function of SGS3 can be also modulated by cellular factors. Sun *et al.* recently performed a forward genetic screen aiming to recover mutants compromised in PTGS of transgenes. Surprisingly, a new allele of *FVE* was identified. FVE is a well-known component that regulates



Trends in Genetics

Figure 4. A proposed model for SGS3 condensate in plant 21–22 nt siRNA biogenesis in post-transcriptional gene silencing (PTGS). SGS3 is an intrinsically disordered protein (IDP) that can undergo liquid–liquid phase separation (LLPS). RDR6, single-stranded (ss)RNAs, and other siRNA processing components are recruited by SGS3 to form 21–22 nt siRNA bodies in the cytoplasm. Phosphorylation (P) of SGS3 (at Ser36 and Ser37) impairs its phase separation and siRNA-body assembly. The WD40 protein FVE triggers SGS3 dimerization and binding to ssRNA substrates. Subsequently, ssRNAs are converted to double-stranded (ds)RNAs by RDR6, and the dsRNAs are relayed to DCL2/4–DRB4 complexes for efficient cleavage into siRNA by DCL2/4, which is hypothesized to occur in the siRNA bodies. The siRNAs are subsequently loaded into corresponding AGO proteins where they direct the cleavage and silencing of their RNA targets.

flowering time and cold response at an epigenetic level in *A. thaliana* [112,113]. However, Sun observed that FVE could be localized in the nucleus and cytoplasm, while the cytoplasmic portion of FVE could fully rescue the defect of the *fve* mutant in PTGS, but not the defect in the flowering time. These results indicate that FVE might have a noncanonical role in cytoplasmic RNA silencing. Further mechanistic studies show that FVE targets SGS3 and promotes its homodimerization, which is prerequisite for its function. FVE can also bind to ssRNAs and dsRNAs with moderate affinities. These features might promote the launching of SGS3, which itself has a very low affinity for ssRNAs, onto the ssRNA substrates. FVE binding to RNA substrates can also promote the routing of SGS3/RNA to DRB4/DCL2/4 complexes for further processing. Indeed, FVE can also interact with DRB4 and promote DRB4/DCL2/4 activity in generating siRNAs from dsRNA substrates. Thus, the very epigenetic factor FVE can synchronize the RDR6/SGS3–DRB4/DCL2/4 channel to promote siRNA production and transgene silencing (Figure 4). This notwithstanding, as FVE harbors six tandemly repeated WD40 domains and a small portion of low-complexity regions. Whether FVE contributes to SGS3 dimerization and association with RNA substrates through altering its LLPS, or vice versa, would be an interesting topic for future studies [109].

## Concluding remarks and future perspectives

Recent studies have witnessed the increasingly important roles of IDPs in diverse biological processes in plants and animals. Many IDPs can serve as scaffolds for various RNP complexes such

## Outstanding questions

Systemic identification of the *bona fide* IDPs in the plant proteome, and of their substrates or partners related to gene silencing and other biological functions, remains a major issue. Numerous IDPs serve as scaffolds or hubs to generate various membraneless organelles such as D-bodies, P-bodies, and siRNA bodies. It would be interesting to know how many other types of bodies are present *in vivo*, the identities of the IDPs that contribute to their formation, and their functions.

Behavioral characterization of IDPs will require comprehensive and quantitative studies *in vitro* and *in vivo*, and will need to be validated by multiple independent lines of evidence. Such advanced studies would not only provide a complete picture of the features of the IDPs but also enable us to compare their strengths or propensities to form condensates in different physiological contexts.

Studies on how PTMs alter IDP behaviors, and possibly their condensation, are likely to uncover new regulatory layers that control biological processes. Many IDPs may not harbor specific motifs or domains in their primary sequences, and might evolve neofunctions when bound to their substrates or partners. Furthermore, IDPs might act in multiple niches of condensates or complexes, and precise characterization of their specific functions in these niches will entail new ideas and approaches.

Biological condensation can entail positive or negative regulation of the proteins involved. In addition, IDPs could display diverse molecular mechanisms which are not necessarily implicated in LLPS/biological condensation. Conversely, LLPS/condensation is not exclusively linked to IDPs because some non-IDP proteins can form condensates under specific physiological conditions. Future investigations should embrace objective views, creative ideas, and novel approaches to uncover potential new paradigm mechanisms.

Finally, manipulation of the IDP-regulated processes to control gene silencing offers new strategies to control biological processes for biotechnological applications and crop improvement.

as Cajal bodies, D-bodies, siRNA bodies, P-bodies, and SGs, among others, whereas other IDPs can act as regulatory hubs for molecular functions. Numerous IDPs tend to form condensates via phase separation to orchestrate transcription regulation and gene silencing, and among the proteins are FRI, SE, and SGS3 that are the landmarks of epigenetic regulation, miRNA, and siRNA pathways, respectively (Table S3 in the supplemental information online). These proteins can form condensates to recruit partners and substrates, thus impacting on enzymatic or processing events, or changing downstream regulatory cascades. The condensation behaviors of these proteins themselves can also be modulated by additional IDPs, substrates, or even PTMs. These might be the tip of the iceberg for the IDP field, and new regulatory modes of action are being revealed (see Outstanding questions).

This notwithstanding, hundreds or even thousands of proteins have been predicted to be IDPs according to different computational analyses, one outstanding issue would be how we can separate the wheat from the chaff and pinpoint the *bona fide* IDPs among the hypothetic candidates. Alphafold2 might be a powerful tool to complement current computational predictions to further screen potential IDPs *in silico* [93]. A related question would be to precisely decode the compositions and interactomes of condensates in physiological contexts. Proximity labeling-based MS would greatly advance the field in this objective [114].

Second, although many of the identified IDPs appear to undergo LLPS or have been interpreted to participate in condensation, the inference that all IDPs have the potential to form condensates is probably an oversimplification. Behavioral characterization of IDPs entails comprehensive studies *in vitro* and *in vivo* and needs to be validated by multiple independent lines of evidence. A combination of *in vitro* quantitative biochemistry and biophysical studies with *in vivo* investigation via cell biology and genetics could provide a more complete picture of the candidate proteins. These include *in vitro* thermodynamic parameters of enthalpy and entropy, surface tension parameters of condensate viscosity, inverse capillary velocity, molecular diffusion, and binding affinity, among others. In parallel, integration of fluorescence recovery after photobleaching (FRAP) [115], single-photon microscopy [116], single-molecule Förster resonance energy transfer (FRET) [1], and proximity labeling [117,118] would be powerful to decipher the propensities and organizing principles of condensates *in vivo*. Multiple stringent criteria and integration of advanced methodologies to dissect the composition and properties of biomolecular condensates quantitatively are greatly needed for our understanding of the biological functions of protein condensation and the underlying mechanisms. However, results from *in vitro* reconstitution systems frequently do not perfectly reiterate *in vivo* observations under different experimental or biological conditions. In these scenarios, physiological relevance is a predominant factor to reconcile inconsistencies. For instance, whether mutagenesis to disrupt condensation causes functional defects of the proteins *in vivo*, and whether the process can be rescued by exchange with known LCDs, will be crucial to determine whether condensation is a central feature of a particular protein.

Third, IDPs, like other folded proteins, are subject to various types of PTMs in response to physiological needs. Studies on how the PTMs alter the behaviors and fates of IDPs, including protein condensation, would be likely to uncover new regulatory layers in the control of biological processes. Several proteomic datasets including phosphorylation are available for the model organisms, and these will be invaluable resources for studying the PTMs of IDPs.

Fourth, numerous IDPs are crucial components of RNPs. Many IDPs may not be RNA-binding proteins or lack canonical RNA recognition motifs in their primary amino acid sequences. However, some IDPs exemplified by FVE and SAID1/2 appear to have strong affinity for RNA substrates [91,109]. Whether and how these IDPs have evolved neofunctions in RNA processing or metabolism awaits future in-depth investigation.

Finally, given that several IDPs are known to be master players in epigenetic and PTGS pathways, manipulation of IDP-regulated processes to control gene-silencing pathways offers new strategies to manipulate biological processes for biotechnological applications and crop improvement. Future efforts in this direction are certain to be rewarding.

## Acknowledgments

We apologize to colleagues whose work could not be cited owing to space limitations. We thank members of the laboratory of X.Z. for helpful discussions on the manuscript. Figures were created with BioRender. This work was supported by the National Institutes of Health (NIH; grant R35 GM151976) and the National Science Foundation (NSF; grant MCB 2139857) to X.Z.

## Declaration of interests

The authors declare no conflicts of interests.

## Supplemental information

Supplemental information associated with this article can be found, in the online version, at https://doi.org/10.1016/j.tig.2023.12.009.

## References

1. Metskas, L.A. and Rhoades, E. (2020) Single-molecule FRET of intrinsically disordered proteins. *Annu. Rev. Phys. Chem.* 71, 391–414
2. Dignon, G.L. *et al.* (2020) Biomolecular phase separation: From molecular driving forces to macroscopic properties. *Annu. Rev. Phys. Chem.* 71, 53–75
3. Zarin, T. *et al.* (2019) Proteome-wide signatures of function in highly diverged intrinsically disordered regions. *eLife* 8, e46883
4. Cermakova, K. and Hodges, H.C. (2023) Interaction modules that impart specificity to disordered protein. *Trends Biochem. Sci.* 48, 477–490
5. Cohan, M.C. and Pappu, R.V. (2020) Making the case for disordered proteins and biomolecular condensates in bacteria. *Trends Biochem. Sci.* 45, 668–680
6. Biesaga, M. *et al.* (2021) Intrinsically disordered proteins and biomolecular condensates as drug targets. *Curr. Opin. Chem. Biol.* 62, 90–100
7. Mitrea, D.M. *et al.* (2022) Modulating biomolecular condensates: a novel approach to drug discovery. *Nat. Rev. Drug Discov.* 21, 841–862
8. Powers, S.K. *et al.* (2019) Nucleo-cytoplasmic partitioning of ARF proteins controls auxin responses in *Arabidopsis thaliana*. *Mol. Cell* 76, 177–190
9. Dorone, Y. *et al.* (2021) A prion-like protein regulator of seed germination undergoes hydration-dependent phase separation. *Cell* 184, 4284–4298
10. Ouyang, M. *et al.* (2020) Liquid–liquid phase transition drives intra-chloroplast cargo sorting. *Cell* 180, 1144–1159
11. Fang, X. *et al.* (2019) *Arabidopsis* FLL2 promotes liquid–liquid phase separation of polyadenylation complexes. *Nature* 569, 265–269
12. Zhu, P. *et al.* (2021) Cold-induced *Arabidopsis* FRIGIDA nuclear condensates for *FLC* repression. *Nature* 599, 657–661
13. Zavaliev, R. *et al.* (2020) Formation of NPR1 condensates promotes cell survival during the plant immune response. *Cell* 182, 1093–1108
14. Jung, J.H. *et al.* (2020) A prion-like domain in ELF3 functions as a thermosensor in *Arabidopsis*. *Nature* 585, 256–260
15. Chen, D. *et al.* (2022) Integration of light and temperature sensing by liquid–liquid phase separation of phytochrome B. *Mol. Cell* 82, 3015–3029
16. Kulkarni, P. *et al.* (2022) Intrinsically disordered proteins: critical components of the wetware. *Chem. Rev.* 122, 6614–6633
17. Oldfield, C.J. and Dunker, A.K. (2014) Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* 83, 553–584
18. Alberti, S. *et al.* (2009) A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell* 137, 146–158
19. Sugase, K. *et al.* (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447, 1021–1025
20. Wright, P.E. and Dyson, H.J. (2009) Linking folding and binding. *Curr. Opin. Struct. Biol.* 19, 31–38
21. Gomes, E. and Shorter, J. (2019) The molecular language of membraneless organelles. *J. Biol. Chem.* 294, 7115–7127
22. Hirose, T. *et al.* (2023) A guide to membraneless organelles and their various roles in gene regulation. *Nat. Rev. Mol. Cell Biol.* 24, 288–304
23. Kurgan, L. *et al.* (2023) Tutorial: a guide for the selection of fast and accurate computational tools for the prediction of intrinsic disorder in proteins. *Nat. Protoc.* 18, 3157–3172
24. Vendruscolo, M. and Fuxreiter, M. (2023) Towards sequence-based principles for protein phase separation predictions. *Curr. Opin. Chem. Biol.* 75, 102317
25. Chen, Z. *et al.* (2022) Screening membraneless organelle participants with machine-learning models that integrate multimodal features. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2115369119
26. Bolognesi, B. *et al.* (2016) A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression. *Cell Rep.* 16, 222–231
27. Lancaster, A.K. *et al.* (2014) PLAAC: a web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics* 30, 2501–2502
28. Vernon, R.M. *et al.* (2018) Pi–Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* 7, e31486
29. Zhang, H. *et al.* (2023) Large-scale identification of potential phase-separation proteins from plants using a cell-free system. *Mol. Plant* 16, 310–313
30. Eick, D. and Geyer, M. (2013) The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem. Rev.* 113, 8456–8490
31. Boehning, M. *et al.* (2018) RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* 25, 833–840
32. Guo, Y.E. *et al.* (2019) Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* 572, 543–548
33. Brodsky, S. *et al.* (2020) Intrinsically disordered regions direct transcription factor in vivo binding specificity. *Mol. Cell* 79, 459–471
34. Whittaker, C. and Dean, C. (2017) The FLC locus: a platform for discoveries in epigenetics and adaptation. *Annu. Rev. Cell Dev. Biol.* 33, 555–575
35. Yang, H. *et al.* (2014) Antagonistic roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at *Arabidopsis FLC*. *Curr. Biol.* 24, 1793–1797
36. Zhu, P. and Dean, C. (2023) Reply to: Cold induction of nuclear FRIGIDA condensation in *Arabidopsis*. *Nature* 619, E33–E37

37. Zhang, Z. *et al.* (2023) Cold induction of nuclear FRIGIDA condensation in *Arabidopsis*. *Nature* 619, E27–E32
38. Shvedunova, M. and Akhtar, A. (2022) Modulation of cellular processes by histone and non-histone protein acetylation. *Nat. Rev. Mol. Cell Biol.* 23, 329–349
39. Chow, H.T. and Mosher, R.A. (2023) Small RNA-mediated DNA methylation during plant reproduction. *Plant Cell* 35, 1787–1800
40. Bieluszewski, T. *et al.* (2023) The role and activity of SWI/SNF chromatin remodelers. *Annu. Rev. Plant Biol.* 74, 139–163
41. Buttress, T. *et al.* (2022) Histone H2B.8 compacts flowering plant sperm through chromatin phase separation. *Nature* 611, 614–622
42. Piunti, A. and Shilatifard, A. (2021) The roles of Polycomb repressive complexes in mammalian development and cancer. *Nat. Rev. Mol. Cell Biol.* 22, 326–345
43. Bieluszewski, T. *et al.* (2021) PRC2 activity, recruitment, and silencing: a comparative perspective. *Trends Plant Sci.* 26, 1186–1198
44. Wang, Y. *et al.* (2023) Emerging roles of phase separation in plant transcription and chromatin organization. *Curr. Opin. Plant Biol.* 75, 102387
45. Plys, A.J. *et al.* (2019) Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes Dev.* 33, 799–813
46. Berry, S. *et al.* (2017) Disruption of an RNA-binding hinge region abolishes LHP1-mediated epigenetic repression. *Genes Dev.* 31, 2115–2120
47. Zhang, Y. *et al.* (2020) Phase separation of *Arabidopsis* EMB1579 controls transcription, mRNA splicing, and development. *PLoS Biol.* 18, e3000782
48. Sanulli, S. *et al.* (2019) HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature* 575, 390–394
49. Zhao, S. *et al.* (2019) Plant HP1 protein ADCP1 links multivalent H3K9 methylation readout to heterochromatin formation. *Cell Res.* 29, 54–66
50. Clapier, C.R. *et al.* (2017) Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* 18, 407–422
51. Zhang, H. *et al.* (2018) Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.* 19, 489–506
52. Zemach, A. *et al.* (2013) The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* 153, 193–205
53. Gallego-Bartolome, J. *et al.* (2019) Co-targeting RNA polymerases IV and V promotes efficient *de novo* DNA methylation in *Arabidopsis*. *Cell* 176, 1068–1082
54. Huang, K. *et al.* (2021) Pol IV and RDR2: a two-RNA-polymerase machine that produces double-stranded RNA. *Science* 374, 1579–1586
55. Mishra, V. *et al.* (2021) Assembly of a dsRNA synthesizing complex: RNA-DEPENDENT RNA POLYMERASE 2 contacts the largest subunit of NUCLEAR RNA POLYMERASE IV. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2019276118
56. Singh, J. *et al.* (2019) Reaction mechanisms of Pol IV, RDR2, and DCL3 drive RNA channeling in the siRNA-directed DNA methylation pathway. *Mol. Cell* 75, 576–589
57. Wang, Q. *et al.* (2021) Mechanism of siRNA production by a plant Dicer–RNA complex in dicing-competent conformation. *Science* 374, 1152–1157
58. Sigman, M.J. *et al.* (2021) An siRNA-guided ARGONAUTE protein directs RNA polymerase V to initiate DNA methylation. *Nat. Plants* 7, 1461–1474
59. Xie, G. *et al.* (2023) Structure and mechanism of the plant RNA polymerase V. *Science* 379, 1209–1213
60. Xue, Y. *et al.* (2021) *Arabidopsis* MORC proteins function in the efficient establishment of RNA directed DNA methylation. *Nat. Commun.* 12, 4292
61. Taliansky, M.E. *et al.* (2023) Cajal bodies: evolutionarily conserved nuclear biomolecular condensates with properties unique to plants. *Plant Cell* 35, 3214–3235
62. Courchaine, E. *et al.* (2022) The coilin N-terminus mediates multivalent interactions between coilin and Nopp140 to form and maintain Cajal bodies. *Nat. Commun.* 13, 6005
63. Hearst, S.M. *et al.* (2009) Cajal-body formation correlates with differential coilin phosphorylation in primary and transformed cell lines. *J. Cell Sci.* 122, 1872–1881
64. Hebert, M.D. *et al.* (2002) Coilin methylation regulates nuclear body formation. *Dev. Cell* 3, 329–337
65. Pontes, O. *et al.* (2006) The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell* 126, 79–92
66. Li, C.F. *et al.* (2006) An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* 126, 93–106
67. Li, C.F. *et al.* (2008) Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in *Arabidopsis thaliana*. *PLoS Genet.* 4, e27
68. Lei, Z. *et al.* (2021) Phase separation of chromatin and small RNA pathways in plants. *Plant J.* 108, 1256–1265
69. Ma, Z. and Zhang, X. (2018) Actions of plant Argonautes: predictable or unpredictable? *Curr. Opin. Plant Biol.* 45, 59–67
70. Li, S. *et al.* (2017) The functions of plant small RNAs in development and in stress responses. *Plant J.* 90, 654–670
71. Tang, J. and Chu, C. (2017) MicroRNAs in crop improvement: fine-tuners for complex traits. *Nat. Plants* 3, 17077
72. Pelletier, D. *et al.* (2023) miRNA biogenesis and inherited disorders: clinico-molecular insights. *Trends Genet.* 39, 401–414
73. Partin, A.C. *et al.* (2020) Cryo-EM structures of human Drosha and DGCR8 in complex with primary microRNA. *Mol. Cell* 78, 411–422
74. Jin, W. *et al.* (2020) Structural basis for pri-miRNA recognition by Drosha. *Mol. Cell* 78, 423–433 e425
75. Lee, Y.Y. *et al.* (2023) Structure of the human DICER–pre-miRNA complex in a dicing state. *Nature* 615, 331–338
76. Cui, Y. *et al.* (2021) Global miRNA dosage control of embryonic germ layer specification. *Nature* 593, 602–606
77. Zhu, H. *et al.* (2013) Bidirectional processing of pri-miRNAs with branched terminal loops by *Arabidopsis* Dicer-like1. *Nat. Struct. Mol. Biol.* 20, 1106–1115
78. Fang, Y. and Spector, D.L. (2007) Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living *Arabidopsis* plants. *Curr. Biol.* 17, 818–823
79. Liang, C. *et al.* (2022) *Arabidopsis* RBV is a conserved WD40 repeat protein that promotes microRNA biogenesis and ARGONAUTE1 loading. *Nat. Commun.* 13, 1217
80. Fan, L. *et al.* (2022) *Arabidopsis* AAR2, a conserved splicing factor in eukaryotes, acts in microRNA biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2208415119
81. Gonzalo, L. *et al.* (2022) R-loops at microRNA encoding loci promote co-transcriptional processing of pri-miRNAs in plants. *Nat. Plants* 8, 402–418
82. Zhu, J. *et al.* (2022) The R-loop influences miRNA birth place. *Nat. Plants* 8, 320–321
83. Li, Y. *et al.* (2020) Degradation of SERRATE via ubiquitin-independent 20S proteasome to survey RNA metabolism. *Nat. Plants* 6, 970–982
84. Jozwiak, M. *et al.* (2023) SERRATE: a key factor in coordinated RNA processing in plants. *Trends Plant Sci.* 28, 841–853
85. Machida, S. *et al.* (2011) Molecular insights into miRNA processing by *Arabidopsis thaliana* SERRATE. *Nucleic Acids Res.* 39, 7828–7836
86. Xie, D. *et al.* (2021) Phase separation of SERRATE drives dicing body assembly and promotes miRNA processing in *Arabidopsis*. *Nat. Cell Biol.* 23, 32–39
87. Wang, L. *et al.* (2022) PRP4KA phosphorylates SERRATE for degradation via 20S proteasome to fine-tune miRNA production in *Arabidopsis*. *Sci. Adv.* 8, eabm8435
88. Borcherds, W. *et al.* (2021) How do intrinsically disordered protein regions encode a driving force for liquid–liquid phase separation? *Curr. Opin. Struct. Biol.* 67, 41–50
89. Babu, M. *et al.* (2022) Peptidyl prolyl isomerase A modulates the liquid–liquid phase separation of proline-rich IDPs. *J. Am. Chem. Soc.* 144, 16157–16163
90. Zhao, G. *et al.* (2023) Peptidyl-prolyl isomerase Cyclophilin71 promotes SERRATE phase separation and miRNA processing in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2305244120
91. Shang, B. *et al.* (2023) Intrinsically disordered proteins SAID1/2 condensate on SERRATE for dual inhibition of miRNA biogenesis in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 120, e2216006120
92. Tomassi, A.H. *et al.* (2020) The intrinsically disordered protein CARP9 bridges HYL1 to AGO1 in the nucleus to promote microRNA activity. *Plant Physiol.* 184, 316–329

93. Jumper, J. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589
94. Li, Q. *et al.* (2021) DEAD-box helicases modulate dicing body formation in *Arabidopsis*. *Sci. Adv.* 7, eabc6266
95. Chantarachot, T. *et al.* (2020) DHH1/DDX6-like RNA helicases maintain ephemeral half-lives of stress-response mRNAs. *Nat. Plants* 6, 675–685
96. Li, X. *et al.* (2023) Functions and mechanisms of RNA helicases in plants. *J. Exp. Bot.* 74, 2295–2310
97. Wang, Z. *et al.* (2018) SWI2/SNF2 ATPase CHR2 remodels pri-miRNAs via SERRATE to impede miRNA production. *Nature* 557, 516–521
98. Ma, Z. *et al.* (2018) *Arabidopsis* SERRATE coordinates histone methyltransferases ATXR5/6 and RNA processing factor RDR6 to regulate transposon expression. *Dev. Cell* 45, 769–784
99. Sheu-Gruttadauria, J. and MacRae, I.J. (2018) Phase transitions in the assembly and function of human miRISC. *Cell* 173, 946–957
100. Liu, X.M. *et al.* (2021) Selective sorting of microRNAs into exosomes by phase-separated YBX1 condensates. *eLife* 10, e71982
101. Dalmay, T. *et al.* (2000) An RNA-dependent RNA polymerase gene in *Arabidopsis* is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* 101, 543–553
102. Mourrain, P. *et al.* (2000) *Arabidopsis* SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. *Cell* 101, 533–542
103. Xie, Z. *et al.* (2005) DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12984–12989
104. Adenot, X. *et al.* (2006) DRB4-dependent TAS3 trans-acting siRNAs control leaf morphology through AGO7. *Curr. Biol.* 16, 927–932
105. Iwakawa, H.O. *et al.* (2021) Ribosome stalling caused by the Argonaute–microRNA–SGS3 complex regulates the production of secondary siRNAs in plants. *Cell Rep.* 35, 109300
106. Kim, E.Y. *et al.* (2021) Ribosome stalling and SGS3 phase separation prime the epigenetic silencing of transposons. *Nat. Plants* 7, 303–309
107. Tan, H. *et al.* (2023) Phase separation of SGS3 drives siRNA body formation and promotes endogenous gene silencing. *Cell Rep.* 42, 111985
108. Fukunaga, R. and Doudna, J.A. (2009) dsRNA with 5′ overhangs contributes to endogenous and antiviral RNA silencing pathways in plants. *EMBO J.* 28, 545–555
109. Sun, D. *et al.* (2021) The epigenetic factor FVE orchestrates cytoplasmic SGS3-DRB4-DCL4 activities to promote transgene silencing in *Arabidopsis*. *Sci. Adv.* 7, eabf3898
110. Willems, P. *et al.* (2019) The plant PTM viewer, a central resource for exploring plant protein modifications. *Plant J.* 99, 752–762
111. Han, Y. *et al.* (2023) The phase separation of SGS3 regulates antiviral immunity and fertility in *Arabidopsis*. *Sci. China Life Sci.* 66, 1938–1941
112. Ausín, I. *et al.* (2004) Regulation of flowering time by FVE, a retinoblastoma-associated protein. *Nat. Genet.* 36, 162–166
113. Kim, H.J. *et al.* (2004) A genetic link between cold responses and flowering time through FVE in *Arabidopsis thaliana*. *Nat. Genet.* 36, 167–171
114. Qin, W. *et al.* (2021) Deciphering molecular interactions by proximity labeling. *Nat. Methods* 18, 133–143
115. Alberti, S. *et al.* (2019) Considerations and challenges in studying liquid–liquid phase separation and biomolecular condensates. *Cell* 176, 419–434
116. Perego, E. *et al.* (2023) Single-photon microscopy to study biomolecular condensates. *Nat. Commun.* 14, 8224
117. Markmiller, S. *et al.* (2018) Context-dependent and disease-specific diversity in protein interactions within stress granules. *Cell* 172, 590–604
118. Segal, D. *et al.* (2023) A central chaperone-like role for 14-3-3 proteins in human cells. *Mol. Cell* 83, 974–993