HUNGARIAN UNIVERSITY OF AGRICULTURE AND LIFE SCIENCES

Doctoral School of Biological Sciences

ANALYSIS OF ARGONAUTE PROTEINS IN BARLEY: INSIGHTS INTO miRNA LOADING EFFICIENCY AND AGO4 FUNCTIONALITY IN HETEROLOGOUS COMPLEMENTATION

DOI: 10.54598/004940

Doctoral (PhD) dissertation

Fabio Miloro

Gödöllő

2024

The PhD program

Name: Doctoral School of Biological Sciences

Discipline: Plant Biotechnology

Leader of the school: Prof. Dr. Nagy Zoltán, DSc

Head of Biology Sciences Doctoral School

Hungarian University of Agriculture and Life Sciences, Department of Plant Physiology and Plant Ecology, Agroecology Research Group

Supervisors: Dr. Zoltán Havelda (DSc) and Dr. Ágnes Dalmadi

Hungarian University of Agriculture and Life Sciences, Genetic and Biotechnology Institute, Plant Physiology and Developmental Biology Group

……..……

Approval of the PhD School leader

……..……

Approval of the Supervisor

……..…

Approval of the Co-supervisor

TABLE OF CONTENTS

1 Abbreviations

AGO: Argonaute

BiP: Lumenal binding proteins

DCL: Dicer-like

DRM2: Domains Rearranged Methyltransferase 2

dsRNA: Double-stranded RNA

EDTA: Ethylenediaminetetraacetic acid

GFP: Green fluorescent protein

hc-siRNA: Heterochromatic small interfering RNA

HEN1: HUA Enhancer 1

HMW: High molecular weight

hpRNA: Hairpin-structured RNA

HTS: High-throughput sequencing

HYL1/DRB1: Hyponastic Leaves 1

Leu: Leucine

lmiRNA: long miRNA

Met: Methionine

miRNA: microRNA

MS: Murashike – Skoog

Na3VO4: Sodium orthovanadate

NaF: Sodium floride

NRPE1: Nuclear RNA Polymerase E1

PAZ: PIWI/Argonaute/Zwille

phasiRNA: phased secondary siRNA

piRNAs: Piwi-interacting RNA

PIWI: P-element-induced WImpy testis

Pol IV: DNA-dependent RNA Polymerase IV

Pol V: DNA-dependent RNA Polymerase V

RdDM: RNA-directed DNA methylation

RDR: RNA-dependent RNA Polymerase

RISC: RNA-induced silencing complexes

SDS: Sodium dodecyl sulfate

SE: Serrate

SHH1: Sawadee Homeodomain Homolog 1

siRNA: small interfering RNA

sRNA: small RNA

ssRNA: Single-stranded RNA

tasiRNA: trans-acting siRNAs

TE: Transposable element

TGH: Tough

Val: Valine

2 INTRODUCTION

Barley (*Hordeum vulgare* L.) is the fourth most widely cultivated cereal in the world, following wheat, rice, and maize, and serves as a valuable food resource for both humans and animals (Newton et al., 2011). Its close relationship to wheat and its diploid ancestral genome make barley an ideal crop to study cereal genetics (Lü et al., 2015; Rajendran et al., 2022). Barley's phenotypic plasticity and high epigenetic diversity in response to environmental changes, such as increased temperature and decreased precipitation, highlight its importance for epigenetic research (Chano et al., 2021; Hamar et al., 2020). Barley's phenotypic plasticity and high epigenetic diversity in response to environmental changes, such as increased temperature and decreased precipitation, highlight its importance for epigenetic research (Dawson et al., 2015; Rotasperti et al., 2020; Saisho & Takeda, 2011). In addition, control of transposable elements (TEs) by epigenetic mechanisms helps to maintain genome stability in barley (Wicker et al., 2017).

In eukaryotes, small RNA (sRNA)-mediated gene silencing plays a crucial role in developmental regulation, response to environmental cues, and epigenetic control of TEs (Bologna & Voinnet, 2014; Iki, Cléry, et al., 2018). sRNAs, especially microRNAs (miRNAs), are the key molecules of RNA silencing, typically 20-24 nucleotides in length, and are involved in various biological processes (Rogers & Chen, 2013; Baulcombe, 2004; Molnar et al., 2011). miRNAs primarily regulate endogenous gene expression during development and stress responses, while small interfering RNAs (siRNAs) are also involved in genome integrity and biotic stress responses (Manavella et al., 2019; Ruiz-Ferrer & Voinnet, 2009; Tiwari & Rajam, 2022).

sRNAs associate with Argonaute (AGO) proteins to form RNA-induced silencing complexes (RISCs) that directed specific interactions with target transcripts for mRNA cleavage, translational repression, or chromatin modification (Baulcombe, 2015; Ghildiyal & Zamore, 2009; Q. Guo et al., 2016; Hung & Slotkin, 2021). Eukaryotic AGO proteins are divided into the AGO and PIWI subfamilies, while PIWIs and their interacting sRNAs (piRNAs) are predominantly found in animal germlines (Ross et al., 2014), plant genomes encode several AGO proteins, all of which belong to the AGO subfamily (A. Mallory & Vaucheret, 2012; Vaucheret, 2008). This diversity of functions has been made possible by the expansion of gene families encoding RNA silencing components (Hutvagner & Simard, 2008).

The *Arabidopsis* genome encodes ten different AGO proteins, each with unique roles and sometimes overlapping functions. The functionality of these proteins is generally reflected by the sRNA content

associated with them (Baumberger & Baulcombe, 2005; Havecker et al., 2010; Qi et al., 2005, 2006; Zilberman et al., 2003). These AGO proteins have multiple domains including N-terminal, PIWI/Argonaute/Zwille (PAZ), MID, and P-element-induced WImpy testis (PIWI) domains; the latter plays a critical role in recognizing their 5′ end of sRNA and endonuclease activity (Hutvagner & Simard, 2008; W. Liu et al., 2022; X. Zhang et al., 2014). sRNAs are classified into different AGO proteins based on their length and the 5′ nucleotides. For example, AGO1 loads 21-nt sRNAs starting with a uridine (U) residue (miRNAs), while AGO4 and AGO6 prefer 24-nt sRNAs starting with an adenine (A) residue (siRNAs) (Havecker et al., 2010; Mi et al., 2008).

In other flowering plants, the number of AGO proteins is higher than in *Arabidopsis*. For example, soybean contains 22 AGO proteins (*Glycine max*) (X. Liu et al., 2014), in rice 19 have been identified (*Oryza sativa*) (Kapoor et al., 2008), in maize 17 (*Zea mays*) (Qian et al., 2011; L. Zhai et al., 2014), and in barley 11 have been described so far (*Hordeum vulgare*) (Hamar et al., 2020). All flowering AGOs are grouped into three major clades: AGO1/5/10, AGO2/3/7, and AGO4/6/9 (Zhang et al., 2015). This expansion of the plant AGO family suggests functional diversification of AGO proteins, presumably reflecting the expansion of sRNA-directed regulatory pathways.

An important pathway of sRNA is post-transcriptional gene silencing mediated by the action of miRNAs. In miRNA biogenesis, a genome-encoded primary miRNA (pri-miRNA) transcript with a specific stem-loop structure is cleaved to produce 20-24 nucleotide long miRNA duplexes by Dicer-like 1 (DCL1) (Achkar et al., 2016; Rogers & Chen, 2013;Bologna et al., 2013; Bologna & Voinnet, 2014; Xie et al., 2005). These duplexes are then methylated for protection against exonucleases and are thought to be exported to the cytoplasm (Ji & Chen, 2012). The mature strand of the duplexes is then incorporated into RISC, while the miRNA* is ejected and degraded (Borges & Martienssen, 2015). It's been shown that AGO1 containing RISC is mainly assembled in the nucleus and exported as a complex to the cytosol (Bologna et al., 2018). AGO-RISC loaded with a miRNA identifies its RNA targets via complementary base pairing and mediates their repression (Borges & Martienssen, 2015; Ghildiyal & Zamore, 2009). The level of functional AGO1 protein is finely controlled by mechanisms involving both transcriptional and post-transcriptional feedback loops: *AGO1* mRNA expression is controlled by miR168-programmed AGO1-RISC, and the expression of *AGO1* and *MIR168* genes are co-transcriptionally regulated. This complex, multi-layered regulation of AGO1 ensures robust and precise control of this RNA silencing pathway.

Another key pathway is RNA-directed DNA methylation (RdDM), which control endogenous genes and TEs transcription. In *Arabidopsis*, AGO4 binds 24-nt siRNAs associated with repeats and heterochromatin, which play a critical role in RdDM (Havecker et al., 2010; Wu et al., 2010). siRNAs are originated from DNA-dependent RNA Polymerase IV (Pol IV) transcripts, converted into double-stranded RNAs (dsRNAs) by RNA-dependent RNA Polymerase 2 (RDR2), and processed by Dicer-like 3 (DCL3) into 24-nt siRNAs, which are integrated into AGO4 (Huang et al., 2021; S. Li et al., 2015). AGO4-bound siRNAs guide the targeting of nascent scaffold transcripts of DNA-dependent RNA Polymerase V (Pol V) and facilitate *de novo* methylation (Sigman et al., 2021; Wierzbicki et al., 2009; K. Zheng et al., 2021). The proteins effector of the process, AGO4 and its paralogs, have been extensively studied in *Arabidopsis* using mutant models (Duan et al., 2015; Havecker et al., 2010; F. Wang et al., 2023). In monocots, however, experimental studies have been restricted to orthologs in rice and maize (Aubert et al., 2022; Kapoor et al., 2008; Y. Li et al., 2019; Nonomura et al., 2007; Qian et al., 2011; Yang et al., 2013; L. Zhai et al., 2019). TEs significantly influence the epigenetic modifications that regulate plant development and stress adaptation (Casacuberta & González, 2013; Ito et al., 2013; Kazazian, 2004). In addition to its role in maintaining TE silencing, RdDM can also induce transcriptional silencing of foreign DNA, including novel TE insertions, virus-derived sequences, and transgenes (Chan et al., 2004; Marí-Ordóñez et al., 2013).

The aim of this study was to identify and analyze AGO proteins in barley using bioinformatics approaches and to investigate the function of *Arabidopsis* AGO1 and barley AGO4 *in vivo*. Phylogenetic analysis of barley, rice, and *Arabidopsis* AGO proteins revealed orthologous relationships. The effect of the miR168/miR168* duplex structure on AGO1 loading efficiency and its impact on AGO1 protein levels and phenotypic outcomes was also investigated. In addition, two active paralogous barley *AGO4* genes were identified that exhibited distinct properties in sRNA binding and functionality when introduced into an *Arabidopsis ago4* mutant, confirming their roles in regulating RdDM targets.

3 OBJECTIVES

Barley, an important cereal crop, has a complex and intriguing set of Argonaute (AGO) proteins. These AGO proteins are key components of the RNA silencing machinery in plants and play critical roles in gene regulation, transposon control, and defense against viruses.

The objectives of this study were:

- Genome-wide identification and expression analysis of AGO proteins in barley using *in silico* analysis.

- *In silico* analysis of barley AGO1 expression levels and miR168 target site.

- Investigation of the factors that influence the role of miR168 on the loading efficiency of RISC complexes.

- Exploring the potential of miR168 precursor structure modification to alter the AGO1 loading efficiency.

- Identification and bioinformatic analysis of putative *AGO4* genes in barley, including determining their expression levels and phylogenetic relationships to AGO4 proteins of other plants.

- Assessment of the functionality of the identified barley AGO4 proteins through complementation of the *Arabidopsis ago4-3* mutant.

- Analyze barley AGO4-associated small RNAs in heterologous complementation.

4 LITERATURE OVERVIEW

4.1 The dual nature of barley: crop plant and model organism

Barley (*Hordeum vulgare* L.) is a versatile cereal that has played an important role in human and animal nutrition for thousands of years (Newton et al., 2011; Pourkheirandish & Komatsuda, 2007). It is recognized as one of the oldest cultivated grains, with archaeological evidence indicating its use by ancient farmers as one of their first crops. Barley is an important crop worldwide both in terms of cultivated area (49 million hectares) and grain yield produced (145 Megatonnes) (FAOSTAT , 2023). Its resilience and adaptability to different climates and terrains have contributed to its widespread cultivation (Rotasperti et al., 2020; Saisho & Takeda, 2011). Studying the genetics of barley and its domestication process provides valuable insights into the evolution of crops. It provides a window into how these vital food sources have adapted to different environments over time, ensuring food security (Dawson et al., 2015; Newton et al., 2011; Saisho & Takeda, 2011). This research is not only crucial for historical understanding, but also has significant implications for future crop development and agricultural sustainability.

Both barley and rice are important model organisms in the scientific study of the genetics and epigenetics of cereal crops. Rice, with its smaller genome size and extensive genetic resources, is a commonly used model (Izawa & Shimamoto, 1996; Rensink & Buell, 2004; Shimamoto & Kyozuka, 2002). However, barley has certain advantages that make it particularly suitable for certain areas of research. Particularly, it is highly related to wheat, in whose case the direct molecular studies are over-complicated by the highly complicated genome organization. The barley genome has been completely sequenced, leading to the identification of many critically important genes. These include genes involved in the domestication process, disease resistance, and tolerance to various stresses (Kapazoglou et al., 2013; Kong et al., 2020). Epigenetic mechanisms, such as DNA methylation and histone modifications, are essential for the regulation of gene expression in both barley and rice (Hamar et al., 2020; Wu et al., 2010; H. Zhang et al., 2015). These processes can influence plant development and response to environmental stresses (Dawson et al., 2015; Saisho & Takeda, 2011). However, genes related to epigenetic mechanisms in barley have not been thoroughly annotated. Interestingly, barley's larger genome size and higher proportion of repetitive DNA sequences make it an excellent model for studying the epigenetic regulation of gene expression in more complex genomes (Fig. 1) (Wicker et al., 2017). This complexity provides a broader context for understanding

gene regulation and can provide insights into the evolution of genome structure and function. In addition, the inherent tolerance of barley to abiotic stresses, including cold and drought, makes it a valuable model for studying the genetic and epigenetic basis of stress tolerance in cereals (Wiegmann et al., 2019). Such studies can contribute to the development of more resilient crop varieties, enhancing food security and agricultural sustainability in the face of climate change.

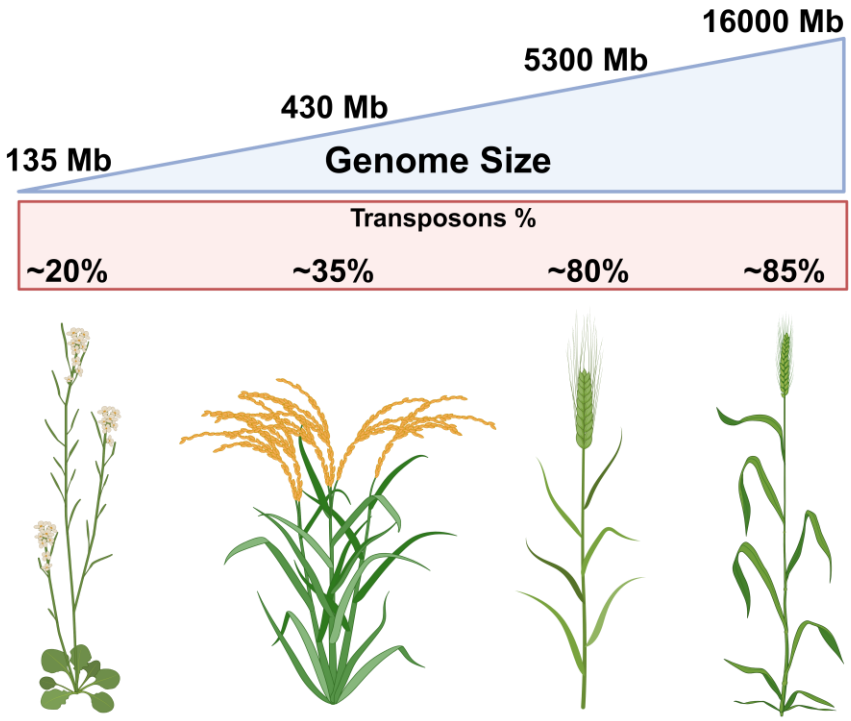

Figure 1. Schematic representation of genome sizes and transposons abundance of model plants compared to barley and wheat. In particular, the plants are from left to right: *Arabidopsis thaliana*, *Oryza sativa*, *Hordeum vulgare* and *Triticum aestivum*. Above each plant there are the value for the genome size in Mb (millions of base pairs) and the percentage of transposons.

4.2 Small RNA-mediated RNA silencing

RNA silencing, a mechanism shared by eukaryotes including plants and animals, is critical for gene regulation and maintaining genome integrity (Baulcombe, 2004; Molnar et al., 2011). RNA silencing is a nucleotide sequence-specific gene regulatory mechanism that controls developmental processes, heterochromatin maintenance, and responses to abiotic and biotic stresses (Kapazoglou et al., 2013; Kong et al., 2020; Kryuvrysanaki et al., 2021; Voinnet, 2005). This process regulates gene expression

either by transcriptional silencing, which involves DNA and histone methylation, or by post-transcriptional action, which involves mRNA cleavage or translational inhibition of the target sequence (Baulcombe, 2015; Ghildiyal & Zamore, 2009; Q. Guo et al., 2016; Hung & Slotkin, 2021). The RNA silencing mechanism is initiated by the recognition of double-stranded RNA (dsRNA) or hairpin-structured RNA (hpRNA) by Dicer-like 1 (DCL1) to produce small RNAs. These are incorporated into the effector complex known as the RNA-induced silencing complex (RISC), which contains Argonaute (AGO) proteins that play a central role in the regulation of gene expression through transcriptional and post-transcriptional gene silencing (Baulcombe, 2004; Q. Guo et al., 2016). In the plant kingdom, RNA silencing is triggered by dsRNA or hpRNA originated from viral genomes or TE sequences (W. Guo et al., 2021; Kryuvrysanaki et al., 2021; Merai et al., 2005; Nuthikattu et al., 2013; Voinnet, 2005). This not only prevents the production of undesired proteins, but also triggers the degradation or silencing of the viral genome or TE sequences, thereby neutralizing the risk of a pathogenic attack. The sequence specificity provided by the AGO-loaded sRNA ensures that only the intended mRNA or DNA sequences are targeted, thereby reducing the risk of off-target effects. The RNA silencing process is also highly adaptable, as it can be induced or suppressed in response to fluctuating environmental conditions or the presence of pathogens (Baulcombe, 2004; Ruiz-Ferrer & Voinnet, 2009; Voinnet, 2005). In addition to its defensive role against pathogens, RNA silencing also plays a central role in gene regulation during plant development and stress responses. It provides a versatile mechanism for regulating the expression changes required for transitions into different developmental stages, such as the vegetative-generative transition. Thus, RNA silencing has been implicated in the regulation of flowering time, leaf senescence, and responses to abiotic stresses such as drought and cold (Y. Liu et al., 2020; H. Zhang et al., 2015).

4.3 Argonaute proteins as central players in RNA silencing

Argonaute (AGO) proteins are a family of proteins that play a critical role in regulating gene expression in plants. The family of AGO proteins was originally identified as PAZ proteins, named after a unique domain, found in the central part of the three pioneering members of the family: PIWI from *Drosophila melanogaster*, AGO1 and ZLL from *Arabidopsis*. To distinguish AGO proteins from DICER proteins, which were later discovered to also contain a PAZ domain, AGO proteins were renamed PPD (PAZ PIWI Domains) proteins due to the presence of a PIWI domain, a feature not

found in Dicer proteins. The revelation that the PIWI domain of several PPD proteins has an RNaseH-like activity that cleaves single-stranded target RNA (ssRNA) in the region complementary to small RNA led to the renaming of PPD proteins as Slicer proteins. This was done to emphasize the contrast with Dicer, which cleaves dsRNA. However, since not all PPD proteins exhibit this cleavage activity, the term Slicer was considered inappropriate (Tolia & Joshua-Tor, 2006). The AGO family, named after AGO1 in *Arabidopsis thaliana*, loss of which leads to tubular shaped leaves resembling small squids (*Argonautus*), is currently the preferred term for this gene family (Bohmert et al., 1998; Vaucheret, 2008).

All AGO proteins possess PAZ, MID (middle) and PIWI domains (Fig. 2A-B), but are classified into three groups based on their phylogenetic relationships and their ability to bind small RNAs (Fig. 2C) (Tolia & Joshua-Tor, 2006; Vaucheret, 2008). Group 1 members bind to miRNAs and siRNAs and are therefore classified as AGO proteins. Members of group 2 bind to PIWI-interacting RNAs (piRNAs) and are therefore termed PIWI proteins (Fang & Qi, 2016). Group 3 members, which have only been identified in worms, bind to secondary siRNAs (Yigit et al., 2006). The amount and diversity of AGO proteins varies widely among organisms. For example, fission yeast (*Schizosccharomyces pombe*) has a single AGO protein, while insects (*Drosophila melanogaster*) have two AGO and three PIWI proteins. Mammals (*Homo sapiens*) have four AGO and four PIWI proteins, and worms (*Caenorhabditis elegans*) have a total of 27 functional AGO and PIWI proteins (Tolia & Joshua-Tor, 2006; Vaucheret, 2008; Yigit et al., 2006). This diversity reflects the complexity and adaptability of these proteins in different organisms.

As mentioned before, AGO proteins consist of the following domains: N-terminal; PIWI/Argonaute/Zwille (PAZ); MID; P-element-Induced WImpy testis (PIWI) domains and two linkers (Hutvagner & Simard, 2008). The N-terminal domains are rather variable, whereas the C-terminal ones (PAZ, MID and PIWI) appear to be more conserved (Tolia & Joshua-Tor, 2006). The N-terminal domain is the most variable of the AGO protein domains, and its exact function is poorly understood and difficult to identify due to its variability (Fang & Qi, 2016). The PAZ domain, a key component of AGO proteins, contains an OB (oligonucleotide/oligosaccharide binding) fold. This structural feature enables AGO proteins to bind to single-stranded nucleic acids. The PAZ domain achieves this by embracing the 3′ end of the guide strand into a specialized binding pocket, thereby securing sRNAs (Yan et al., 2003). In addition, the PAZ domain plays an important role in duplex unwinding, a process that is independent of slicer activity (Gu et al., 2012). A rigid loop within the MID domain, known as the nucleotide specificity loop, is designed to specifically identify the 5′

nucleotide of sRNAs (X. Zhang et al., 2014). This feature explains the binding preferences of different AGO proteins for sRNAs with different 5′ nucleotides. The interface between the MID and PIWI domains contains a basic binding pocket that facilitates the binding and anchoring of the 5′ phosphate of sRNAs (W. Liu et al., 2022). The PIWI domain, which folds in an RNase H-like structure, enables some, but not all, AGO proteins to cleave target RNAs that are complementary to the bound sRNAs. A catalytic trio, typically referred to as Asp-Asp-His/Asp or DDH/D, is widely believed to be the driving force behind slicer activity (Baumberger & Baulcombe, 2005; F. Wang et al., 2023). To date, there are no available crystal structures for complete plant AGOs. However, advances in protein 3D structure prediction technology may provide valuable insights into the unique properties of AGO proteins (Abramson et al., 2024).

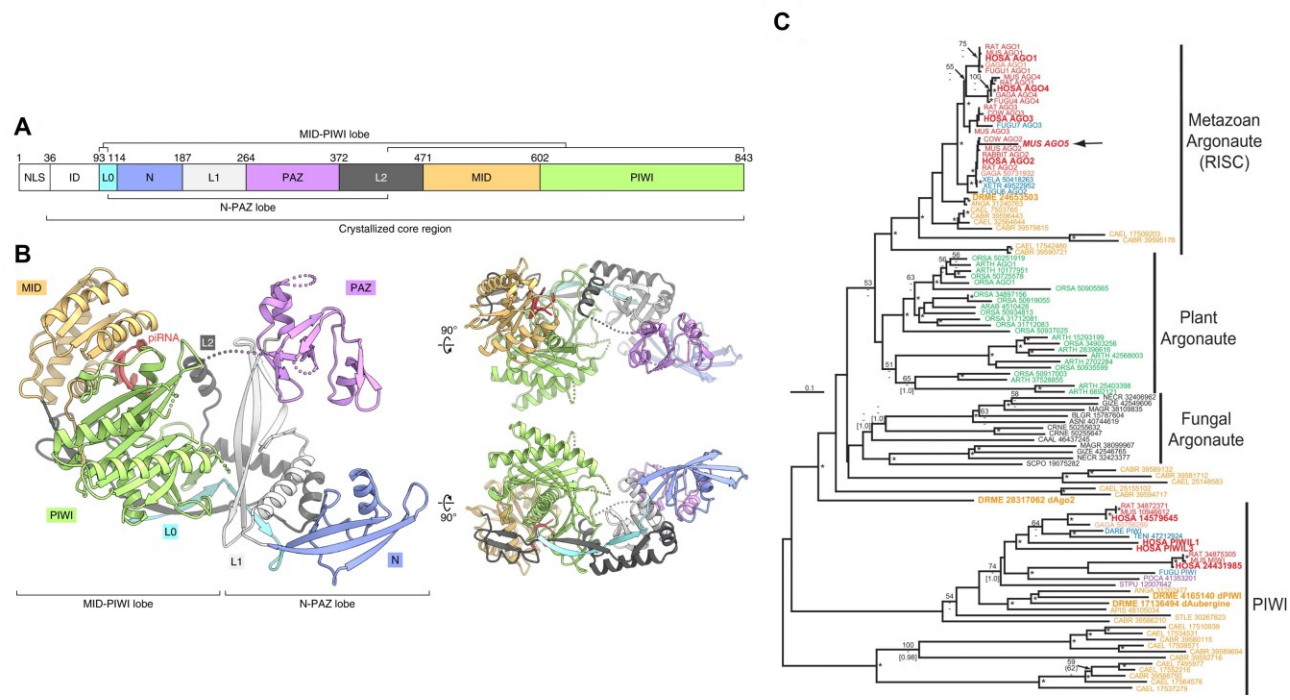

Figure 2. A Domain organization of AGO and Piwi proteins. NLS, nuclear localization signal; ID, intrinsically disordered region. **B** Crystal structure of the Piwi–piRNA complex. **C** Neighbor-joining phylogenetic tree of the Argonaute/PIWI protein family. Major protein subgroups are indicated. The tree is unrooted. Phylogenetic reconstruction method. Branches are labeled with a four-letter species identifier (the first two letters of the genus and species name) and the GenBank accession number.) The only unlabeled subgroup is that of worms (Yamaguchi et al., 2020; Murphy et al., 2008).

The evolution of AGO proteins in the plant kingdom has received considerable attention in recent years (Fang & Qi, 2016; Vaucheret, 2008; Zaheer et al., 2024; H. Zhang et al., 2015). Research indicates that the AGO protein family has undergone substantial duplication and diversification during

plant evolution. The green algae *Chlamydomonas reinhardtii* has three AGOs (T. Zhao et al., 2007), while moss (*Physcomitrella patens*) has six (Arif et al., 2013). However, the number of AGOs further increases in flowering plants: there are 10 in *Arabidopsis*, 15 in poplar (*Populus trichocarpa*), 17 in maize (*Zea mays*) and 19 in rice (*Oryza sativa*) (Kapoor et al., 2008; Morel et al., 2002; Qian et al., 2011; L. Zhai et al., 2014; H. Zhang et al., 2015; K. Zhao et al., 2015). Based on phylogenetic relationships, plant AGO proteins can be grouped into three major clades named after *Arabidopsis* AGOs: AGO1/5/10, AGO2/3/7, and AGO4/6/8/9 (Fig. 3) (Fang & Qi, 2016; Vaucheret, 2008; H. Zhang et al., 2015). In addition, grasses have another subclade, AGO18, a deep branch of the AGO1/5/10 clade (Fang & Qi, 2016; Z. Li et al., 2022; H. Zhang et al., 2015). This has led to the emergence of specialized roles for each family member, increasing the complexity and adaptability of plant responses to various environmental and developmental stresses (Chang et al., 2020; Chinnusamy et al., 2014; Ito, 2022; J. M. Kim et al., 2015; Ramakrishnan et al., 2022). This diversification is a testament to the dynamic nature of plant evolution and the intricate mechanisms that underpin plant biology. The study of AGO proteins therefore provides valuable insights into the molecular mechanisms that have shaped plant evolution and diversity.

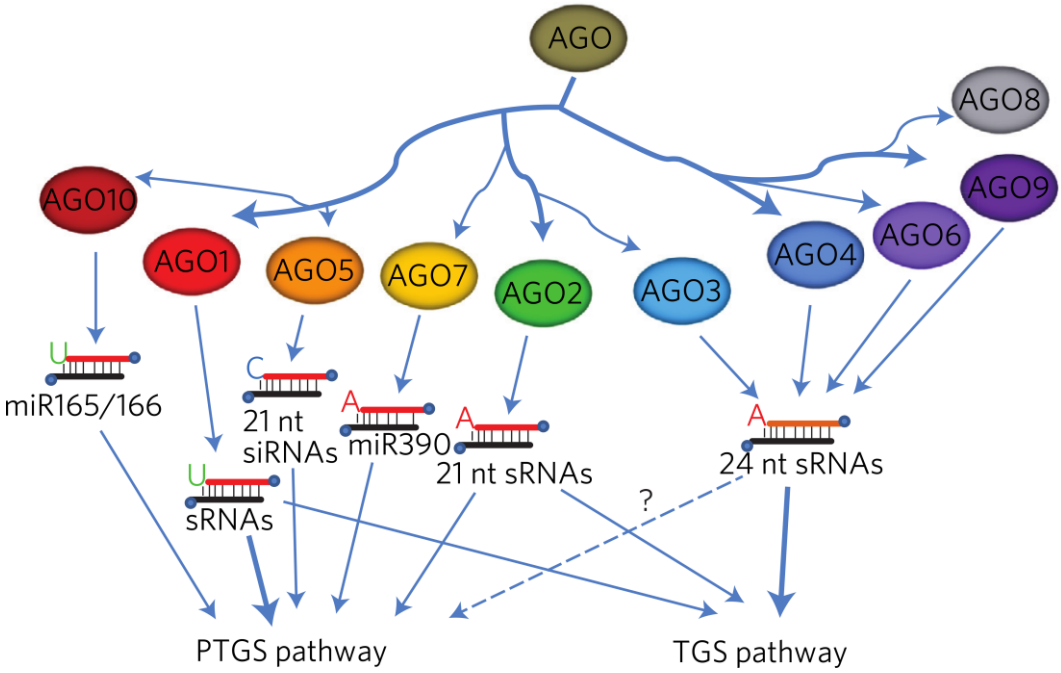

Figure 3. Functional phylogenetic tree of *Arabidopsis* AGO proteins and their associated sRNAs in the TGS (transcriptional gene silencing) and PTGS (post-transcriptional gene silencing) pathways. From the subgrouping observed in the phylogenetic tree, the proteins were grouped into three major clades: AGO1/5/10, AGO2/3/7 and AGO4/6/8/9 (Z. Zhang at al., 2016).

4.4 Small RNAs

In plant biology, small RNAs play a central role in controlling aspects such as growth, reproduction, and responses to biotic and abiotic stresses. They exert their influence at the molecular level by regulating chromatin modifications, transcript levels and translation through a variety of mechanisms. The main types of plant sRNAs are miRNAs and small interfering RNAs (siRNAs). The latter can be further subdivided into heterochromatic small interfering RNAs (hc-siRNAs) and secondary siRNAs. Over the past decades, numerous other types of plant sRNAs have been identified, but their generation processes and roles remain either controversial or not fully understood. The production of sRNAs typically involves the generation of 20 to 24 nucleotide sRNA pairs from larger RNA molecules by Dicer-like (DCL) endoribonucleases. The precursor molecules are either self-complementary RNAs that fold into hairpin-like structures or dsRNAs produced by RNA-dependent RNA polymerases (RDRs). Often, one strand of the short dsRNA molecules produced by DCL is incorporated into AGO proteins. These proteins then target either internal or external RNAs based on sequence complementarity, resulting in the silencing of genes or TEs (Fig. 4). The different pathways are often characterized by protein factors of the same gene families that are specialized for the production or action of the actual sRNA, or the site of action, which is usually in the cytoplasm, except in the case of transcriptional gene silencing.

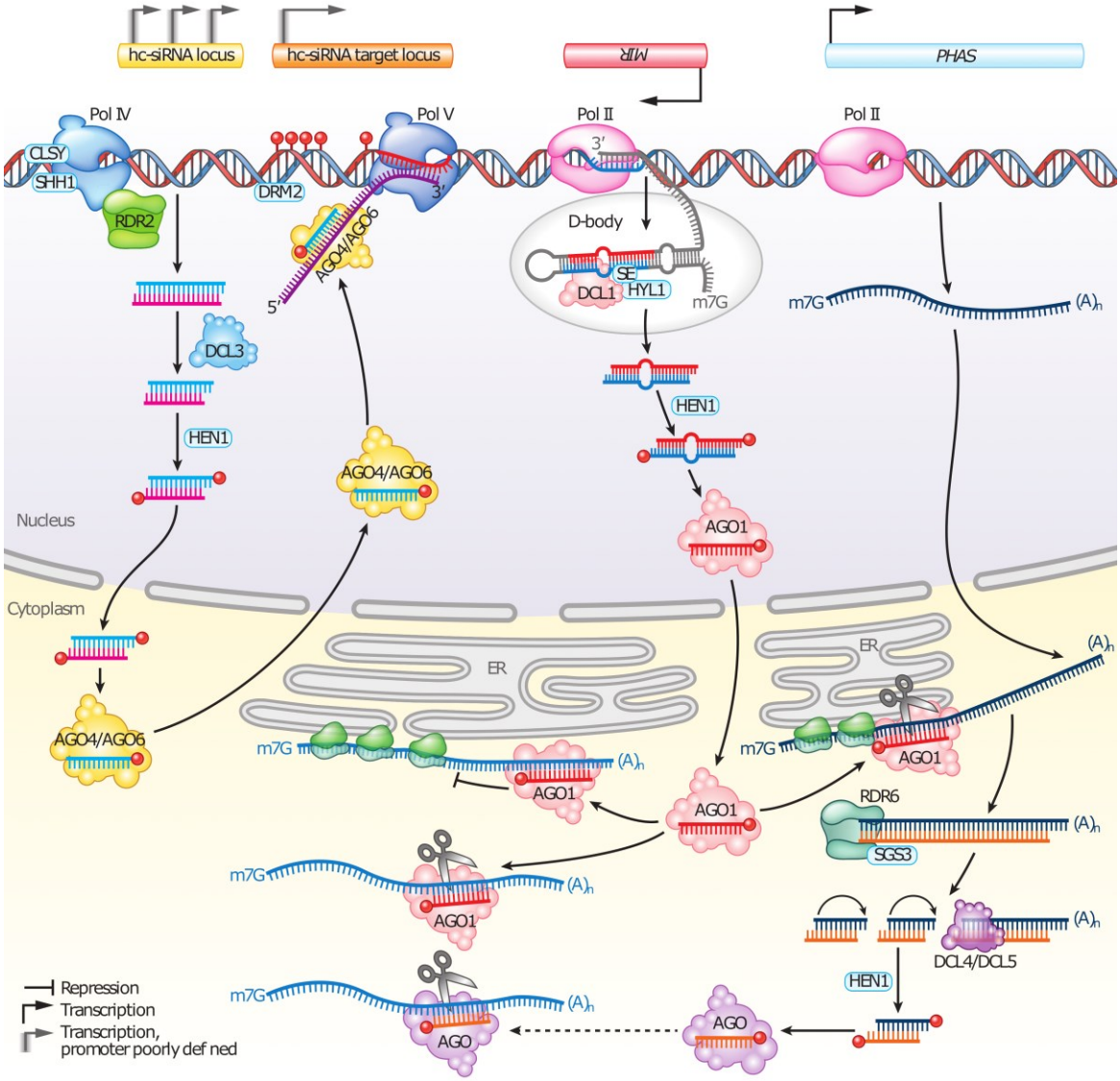

Figure 4. Biogenesis and modes of action of canonical plant sRNAs. The hc-siRNAs are generated mainly through the activities of Pol IV, CLSY, RDR2, and DCL3. The hc-siRNAs form RISCs with AGO4/AGO6 proteins and direct DNA methylation at Pol V–transcribed loci via DRM2. The blurry arrow for the transcriptional start of the hc-siRNAs and targets indicates the poorly characterized Pol IV/Pol V transcription start sites. miRNAs are generated through the activities of primarily Pol II, DCL1, HYL1, and SE. miRNAs form RISCs predominantly with AGO1; miRISCs can direct transcript cleavage, mediate translational repression, or trigger phasiRNA biogenesis. phasiRNAs are produced by Pol II, RDR6, and DCL4/DCL5. Other proteins involved in the biogenesis or actions are described in the main text. Figure from (Zhan & Meyers, 2023).

4.4.1 Biogenesis and action of miRNAs

The biogenesis of miRNAs begins with the transcription of primary miRNA (pri-miRNA) transcripts, which have specific stem-loop hairpin-like structures (Fig. 5). These pri-miRNA are more variable in length and structure in plants than in animals. The processing of pri-miRNAs involves two

consecutive cleavage events by Dicer-like1 (DCL1), resulting in a 20-24 nucleotide miRNA duplex with two nucleotides single-stranded at its 3' end (Achkar et al., 2016; Manavella et al., 2019; Vaucheret et al., 2004). This mostly imperfect duplex consists of a guide strand (miRNA) and a passenger strand (miRNA*) with distinct structural features. The precise processing of pri-miRNAs depends on structural signals such as terminal loop structure and base pairing, as well as the action of DCL1 co-factors such as Tough (TGH), Serrate (SE) and Hyponastic Leaves 1 (HYL1/DRB1) (Fig. 5) (J. Wang et al., 2019; Z. Wang et al., 2018). The resulting miRNA/miRNA* duplexes are then methylated by HUA Enhancer 1 (HEN1) to protect them from exonucleases (Ji & Chen, 2012). These duplexes are then incorporated into the RNA-induced silencing complex (RISC) where the guide strand is loaded onto the Argonaute protein while the passenger strand is degraded. AGO1 is the most important effector protein for miRNAs among the ten AGO proteins encoded in *Arabidopsis*, which are specialized for different RNA silencing pathways but often show functional redundancy (Bohmert et al., 1998).

The selection of the guide strand from the duplex is influenced by the thermodynamic stability of the miRNA duplex ends, and the 5'-end nucleotide of the guide strand plays a key role in determining the sorting preference of miRNAs into specific AGO proteins (Borges & Martienssen, 2015). As AGO1 shows a strong preference for miRNAs with a 5'-uridine, most miRNAs have this feature (Mi et al., 2008). The structural characteristics of certain miRNA duplexes, such as miR165/166 and miR390, order them specifically to AGO10 and AGO7, respectively (Endo et al., 2013; H. Zhu et al., 2011).

The importance of the miRNA pathway is underscored by the severity of mutants that are defective in key components of this pathway. The malfunction of AGO1 disrupts all aspects of plant life, as both the hypomorphic mutation and the overproduction of the protein cause severe malformation, and knockdown of the gene is lethal (Rogers & Chen, 2013). Thus, the amount of functional AGO1 protein is tightly controlled by both transcriptional and post-transcriptional feedback mechanisms. The expression of *AGO1* mRNA is controlled by AGO1-RISC directed by miR168 (Vaucheret et al., 2004). *AGO1* mRNA cleavage leads to the production of secondary siRNAs, which further suppress AGO1 protein expression (A. C. Mallory & Vaucheret, 2009). The expression of AGO1 and MIR168 genes is co-transcriptionally regulated (Vaucheret et al., 2006). The presence of two MIR168 genes, which produce primary transcripts (*pri-MIR168a* and *pri-MIR168b*) with both unique and overlapping functions, adds an additional layer of control over AGO1 (Vaucheret, 2009). Interestingly, miR168 is stabilized by AGO1 binding, and its production remains unaffected by suppressed DCL1 levels (Vaucheret et al., 2006). This intricate multi-layered regulation of AGO1 ensures the robust and

precise functioning of the miRNA pathway. The abundance of AGO1 proteins plays a critical role in the incorporation of mature miRNAs into the plant RISC (Bajczyk et al., 2019; Dalmadi et al., 2019). AGO1 is a protein that shuttles between the nucleus and the cytoplasm. In the nucleus, it undergoes structural changes to bind miRNA/miRNA* duplexes, facilitated by the Hsp90 chaperone (Bajczyk et al., 2019; Bologna et al., 2018). Once loaded, AGO1 is transported to the cytoplasm where it facilitates cleavage of the target mRNA or inhibits its translation. Factors associated with pre-miRNAs influence the efficiency of RISC loading for individual miRNAs. The abundance of AGO1 can act as a bottleneck, as overexpression of miR168-resistant AGO1 increases the loading of miR168 and other miRNAs into RISC (Dalmadi et al., 2019; Vaucheret et al., 2006). The 3' to 5' exonuclease Small RNA Degrading Nuclease 1 (SDN1) prefers single-stranded, non-uridylated miRNA substrates and can degrade methylated miRNAs, while the nucleotidyl transferases HESO1 uridylate miRNAs, leading to their degradation (Rogers & Chen, 2013; J. Wang et al., 2019). The regulation of miRNA biogenesis and activity involves multiple checkpoints. The availability of AGO1 serves as a central control point that determines which miRNAs are loaded into RISC and become biologically active. This allows for rapid adaptation of the miRNA pathway in response to changes in cellular conditions (Dalmadi et al., 2019). This mechanism can control the biologically active fraction of produced miRNAs in a specific cellular environment, and the RISC loading efficiencies of different canonical 5'-U miRNAs are predominantly controlled by their diverse precursor RNAs.

The relationship between sRNAs and RISC complexes of different molecular weights may provide a better understanding of this pathway. It was observed that many standard miRNAs were associated with high molecular weight AGO1-containing RISCs (HMW RISC), whereas the majority of 24-nucleotide siRNAs were associated with low molecular weight AGO4-containing complexes (LMW RISC). Interestingly, a substantial group of cytoplasmic sRNAs, including mature miRNA sequences, were identified in the low molecular weight range, suggesting their existence as protein-unbound sRNAs. A comparison between the RISC-loaded and protein-unbound miRNA pools revealed miRNAs with a wide range of loading efficiencies. Functional validation experiments using transient and transgenic systems confirmed the altered loading capabilities of certain miRNAs. This suggests that this process is controlled by information related to the diverse miRNA precursors (Dalmadi et al., 2019). There is a regulatory checkpoint that controls the RISC-loading efficiency of different miRNAs by selectively incorporating a subset of the produced miRNAs into the biologically active RISCs. These findings further our understanding of the post-transcriptional regulation of gene expression mediated by sRNA pathways in plants.

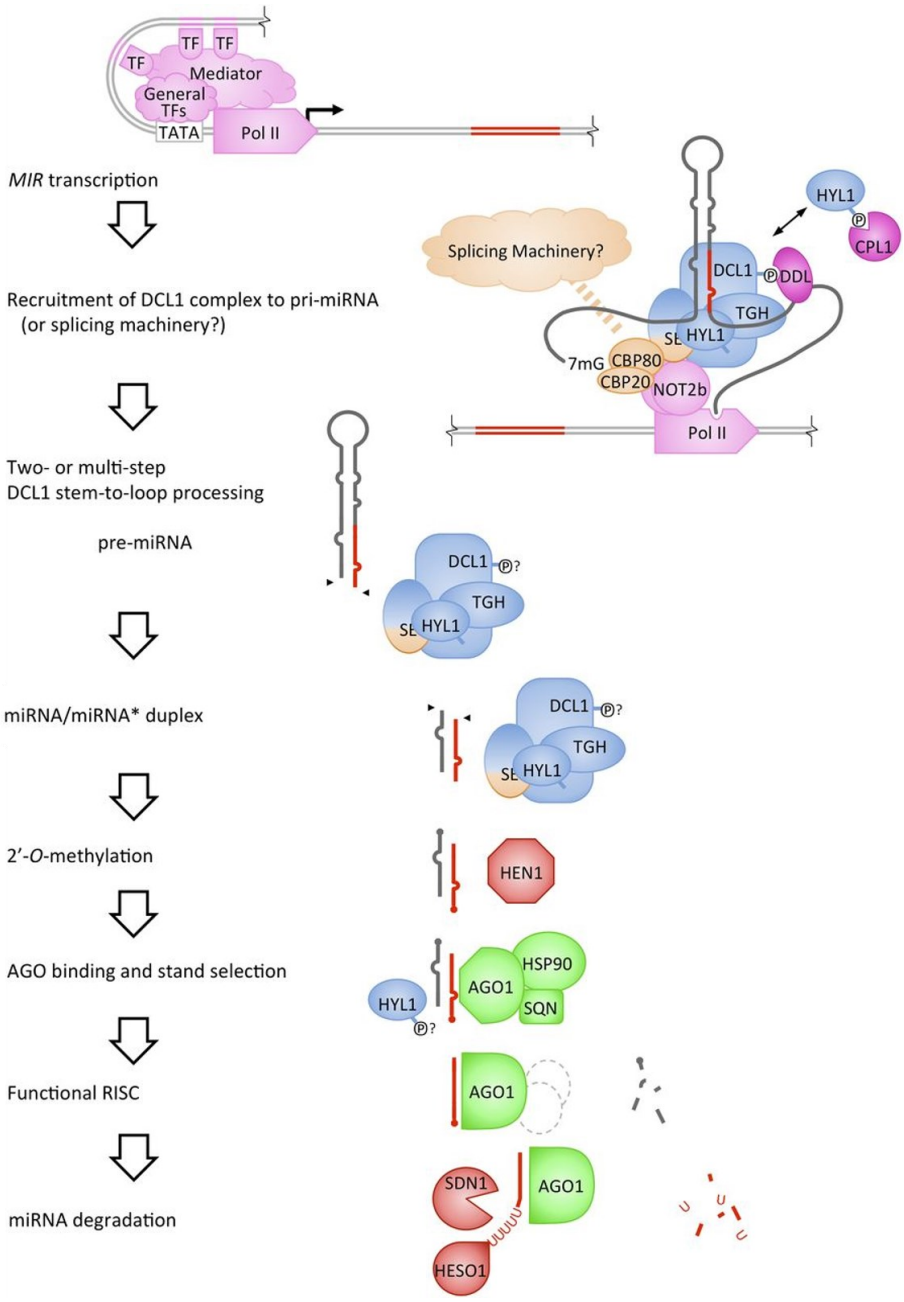

Figure 5. Summary of the Major Steps in miRNA Biogenesis and Turnover. Multiple transcription factors (TFs) control the transcription of MIR genes. The formation of the DCL1 complex and its recruitment to pri-miRNA are mediated by protein–protein interactions and structural features of the pri-miRNA. Some proteins may have dual roles in the recruitment of either the DCL1 complex or the spliceosome to pri-miRNA. As described in the text, protein phosphorylation status may affect the DCL1–DDL interaction and the enhancement of DCL1 processing accuracy by HYL1. Predicted protein phosphorylation (indicated by Ⓟ) may affect distinct steps in pri-miRNA processing, but the role of phosphorylation in other steps (question marks) remains unknown. Functional RISC may include additional unknown AGO interacting proteins (dashed outline). Proteins are color coded according to known functions in MIR transcription (pink), splicing (orange), DCL processing (light blue), phospho-regulation (purple), RISC assembly (green), and miRNA stabilization and turnover (red) (Rogers & Chen, 2013).

4.4.2 The origin of different small interfering RNA classes

The hc-siRNAs, which originate from repetitive regions and TEs in heterochromatic regions of the genome, are 24 nucleotides long and play a role in RNA-directed DNA methylation (RdDM). They represent the majority of siRNAs in plants, accounting for more than 90% of the siRNA population in *Arabidopsis* (Mosher et al., 2008; X. Zhang et al., 2007). The biogenesis of hc-siRNAs starts with transcription by Pol IV, a plant-specific DNA-dependent RNA polymerase, followed by synthesis of a second RNA strand by RDR2(Huang et al., 2021; Law et al., 2013; X. Zhang et al., 2007). These double-stranded RNAs are then cut into hc-siRNA pairs by DCL3 and undergo methylation. A single strand is then incorporated into AGO4, AGO6 or AGO9 to form RISCs (Fig. 1 and 3) (Havecker et al., 2010; Zilberman et al., 2003).

Recent structural studies have shown that the functions of Pol IV and RDR2 are tightly linked (Haag et al., 2012; Huang et al., 2021; Law et al., 2011). Specifically, Pol IV-generated transcripts are directly targeted by RDR2, and Pol IV backtracking allows RDR2 to generate the second strand of hc-siRNA precursors, known as Pol IV-RDR2 (P4R2) RNAs (Fukudome et al., 2021; Huang et al., 2021). These double-stranded P4R2 RNAs are then cut by DCL3 to produce one hc-siRNA for each precursor. The cutting of P4R2 RNAs results in both 24-nt hc-siRNAs and 23-nt siRNAs (Singh et al., 2019). However, immunoprecipitated AGO4 are predominantly companied with 24-nt siRNAs, suggesting that the 23-nt siRNAs are likely passenger strands of hc-siRNA pairs (Havecker et al., 2010; Mi et al., 2008; Qi et al., 2006). Processing of P4R2 RNA by DCL3 has shown that which end of the precursor is cleaved and whether 24- and/or 23-nt siRNAs are produced depending on the first nucleotide of the Pol IV strand of P4R2 RNAs and the start position of RDR2 within the ends of Pol IV-generated transcripts (Loffer et al., 2022). Another recent study found that the Piwi/Argonaute/Zwille (PAZ) and RNase IIIb domains of DCL3 works like a molecular ruler and mark a distance of 24 nt between the 5′-terminal nucleotide and the cleavage site of the RNase IIIb domain for only one strand of P4R2 RNA. The RNase IIIb domain cleaves the measured strand to produce 24-nt siRNAs, while the RNase IIIa domain cleaves the unmeasured strand to produce 23-nt siRNAs (Q. Wang et al., 2021).

Another type of siRNAs, the phased secondary siRNAs (phasiRNAs) are triggered by other miRNAs or siRNAs and generated from long RNA precursors. These siRNAs have a precise complementarity with their precursors. The precursors are initially cleaved by a programmed RISC, made double-

stranded, and then sequentially diced by a DCL protein. The resulted phasiRNAs are loaded into RISCs to target mRNAs (Fig. 1). The genomic loci that give rise to phasiRNAs, known as PHAS loci, can be protein-coding or non-coding (Y. Liu et al., 2020; Montgomery et al., 2008). The *Arabidopsis* genome contains eight known *trans*-acting siRNAs (tasiRNA)-producing loci, *TAS1-TAS4* (Allen et al., 2005; Rajagopalan et al., 2006). The biogenesis of tasiRNA occurs via two pathways, both starting with the cleavage of a TAS precursor by a miRNA. A fragment is converted into dsRNA by RDR6 and iteratively cleaved by DCL4 to produce 21 nt phasiRNA duplexes (Y. Liu et al., 2020). The two pathways differ in several aspects, including the length of miRNA initiators and the pairing of initiators and TAS precursors. Reproductive phasiRNAs, first characterized in grass species and then in other angiosperms, are enriched in anthers and classified into two sizes, 21-nt and 24-nt. In rice and maize, these two size classes accumulate at different stages of anther development (Johnson et al., 2009; J. Zhai, Zhang, et al., 2015). The biogenesis of reproductive phasiRNAs follows the one-hit model, with different miRNA families inducing phasiRNAs of different sizes.

4.5 RNA-directed DNA methylation pathway

In plants, the primary siRNA-mediated epigenetic pathway is RNA-directed DNA methylation (RdDM), which was first identified in plants infected with RNA viruses (Jones et al., 1998; Wassenegger et al., 1994). RdDM requires siRNAs and core RNA silencing proteins and relies on a specialized transcriptional machinery centered around two plant-specific enzymes, Pol IV and Pol V (Chan et al., 2004; Eun et al., 2012; Haag & Pikaard, 2011; Mette et al., 2000). Processed siRNAs loaded on into AGO4 recruit DNA methyltransferase activity, leading to *de novo* methylation of cytosines in all sequence contexts. This results in transcriptional silencing, particularly of TEs and other repetitive DNA. Components of the RdDM pathway have been identified by genetic and biochemical methods and are associated with siRNA biogenesis, *de novo* DNA methylation and chromatin modifications (Erdmann & Picard, 2020; Matzke & Mosher, 2014; H. Zhang et al., 2013).

Pol IV, a polymerase critical for generating most of the 24-nucleotide siRNAs involved in directing methylation within the RdDM pathway, relies on Sawadee Homeodomain Homolog 1 (SHH1) for recruitment to specific sequences, primarily TEs and repeats (Law et al., 2013; Mosher et al., 2008; X. Zhang et al., 2007). SHH1, which binds to H3K9me (Histone H3 Lysine 9 methylated) and unmethylated H3K4 (Histone H3 Lysine 4), facilitates the activity of Pol IV (Smith et al., 2007).

Notably, Pol IV initiates transcription similarly to Pol II, showing a preference for pyrimidines (Y) upstream and purines (R) at the transcription start site, indicating shared preferences despite their distinct genomic roles (Blevins et al., 2015; J. Zhai, Bischof, et al., 2015). This polymerase transcribes ssRNAs at target sites, which are then duplicated into dsRNAs by RDR2 with the help of the chromatin remodeler Classy 1 (CLSY1) (Haag et al., 2012; Law et al., 2011; Smith et al., 2007; Zhou et al., 2018). DCL3 processes the dsRNAs into 24-nucleotide siRNAs, which are methylated by HEN1 for stability before being loaded into AGO4 (Fig.1 and 3) (Havecker et al., 2010; F. Wang et al., 2023; F. Wang & Axtell, 2017). The AGO4 family members, including AGO6 and AGO9, exhibits partial redundancy and tissue specificity, particularly in reproductive tissues (Duan et al., 2015; Havecker et al., 2010).

Pol V-generated transcripts are thought to serve as scaffold RNAs that interact with siRNAs and recruit silencing factors (Wierzbicki et al., 2008). Like Pol IV, the recruitment of Pol V to its target sequences is not fully understood. Genome-wide studies are shedding light on the binding site preferences for Pol V-mediated RdDM. ChIP-seq experiments have shown that most Pol V is located at TEs and repeats associated with 24-nucleotide siRNAs and cytosine methylation, suggesting its role in RdDM at these sites (Niu et al., 2022; Wierzbicki et al., 2012). However, about a quarter of the genomic sites occupied by Pol V lack these features, suggesting that Pol V occupancy alone does not guarantee RdDM (Wierzbicki et al., 2012). These unmethylated sites are often genes, some with repetitive sequences in their coding region. Pol V-mediated RdDM operates across the genome, favoring euchromatic regions, especially small, recently acquired intergenic TEs, and genes with TEs or repeats in their promoters, introns, or coding regions (Niu et al., 2022; Nozawa et al., 2022; Sigman et al., 2021; Zhong et al., 2012). This is consistent with the proposed evolution of Pol IV and Pol V from Pol II, which primarily transcribes genes in euchromatic contexts. The enrichment of Pol V at gene promoters may be an inheritance from the preference of Pol II for upstream gene regions (Zhong et al., 2012). Pol V recruits AGO4 through the AGO hook region in its largest subunit, Nuclear RNA Polymerase E1 (NRPE1) (Böhmdorfer et al., 2014; He et al., 2009). During Pol V transcription, AGO4-bound siRNA is thought to pair with the emerging Pol V transcript and bring in Domains Rearranged Methyltransferase 2 (DRM2), a DNA methyltransferase of the DNMT3 family, to initiate *de novo* methylation at targeted genomic sites (Fig. 6) (Zhong et al., 2014). While RdDM helps to silence TEs, it can also create a dynamic modification that can be removed from euchromatic targets via passive or active demethylation (Lei et al., 2015; Zhong et al., 2012). This is supported by ROS1, a DNA glycosylase involved in active demethylation, which tends to counteract with RdDM-induced

methylation (Tang et al., 2016). The antagonistic relationship between active demethylation and RdDM is further supported by the discovery of numerous known RdDM components in genetic screens for *ros1* mutants suppressors and the reduced expression of *ROS1* in RdDM mutants (Córdoba-Cañero et al., 2017). DNA methylation in plants occurs primarily at cytosine residues, forming 5-methylcytosine (m5C) and 5-hydroxymethylcytosine (m5hC) (Gallego-Bartolomé, 2020; H. Zhang et al., 2018). The m5C is the most abundant DNA modification in plants and occurs mainly in three sequence contexts: CpG, CHG (where H=A, T or C, which are symmetric sites) and CHH (which are asymmetric sites). Plants have a more diverse and abundant cytosine DNA methylation system compared to animals (Gallego-Bartolomé, 2020; H. Zhang et al., 2018).

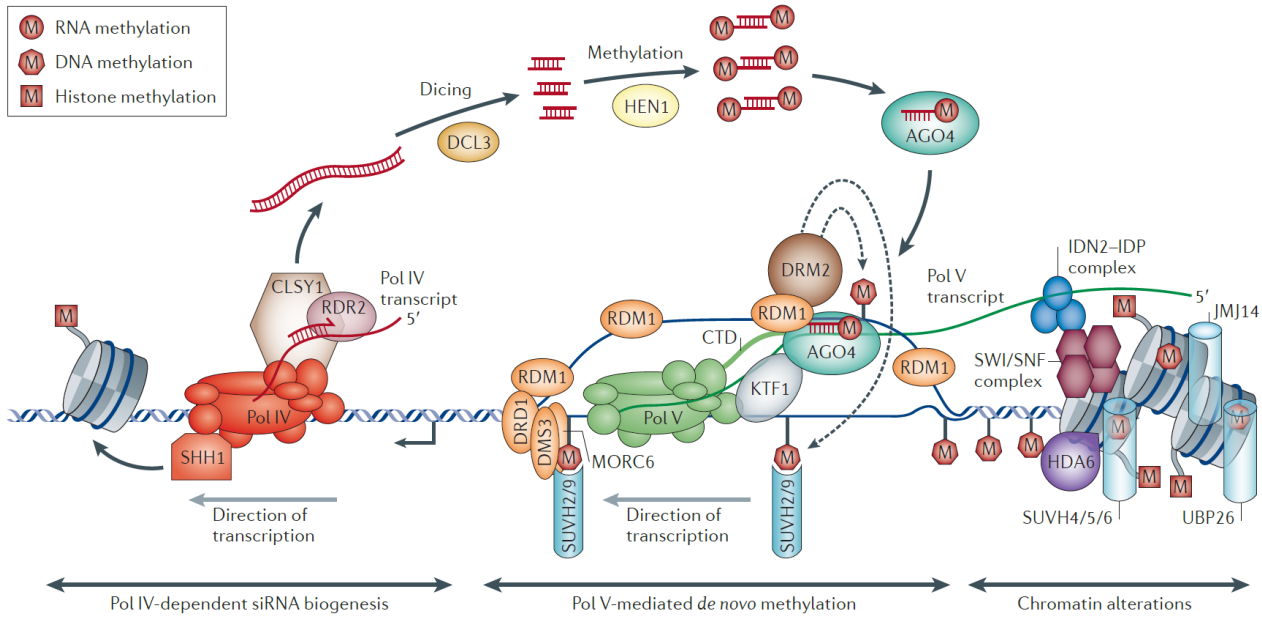

Figure 6. A transcription fork model for RNA-directed DNA methylation (RdDM) is shown. In Pol IV-dependent siRNA biogenesis, Pol IV transcribes an ssRNA, which is copied into a dsRNA by RDR2 with the help of CLSY1. The dsRNA is processed by DCL3 into 24-nucleotide siRNAs, which are methylated at their 3′ ends by HEN1 and incorporated into AGO4. SHH1, which binds to H3K9me, interacts with Pol IV and recruits it to some target loci. In Pol V-mediated *de novo* methylation (middle panel), Pol V transcribes a scaffold RNA that base-pairs with AGO4-bound siRNAs. AGO4 is recruited through interactions with the AGO hook regions in the carboxy-terminal domain of the largest subunit of Pol V and with KTF1. RDM1 links AGO4 to DRM2, which catalyzes *de novo* methylation of DNA. Pol V recruitment may be assisted by SUVH2 or SUVH9, both of which bind to methylated DNA. Nucleosome positioning (right panel) is regulated by the SWI/SNF complex, which interacts with the IDN2-IDP complex that binds to Pol V scaffold RNAs. Deposition of repressive histone modifications is facilitated after removal of active marks by HDA6, JMJ14 and UBP26 (Matzke & Mosher, 2014).

4.6 Control of transposable elements

Our understanding of the biological function of RdDM has expanded through detailed studies of reproduction and the effects of various stresses on the genome (Mosher et al., 2008; X. Zhang et al., 2007). RdDM and Pol IV primarily target repetitive elements, especially smaller, more recent TEs (Lee et al., 2012; Zemach et al., 2013; Zhong et al., 2012).

TEs can be classified into two main categories: Class I, or retrotransposons, which use an RNA intermediate and a "copy-and-paste" method of transposition, and Class II, or DNA TEs, which use a "cut-and-paste" transposition mechanism (Wicker et al., 2007; M. Zhao & Ma, 2013). Despite their inherent ability to replicate and relocate within the genome, the majority of TEs are silenced by epigenetic processes (Zhong et al., 2014). It has been shown that a transcriptionally activated TE can produce siRNAs that trigger DNA methylation via the non-canonical RdDM pathway (Fig. 7) (Cuerda-Gil & Slotkin, 2016; McCue et al., 2015; Nuthikattu et al., 2013; Panda et al., 2016). In this process, the enzyme known as RNA-Dependent RNA Polymerase 6 (RDR6) plays a critical role in converting TE transcripts generated by Pol II into a double-stranded RNA form. This is then processed by Dicer-Like 2 and 4 (DCL2/4) into siRNAs of 21-22 nucleotides in length (Nuthikattu et al., 2013). These siRNAs, also known as secondary siRNAs, direct the methylation of TEs in a homology-dependent manner through a non-canonical RdDM pathway. Argonaute 6 (AGO6) is another key player in this mechanism, responsible for loading the 21-22 nt siRNAs and directing the methylation of DNA at TE sites (McCue et al., 2015). This pathway acts as an important auxiliary system to initiate TE silencing, especially when the traditional 24 nt siRNA-directed RdDM pathway is unable to detect and silence newly inserted or escaped TEs (Cuerda-Gil & Slotkin, 2016; B. Liu & Zhao, 2023; X. Zheng et al., 2007). Once initial methylation of the TE is achieved, it undergoes canonical RdDM to enhance its silencing (Matzke & Mosher, 2014).

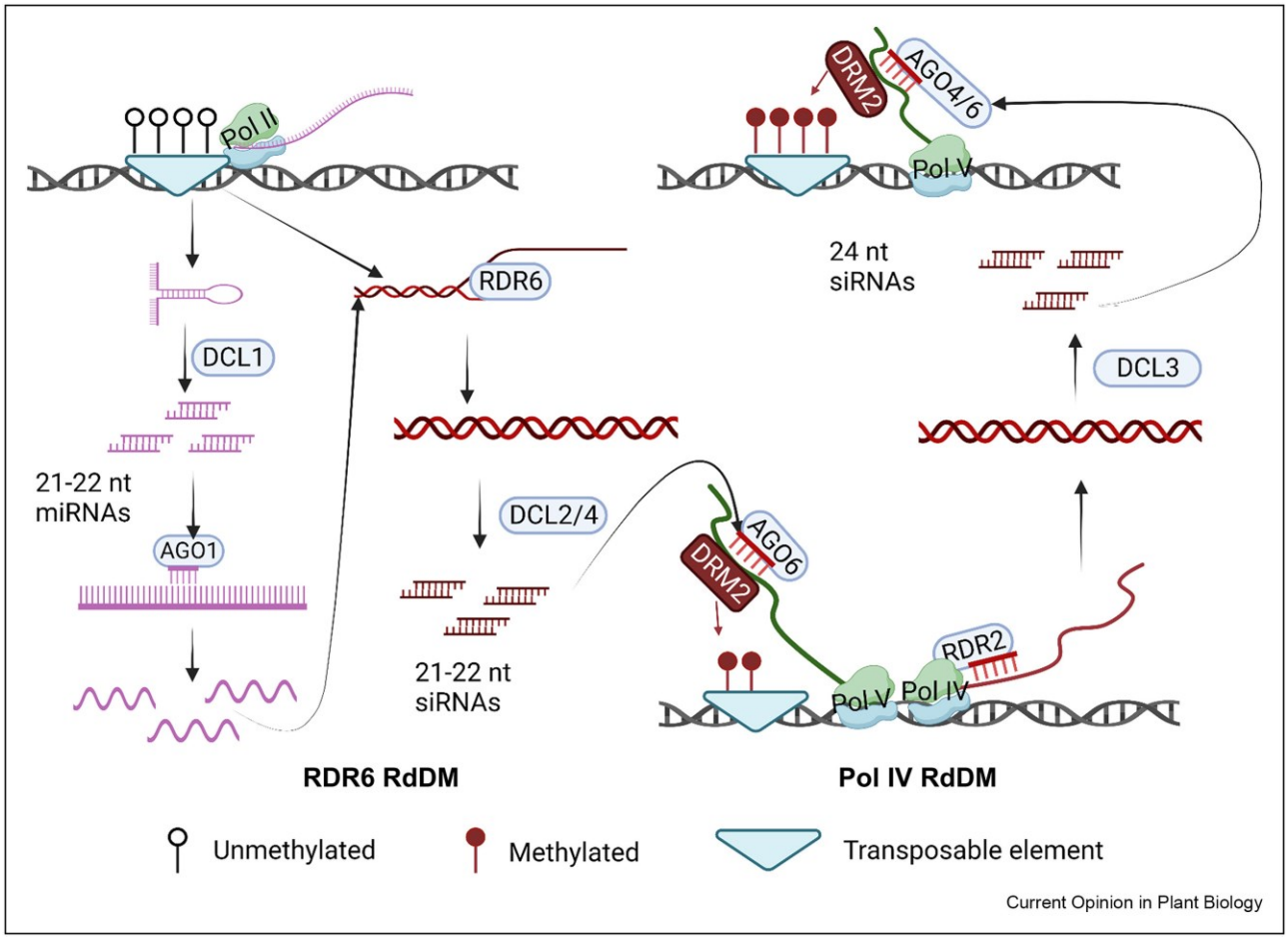

Figure 7. Initiation and establishment of TE silencing through RDR6 and Pol IV RdDM pathways. Active TEs can be transcribed by Pol II and copied into double-stranded RNA by RDR6, which are then cleaved into 21 or 22-nt siRNAs by DCL2/4. These 21 or 22- nt siRNAs, loaded on AGO6, target Pol V scaffolding transcripts in the nucleus, initiating RdDM. Once a low level of methylation is established at the TE, Pol IV is recruited to generate 24-nt siRNAs to establish canonical RdDM (B. Liu & Zhao, 2023).

In *A. thaliana*, proliferation of an introduced retrotransposon triggers RdDM after it has inserted into the genome (Pérez-Hormaeche et al., 2008; Sigman et al., 2021). Similarly, an increase in the copy number of a reactivated endogenous element attracts the RdDM machinery (Marí-Ordóñez et al., 2013). Compared to the Decrease in DNA Methylation 1 (DDM1) pathway, a chromatin remodeler important for maintaining DNA methylation and silencing TEs in heterochromatin (Zemach et al., 2013), RdDM plays a lesser role in DNA methylation of TEs. One study found that more than 2,000 TEs were reactivated in a *ddm1* mutant, while about forty were reactivated in an RdDM-defective mutant (Zemach et al., 2013). This is consistent with previous findings of infrequent TE mobilization in RdDM mutants compared to *ddm1* and *met1* mutants (Ito et al., 2011). The limited activation of endogenous TEs when RdDM is compromised is likely due to the independent maintenance of CG

and CHG methylation (Zemach et al., 2013). In some cases, TEs may require an environmental trigger for activation (Fig. 8) (Chang et al., 2020; Chinnusamy et al., 2014; Ito, 2022; J. M. Kim et al., 2015; Lang-Mladek et al., 2010; Ramakrishnan et al., 2022; Ruiz-Ferrer & Voinnet, 2009; Sun et al., 2020). For example, the retrotransposon *ONSEN* is activated during heat stress, an effect that is enhanced in RdDM-deficient mutants (Hayashi et al., 2020; Ito et al., 2013; Nozawa et al., 2022). Interestingly, although *ONSEN* is transcriptionally active and the presence of extrachromosomal copies suggests reverse transcription, new genomic insertions by retrotransposition occur only in mutants lacking either Pol IV or RDR2 (Fig. 8) (Hayashi et al., 2020). This suggests that the RdDM-associated small RNAs have additional roles in genome defense beyond transcriptional gene silencing (TGS) pathway.

Reduced functionality of the RdDM pathway during stress may result in decreased control of TEs (Ito, 2022; Matzke & Mosher, 2014; Ramakrishnan et al., 2021). TEs can affect gene expression locally and even at distant genomic locations, affecting transcriptional and post-transcriptional levels of gene regulation (Fig. 8) (Casacuberta & González, 2013; Makarevitch et al., 2015; Quesneville, 2020). TEs can affect gene expression by providing cis-regulatory regions with intrinsic regulatory features that can control the expression of nearby genes (Nosaka et al., 2012; Sun et al., 2020). TEs contain regulatory elements such as promoters and enhancers that can act as binding sites for trans-acting factors, thereby modulating the expression of neighboring genes. These TE-derived cis-regulatory regions contribute to the evolutionary novelty of gene regulation and may have tissue-specific functions, adding another layer of complexity to the control of gene expression.

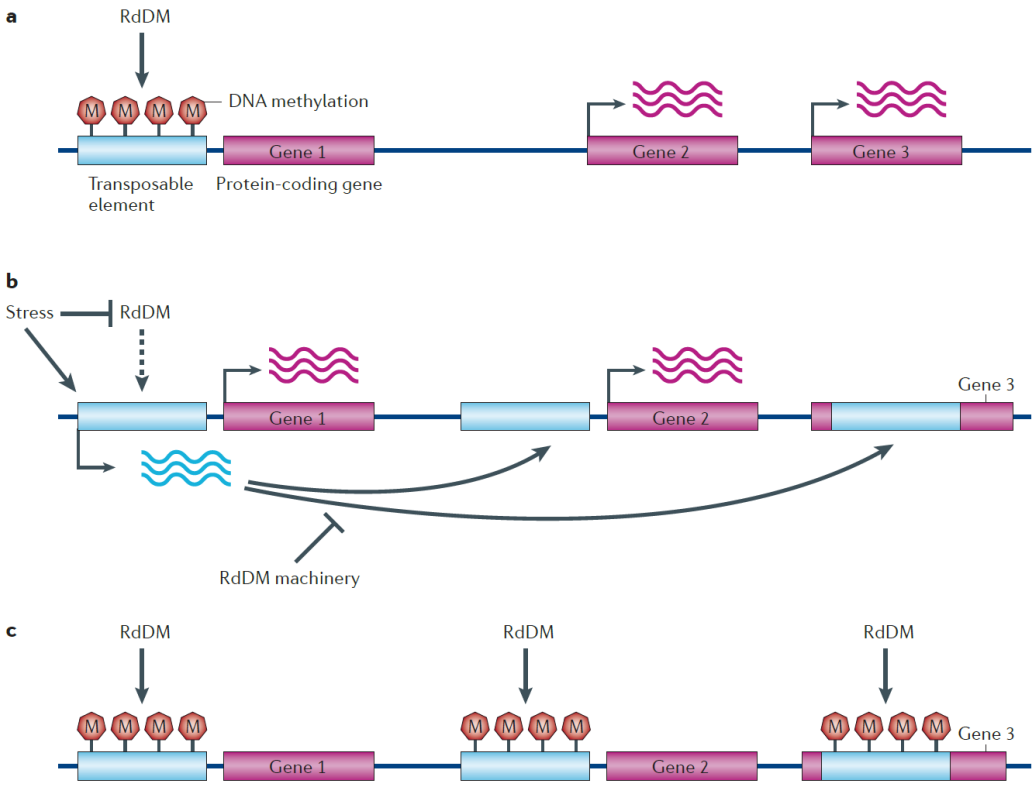

Figure 8. Release of RdDM during stress. **A.** A schematic locus is shown, in which a silenced TE leads to the silencing of a nearby protein-coding gene (Gene 1). By contrast, the activity of other protein-coding genes (Gene 2 and Gene 3) is not under control of the TE. **B.** TEs are activated during biotic and abiotic stress responses through a combination of loss of RdDM and transcriptional responses to stress. The nearby Gene 1 is also activated owing to loss of methylation in promoter regions. Reactivated TEs can integrate into new genomic loci, although the RdDM machinery inhibits the reinsertion of some elements through an unknown mechanism. **C.** New TE insertions can establish stress-responsive transcription at additional protein-coding genes (Gene 2) or might permanently disrupt gene function (Gene 3) (Matzke & Mosher, 2014).

5 MATERIALS AND METHODS

5.1 Plant material and growth conditions

In this study, *Arabidopsis thaliana* plants were grown after incubation for three days at 4°C, transgenic, mutant, and wild-type Columbia (Col-0) plants were germinated at 21°C on MS agar medium supplemented with 1% sucrose with or without the presence of appropriate antibiotics, respectively. Seedlings were initially grown in a controlled environment with 8 hours of light and 16 hours of dark at a temperature of 21°C, and after 3 weeks they were transferred to 16 h light and 8 h dark at 21°C.

The *ago4-3* mutant, which is homozygous, was derived from the WiscDSLox338A06.0 T0 line (Havecker et al., 2010). Periodic checks were performed on this line by growing its seeds on MS agar plates with phosphinothricin (10 mg/L), evaluating the AtAGO4 protein level using an anti-AtAGO4 antibody (Agrisera, AS09 617), and performing PCR amplification of the T-DNA region inserted approximately 170 bp before translation initiation (Fig. S1A-B-C). Seedlings, leaves, and inflorescences at different stages were collected for DNA and RNA extraction.

For the heat stress experiment, 1-week-old *Arabidopsis* seedlings grown on agar plates (1/2 MS, 0.8% agar, and 0.5% sucrose) were transferred to a growth chamber set to 37°C for 24 h. Light conditions were the same as for control plants (21°C), following a modified protocol described previously (Szádeczky-Kardoss et al., 2022). These seedlings were then harvested in bulk for DNA and RNA extraction.

Nicotiana benthamiana plants were grown under a 16-h photoperiod at 21°C and the plants were used for agroinfiltration at the 3-leaf stage.

Hordeum vulgare cv. Golden Promise was grown under a 16-hour light (400 µmol/m^2/s) and 8-hour dark photoperiod at temperatures of 20°C during the day and 16°C at night. The main tillers of these plants, which were about 3 months old, were dissected to collect the developing inflorescences. Only inflorescences between 15 and 25 mm in length (white anther stage) were collected for RNA extraction.

5.2 Phylogenetic analysis and i*n silico* predictions

The full peptide sequences of AGO proteins from *Arabidopsis thaliana*, *Oryza sativa ssp. japonica* and *Hordeum vulgare* were retrieved from Ensembl Plants and Uniprot databases. The sequences were aligned using ClustalW software and used to construct a phylogenetic tree with the neighbor-joining method (Saitou & Nei, 1987). The percentage of replicate trees in which the associated taxa clustered together in bootstrap test (1000 replicates) is shown above the branches. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Evolutionary distances were calculated using the Poisson correction method. Evolutionary analyses were conducted in MEGA11 (Tamura et al., 2021).

The miRNA sequences were used to predict the target site on the *AGO1* gene cDNAs using psRNATarget (version 2017). The psRNATarget tool was accessed through its web interface (https://www.zhaolab.org/psRNATarget/). The miRNA sequences were entered in FASTA format, and the corresponding transcript sequences of *Arabidopsis thaliana* and *Hordeum vulgare* were retrieved from the TAIR (The *Arabidopsis* Information Resource) and EnsemblPlants databases, respectively, and uploaded in FASTA format. The scoring scheme settings were the default (Schema V2), with different values for mismatches, wobble base pairs (G-U), and gaps.

5.3 Plasmid construction and plant transformation

All plant expression plasmids were constructed using the pGreen 0029 vector (kanamycin resistance) (www.pgreen.ac.uk) and all the amplifications were performed using Phusion Hot Start II DNA polymerase (Thermo Fisher Scientific) following the manufacturer's instructions (primer sequences in Supplementary Table S2).

The miRNA constructs for *MIR168a*, *MIR168b*, *MIR156a*, and *MIR171a* incorporated a 10-base pair (bp) flanking region on both sides of their respective stem-loop structures. For the artificial microRNA (amiRNA) constructs, the modified *hvu-MIR171* stem-loop was used, as previously described (Kis et al., 2016). PCR mutagenesis was used to introduce specific nucleotide changes within the passenger strand of the *MIR168* stem-loop. Additionally, to convert the duplex portion of *hvu-MIR171a* to *MIR168a* and its corresponding passenger strand. All the constructs were inserted in a pGreen0029

containing a 35S expression cassette. For the *AGO1* sensor, a 558 bp cDNA fragment containing the miR168 target site was amplified and annealed to the 5′ portion of the *GFP* in-frame.

To investigate heterologous complementation of barley AGO4 proteins, vectors were developed comprising the *AtAGO4* promoter (~2.5 kb) and terminator (~0.5 kb) sequences, encompassing the 5' UTR and 3' UTR sequences, in addition to the coding sequences (cDNAs) of *HvAGO4a* (2766 bp) and *HvAGO4b* (2757 bp). The ATG start codon was removed from both cDNAs. An HA epitope tag with its own start codon and a MluI restriction site were introduced at the 5' end. Similarly, a TAA stop codon and XbaI restriction site were added downstream of the existing TGA stop codon at the 3' end. The pGreen0029 vector was used for directional cloning. The *AtAGO4* promoter was first ligated into the vector, followed by the terminator, and finally the modified cDNAs using the specific restriction sites.

Initially, the plasmids were introduced into *Escherichia coli* DH5α competent cells through the heat shock transformation technique. Then, *E. coli* were cultivated on LB plates and in liquid media supplemented with 50 mg/L kanamycin. Subsequently, the plasmids were isolated, and their sequences were validated through sequencing procedures. All plasmids were transformed into *Agrobacterium tumefaciens* AGL1 strain with electroporation (360 Ω, 25 μF and 2,5 kV) in the presence of pSoup helper plasmid. *A. tumefaciens* was grown on YEB plates and liquid media containing 25 mg/L rifampicin and 50 mg/L kanamycin. A single colony was selected for the transformation of *Arabidopsis* Col-0 or *ago4-3* mutant (phosphinothricin-resistant) plants using the floral dip method (X. Zhang et al., 2006). The developing *Arabidopsis* inflorescences were dipped into the *A. tumefaciens* suspension containing 5% sucrose and 0.05% (vol/vol) Silwet L-77 for 1 min. Transformed plants were screened on agar plates (1/2 MS, 0.8% agar and 0.5% sucrose) supplemented with kanamycin (50 mg/L) and phosphinothricin (10 mg/L) as required.

5.4 Transient assay

Six-week-old *Nicotiana benthamiana* leaves were infiltrated with a specific mixture of *Agrobacterium tumefaciens* (AGL1) suspensions. These suspensions had an optical density of 1.0 at 600 nm [OD600] and contained sensor, miRNA-producing, and P14 constructs, as previously described (Vargason et al., 2003). P14, which acts as an inhibitor of the siRNA pathway, was consistently included in each

experiment at a standard concentration (Szádeczky-Kardoss et al., 2018). This ensured that it did not interfere with the monitoring of relative signal loss. To ensure consistent miRNA production across constructs, normalized amounts were used and the mixtures were supplemented with *A. tumefaciens* (AGL1) containing the empty pGreen0029 vector (Dalmadi et al., 2021).

Samples were collected on the third day after infiltration; four 1 cm diameter discs were pooled from different leaf areas for each construct. Sampling was performed simultaneously from both sides of the same leaf. Each miRNA-producing sensor construct combination was evaluated on four to five plants, with each experiment repeated at least three times.

5.5 RNA isolation and RT-qPCR

Total RNA was extracted from *A. thaliana* seedlings, young rosette leaves, mixed-stage flowers and barley developing inflorescences, using the standard phenol-chloroform method as described previously (Dalmadi et al., 2021). Briefly, the samples were collected both fresh or frozen and homogenized in an ice-cold mortar with 650 µL of extraction buffer (0.1 M glycine–NaOH, pH 9.0, 100 mM NaCl, 10 mM EDTA, 2% SDS). The homogenate was then subjected to a series of centrifugation steps, following a standard phenol-chloroform method. The RNA precipitation step was performed in 1 mL of ethanol added with 20 µL of 4 M Na-acetate for 30 minutes in an -80°C ultra-low-temperature freezer.

For RT-qPCR assays, 4 µg total RNA was treated with DNaseI, re-isolated by the phenol–chloroform method, and resuspended in sterile water. 2 µg of DNaseI-treated total RNA and a combination of random hexamer primers and oligo(dT)$_{18}$ primers were used to synthesize first-strand cDNA using the Maxima H Minus First Strand cDNA Synthesis Kit with dsDNase (Thermo Fisher Scientific) according to the manufacturer's instructions. qPCR was performed using the Luminaris Color HiGreen qPCR Master Mix (Thermo Fisher Scientific) and performed on a LightCycler 96 Instrument (Roche) real-time PCR machine. Data were obtained from three independent biological replicates and were normalized to *AtUBC9*, *AtACT2*, and *AtPP2AA3* using LigthCycler 96 software ($2^{-\Delta\Delta Ct}$ method). Graphs and statistical analysis were generated using GraphPad Prism 8. For primer sequences, please see Supplementary Table S1.

5.6 Gel filtration assay

Size separation gel-filtration experiments were carried out using Superdex 200 10/300 column (Akta-FPLC, GE Healthcare) or the same size Sephacryl S-300 High Resolution (Pharmacia LKB) columns. The gel-filtration experiments were carried out as it was described previously (Lakatos et al., 2004) with minor modifications. In summary, plant tissues were homogenized in liquid nitrogen, adding 600 µl of elution buffer (comprising 50 mM Tris–HCl at pH 7.5, 10 mM NaCl, 5 mM $MgCl_2$, and 4 mM DTT) for 0.3 g of plant material. The resulting homogenized crude extracts were consistently maintained on ice and subjected to centrifugation three times, each for a duration of 10 minutes at 12,000 x g and a temperature of 4°C, to effectively remove any tissue debris. Following this, 200 µl of the purified crude extract were injected in the gel-filtration column that had been equilibrated with a cold buffer. The process of size-separation chromatography was conducted in a cold room maintained at 4°C. 48 fractions of gel-filtration were divided into two and RNA was extracted with phenol-chloroform method from the odd samples, while even samples were used for protein purification using acetone precipitation.

5.7 miRNA detection and quantification

Small RNA northern blot analyses were performed using either 4 µg of total RNA or gel-filtration samples. The samples were separated on denaturing 12% polyacrylamide gels containing 8 M urea, and then the gels were transferred to Hybond NX membrane (GE Healthcare) using Biometra Fastblot™ (Analytik Jena) semi-dry blotting system. Membranes were chemically cross-linked (Pall & Hamilton, 2008) and probed using biotinylated locked nucleic acid (LNA) oligonucleotide probes or DNA probes. Signal was detected using Chemiluminescent Nucleic Acid Detection Module Kit (ThermoFischer) and visualized with ChemiDoc™ MP Imaging System (Biorad).

For the analysis of gel-filtration blots, the volume intensity of the four most prominent RISC loaded and unbound fractions were measured and summarized. Loading efficiency (LE) was calculated as RISC loaded volume intensity divided with the total sum of RISC loaded plus unbound volume intensities and was represented as a percentage.

5.8 Protein extraction and western blotting

Arabidopsis seedlings, young rosette leaves and mixed-stage flowers were homogenized in extraction buffer (see Section 5.5) and an equal volume of 2 × Laemmli buffer (65.8 mM Tris-HCl, pH 6.8, 2.1% SDS, 26.3% (w/v) glycerol, 0.01% bromophenol blue) was added. Following this, samples were boiled for 5 min, and centrifuged at 20,000 × g at 4 °C for 5 min to remove debris.

Protein samples from *Arabidopsis* were loaded on an 8% or 10% SDS-polyacrylamide gel, subsequently transferred to an Amersham Hybond P 0.45 PVDF blotting membrane (manufactured by GE Healthcare) via wet tank transfer, and then subjected to western blot analysis. Antibodies were used in following concentrations: anti-AGO1 1:7500 (rabbit AS09 527, Agrisera), anti-HA-peroxidase 1:2000 (rat 3F10, Roche), anti-actin (plant) 1:2000 (mouse 10-B3, Sigma-Aldrich), anti-AtAGO4 1:5000 (rabbit AS09 617, Agrisera), and anti-BiP 1:10000 (rabbit AS09 481, Agrisera). The secondary antibodies used were goat anti-rabbit HRP-conjugated (for anti-AGO1, anti-AtAGO4 and anti-BiP, AS09 602, Agrisera) and goat anti-mouse HRP-conjugated (for anti-actin, A4416, Sigma-Aldrich) in the dilution of 1:10000. The blocking process was conducted using 5% milk powder in 1X Phosphate-Buffered Saline, 0.1% Tween® 20 Detergent (PBST) for a duration of 1 hour. Following this, the primary antibodies were incubated in a solution of 1% non-fat milk powder in PBST for a period ranging from 1 to 2.5 hours. Subsequently, the secondary antibodies, which were diluted with PBST, were incubated on the membrane for 1 hour. Blots were washed three times for 5 min with PBST between the two solutions and after the secondary antibody, and finally developed using High Clarity Western ECL (Biorad) on ChemiDoc™ MP Imaging System (Biorad). Volume intensity of the signal was quantified using Image Lab 6.1; protein signals were normalized to either actin or BiP.

5.9 Chop-qPCR analysis

In the Chop-PCR assay, the ZenoGene Plant DNA Purification Kit (ZENON Bio) was used to extract genomic DNA from the mixed-stage inflorescence of *Arabidopsis* T3, following the guidelines provided by the manufacturer. The amount of genomic DNA was determined using a Nanodrop ND-1000 spectrophotometer. The enzymatic digestion of DNA, sensitive to methylation, was carried out using MspJI (New England BioLabs). This enzyme, which is dependent on modifications, specifically

identifies cytosine modifications such as C5-methylation (5-mC) and C5-hydroxymethylation (5-hmC), and was employed to digest the genomic DNA (Dasgupta & Chaudhuri, 2019; H. Zhang et al., 2014; Y. Zheng et al., 2010). It can cleave asymmetric methylation sites (CHH), along with CpG and CHG regions.

The reaction was performed in a 30 µL reaction mix containing 10× rCutSmart™ Buffer (New England BioLabs), 1 µL of MspJI enzyme, 1 µL of Enzyme Activator Solution (New England BioLabs) and 600 ng of genomic DNA over a period of for 4 h at 37 °C. MspJI was omitted exclusively in the control reaction. Following heat inactivation, 1 µL of either digested or undigested genomic DNA served as a template for the Chop-qPCR assay. Measurements were prepared using the Luminaris Color HiGreen qPCR Master Mix (Thermo Fisher Scientific) and performed using a LightCycler 96 Instrument (Roche) real-time PCR machine. Data were obtained from three independent biological replicates and were normalized to undigested AtSN1 using LigthCycler 96 software. Graphs and statistical analysis were generated using GraphPad Prism 8. For primer sequences, please see Supplementary Table S1.

5.10 Relative copy number assessment of transposable elements

Genomic DNA was extracted from *Arabidopsis* non-treated (NT) and heat-stressed (HS) seedlings using the ZenoGene Plant DNA Purification Kit (ZENON Bio) according to the manufacturer's instructions. Quantification of genomic DNA was performed using a Nanodrop ND-1000 spectrophotometer.

The relative quantification of *Ty1/copia*-like retrotransposon *ATCOPIA78* (*ONSEN*) copies was performed using non-treated *Arabidopsis* wild-type DNA as a control, and in relation to its quantification consisting of eight *ONSEN* copies in the Col-0 ecotype genome (Hayashi et al., 2020; Ito et al., 2011, 2013; Nozawa et al., 2022). All the relative amounts of the other non-treated and heat-stressed samples were measured using qPCR. Analysis by qPCR was performed using the Luminaris Color HiGreen qPCR Master Mix (Thermo Fisher Scientific) with 20 ng of genomic DNA per reaction and the LightCycler 96 Instrument real-time PCR machine (Roche). Data were obtained from three independent biological replicates and were normalized to *AtUBC9* using LightCycler 96 software. For

primer sequences and the list of the eight *ONSEN* copies locus names, please see Supplementary Table S1.

5.11 RNA-seq of developing barley inflorescences

RNA extracted from developing barley inflorescences was quantified using the Qubit RNA HS Assay and a quality check was carried out using the LabChip GX Touch Nucleic Acid Analyzer with the DNA 5K/RNA/CZE chip, resulting in an RNA quality score of 10 in all analyzed samples.

The libraries and the RNA-seq were conducted in collaboration with Dr. Krisztián Frank and Dr. Taller János (MATE). Briefly, NEXTFLEX® Rapid Directional RNA-Seq Kit 2.0 was used to prepare libraries for Illumina sequencing instruments according to the manufacturer's protocol. Briefly, poly-A-containing mRNA was purified from 1250 ng total RNA using the Nextflex polyA beads 2.0 kit. The purified mRNA was fragmented and the first and second cDNA strands were synthesized. Adapter ligation (NextFlex Unique Dual Index Barcodes) and amplification by PCR were then performed. The XMark HT chip (Labchip GX Touch Nucleic Acid Analyzer) was used for the quantification and the quality control of the purified cDNA libraries, which were then normalized and pooled equimolarly. The flow cell was loaded onto the Illumina NovaSeq 6000 for paired-end sequencing using the S4 Reagent Kit v1.5 (300 cycles) according to the manufacturer's instructions.

Analysis was performed on paired-end fastq raw data with salmon (version 1.10.1) using the Morex V3 (GCA_904849725) reference genome, transcripts fasta file and a file containing a mapping of transcripts to genes (GTF). This analysis enabled the obtainment of the number of reads mapped on each transcript as well as the relative abundance of the transcript in Transcripts per Million (TPM).

5.12 Immunoprecipitation

For the preparation of crude extracts, 0.4 g of either seedlings, rosette leaves or mixed inflorescences were homogenized in a solution containing four volumes of lysis buffer (10 mM Tris–HCl pH 7.6; 1 mM EDTA; 150 mM NaCl; 10% glycerol; 0.5% Nonidet P-40; 5 mM NaF; 1 mM dithiothreitol; 0.5 mM Na3VO4; 1 mM phenylmethylsulfonyl fluoride). This mixture was then centrifuged three times

at 4°C, with the supernatant transferred to fresh tubes after each centrifugation to effectively remove any cellular debris. Then, the supernatant was processed using the Dynabeads Protein G Immunoprecipitation Kit (Thermo Fisher Scientific) with anti-AGO1 (Agrisera) or anti-HA-peroxidase antibody (rat 3F10, Roche) according to the manufacturer's instructions. The extracts were used for RNA purification by the phenol-chloroform method and for protein extraction by adding a volume of Laemmli buffer (2×) equal to the volume of the extract. Protein samples were used to check for the corresponding transgenic AGO4 by western blot using anti-HA-peroxidase antibody (rat 3F10, Roche). Contamination of total proteins was checked by visualization provided by the TGX technology of BioRad gels (Fig. S2A-B).

5.13 Small RNA library preparation and analysis

To construct cDNA libraries for sequencing, we used RNA samples obtained from bulked seedlings of transgenic lines overexpressing different miR168 precursors and from immunoprecipitation of complementation lines for HvAGO4A and HvAGO4B. Briefly, samples were loaded individually onto polyacrylamide gels, from which the small RNA fraction enriched in the 21-22 nt range was isolated. Libraries were then prepared exclusively from this fraction using the Truseq Small RNA Library Preparation Kit (Illumina) according to a modified protocol described previously (Czotter et al., 2018). The libraries were sequenced on an Illumina NextSeq 500 system using the NextSeq 500/550 v2.5 sequencing reagent kit. The sequencing mode was single-end 50 bp reads (UD-GenoMed).

Raw data of AtAGO4 sRNA-IP sequencing were retrieved from NCBI (SRX11482423, SRX11482424, SRX11482425) (Sigman et al., 2021). Analysis of the sRNA-IP sequencing was performed using the Galaxy platform (Afgan et al., 2022) to control quality, trim and map to the *A. thaliana* reference genome (TAIR10.1) using hisat2 (version 2.2.1) (D. Kim et al., 2015). The resulting reads were mapped on the different genomic sequences categories using the sRNAPipe pipeline (version 1.1.1) (Pogorelcnik et al., 2018): TEs, gene transcripts, microRNAs (miRNAs), small nuclear RNAs (snRNAs), ribosomal RNAs (rRNAs), and transfer RNAs (tRNAs). sRNAPipe allowed for the selection of the size range (18–27 nt) of the sRNAs and the generation of "bonafide" reads mapped on the genome, excluding reads that map miRNAs, rRNAs, tRNAs or snRNAs for the

normalization as RPM (Reads Per Million) and RPKM (Reads Per Kilobase per Million mapped reads).

Visualization of sRNAs mapped to chromosomes or loci was performed using IGV (Thorvaldsdóttir et al., 2013) after alignment. Protein alignments were found using ESPript 3 (Robert & Gouet, 2014). GraphPad Prism 8 was used to generate graphs and for statistical analysis.

6 RESULTS

6.1 Bioinformatic analysis of putative AGO genes in barley

To identify putative *AGO4* gene(s) and to clarify the phylogenetic linkage of barley AGOs with the previously described rice and *Arabidopsis* orthologs, detailed *in silico* analysis was carried out. The analysis of the whole barley genome (both Morex V3 and Golden Promise v1) identified 21 putative candidate genes belonging to the Argonaute clade using sequences from *A. thaliana* and *O. sativa*. Following a BLAST search on Ensembl Plants, the translated protein sequences of these genes were analyzed using InterPro (Paysan-Lafosse et al., 2023) to determine the presence of characteristic domains of functional AGO proteins. To keep the nomenclature consistent, the corresponding proteins were named based on the phylogenetic relationship to their *Arabidopsis* and rice counterparts.

According to the phylogenetic tree generated from the AGO proteins of *A. thaliana*, *O. sativa ssp. japonica*, and *H. vulgare*, 3 main clades were distinguished: AGO1/5/10, AGO2/3/7, and AGO4/6/9 (Zhang et al., 2015) (Fig. 9). In addition, there is also AGO18, previously found only in grasses as an extra subclade, which has a significantly different mid-domain and N-terminal ends compared to other AGOs, for this reason it is not inserted in any other clade (Das et al., 2020). The *Arabidopsis* genome encodes 10 AGO proteins, named AGO1 to AGO10, but this number is higher in the two other species studied here: 19 in rice (Zhang et al., 2015) and 21 in barley (Table S3). Previously, it has been shown that *AGO* genes have duplicated and diverged from this set, acquiring new or specific functions and different expression pathways. Within the AGO1/5/10 clade, there is an expansion in terms of size compared to *A. thaliana*, specifically talking about the gene duplication of *AGO1* and *AGO5*. For example, if we consider *AGO1* in *A. thaliana*, which has a single copy, there are four copies in rice and five copies in barley. A similar situation can be seen for the AGO5 subclade, specifically with MEL1, which gains germline-specific expression and the function to bind phasiRNAs. In contrast, the other two clades remained more similar in size between *Arabidopsis* and rice or barley. Clade AGO2/3/7 is quite similar between the three different plants, representing a stable and conserved set of proteins that have maintained their original function over time.

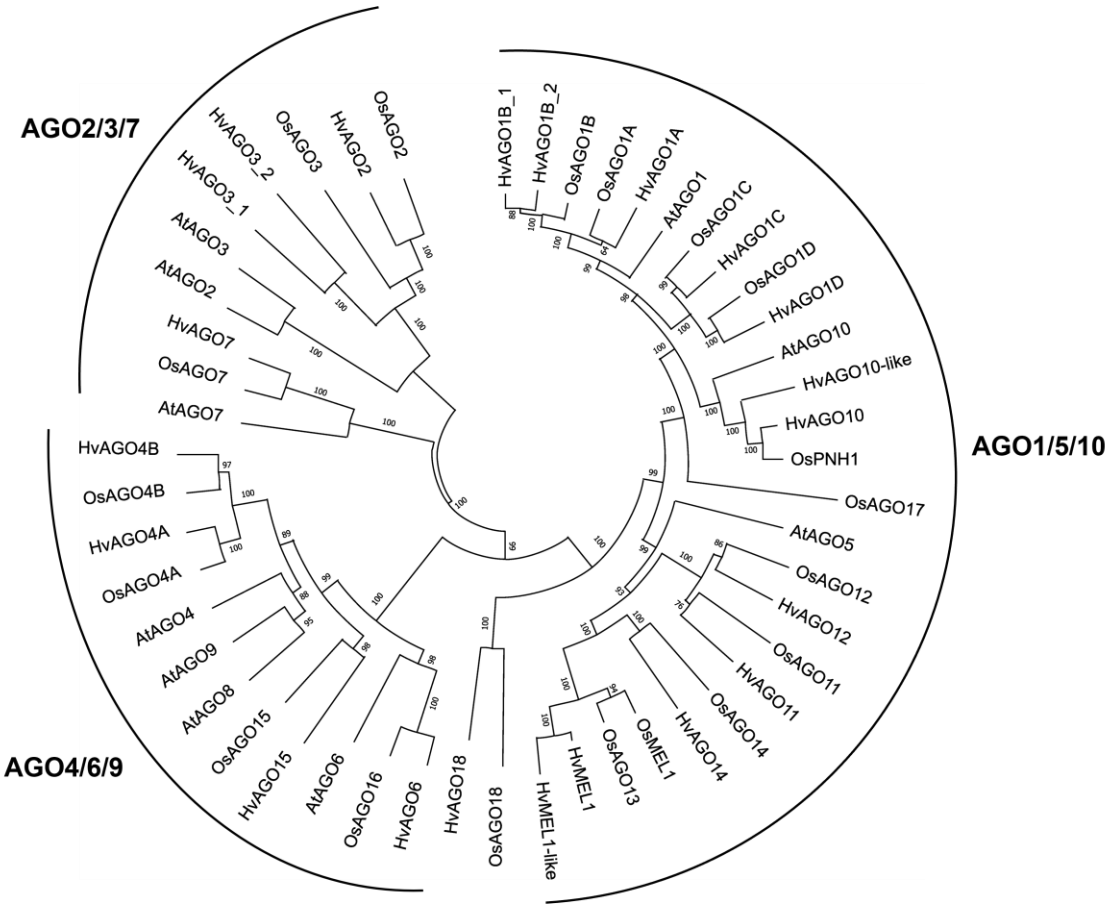

Figure 9. Phylogenetic tree of AGO proteins in *A. thaliana*, *O. sativa* and *H. vulgare* was inferred using Maximum Likelihood method (1000 bootstrap repetitions) and JTT matrix-based model. The names of the proteins are based on the rice nomenclature, except for AGO6 and AGO10, where the names are taken from the *Arabidopsis* counterpart. The phylogenetic tree is divided into 3 clades: AGO1/5/10, AGO2/3/7, and AGO4/6/9. In addition, there is the subclade AGO18, which belongs to the AGO5/MEL1 subclade, but is considered to be independent because it is exclusive to grasses.

6.1.1 Barley *AGO1* genes

In barley, five different AGO1 orthologous proteins were identified, mirroring the findings in rice, which has 4 copies of AGO1 (Wu et al., 2009, 2010). To investigate the gene expression of *AGO1* in barley, the BarleyExpDB database was analyzed (T. Li et al., 2023), focusing on the differential gene expression in different plant tissues (Fig. 10A). Each of the five *AGO1* genes showed a different expression pattern. In particular, *HvAGO1B_1* emerged as the most highly expressed gene in inflorescences at both 5 mm (INF1) and 1.5 cm (INF2). Interestingly, the two *HvAGO1B* genes showed similar expression patterns in all tissues except inflorescences, where *HvAGO1B_1* showed more than twice the expression level of its counterpart. *HvAGO1A* showed consistently low gene expression in all tissues, notably lacking expression in embryo (EMB) and inflorescence, where the

other four *AGO1* genes showed higher expression levels. *HvAGO1C* showed low but widespread expression in all tissues, with levels comparable to *HvAGO1B_2* in INF2. Conversely, *HvAGO1D* showed minimal expression in most tissues, with low expression in INF1, but emerged as the second most expressed *AGO1* gene in INF2. This indicates a precise and specific expression pattern in a single tissue at a particular stage of development.

Following the expression pattern analysis, it was also investigated whether the target site of miR168 is conserved among the five *AGO1* genes in barley, and thus whether it plays a role in modulating the expression of these genes in different tissues. Using psRNATarget, the target site of miR168, whose sequence is also conserved in barley, was predicted in the barley *AGO1* genes together with the *Arabidopsis AGO1* (Fig. 10B). The predictions indicate an expectation value, which is the penalty for mismatches between the mature small RNA and the target sequence. A higher value indicates less similarity (and possibility) between the small RNA and the candidate target. In the case of *AtAGO1*, which is targeted by miR168 resulting in mRNA cleavage, it has an expectation value of 3.0 with two mismatches (at the positions 2 and 14 of miR168) and one wobble base pair (G-U) at 8th nucleotide of the miRNA. For barley *AGO1*, *HvAGO1C* has an expectation value of 2.0 with two mismatches (at positions 2 and 21) and one wobble base pair at position 9. Other *AGO1* genes have expectation values between 3.0 and 3.5. Notably, a consistent mismatch at position 14, as seen in *AtAGO1*, is observed in all other barley *AGO1* genes. *HvAGO1D* also retains a mismatch at position 2 like *AtAGO1*. Similar to *HvAGO1C*, *HvAGO1D* has one mismatch at position 21 and one wobble base pair at position 9. However, *HvAGO1A*, *HvAGO1B_1* and *HvAGO1B_2* share the same alignment with miR168. They have two mismatches at positions 1 and 2 and the conserved mismatch at position 14 of *AtAGO1*, but no wobble base pairs (Fig. 10B).

Furthermore, all AGO1 proteins from *Arabidopsis*, rice and barley were compared by alignment to identify similarities and differences between the proteins. It is interesting to observe how many amino acids are conserved between different plants and different proteins, such as in the region of AGO1 involved in binding to the 5' end of miRNAs (Fig. 10C), where entire regions are observed with complete amino acid identity between all proteins (regions 725-738 and 753-765). Focusing on the amino acids responsible for effective binding to the 5' end of miRNAs, seven single amino acids and one site consisting of four amino acids (QCCx) were identified. By analyzing the differences between the proteins, the alignment shows that the proteins can be divided into two groups: AGO1_QCCL and AGO1_QCCC. In fact, the proteins belonging to the first group AGO1_QCCL, such as AtAGO1, OsAGO1A, OsAGO1B, HvAGO1A and HvAGO1B, conserved all the sites for binding to the 5' end

of miRNAs. While the proteins in rice and barley AGO1C and AGO1D belong to the AGO1_QCCC group and show differences in three different binding amino acids, L708C, H711Q and Y719I (Fig. 10C). The conserved amino acid differences are likely to play an important role in the structure and function of the proteins, especially in their interaction with miRNAs. These differences, based on the side chains and properties of the amino acids, for example the presence of tyrosine and histidine in QCCL proteins, could enhance these interactions compared to the more hydrophobic and less interactive residues (isoleucine and glutamine) in QCCC proteins. All these variations can alter the conformation, stability, and binding affinity of the proteins, leading to potential functional divergence.

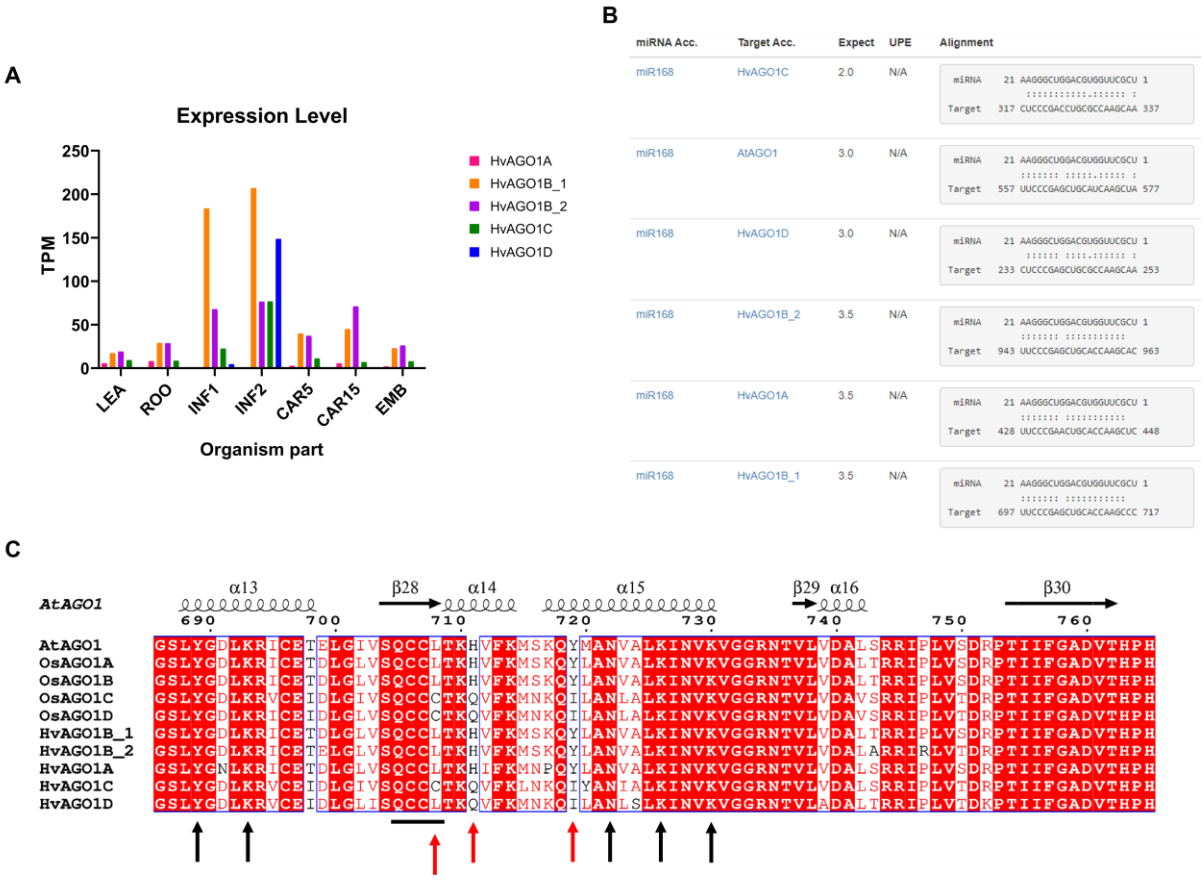

Figure 10. A Expression levels in Transcript Per Million (TPM) retrieved from BarleyExpDB using PRJEB14349 RNA-seq data as samples. The different colors indicate the genes and on the x-axis is the organism part of barley. LEA = Shoot from the seedlings 10cm shoot stage, ROO = Roots from the seedlings 10cm shoot stage, INF1 = Young Inflorescences 5mm, INF2 = Inflorescences 1-1.5cm, CAR5 = Grain with removed bracts 5DPA, CAR15 = Grain with removed bracts 15DPA, EMB = Embryos 4d dissected from germinating grains. **B** Prediction of miR168 target sites for *AGO1* mRNAs in *Arabidopsis* and barley using psRNATarget with default parameters. **C** Protein alignment of the PIWI domain region involved in the 5′ miRNA anchoring of the *Arabidopsis*, rice, and barley AGO1 proteins. Alignment was performed with ClustalW and was visualized using ESPript. The vertical black arrows and the black line indicate the amino acids directly involved in the binding domain, while the red arrow indicates the AA showing variations between *Arabidopsis* (dicot) and the monocots.

6.1.2 Structure and expression analysis of putative *AGO4* genes in barley

A focus on the AGO4/6/9 clade shows how AGO6 did not undergo gene duplication, while AGO4 did, with three different genes: *AGO4a*, *AGO4b* and *AGO15* (Fig. 11A). The similarities to rice and the maintenance of the expression pathway underlines the importance of the study of this gene duplication in more detail in barley. To assess the expression levels of the three barley genes closely associated with the AGO4 clade, RNA-seq was performed on developing barley inflorescences. Within this clade, *HvAGO4a* (HORVU.MOREX.r3.3HG0256890) showed the most pronounced expression, followed by *HvAGO4b* (HORVU.MOREX.r3.1HG0095310), while *HvAGO15* (HORVU.MOREX.r3.7HG0736980) showed an exceptionally low level of expression. Interestingly, the expression of *HvAGO6* (HORVU.MOREX.r3.5HG0468930) was significantly lower compared to the two *AGO4* genes (Fig. 11B). The putative orthologous AGO4 proteins from barley showed higher identity to the corresponding rice proteins (85% and 82% for AGO4A and AGO4B, respectively) than to each other (barley 76% and rice 79%). Sequence analysis revealed that HvAGO4A and HvAGO4B contained 921 and 918 amino acids, respectively, and showed an identical gene structure with respect to the number of coding exons and the position of the PAZ and PIWI domains (Fig. 11C). Specifically, the PIWI domains of the three AGO4 proteins from barley, *Arabidopsis*, and rice were analyzed to highlight their similarities and differences (Fig. 11D). Within the subdomains that define the PIWI domain, one region is involved in anchoring the 5′ end of sRNAs. A comparison of the AGO4 proteins in this subdomain showed a significant level of conservation, with almost all amino acids involved in anchoring the 5′ end of sRNAs being strictly conserved across the three plant species. However, an interesting deviation was observed at the site consisting of four amino acids (QCxA), where the monocot protein sequences showed a single amino acid change that was not present in the *Arabidopsis* version, highlighting the inherent differences between the two AGO4 proteins. Interestingly, the identified barley protein sequences share this feature with their rice counterparts, suggesting a similar function for HvAGO4A and HvAGO4B.

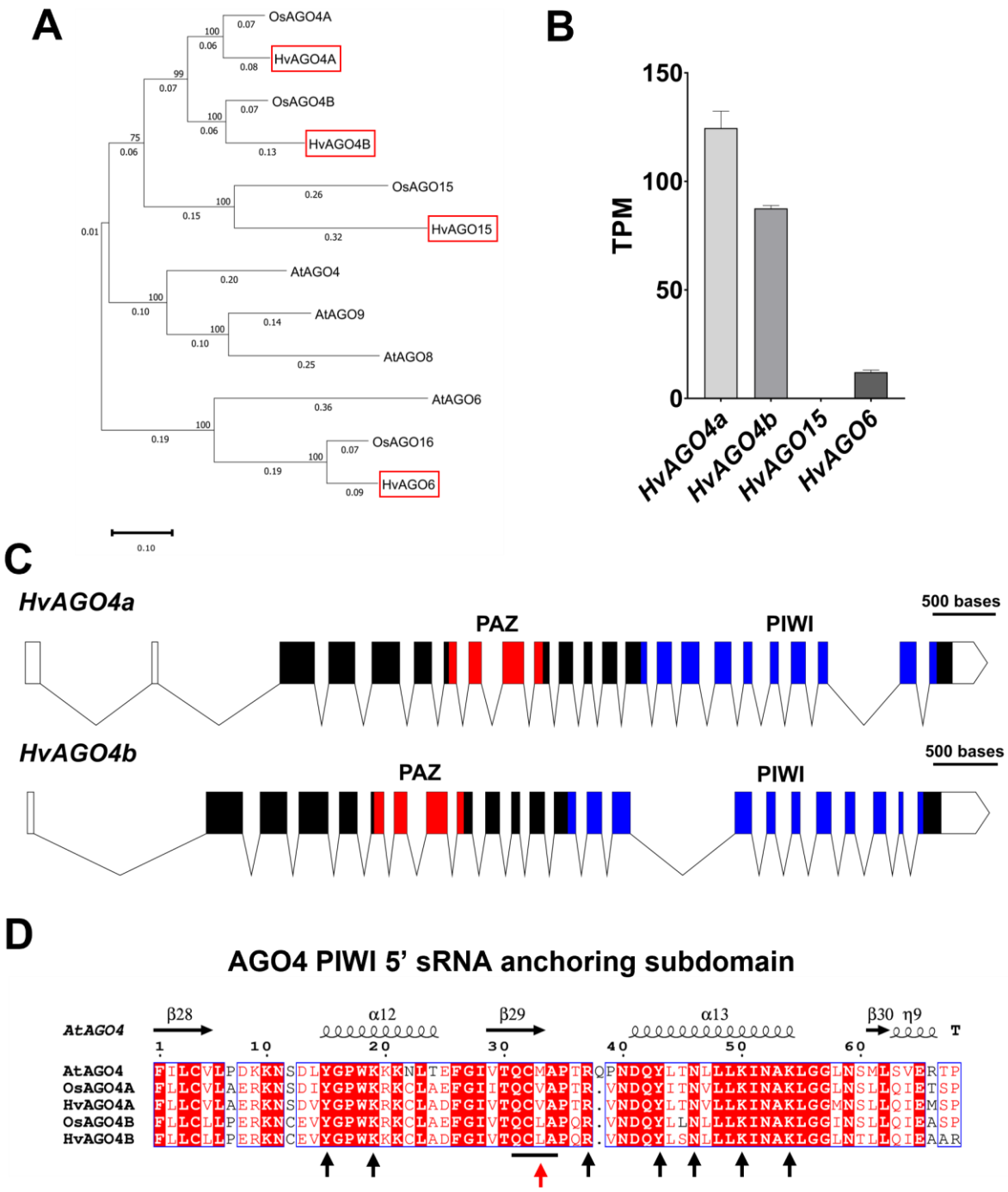

Figure 11. A Phylogenetic tree of AGO4-clade protein sequences from *Arabidopsis thaliana*, *Oryza sativa* subsp. *japonica* and *Hordeum vulgare* was inferred using Maximum Likelihood method (1000 bootstrap repetitions) and JTT matrix-based model. Barley AGO4-like translated protein sequences were marked with a red square. **B** Transcript abundancy of putative AGO4-clade genes in *H. vulgare* cv. Golden Promise (developing inflorescences) expressed in TPM (calculated using Salmon). Error bars represent the mean ± SD, n = 3. **C** Visualization of the *HvAGO4a* and *HvAGO4b* gene structure, including the 5′ and 3′ UTR. Solid rectangles and lines indicate exons of the coding regions and introns, respectively. The UTR regions are marked with empty shapes. Within the coding region PAZ and PIWI domains were labeled with red and blue, respectively. The scale bar represents 500 bp. **D** Protein alignment of the PIWI domain region involved in the 5′ sRNA anchoring of the *Arabidopsis*, rice, and barley AGO4 proteins. Alignment was performed with ClustalW and was visualized using ESPript. The vertical black arrows and the black line indicate the amino acids directly involved in the binding domain, while the red arrow indicates the AA showing variations between *Arabidopsis* (dicot) and the monocots.

6.1.3 PIWI structure of AGO6 in barley and rice

An analogous analysis was performed for AGO6, which revealed that the anchor site for the 5′ end of the sRNAs is located at the same position as in AGO4 (α12-β29-α13). Furthermore, this analysis revealed a significant conservation among the sites across different plants, with the exception of one amino acid in the four-amino acid site (QCIx), which represented a variation at a different position compared to that observed in AGO4 (Fig. 12).

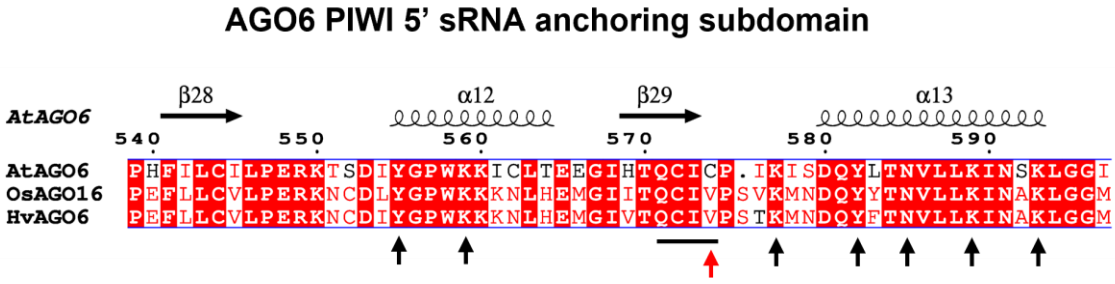

Figure 12. Protein alignment of the PIWI domain region involved in the 5' sRNA anchoring of the *Arabidopsis*, rice and barley AGO6 proteins. Alignment was performed with ClustalW and was visualized using ESPript. The vertical black arrows and the black line indicate the amino acids directly involved in the binding domain, while the red arrow indicates the AA showing variations between *Arabidopsis* (dicot) and the monocots.

6.1.4 *HvAGO15* as a pseudogene

Unlike rice, the *AGO15* gene in barley does not appear to be a product of an *AGO4a* tandem duplication, given its location on a separate chromosome and the absence of intronic TEs (Wu et al., 2010). However, similar to rice, the expression level of *HvAGO15* was undetectable, suggesting that it may be a pseudogene or that its expression may be restricted to certain conditions and/or tissue types. Upon examining this gene, we found six tandem repeats in exon 1, aligned with the coding region and potentially different start codons (Fig. 13). In addition, we faced challenges in amplifying *HvAGO15* as a complete gene from both leaf and inflorescence through PCR, resulting in only fragments, implying its pseudogene status.

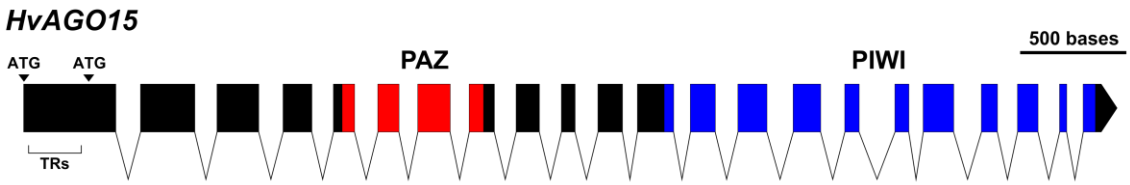

Figure 13. Visualization of the HvAGO15 gene structure. Solid rectangles and lines indicate exons of the coding regions and the introns, respectively. The different possible transcriptional starts are indicated with ATGs, and the region with possible regulatory elements to control expression is marked with TRs (tandem repeats). Within the coding region PAZ and PIWI domains were labelled with red and blue, respectively. The scale bar represents 500 bp.

6.2 Investigation of mechanisms behind the AGO1-miR168 feedback regulatory loop

As mentioned before AGO1 is under the control of miR168 in an autoregulatory loop, and this miRNA was defined as a low loaded sRNA, and accumulates mostly as AGO un-loaded free miRNA in the cytoplasm (Dalmadi et al. 2019). Previous data suggested a regulatory function of miRNA precursor structure in the determination of the miR168 loading efficiency. In the current work we wanted to investigate the role of miRNA precursor elements, like duplex structure, on the loading efficiency of miR168. First, we examined the efficiency of AGO1-loading of miR168, which was overproduced from the wild-type *A. thaliana MIR168a* precursor fragment in *N. benthamiana* leaves infiltrated with *Agrobacterium*. A sensor expressing a fragment of *AGO1* containing the miR168 target site fused to *GFP* was designed (Fig. 14A). The *AGO1* sensor was transiently expressed in *N. benthamiana* leaves, either in the presence or absence of the 35S::*pri-MIR168a* (*MIR168a*) precursor binary construct (Fig. 14B). To prevent siRNA-mediated transgene-induced RNAi, we included a viral P14 silencing suppressor construct in the *Agrobacterium* infiltration mix. Despite the strong overexpression of miR168, it resulted in only a modest reduction of the GFP signal of the AGO1 sensor compared to the empty vector control infiltration. This observation was further supported by western blot analysis.

To further investigate the impact of miR168 overexpression on AGO1 accumulation, transgenic *Arabidopsis* plants were generated to overexpress the *ath-MIR168a* precursor fragment (*MIR168a*) (Fig. 14C). Consistent with previous research (Vaucheret et al., 2006), transgenic overexpression of *MIR168a* induced only subtle phenotypic changes. The general appearance and fertility of the transgenic plants were largely similar to those of the wild-type plants. We observed only minor developmental differences in the transgenic lines, such as serrated rosette leaves and delayed flowering (Fig. 14C). Interestingly, the intensity of these phenotypes was directly proportional to the

level of miR168 accumulation. Indeed, even in transgenic lines, only a moderate downregulation of AGO1 protein levels was observed compared to wild-type plants. Interestingly, the level of accumulation of miR159 was also slightly decreased.

To investigate the efficiency of miR168 loading into AGO1-RISC, a size separation gel filtration assay was used (Fig. 14D). miR159 was predominantly found in high molecular weight (HMW) AGO1-RISC complexes. On the other hand, miR168 was primarily accumulated in a protein-unbound form in the same sample, with only a small fraction of miR168 being loaded into HMW AGO1-RISC. Even with increased levels of miR168, there was only a modest increase in HMW AGO1-RISC loading of miR168. This suggests that the substantial excess of miR168 maturing from the overexpressed wild-type *MIR168a* precursor fragment is not efficiently incorporated into AGO1-RISC. As a result, the limited loading efficiency of miR168 resulted in only a small decrease in AGO1 protein levels. In summary, the finding that miR168 overaccumulation mediated by the *MIR168a* precursor fragment does not correspond to a significant increase in AGO1-RISC loading suggests that the loading efficiency of miR168 into AGO1 is under strict regulation.

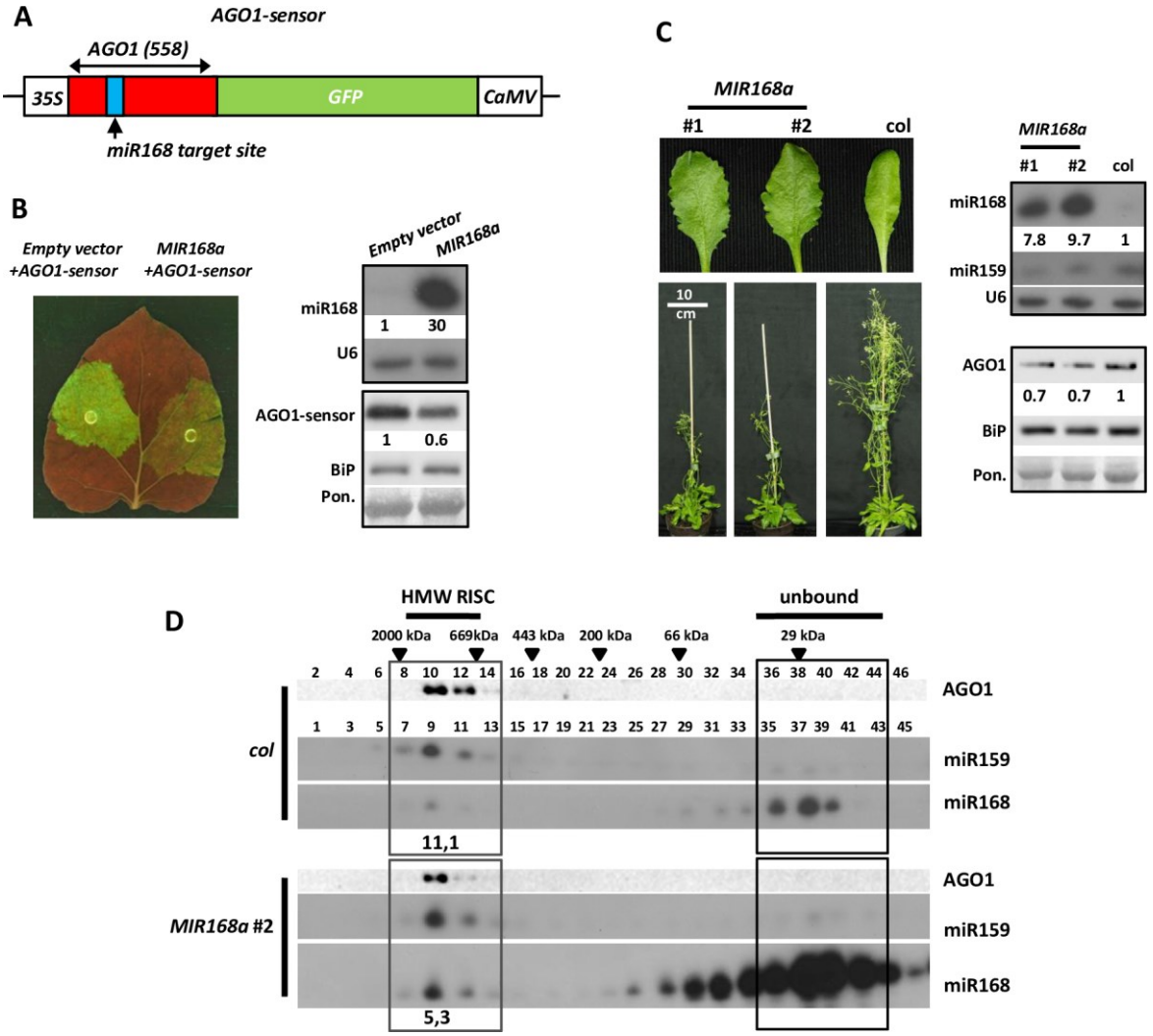

Figure 14. Over-expression of wild type *ath-pri-MIR168a* in transient and transgenic systems. **A** Schematic representation of the *AGO1-sensor* construct. AGO1-derived part of the sensor is indicated with red, miR168 target site is marked with blue rectangle. **B** Transient over-expression of *ath-pri-MIR168a* (*MIR168a* on figures) precursor fragment in the presence of *AGO1-sensor*. Left panel shows the GFP fluorescence in leaves after co-infiltration of *AGO1-sensor* construct containing *Agrobacteria* and either empty vector (*pGreen0029*) or *MIR168a*. Right panels indicate the RNA levels of miR168 and the protein levels of the AGO1-sensor in the infiltrated patches. AGO1 part of the sensor fusion protein was detected with antibody raised against ath-AGO1. For the northern blot U6, while for the western blot BiP (Lumenal binding proteins) and Ponceau staining were used as loading controls. **C** Phenotypes of nine weeks old *MIR168a* precursor fragment overexpressing transgenic lines. Right panels show miR168, miR159 and AGO1 levels in overexpressing and control plants. U6, BiP and Ponceau staining were used as loading controls, respectively. **D** MiR168 distribution among gel filtration fractions of wild type *A. thaliana* Columbia (*col*) and *MIR168a* overexpressing plants. High molecular weight (HMW) RISCs were detected with AGO1 western and miR159 northern blots. Even fractions were used for protein extraction, odd fractions for RNA purification. Black triangles show positions of known molecular weight markers amongst gel-filtration fractions.

6.2.1 Changes in duplex structure can further reduce the AGO1-loading of miR168

Computational analyses of the miR168/miR168* duplex structure encoded by the genomes of different plant species revealed a predominantly conserved nucleotide mismatch at the fourth position

of the duplex. Such structural features could influence the efficiency of AGO1 loading. To investigate the potential impact of this mismatch on AGO loading, a construct (*MIR168-4bp*) was designed in which the base pairing at the fourth position was introduced exclusively by modification of the miR168* strand (Fig. 15A). Transient overexpression of *MIR168-4bp* by agroinfiltration resulted in an increased GFP signal and higher levels of AGO1 sensor protein compared to overexpression of the wild-type *MIR168a* precursor fragment at comparable levels (Fig. 15B). Subsequently, *MIR168-4bp* transgenic lines were generated and *MIR168* wild-type transgenic lines with similar miR168 expression were selected. The selected *MIR168-4bp* lines exhibited a less flowering delay and a higher level of AGO1 protein compared to the corresponding *MIR168a* lines, suggesting a reduced downregulation of AGO1 levels (Fig. 15C). While the intensity of the delayed flowering phenotype and the downregulation of AGO1 correlated with miR168 levels, even the most extreme overexpression of miR168 derived from *MIR168-4bp* could not induce a more pronounced effect than that observed in the *MIR168a* control line. Gel filtration experiments showed that *MIR168-4bp* overexpressing plants had a lower HMW-RISC miR168 loading efficiency than *MIR168a* overexpressing plants (Fig. 15D). AGO1 immunoprecipitation experiments confirmed this with reduced miR168 accumulation in *MIR168-4bp* transgenic plants compared to *MIR168a* transgenic plants (Fig. 15E). These results suggest that base-pairing at the fourth nucleotide base pair in the miR168/miR168* duplex doesn't inhibit miR168 production, but reduces the AGO1-RISC loading capacity of miR168, resulting in a different steady-state level of AGO1 protein compared to *MIR168a* overexpression.

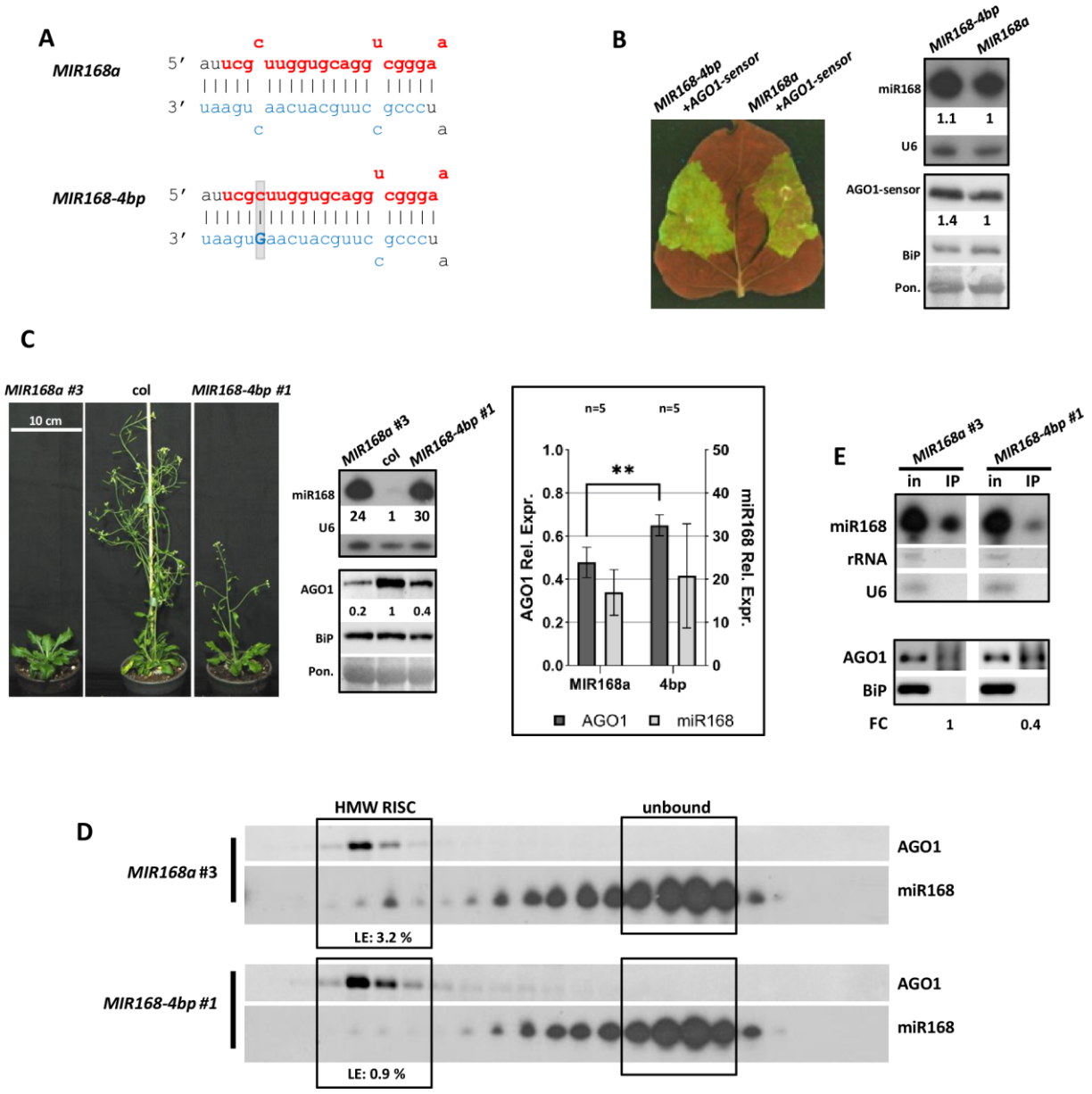

Figure 15. Effect of *MIR168a-4bp* overexpression. **A** Structure of the modified duplex region in *MIR168-4bp*. Modified nucleotide is highlighted with a bold capital letter, grey background shows structural change. Mature miRNAs and star strands were marked with red and blue, respectively. **B** Transient over-expression of *MIR168-4bp* and wild type *MIR168a*. Blots demonstrate miR168 and AGO1 sensor content in infiltrated patches of *N. benthamiana* leaves. AGO1 part of the sensor fusion protein was detected with antibody raised against AtAGO1. For the northern blot *U6*, while for the western blot, BiP and Ponceau staining were used as loading controls. **C** Phenotypes of nine weeks old *MIR168-4bp* and *MIR168a* lines, miR168 and AGO1 content of a transgenic line overexpressing *MIR168-4bp* and a wild type *MIR168a* line with comparably high over-expression rate of miR168. Statistical representation of five *MIR168a* (#: 3, 4, 2, 15, 16) and five *MIR168-4bp* (#: 1, 2, 6, 10, 14) lines. Asterisks represent significant difference (t-test, P = 0,05). **D** MiR168 distribution in gel-filtration fractions of the lines presented. High molecular weight RISC is identified by the presence AGO1 protein. Loading efficiency (LE) is calculated as described in Materials and Methods. **E** AGO1 IP of the investigated *MIR168-4bp* and the corresponding *MIR168a* overexpressing lines. Cytoplasmic contamination in IP samples at RNA and protein level was checked by detection of rRNA, *U6* and BiP, respectively. Fold change (FC) of AGO1 associated miR168 was calculated as the volume intensity ratio of miR168 and AGO1 signals in IP samples and was presented on the basis of corresponding *MIR168a* line.

6.2.2 Altered duplex structure can enhance AGO1 loading of miR168

To further test the capability of duplex structure in regulation of loading efficiency, another precursor of a well-loaded miRNA, the *hvu-MIR171* was employed. In addition, a heterologous precursor from barley was used to show that only the structure is important. Actually, the structure of the *hvu-MIR171* duplex, which contains three mismatches at positions 4, 9 and 12 of the miRNA was mimicked on the MIR168 duplex, with the modification of only the passenger strand (*MIR168-3mm*) (Fig. 16A). As a result, this precursor was expected to produce the wild type miR168 mature miRNA from a modified duplex structure. When *MIR168-3mm* was tested in a similar transient expression assay in *N. benthamiana* as was described in 6.2.1, it showed a visible reduction in AGO1-GFP sensor fluorescence, which was also consistent with protein quantification (Fig. 16B). Interestingly, low miR168 production was observed, but this was sufficient to achieve a reduction in sensor accumulation compared to the *MIR168a* construct in transient expression assay. Plants genetically modified to contain the *MIR168-3mm* construct showed only a modest overexpression of miR168. Interestingly, this modest increase in miR168 resulted in a more pronounced delay in flowering and a more pronounced down-regulation of AGO1 compared to transgenic lines expressing higher levels of miR168 from the original *MIR168a* construct (Fig. 16C). Gel filtration experiments were performed to determine whether the increased activity of *MIR168-3mm*-derived miR168 was due to their improved RISC loading efficiency. Indeed, *MIR168-3mm*-derived miR168 was more efficiently incorporated into HMW-RISC, resulting in enhanced downregulation of AGO1 protein (Fig. 16D). This was further confirmed by immunoprecipitation of AGO1 from the *MIR168-3mm* transgenic line (Fig. 16). These results suggest that the inherent structural features of the wild type miR168/miR168* duplex tightly regulate the AGO1 loading of miR168.

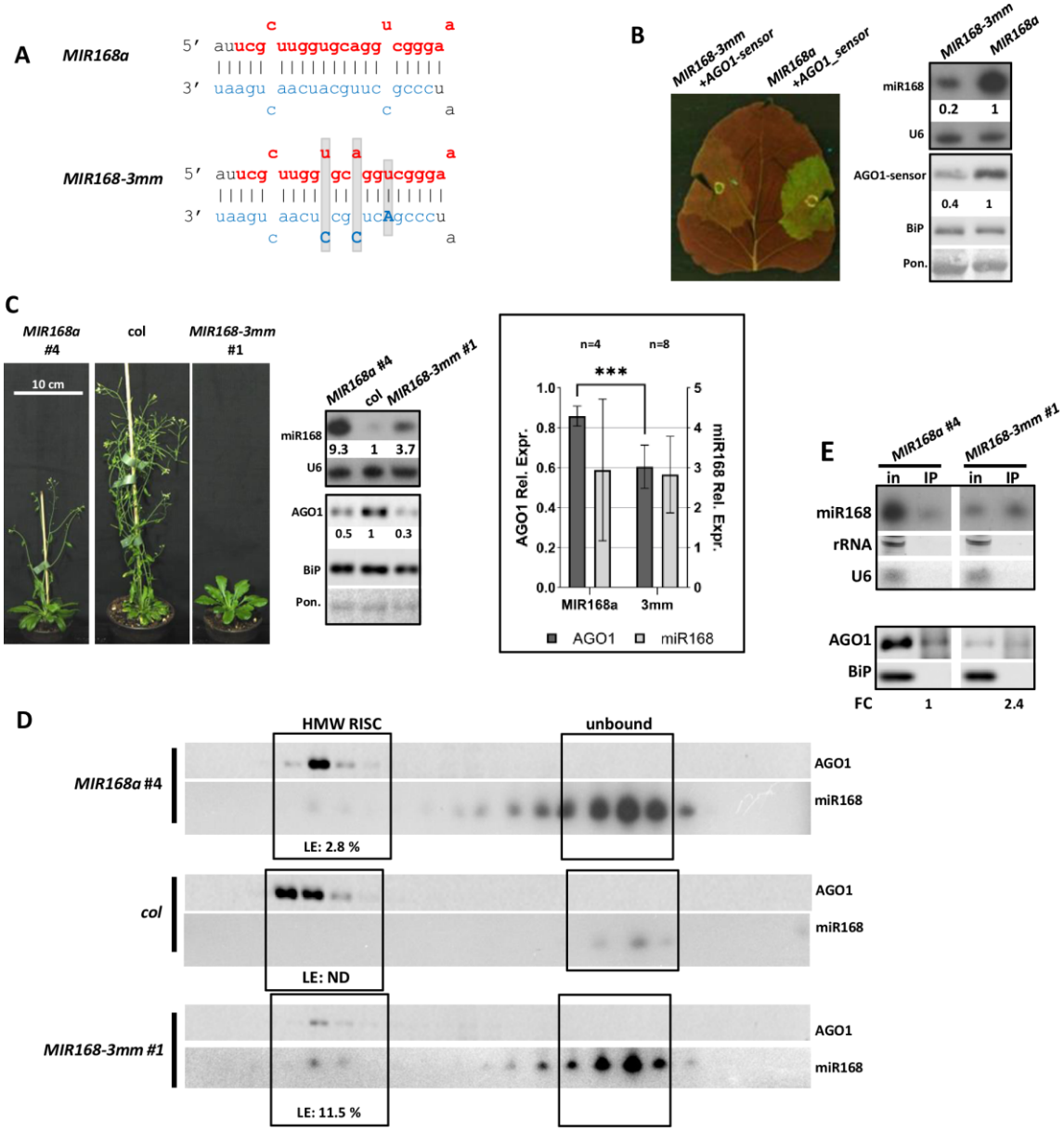

Figure 16. Effect of *MIR168-3mm* overexpression. **A** Structure of the modified duplex region in *MIR168-3mm* in comparison with wild type *MIR168a* duplex. Modified nucleotides of star strand are highlighted with bold capital letters, grey background indicates structural changes. Mature miRNAs and star strands were marked with red and blue, respectively. **B** Transient over-expression of *MIR168-3mm* and wild type *MIR168a* precursor fragment constructs. Blots demonstrate miR168 and AGO1-sensor content of infiltrated patches of *N. benthamiana* leaves. AGO1 part of the sensor fusion protein was detected with antibody raised against ath-AGO1. For the northern blot U6, while for the western blots, BiP and Ponceau staining were used as loading controls. **C** Phenotypes of nine weeks old *MIR168-3mm* and a *MIR168a* line. MiR168 and AGO1 content of a transgenic line overexpressing *MIR168-3mm* and a wild type *ath-pri-MIR168a* line with a comparable over-expression rate of miR168. Statistical presentation of four *MIR168a* (#: 11–14) and eight *MIR168-3mm* (#: 1–8) lines. Asterisks label significant difference (*t*-test, $P = 0{,}01$). **D** MiR168 distribution in gel-filtration fractions of the investigated lines and wild type *Columbia*. High molecular weight (HMW) RISC is presented with AGO1. Loading efficiency (LE) is calculated as described in Methods. **E** AGO1 IP of the investigated *MIR168-3mm* and the corresponding *MIR168a* overexpressing line. Cytoplasmic contamination in IP samples at RNA and protein level was checked by detection of rRNA, U6 and BiP, respectively. Fold change (FC) of AGO1 associated miR168 was calculated as the volume intensity ratio of miR168 and AGO1 signals in IP samples and was reported on the basis of corresponding *MIR168a* line.

6.2.3 Artificial precursor structure enhanced AGO1 loading of miR168

The *MIR168-3mm* construct has a backbone derived from the *MIR168a* precursor fragment but displays *hvu-MIR171*-specific features in the miR168/miR168* duplex region. To further explore the structural features that can influence the AGO1 loading capacity of miR168, we created artificial miR168 precursor (AMIR) constructs based on a modified barley *hvu-MIR171* precursor fragment. Two AMIR variants were designed to express miR168 from the *hvu-MIR171* backbone, maintaining the *hvu-MIR171* stem-loop structure by changing the orientation of the miR168 guide strand and modifying the star strand (Fig. 17A). This ensured that the three mismatches within the duplex remained in the same positions as in the *hvu-MIR171* duplex. The only difference between *AMIR-1* and *AMIR-2* is the identity of the mismatched nucleotides at positions 4 and 9 in the duplex (Fig. 17A). *AMIR-1* and *AMIR-2* were transiently expressed in *N. benthamiana* leaves with AGO1 sensor. Both constructs resulted in a greater reduction in GFP signal and AGO1-sensor protein levels than *MIR168a* (Fig. 17B). Furthermore, small RNA northern blot results showed that the enhanced AGO1 downregulation was associated with a lower level of miR168 overexpression of the 2 *AMIR* constructs compared to *MIR168a* (Fig. 17B). To further investigate the effect of altered stem-loop structures on miR168 HMW-RISC loading efficiency, stable transgenic lines of *AMIR-1* and *-2* were generated. Lines overexpressing similar levels of miR168 were selected for further studies (Fig. 17C). Compared to the *MIR168a* line, *AMIR-1* and *AMIR-2* exhibited more pronounced phenotypic changes, such as delayed flowering and reduced rosette diameter. The severity of these phenotypes correlated with the level of overproduced miR168, resembling hypomorphic *ago1-25* and *ago1-27* mutants in cases of strong miR168 overexpression (Morel et al., 2002). Despite slightly lower miR168 overexpression, AGO1 protein levels were significantly reduced in *AMIR* lines compared to *MIR168a* lines (Fig. 17C). Gel filtration experiments were performed on seedlings of selected *AMIR-1*, *AMIR-2* and control *MIR168a* lines. Despite reduced AGO1 protein accumulation, *AMIR-1* and *AMIR-2* lines showed increased miR168 accumulation in HMW AGO-RISC fractions (Fig. 17D). Immunoprecipitation experiments confirmed increased miR168 incorporation into AGO1-RISC in both *AMIR* lines (Fig. 17E). This suggests that the production of miR168 from alternative stem-loop structures may increase the loading efficiency of AGO1-RISC, thereby enhancing its biological activity.

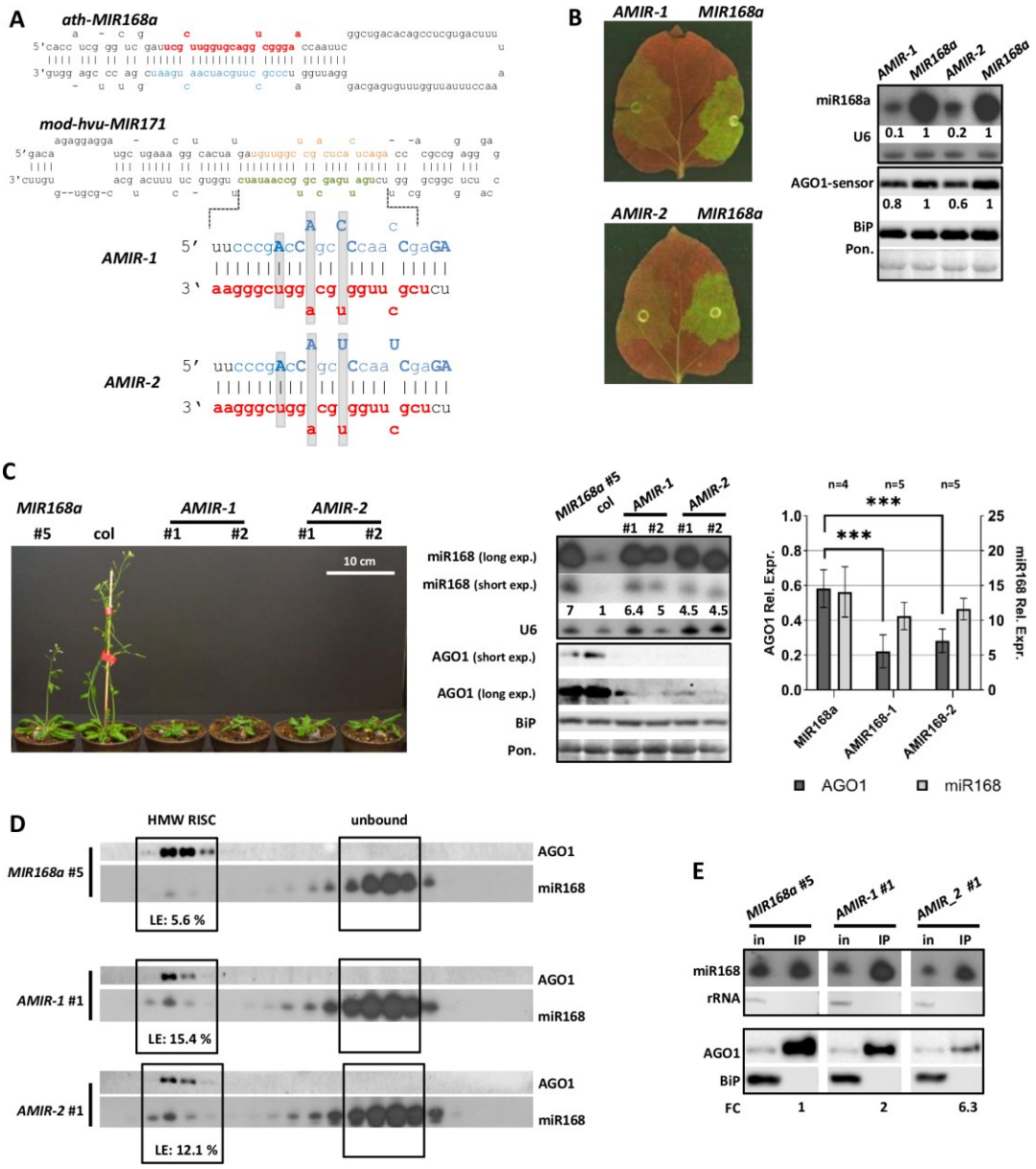

Figure 17. Overexpression of miR168 producing AMIR constructs. **A** Structure of the modified duplex region implemented into *hvu-MIR171* backbone in case of AMIR constructs. Modified nucleotides of star strands compared to wild type *MIR168a* duplex are highlighted with bold capital letters, grey background shows structural changes. Mature miR168, miR168*, miR171 and miR171* strands were marked with red, blue, green and brown, respectively. **B** Transient over-expression of AMIR constructs and wild type *MIR168a* precursor fragment. Blots demonstrate miR168 and AGO1-sensor content of infiltrated patches of *N. benthamiana* leaves. For the northern blot U6, while for the western blot, BiP and Ponceau staining were used as loading controls. **C** Phenotype of eight weeks old *MIR168a*, *AMIR-1* and *AMIR-2* lines. MiR168 and AGO1 content of transgenic lines overexpressing AMIR precursors and a *MIR168a* line with a comparable over-expression rate of miR168. Statistical analysis of four *MIR168a* (#: 2, 6, 8, 9), five *AMIR-1* (#: 3, 4, 5, 6, 7) and five *AMIR-2* (#: 5, 6, 7, 8, 9) lines. Asterisks display significant difference (t-test, P = 0,01). **D** MiR168 distribution in gel-filtration fractions of *MIR168a* #5, *AMIR-1* #1 and *AMIR-2* #1 lines. High molecular weight RISC is presented with AGO1. Loading efficiency (LE) is calculated as described in Materials and Methods. **E** AGO1 IP of the investigated *AMIR-1*, *AMIR-2* and the corresponding *MIR168a* overexpressing line. Cytoplasmic contamination in IP samples at RNA and protein level was checked by detection of rRNA and BiP, respectively. Fold change (FC) of AGO1 associated miR168 was calculated as the volume intensity ratio of miR168 and AGO1 signals in IP samples and was reported on the basis of corresponding *MIR168a* line.

6.2.4 Small RNA-sequencing and miR168 species

Excessive production of miR168 from altered precursors may lead to misprocessing of the miR168 duplex, resulting in differential accumulation of canonical and non-canonical miR168 species (iso-miRs). These miRNA isoforms may be loaded into different AGOs at different rates, potentially perturbing RNA silencing autoregulation. High-throughput sequencing (HTS) analyses were performed on sRNA pools from selected transgenic lines to understand miR168 maturation (Fig. S3A). Reads mapped to miR168 precursors revealed predominantly 21 nt long small RNAs (Fig. S3B). The reads per million (RPM) data showed that all precursor fragments overproduced miR168 compared to the wild-type plant. Besides the dominant overproduction of 5′ U miR168, there was also overproduction of 5′ C and small amounts of 5′ G and 5′ A miR168 species in the transgenic plants (Fig. 18A). This differential accumulation raised the question whether miR168 misprocessing may affect AGO1 loading efficiency. The relative accumulation of different miR168 species at the 5′ end was calculated (Fig. 18B). Most 5′ U reads represented the canonical miR168 species, except in *AMIR-2* where a truncated version of miR168 was most abundant (Fig. 18C). The 5′ C iso-miRNA content was similar between the different overexpressing lines. *MIR168a* and *AMIR-1* lines showed a similar distribution of 5′ U and 5′ C miR168 species. In *MIR168-3mm*, an increased ratio of 5′ U miR168 to 5′ C miR168 was detected, like the *AMIR-2* line. However, *MIR168-3mm* was associated with more efficient AGO1 loading with less than 20% miR168 overexpression compared to other lines. *MIR168-4bp*, which overproduced miR168 but showed inefficient AGO loading, was associated with a higher 5′ C miR168 ratio. This suggests that the generated 5′ C miR168 species do not significantly interfere with or contribute to differential AGO1 loading. However, the differences in iso-miR168 species accumulation probably don't play a significant role in the increased AGO loading efficiency. By comparing the miR168 species analysis in previously published AGO1 immunoprecipitation data with the input total RNA of the wild-type sample, most of the 5′ C miR168 species produced are biologically active and follow the same AGO loading rules as 5′ U miR168 (Fig. S3C). In conclusion, the structural features of the miRNA duplex are mainly responsible for the changes in miR168 loading efficiency. However, the altered production of miR168 species may also contribute to the phenomenon to a lesser extent.

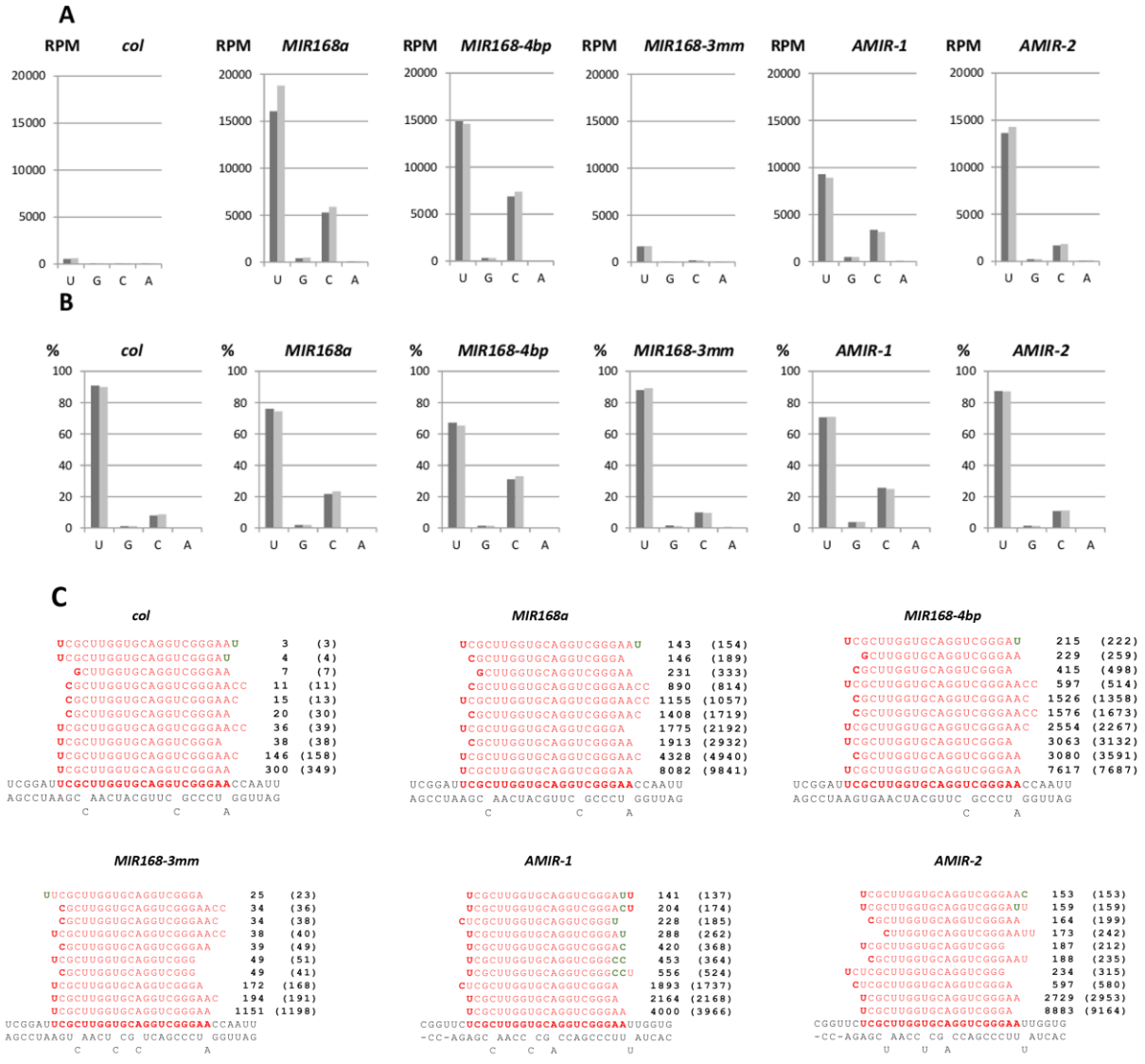

Figure 18. HTS data analyses and miR168 loading rate under competitive circumstances. Distribution of 5′ U/G/C/A miR168 species in HTS data of *A. thaliana* Col-0, *MIR168a*, *MIR168-4bp*, *MIR168-3mm*, *AMIR-1* and *AMIR-2* overexpressing plants displayed as **A** read per million (RPM) and **B** the percentage of whole miR168 homologous sequences. The two replicas were distinguished with different shades of grey. **C** Top ten most abundant sequences detected amongst the reads mapped to the precursor arm containing the mature miRNA. Numbers show the read per million (RPM) of the given sequence, numbers in brackets correspond to replica experiment.

6.3 Investigation of barley *AGO4* genes functionality in *Arabidopsis* complementation studies

As our work demonstrated, the model plant *Arabidopsis thaliana* is a suitable organism to test AGO functionality, as there are molecular methods and mutants described which enable rapid functional testing of exogenous genes. To validate the functionality of the two potential barley *AGO4* genes, a heterologous complementation assay was developed. Using the *AtAGO4* promoter and terminator, a 5′ HA-tagged version of each barley gene was introduced into *A. thaliana ago4-3* mutants (WiscDSLox338A06). The initial selection of T0 plants was based on kanamycin resistance, and subsequent generations (T1 onwards) were analyzed for mRNA and protein expression levels in the inflorescence. Given the lack of phenotypic differences between wild-type Col-0 and the *ago4-3* mutant (Havecker et al., 2010), transgenic lines resembling the wild type were selected for further study to avoid positional effects of the transgene insertion.

The transformants showed different levels of transgene expression at the RNA level, as determined by primers specific for the respective barley genes. When *HA-HvAGO4a* and *HA-HvAGO4b* expression was compared to *AtAGO4* expression in wild-type Col-0, some transformed lines mirrored the endogenous *AGO4* level, in particular two HvAGO4B lines (#1 and #17) (Fig. 19A). To further compare AGO4 protein levels in the transformant lines, a western blot was performed using an anti-HA antibody, which revealed a correlation between mRNA and protein levels (Fig. 19B). Three lines with high levels of AGO4 protein for both barley genes were then selected with the assumption that their different expression states would facilitate a transgene-level dependent analysis of the complementation effect.

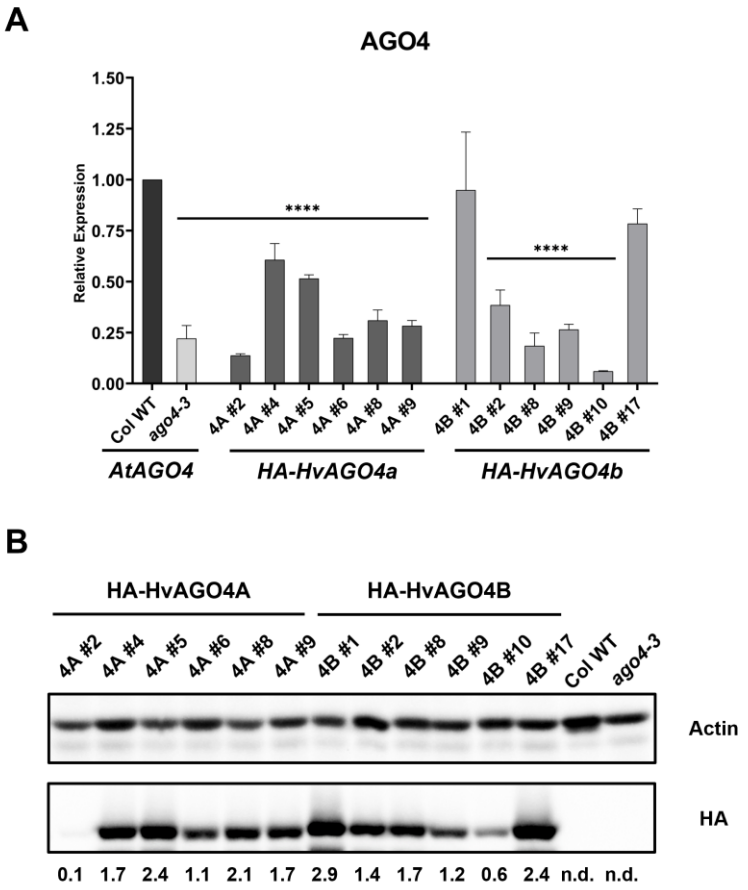

Figure 19. A *AtAGO4* and HA-tagged barley *AGO4* genes expression levels determined by RT-qPCR of T1 mixed-stage inflorescences. Data were normalized using *AtUBC9* and *AtACT2*. For individual primer pairs used to detect the three different *AGO4* genes see Table S1. The average of 3 independent biological replicates was calculated and statistically significant differences from wild-type *A. thaliana* Col-0 (Col WT) are indicated with asterisks (Anova one-way with Dunnett's post-hoc test, * < 0.05, ** < 0.01, *** < 0.001 and **** < 0.0001). Error bars represent the mean ± SD, n = 3. **B** HA-HvAGO4 protein level in mixed-stage inflorescences of T1 transgenic plants. To quantify the HA-HvAGO4, volume intensity of each sample was referred to the corresponding actin signal and was presented as the ratio of HA and Actin signal.

6.3.1 Functional complementation properties of barley *AGO4* genes in relation to *Arabidopsis ago4* mutant

The functionality of the barley *AGO4* genes was investigated in the heterologous complementation lines of *HA-HvAGO4a* and *HA-HvAGO4b.* The initial assay employed the retrotransposon *AtSN1*, which is regulated by AGO4 in *Arabidopsis* through the maintenance of methylation at the locus (Duan et al., 2015; Havecker et al., 2010; Zilberman et al., 2003). Consistent with previous studies, a substantial upregulation (30-fold change) of *AtSN1* expression was observed in the *Arabidopsis ago4-3* mutant, which was effectively reduced by transgenic HvAGO4A and HvAGO4B. This

complementation effect was directly proportional to the transgenic AGO4 level in all lines (Fig. 20A). Plants with elevated transgenic AGO4 levels exhibited *AtSN1* expression levels more similar to the wild type. Chop-PCR revealed that the methylation level of *AtSN1* in the *ago4-3* mutant was less than 50% of that in wild-type *Arabidopsis* plants. As expected, increased levels of both barley AGO4 proteins successfully restored *AtSN1* methylation in the complemented *Arabidopsis ago4-3* plants, and *AtSN1* expression reverse correlated with the methylation level of the different transgenic lines. This led to the conclusion that the reduced *AtSN1* expression levels in the complemented lines were due to the increased methylation state at the locus (Fig. 20B). To further validate the functionality of the barley *AGO4* genes, the expression of *AtROS1*, a DNA glycosylase/lyase, was analyzed given its known dependence on RdDM (Córdoba-Cañero et al., 2017; Lei et al., 2015; Tang et al., 2016). A 65% reduction in *ROS1* expression was observed in the *ago4-3* mutant relative to the wild type, which was restored in the complementation lines (Fig. 20C). Notably, HvAGO4A plants exhibited *AtROS1* expression levels above those of the wild type, particularly line 4A #4, which showed a significant upregulation (Fig. 20C). Furthermore, both barley AGO4 proteins were able to compensate for the impairment caused by the *ago4-3* mutation.

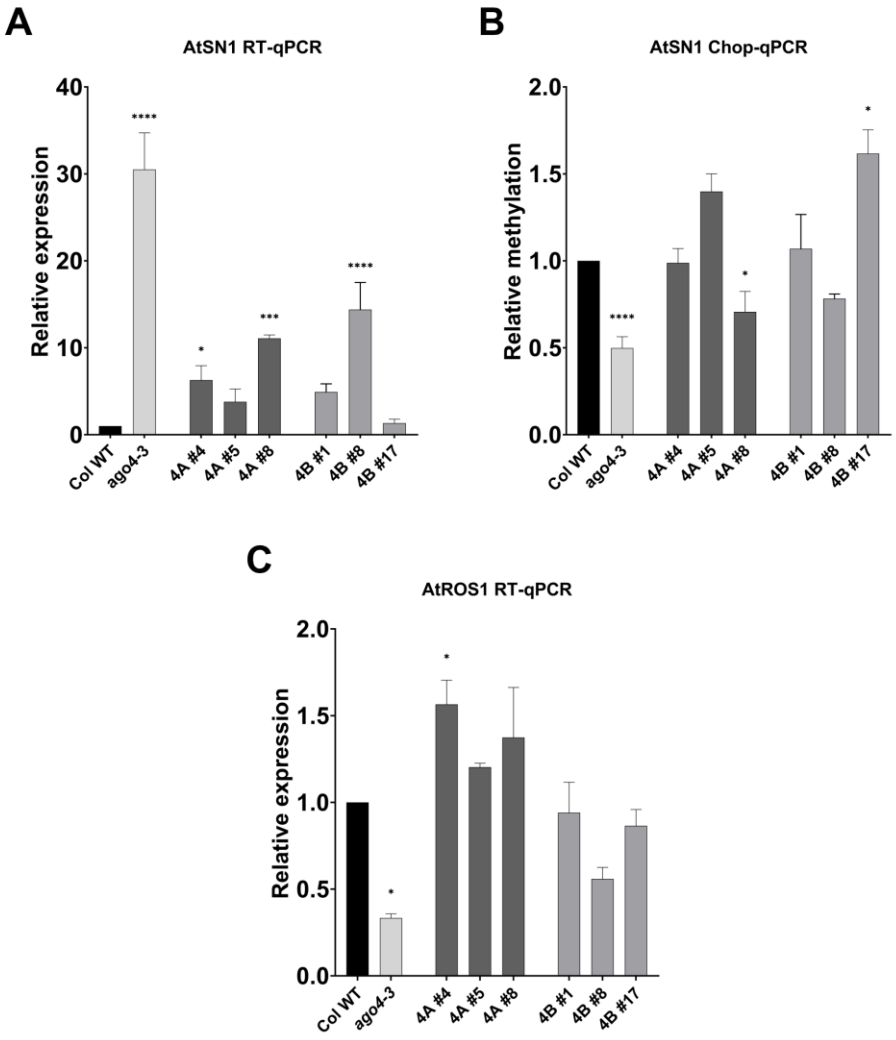

Figure 20. **A** *AtSN1* expression levels determined by RT-qPCR analysis of mixed-stage inflorescences. Data were normalized using *AtUBC9* and *AtACT2*. **B** Methylation levels of *AtSN1* locus determined using Chop-qPCR on digested and undigested DNA. Digestion was performed with MspJI, modification-dependent restriction endonuclease, and data were normalized to *AtSN1* level from the undigested DNA and then the reciprocal was calculated to show the relative methylation levels. **C** *AtROS1* expression levels determined by RT-qPCR analysis of mixed-stage inflorescences. Data were normalized using *AtUBC9* and *AtACT2*. All the results show the average of 3 independent biological replicates and were statistically analyzed against wild-type *A. thaliana* Col-0 (Col WT) and significant differences are indicated with asterisks (Anova one-way with Dunnett's post-hoc test, * < 0.05, ** < 0.01, *** < 0.001 and **** < 0.0001). Error bars represent the mean ± SD, n = 3.

6.3.2 Effect of barley AGO4 proteins on transposable element activation under heat stress conditions

Transposable elements (TEs) and nearby genes can undergo transcriptional activation in response to stress conditions (Ito, 2022; Makarevitch et al., 2015). The influence of barley AGO4 proteins on TE activation under heat stress was investigated by subjecting 1-week-old *Arabidopsis* seedlings to 24 h

of heat stress at 37°C. The focus was on extrachromosomal copy formation and transcript levels of *ONSEN*, a heat-responsive *Ty1/copia*-like retrotransposon *ATCOPIA78*. Under control conditions, the extrachromosomal copy number of the 8 *ONSEN* genes remained constant in both mutant and transformant *Arabidopsis* lines compared to the wild type (Fig. 21A-B). Consistent with previous studies, heat stress induced an increase in TE activity in all samples (Fig. 21A-B), with a notable 5-fold increase in extrachromosomal DNA (ecDNA) copy number in the *ago4-3* mutant compared to wild-type samples. Both HvAGO4A and HvAGO4B reduced TE activation under stress conditions compared to the *ago4-3* mutant, keeping *ONSEN* ecDNA levels comparable to wild type, even in lower expression transformant lines. To determine whether this phenomenon was directly related to the expression rate of *ONSEN* transcripts, it was examined in non-treated (NT) and heat-stressed (HS) seedlings. RT-qPCR analysis revealed a significant up-regulation of *ONSEN* under heat stress conditions in the *ago4-3* mutant compared to the wild type, while all transformant lines exhibited *ONSEN* expression levels closer to the wild type (Figs. 18C-D). Interestingly, lines with higher transgenic protein levels, such as 4A #5, 4B #1, and 4B #17, showed a significant downregulation of *ONSEN* compared to the wild type. Under control conditions, a significant difference in *ONSEN* expression was observed between transgenic, *ago4-3* mutant, and wild-type plants. However, this difference was negligible when compared to heat-stressed plants (Fig. 21C-D).

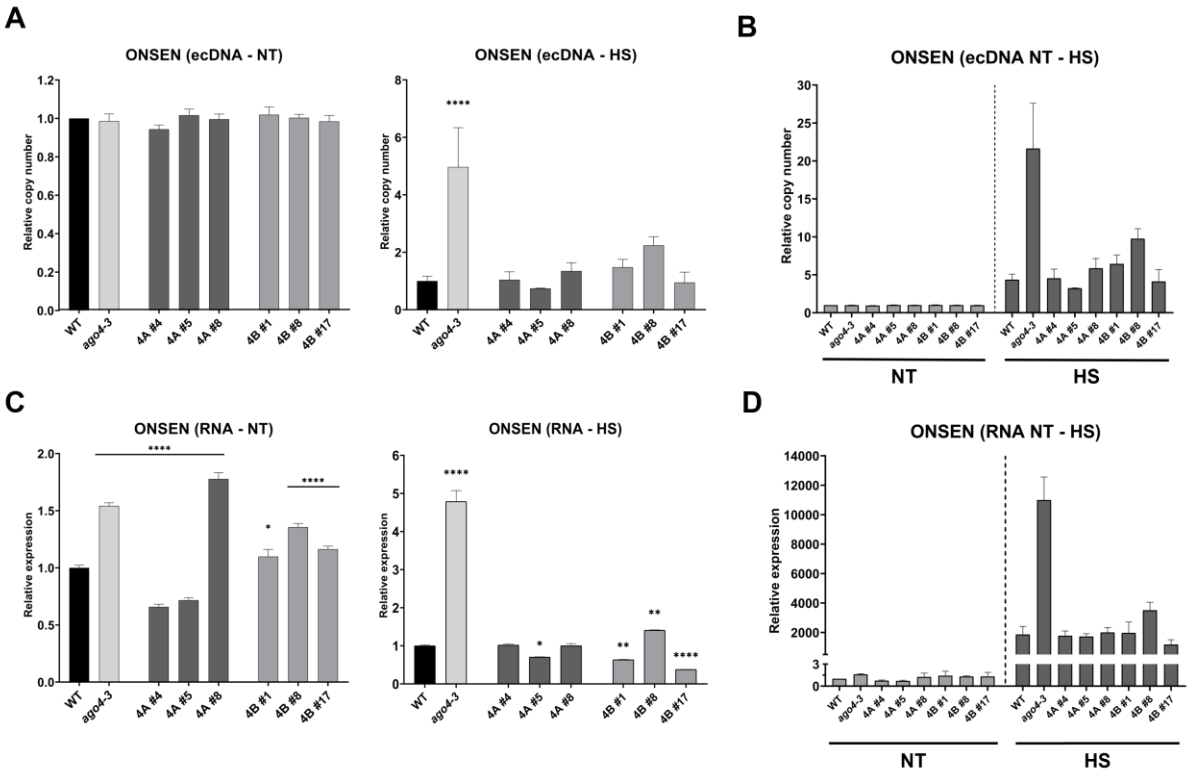

Figure 21. A Relative copy number of *ONSEN* extrachromosomal DNA (ecDNA) in non-treated (NT) and heat stress (HS—24 h at 37 °C) 1-week-old *Arabidopsis* seedlings. Data were normalized using *AtUBC9*. **B** Relative copy number of *ONSEN* extrachromosomal DNA (ecDNA) in non-treated (NT) and heat stress (HS - 24 hours at 37 °C) 1-week old *Arabidopsis* seedlings. Data was normalized using *AtUBC9*. **C** Relative expression of *ONSEN* before and after the heat stress activation measured by RT-qPCR and normalized on *AtUBC9* and *AtPP2AA3*. **D** Relative expression of *ONSEN* before and after the heat stress activation measured by RT-qPCR and normalized on *AtUBC9* and *AtPP2AA3*. For the graphs **B** and **D**, WT in NT group was put as value 1 for all the other samples and the dotted line split the graph between NT and HS samples. Statistically significant differences compared to Col WT are indicated by asterisks (Anova one-way with Dunnett's post-hoc test, * < 0.05, ** < 0.01, *** < 0.001 and **** < 0.0001). Error bars represent the mean ± SD, n = 3.

6.3.3 Differential sRNA-binding preferences of barley AGO4 proteins

While the sRNA binding preference of AGO4 has been extensively studied in *Arabidopsis* and rice, such information is currently lacking for barley. This is particularly relevant as rice have AGO4 proteins shown remarkable differences in sRNA binding preferences based on their 5′ end nucleotide (Wu et al. 2010). To elucidate the sRNA binding capabilities of the putative HvAGO4A and HvAGO4B proteins and to identify any similarities or differences, sRNA-IP sequencing was performed on three *Arabidopsis* complementation lines of HvAGO4A and HvAGO4B. Any contamination in the IP samples was detected by an HA-specific western blot (Fig. S2). As a control, raw data from an AtAGO4 sRNA-IP sequencing, obtained by a method similar to the one described

in this study, were used and re-analyzed with the same parameters as the data produced in this study (Sigman et al. 2021). The striking consistency in the size distribution pattern of the mapped sRNA reads and the robust affinity of the putative barley AGO4 proteins for the 24-nt sRNAs suggested a function analogous to AtAGO4 (Fig. 22A). Furthermore, a comparison of the origin of the reads categories revealed a similar distribution pattern across all libraries. In all samples, the majority of reads derived from TEs (45-48%), while approximately 12% and 40% derived from transcripts and unannotated genomic regions, respectively (Fig. 22B). However, when the 5′-end nucleotide distribution of the sRNAs was examined, a significant difference between HvAGO4A and HvAGO4B was observed. While both barley AGO4 proteins bind 24 nt sRNAs with an adenine at the 5′ end (24A sRNAs), only HvAGO4B showed a distribution pattern similar to that of AtAGO4, which also binds sRNAs with a C, G or U residue at the 5′ end. In contrast, HvAGO4A showed an almost exclusive preference for sRNAs starting with an A residue (Fig. 22C).

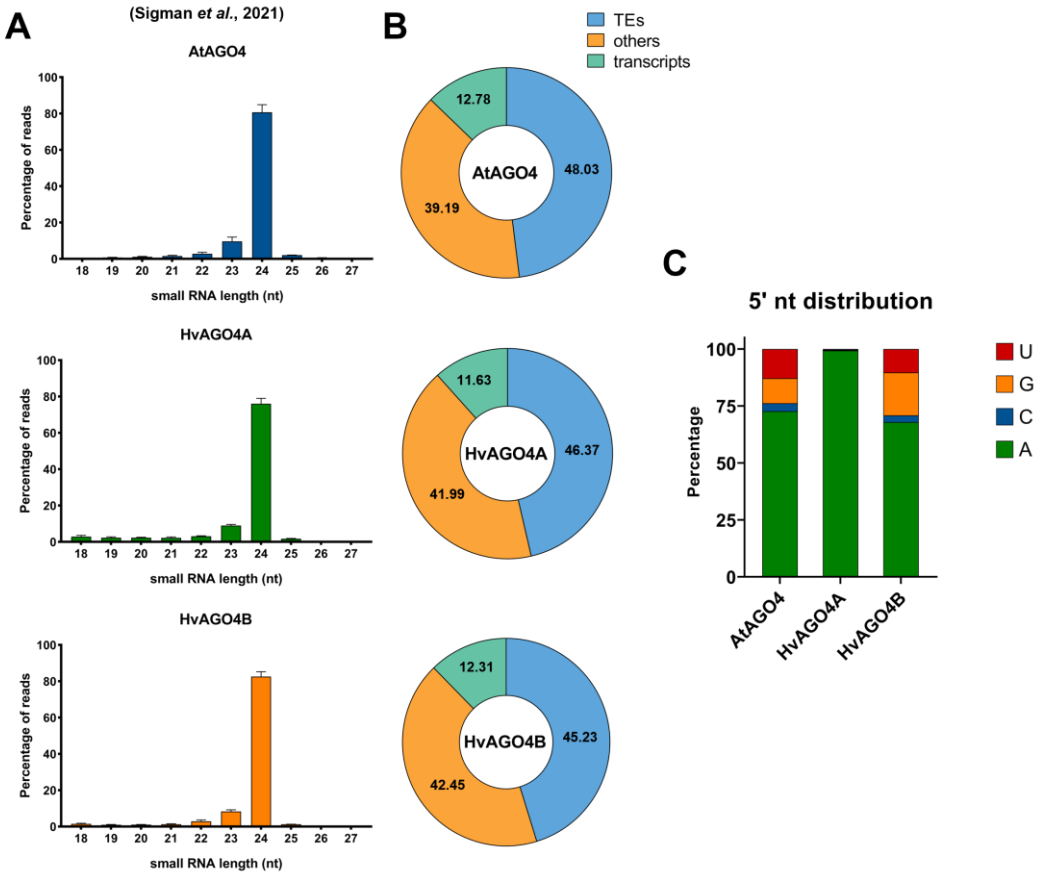

Figure 22. Sequencing of the HA-HvAGO4A- and HA-HvAGO4B-associated sRNA pools of *Arabidopsis* complementation plants. **A** Size distribution profile of filtered sRNA-IP data sets derived from the mean of 3 independent biological replicates. AtAGO4 IP raw data were retrieved from Sigman et al. (2021). Error bars represent

the mean ± SD, n = 3. **B** Percentage of sRNA-IP read distribution based on their origin. **C** Percentage of the sRNA read distribution according to the 5′ nucleotide identity.

Based on the sequence conservation analysis of the 24-nt sRNA pools obtained from the sRNA-IP sequencing data, minor differences were detected between AtAGO4 and the two HvAGO4 proteins (Fig. 23A-B-C). While HvAGO4A exclusively loaded 24A sRNAs, HvAGO4B also interacted with 24nt sRNAs with G, C or U residues at the 5′ end. Interestingly, the AGO4 proteins also showed a lower degree of conservation at the 3′ end. AtAGO4 showed a loading preference for sRNAs with a U residue at the 3′ terminus, HvAGO4B favored sRNAs with a C residue at the 23rd nucleotide, while HvAGO4A showed no conserved position (Fig. 23A-B-C).

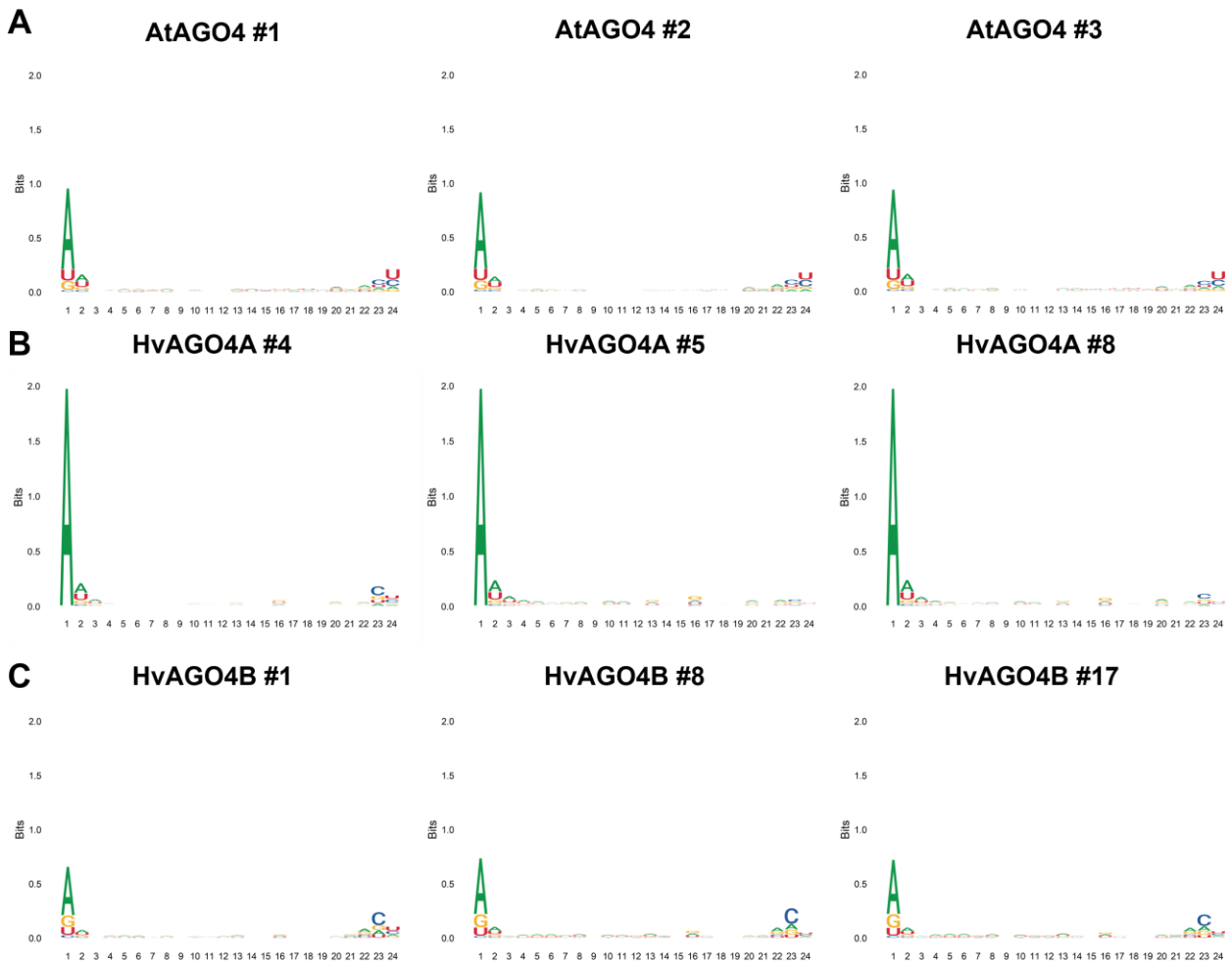

Figure 23. Graphical representation of the 24-nt long sequences conservation of nucleotides using sequence logos. The graphs represent the sequence conservation of three independent replicate of **A** AtAGO4 sRNA-IP sequencing, **B** HA-HvAGO4A and **C** HA-HvAGO4B. Maximum value in bits is 2 on the Y axis. Higher value for a nucleotide indicates a higher conservation.

6.3.4 Detailed analysis of the role of barley AGO4 proteins in transposable element regulation

TE-derived sRNAs associated with HvAGO4A and HvAGO4B were compared to those of AtAGO4 to identify TEs with different sRNA mapping abundances. Interestingly, 1877 TEs associated with HvAGO4A were identified, revealing a significant difference in the number of TE-derived sRNA reads compared to AtAGO4 (Fig. 24). Out of these, 591 had higher read counts for HvAGO4A, while 1286 had fewer reads compared to AtAGO4. In contrast, a smaller number of distinct TEs were associated with HvAGO4B (1454 TEs), where mapped sRNAs showed either a higher (401 TEs) or lower (1053 TEs) abundance in the barley AGO4 proteins immunoprecipitated samples compared to AtAGO4. Furthermore, considering all TEs that showed changes in sRNA abundance in barley proteins compared to AtAGO4, 128 and 449 TEs with increased and decreased sRNA content, respectively, were identified in both HvAGO4 proteins (Fig. 24). Interestingly, 17 TEs showed an inverse behavior, with higher representation in HvAGO4B and lower sRNA abundance in HvAGO4A compared to AtAGO4. However, no TEs showed the opposite trend (Fig. 24).

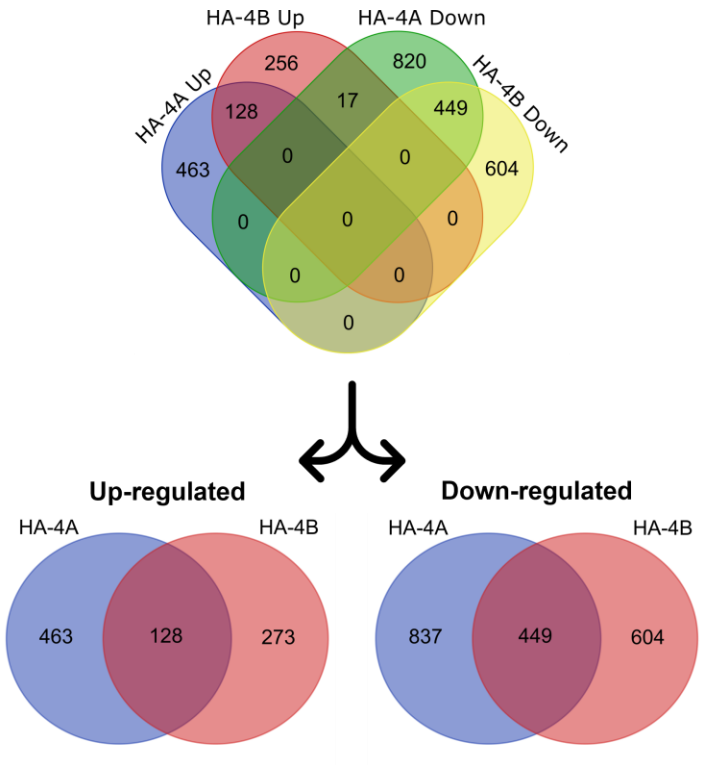

Figure 24. Venn diagram showing the number of TEs showing at least a 2-fold statistically significant (p-value < 0.05) change in the amount of sRNAs compared to AtAGO4. The up-regulated TEs are shown on the left and the down-regulated TEs on the right.

To further investigate the differential effects of these proteins on TEs, an analysis of the sRNAs mapped to the loci of specific TEs was performed. Detailed analysis of TE-derived sRNAs revealed that *AtSN1* had significantly fewer mapped sRNAs in HA-HvAGO4A compared to both AtAGO1 and HA-HvAGO4B. Interestingly, this discrepancy was accompanied by a shift in the distribution of reads mapped to the retrotransposon. AtAGO4 and HvAGO4B predominantly bind to sRNAs from all four regions of the locus, whereas HvAGO4A showed a preference for sRNAs originating only from the central regions, thereby reducing the total amount of *AtSN1*-derived sRNAs (Fig. 25A). When sRNAs from different regions were analyzed, a unique distribution of 5′-end nucleotides was observed, with the central regions predominantly producing sRNAs starting with an A residue and the two lateral regions predominantly producing sRNAs with a G residue at the 5′ end. Consistent with the distribution pattern of the total pools (Fig. 22C), *AtSN1*-derived sRNAs also showed differences in 5′-end nucleotide preference among the different proteins. While HvAGO4A binds almost exclusively to *AtSN1*-derived sRNAs starting with A, AtAGO4 and HvAGO4B showed a stronger affinity for 5′ G (about 60% for both) but retained the ability to bind sRNAs with any nucleotide at the 5′ end (Fig. 25B). Closer examination of other TEs also revealed changes in the distribution pattern of the 5′-end

nucleotides of sRNAs associated with TE-derived sRNAs. For example, analysis of sRNAs derived from a *RathE3* TE (*AT5TE27090*) identified three regions sRNAs producing in case of AtAGO4 and HvAGO4B, but no significant amount of sRNAs was detected in HvAGO4A (Fig. 25C). A detailed analysis of the type of sRNAs derived from this TE revealed that these regions exclusively produced sRNAs with a G residue at the 5′ end and thus could not be loaded by HvAGO4A. The investigation of *AtSN1* and *AT5TE27090* revealed that despite the similar regulatory role of the three AGO4 proteins (Fig. 20A-B), there are significant differences in their mode of action.

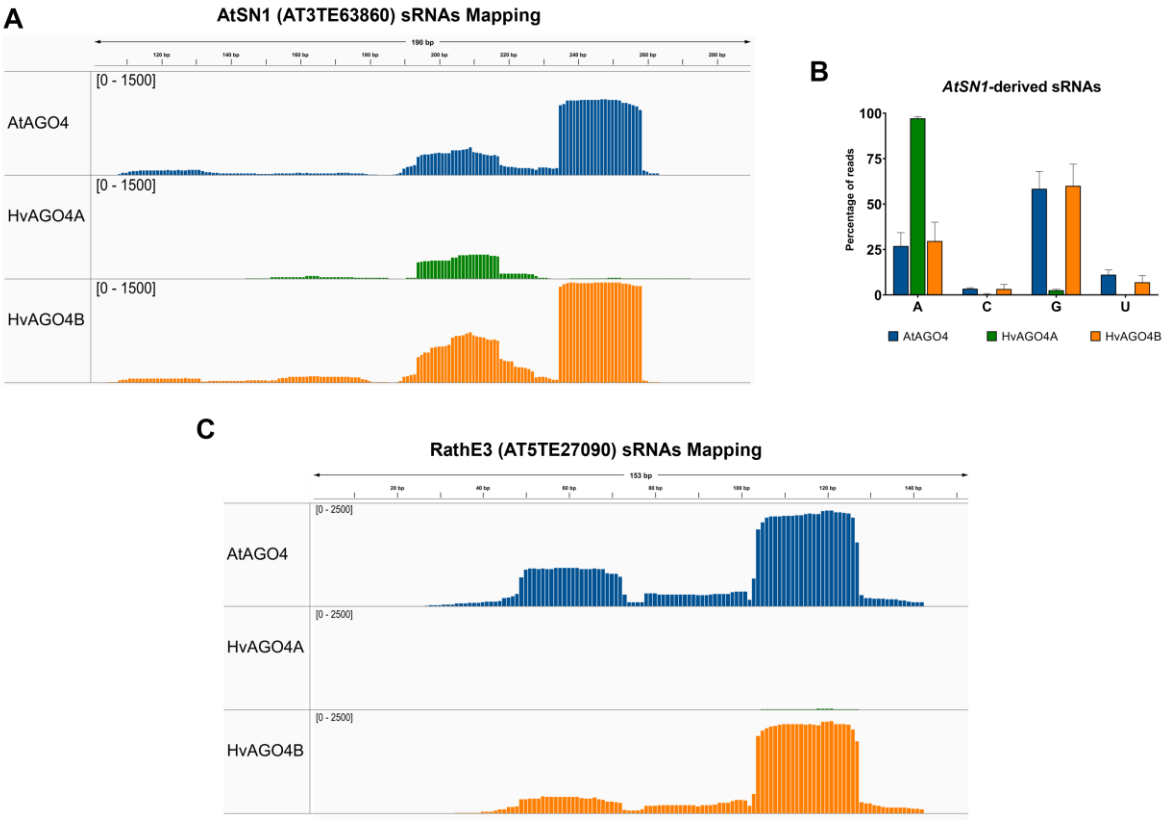

Figure 25. A sRNAs mapped on *AtSN1* genomic locus. The reads are merged from the three independent biological replicate IP datasets. The quantity of mapped reads (counts) is written on the left side of the image inside the square brackets. Different colors indicate individual IP datasets: AtAGO4 (blue), HA-HvAGO4A (green) and HA-HvAGO4B (orange). **B** Percentage of the *AtSN1*-derived sRNAs in the categories based on 5′-end nucleotides. Error bars represent the SD of the three sequenced lines. **C** sRNAs mapped on *RathE3 TE* (*AT5TE27090*) genomic locus. The reads are merged from the three independent biological replicate IP datasets. The quantity of mapped reads (counts) is written on the left side of the image inside the square brackets. Different colors indicate individual IP datasets: AtAGO4 (blue), HA-HvAGO4A (green) and HA-HvAGO4B (orange).

6.3.5 Prediction of 3D structure of barley AGO4 proteins bound to sRNAs

The different binding affinity of HvAGO4A and HvAGO4B was confirmed by sRNA-IP sequencing, in which a specific affinity for 24 nt 5′ A sRNAs was observed in the case of HvAGO4A, while HvAGO4B showed a less strict selection. To demonstrate that this affinity is indeed due to the PIWI domain region involved in 5′ sRNA anchoring, predictions were made using AlphaFold3 (Abramson et al., 2024). The web service AlphaFold Server was selected to use as it can generate highly accurate biomolecular structure predictions of complexes containing proteins, DNA, RNA, ligands, ions, and model chemical modifications for proteins and nucleic acids in one platform. In fact, this new platform has made it possible to simultaneously predict the 3D structure of proteins, for which there is no 3D structure derived from cryo-EM, and of a sRNA, while predicting their interaction. Thus, it was possible to predict the conformation of the AGO4-sRNA complex and observe how they interact (Fig. 26A-B-C). Interestingly, depending on the sequence of the sRNA used, it was observed that the sRNA automatically positioned itself with the 5' end directly in contact with the MID-PIWI protein domains, confirming that this protein pocket is indeed responsible for binding to the 5' end of the sRNA and that the sequence allows the correct automatic positioning of the sRNA. Focusing only on the region of the PIWI domain involved in the 5′ sRNA anchoring, the domain, composed of 4 amino acids with a single AA difference between the 3 proteins (QCxA), is in close contact with the effective 5′ end of the sRNA (Fig. 26D-E-F). Indeed, intermolecular hydrogen bonds can be observed between this region of the proteins and the sRNA; these bonds are important for the establishment of sequence specificity in ssRNA (Auweter et al., 2006). In the predictions, the variable amino acid is in the central part of the sequence responsible for the interaction with the sRNA 5'-end, with a methionine (MET 609) in AtAGO4, a valine (VAL 606) in HvAGO4A and a leucine (LEU 603) in HvAGO4B (Fig. 26D-E-F). Met, Val, and Leu are all non-polar, aliphatic, and hydrophobic amino acids. However, only Val has a branched side chain, whereas Met and Leu have linear side chains. In addition, only Met has a sulfur-containing side chain and Leu has a longer side chain with a terminal methyl group. The variable amino acid between different AGO4 proteins could affect the binding between this region and sRNA due to differences in the side chain size and properties of individual amino acids; these features have already been shown to alter binding affinity (P. Wang et al., 2023).

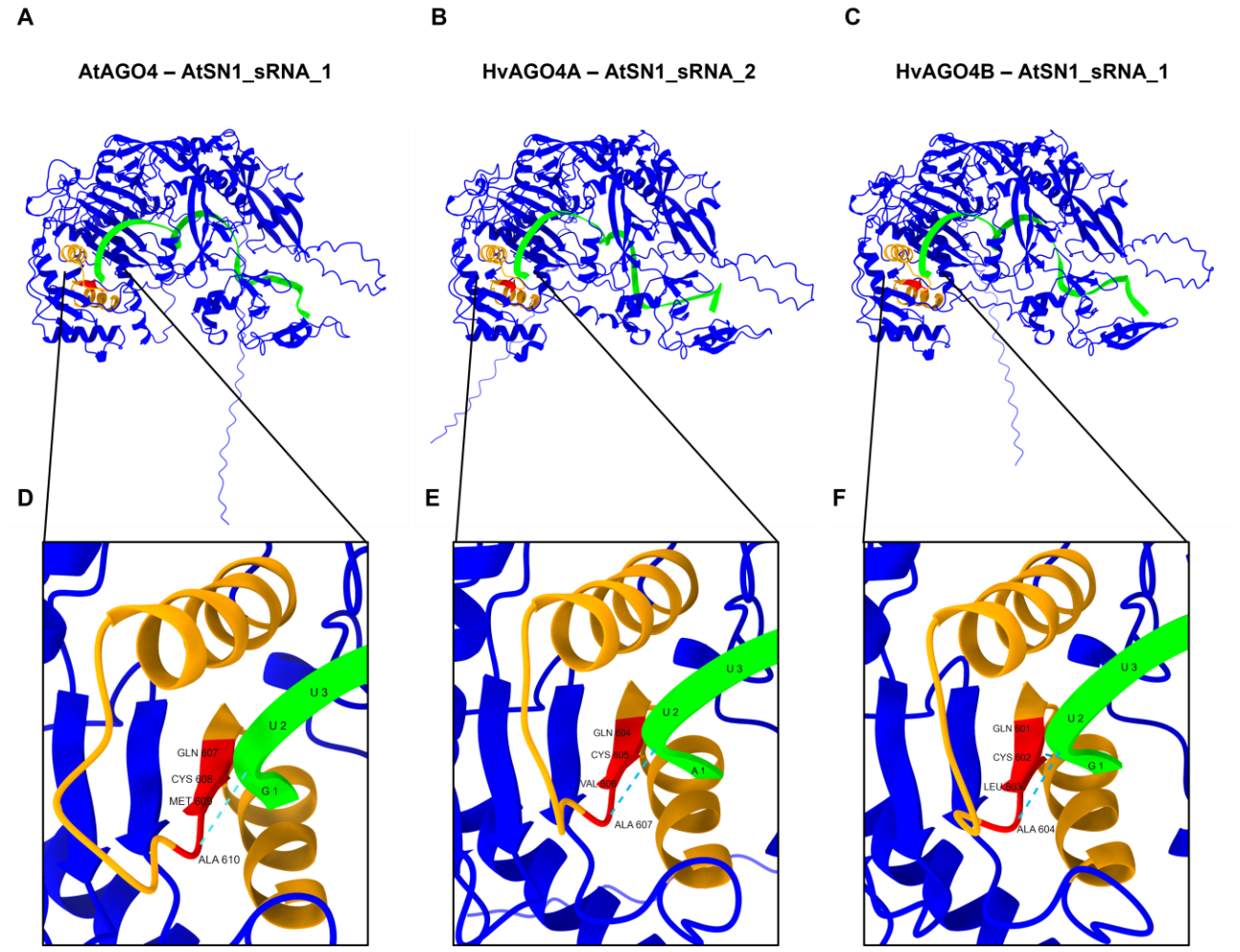

Figure 26. 3D structure of the **A.** AtAGO1, **B.** HvAGO4A and **C.** HvAGO4B bound to different *AtSN1*-derived sRNAs. Each protein structure was predicted using AlphaFold3 (Abramson et al., 2024) together with the more abundant *AtSN1*-derived sRNA found in the specific sRNA-IP sequencing. The protein structure is shown in blue, the PIWI domain region involved in 5′ sRNA anchoring is shown in orange, the 4 amino acids peptide showing a single AA difference between the 3 proteins (QCxA) are shown in red, and the *AtSN1*-derived sRNA is shown in green. *AtSN1*-derived sRNA 1: GUUGUUGGCCCAGUGGUAAAUCUC. *AtSN1*-derived sRNA 2: AUUCGAGACACGUUGGGAAGGAUC.

7 DISCUSSION

This study focused on the understanding of important aspects of the role and regulation of AGO proteins in RNA silencing pathways with a focus on barley. AGO proteins play a crucial role in the recognition and binding of small RNAs, which are essential for the regulation of gene expression in almost all aspects of plant development and response to stimuli (Hutvagner & Simard, 2008; Zaheer et al., 2024). In addition, investigating how the regulatory feedback of AGO1-miR168 is fine-tuned and highly regulated to avoid imbalances in the plant will allow refinement of the model of RNA silencing autoregulation.

The comprehensive analysis of the barley genome has revealed 21 putative candidate genes belonging to the Argonaute (AGO) clade, providing new insights into the role and evolution of AGO proteins in plants. The phylogenetic relationships of these proteins to their counterparts in *Arabidopsis* and rice has provided a consistent nomenclature and a basis for comparative analysis. Three major clades (AGO1/5/10, AGO2/3/7, and AGO4/6/9) were distinguished, along with an additional subclade, AGO18, unique to grasses (Das et al., 2020; Vaucheret, 2008; H. Zhang et al., 2015). This diversity suggests a complex evolutionary history of AGO proteins. Notably, the number of AGO proteins in rice and barley exceeds that in *Arabidopsis*, especially within the AGO1/5/10 clade. This expansion, particularly the specific duplication of AGO1 and AGO5, could indicate the acquisition of new or specific functions and different expression pathways in these species (Zaheer et al., 2024). The identification of five different AGO1 orthologous genes in barley, reminiscent of the four copies found in rice (Kapoor et al., 2008; Wu et al., 2010), underscores the complexity and diversity of the expansion of the *AGO* genes family across plant species. An in-depth exploration of *AGO1* genes expression in barley using the BarleyExpDB database (T. Li et al., 2023) provided a detailed understanding of differential gene expression patterns in different plant tissues. Each of the five *AGO1* genes exhibited unique expression profiles, with *HvAGO1B_1* emerging as the most highly expressed gene in inflorescences at different developmental stages, demonstrating tissue-specific expression dynamics. The differential expression patterns observed among the *AGO1* genes in barley suggest intricate regulatory mechanisms governing gene expression in specific tissues and developmental stages. Furthermore, the conservation of the miR168 target site among the five *AGO1* genes in barley sheds light on the potential regulatory role of this microRNA in modulating gene expression in different tissues. Analysis using psRNATarget revealed different expectation values for the miR168 target site in barley *AGO1* genes, indicating the degree of similarity between the miRNA and its target

sequence. In particular, the near-perfect complementarity between miR168 and *AGO1* mRNAs and the presence of conserved mismatches at specific positions, such as position 14 shared between barley *AGO1* and *Arabidopsis AGO1* genes, suggest potential functional conservation in miR168-mediated regulation across these species. Alignment of the region of AGO1 proteins involved in binding to the 5' end of miRNA showed a large degree of conservation among AGO1 of different plants (Fig. 10C), but also showed how AGO1A and AGO1B proteins of rice and barley, in agreement with the phylogenetic tree, have a higher degree of conservation with AGO1 of *Arabidopsis*. In fact, all these proteins have complete identity of all binding sites with the 5' end of the miRNA. In contrast, AGO1C and AGO1D proteins showed a lower degree of conservation with the presence of three different amino acids in the binding sites with 5' end of the miRNA. In examining the role of specific amino acid differences in orthologous proteins and their potential impact on the binding of miRNAs, the peptide segment spanning positions 705 to 708 (QCCx) was investigated. Notably, there is a variation at the third amino acid position within this segment, where leucine (L) is replaced by cysteine (C) in some proteins. Additionally, conserved differences at positions 711 and 719 further suggest a functional divergence among these proteins. The substitution of leucine for cysteine at the third position in the peptide (QCCL vs. QCCC) is significant because of the different chemical properties of these amino acids (Cléry et al., 2008; Ray et al., 2009). Leucine, a hydrophobic residue, helps stabilize hydrophobic interactions within the protein. In contrast, cysteine contains a thiol (-SH) group capable of forming disulfide bonds that can induce conformational changes and potentially form intra- or intermolecular bridges (Lunde et al., 2007). At position 711, proteins with QCCL have histidine (H), while those with QCCC have glutamine (Q). Histidine's imidazole side chain allows for a variety of interactions, including hydrogen bonding, metal ion coordination, and acting as a proton donor or acceptor (Glisovic et al., 2008). Glutamine, with its amide side chain, forms hydrogen bonds but lacks the broader interaction capabilities of histidine. The variation at position 719, where QCCL proteins have tyrosine (Y) and QCCC proteins have isoleucine (I), further underscores potential functional differences. The polar, aromatic side chain of tyrosine can participate in hydrogen bonding and stacking interactions with nucleic acids, thereby increasing RNA binding affinity (Cléry et al., 2008; Glisovic et al., 2008; Lunde et al., 2007). In contrast, isoleucine is hydrophobic and contributes to protein stability primarily through hydrophobic interactions. The lack of hydrogen bonding and stacking capabilities in isoleucine suggests that QCCC proteins may exhibit different binding affinities and interaction dynamics compared to their QCCL counterparts. The combination of these specific amino acid differences likely results in different structural conformations and binding properties

between QCCL and QCCC proteins. These differences are conserved between rice and barley, though only rice AGO1 proteins have been analyzed (Wu et al., 2009, 2010). Specifically, rice AGO1A, AGO1B, and AGO1C exhibit slicing activity and preferentially bind sRNAs with a 5′ U, similar to their *Arabidopsis* counterparts. Deep sequencing of sRNAs associated with each rice AGO1 revealed that all predominantly recruit miRNAs, suggesting a general redundancy in the miRNA pathway. However, some miRNAs are predominantly recruited by specific AGO1s, indicating evolved specificity. This specificity could arise from differences in spatial or temporal expression, or subcellular localization, and may also be present in barley.

In contrast to the duplication of *AGO1*, the relative size similarity of the other two clades (AGO2/3/7 and AGO4/6/9) across *Arabidopsis*, rice, and barley suggests conservation of ancestral functions over time. The multiple duplication of *AGO* genes in monocots, in contrast to the single copies of *AGO1* and *AGO4* in *Arabidopsis*, suggests a selective advantage in gene regulation and environmental adaptation (Z. Li et al., 2022; A. Mallory & Vaucheret, 2012; H. Zhang et al., 2015). For example, *MEL1* in the AGO5 subclade has acquired germline-specific expression and the ability to bind phasiRNAs (Y. Liu et al., 2020; Nonomura et al., 2007; J. Zhai, Zhang, et al., 2015). This study found that the duplication of *AGO1* and *AGO4* in barley led to specialization in tissue localization, expression levels, and small RNA binding affinity. This supports the idea that gene duplication has played a key role in the evolution of gene regulation in monocots, a phenomenon also observed in other plants (Kapoor et al., 2008; Qian et al., 2011; Yang et al., 2013; L. Zhai et al., 2019). This comprehensive analysis not only elucidates the intricate gene expression patterns and regulatory mechanisms governing *AGO1* genes in barley, but also highlights the evolutionary conservation of regulatory elements such as miR168 in fine-tuning gene expression across plant species. The precise tissue-specific expression patterns and conservation of miRNA target sites underscore the sophisticated regulatory networks orchestrating gene expression in plants and provide valuable insights into the molecular mechanisms underlying plant development and stress responses.

In plants, miRNAs serve as key regulators of growth and development, as well as adaptation to biotic and abiotic stresses and other physiological processes. They do this by controlling the expression of target transcription factors and proteins associated with stress response (S. Li et al., 2017). These multiple regulatory functions make the miRNA pathway one of the most adaptable and diverse regulatory mechanisms. Given these characteristics, the miRNA pathway is meticulously adjusted in response to environmental changes. This adaptation involves transcriptional regulation of miRNA-encoding genes, tissue-specific expression of biogenesis cofactors, post-translational modifications,

and management of miRNA stability and processing (Manavella et al., 2019). The secondary structure of the pre-miRNA, which includes the miRNA/miRNA* duplex region, is crucial in determining the efficiency of miRNA biogenesis (Meyers et al., 2010; Moro et al., 2018; Z. Wang et al., 2018; J. Zhu et al., 2021). In addition, both the 5'-nucleotide and the structural motifs of the precursors can influence the specification of AGO proteins, thereby determining the sorting of miRNAs/sRNAs into the appropriate executor complexes (Iki, 2017; Iki et al., 2018; Mi et al., 2008; X. Zhang et al., 2014). The key effector protein in the miRNA pathway, AGO1, is regulated by a feedback mechanism involving the conserved miR168 family, which specifically targets *AGO1* mRNA for regulation (A. Mallory & Vaucheret, 2012). The importance of the miR168-driven auto-regulatory loop was highlighted when overexpression of a miR168-resistant version of *AGO1* mRNA led to various developmental defects and ultimately to plant death (Vaucheret et al., 2004). These findings underscore that unbalanced overaccumulation of AGO1 protein is a serious threat to the proper functioning of the plant.

Previous studies using size separation gel filtration methods on crude plant extracts revealed a unique characteristic of miR168. It was found that a small fraction of mature miR168 was present in high molecular weight RNA-induced silencing complexes (HMW-RISCs), while the majority accumulated in low molecular weight fractions in a protein-unbound form (Várallyay et al., 2010). Furthermore, it was shown that transient or stable overexpression of different miRNA precursors can lead to different degrees of HMW-RISC loading, ranging from complete (miR159), efficient (miR171) to limited (miR168) in the same cellular context (Dalmadi et al., 2019). These miRNAs mature from precursors with different secondary structures, suggesting that the structural motifs of miRNA precursors could influence the AGO1 loading rate. A significant increase in miR168/AGO1 RISC loading and AGO1 downregulation was not observed in transgenic lines overexpressing the *MIR168a* precursor. The majority of overexpressed miR168 was allocated to the free pool, with only a small subset loaded into the RISC. This observation suggests that the limited balance of miR168 loading into AGO1 plays a critical biological role. We propose that the structural motifs of the *MIR168a* precursor may contribute to the limited loading of miR168 into the AGO1-RISC complex.

A collection of alternative precursor fragments that produce miR168 was developed by altering the wild-type *MIR168a* precursor miRNA duplex region or by expressing miR168 from heterologous constructs containing the *hvu-MIR171* precursor backbone: *MIR168-4bp*, *MIR168-3mm*, *AMIR-1* and *AMIR-2*. The secondary structure of the miRNA/miRNA* duplex regions in these constructs was adjusted by introducing changes only in the miRNA*, while leaving the guide strand unchanged. To

ensure that the miRNA abundance does not influence the loading efficiency, transgenic lines with similar expression patterns were selected from the different modified precursor fragments. These transgenic lines were used to investigate the effect of altered miR168 precursor structures on the loading efficiency of miR168 into AGO1-RISC. The overexpressed miR168 species were primarily produced in the correct size, and no increase in secondary siRNA production from *AGO1* mRNA was observed in the experiments (Dalmadi et al., 2021). However, a change in the 5′ U/C end ratio of miR168 was observed upon overexpression of several modified miR168 precursor fragments. This alternative maturation of miR168 could potentially influence loading and subsequent effects on the AGO1 feedback loop. The similar AGO loading characteristics of 5′ C or 5′ U miR168 species in wild-type plants (Fig. 18B-C) and the association of different AGO1 loading efficiencies with similar 5′ U/C miR168 ratios make this possibility less likely. Nevertheless, we cannot completely exclude the influence of altered miR168 isoform production on differential AGO sorting or loading, such as in floral tissues where AGO5, which prefers 5' C miRNAs, is predominantly expressed (Vaucheret et al., 2004).

The *MIR168a* precursor, when efficiently processed, generates a substantial excess of miR168, but only a small fraction of this is loaded into AGO1-RISC. The unincorporated miR168 species accumulate as miR168/miR168* duplexes in the cytoplasm. The balance between AGO1-RISC-loaded and unbound miR168 is determined by the structural features of the precursor, which includes the miR168/miR168* duplex region. Other miRNA precursors, such as MIR159a or MIR171a, have structural features that allow for more efficient AGO1 loading.

The striking variation in the shape and size of RNA stem loops in plant miRNA precursors compared to their more uniform animal counterparts suggests that structural features may play an important role in the biogenesis and function of plant miRNAs. The regulated loading effect of miR168 guided by RNA structural motifs could potentially be relevant to other miRNAs. This theory is supported by the observation that numerous miRNAs have both AGO-loaded and AGO-unbound forms (Dalmadi et al., 2019). Recent evidence indicates that miRNA loading of AGO1 occurs primarily in the nucleus, suggesting that the loading rate of miR168 is regulated in the nucleus (Bologna et al., 2018). However, it cannot be excluded that cytoplasmic miRNA duplexes serve as a biologically active reserve.

The binding specificity of AGO1 may be influenced not only by the 5' nucleotide, but also by the secondary structure of the miRNA duplex, as previously demonstrated (Bartel, 2009). While the importance of the 5'-nucleotide in AGO protein loading is still under investigation, recent evidence

suggests that *Arabidopsis* AGO4 has a relatively simpler binding specificity for small RNAs compared to other AGO proteins. *Arabidopsis* AGO4 binds predominantly to 24-nucleotide RNAs starting with adenosine but can also interact with RNAs starting with other nucleotides (A, U, C, or G) and efficiently facilitate target RNA cleavage (Mi et al., 2008). This distinct behavior contradicts with findings in rice, where OsAGO4A and OsAGO4B exhibit unique 5'-nucleotide preferences not observed in *Arabidopsis*, providing valuable insights into nucleotide specificity and potential conservation in monocots (Wu et al., 2010). Despite the earlier discovery of these affinities in rice, the study of barley AGO4 proteins could be essential to validate and extend these findings in different plant species. In addition, the study of barley AGO4 proteins provides an opportunity to explore potential variations or similarities in small RNA binding properties and regulatory mechanisms across cereal crops, contributing to a more comprehensive understanding of AGO protein functionality in plant systems.

To identify genes or protein sequences in barley that are orthologous to the RdDM pathway, *Arabidopsis* and rice are often used as model plants because they are among the few plants in which this mechanism has been thoroughly studied. Moreover, as monocotyledons, rice and barley share certain similarities in various processes. We discovered 4 genes in barley that belong to the AGO4/6/9 clade: *HvAGO4a*, *HvAGO4b*, *HvAGO15*, and *HvAGO6*. These genes were previously named using simple computational methods, but their nomenclature contained inaccuracies and their analysis was not complete (Hamar et al., 2020; Madsen et al., 2009; Yao et al., 2021). The primary focus was on *HvAGO4a* and *HvAGO4b*, as the expression level of *HvAGO15* was not detectable and the functionalities of *HvAGO6* were potentially different from those of *AGO4* (Duan et al., 2015; McCue et al., 2015; Trujillo et al., 2018). The two barley *AGO4* genes showed minor variations in expression levels in inflorescence tissues (Fig. 11D). Comparable changes in their expression were observed in several databases, including BaRTv1.0, ePlant, and BarleyExpDB. These databases also confirmed the previously reported observation for *Arabidopsis AGO4* that the expression of these two genes is more pronounced in inflorescences compared to other plant tissues (Chan et al., 2004; T. Li et al., 2023; Mascher et al., 2017; Thiel et al., 2021; Zilberman et al., 2003).

Arabidopsis thaliana was chosen for heterologous complementation of the barley *AGO4* genes to allow a detailed functional comparison between HvAGO4A, HvAGO4B, and AtAGO4. This approach provided a solid framework to evaluate whether the barley AGO4 proteins might function similarly to the well-characterized AtAGO4 in the extensively studied *Arabidopsis* model system. To verify and test the AGO4 functionality of the identified putative barley genes, they were introduced

into the *Arabidopsis ago4-3* mutant. To test the complementation effect of transgenic barley *AGO4* genes, the *AtSN1* and *AtROS1* loci were specifically selected because of their documented transcriptional response to mutations in the RdDM pathway (Córdoba-Cañero et al., 2017; Havecker et al., 2010; Zilberman et al., 2003). Importantly, *ROS1*, a DNA glycosylase/lyase gene that is tightly regulated by AGO4-mediated promoter methylation, plays a critical role in maintaining the methylation balance (Córdoba-Cañero et al., 2017; Havecker et al., 2010; Lei et al., 2015; Tang et al., 2016). The regulatory interaction between AGO4 and ROS1 is crucial given their opposing roles at common target loci. Our research confirmed that the absence of AGO4 function in *Arabidopsis* significantly affects the expression of these loci. Interestingly, barley AGO4 proteins effectively restored the expression levels of both loci to wild-type conditions. Most importantly, methylation levels at the *AtSN1* locus were restored, suggesting a highly regulated and interrelated effect. These findings highlight the broad conservation of the RdDM pathway across plant species, as demonstrated by the ability of barley AGO4 proteins to successfully restore pathway functionality in *Arabidopsis*. This highlights the critical role of AGO4 in controlling DNA methylation and gene expression at specific loci, such as *AtSN1* and *AtROS1*. The fundamental and evolutionarily conserved role of AGO4 in the RdDM pathway is underscored by our results, which are consistent with previous studies identifying AGO4 as a key regulatory protein in RdDM across plant species.

Previous research has highlighted the effects of heat stress on TE activity, suggesting a role for the RdDM pathway (Cavrak et al., 2014; Lang-Mladek et al., 2010). Although not all elements of this pathway are critical for basal heat stress tolerance, the absence of AGO4 in *Arabidopsis* has been found to increase plant susceptibility to this particular type of stress (Popova et al., 2013). The *Ty1/copia*-type retrotransposon *ONSEN* (*ATCOPIA78*), which is known to be activated under heat stress conditions in *Arabidopsis*, showed increased transcript levels and presence of ecDNA in mutants associated with the RdDM pathway (Hayashi et al., 2020; Ito, 2022; Ito et al., 2011, 2013). The activation mechanism specifically involves the identification of a sequence within the long terminal repeat (LTR) by the heat-responsive transcription factor HsfA2 (Cavrak et al., 2014). Interestingly, under normal conditions, *ONSEN* retains its inactivity even in mutants related to the RdDM pathway, due to the lack of CpG and CHG methylation sites in the *ONSEN* promoter; furthermore, the decrease in DNA methylation at the CHH sites is insufficient to activate the element (Cavrak et al., 2014; Ito et al., 2011). When assessing the upregulation of *ONSEN* transcripts after heat stress in all plants, it was observed that the basal level of *ONSEN* was extremely low under controlled conditions, with a 2000-fold increase in the wild type and a significant 7000-fold increase

in the *ago4-3* mutant (Fig. 21D). Our results indicate that both barley AGO4 proteins effectively restored the ecDNA and transcript levels of *ONSEN* to wild-type levels, while the *ago4-3* mutant shows elevated levels in both cases (Fig. 21A-B). Furthermore, the degree of *ONSEN* downregulation was found to be directly proportional to the expression levels of the introduced transgenes. This observation suggests that the presence of barley AGO4 proteins is effective in restoring the repression of specific TEs that are activated in the *ago4-3* mutant during heat stress conditions.

While we have highlighted the similarities and the functionality of the barley AGO4 proteins, a significant divergence was observed in their affinity for the 5′ end nucleotide of sRNAs. HvAGO4B closely resembles AtAGO4 and OsAGO4B in terms of primary properties, both of which tend to bind sRNAs with a 5′ A residue but retain the ability to bind sRNAs with 5′ G or U residues (Fig. 22C) (Havecker et al., 2010; Mi et al., 2008; Wu et al., 2010). Conversely, HvAGO4A appears to bind exclusively to sRNAs with a 5′ A residue. A similar distinction can be seen in rice when OsAGO4A and OsAGO4B are compared (Wu et al., 2010). In addition, comparable differences can be observed with *Arabidopsis* AtAGO6 and AtAGO9, which, in contrast to AtAGO4, both show a strong loading preference for sRNAs with a 5′ A residue, with percentages of 94% and 97% for AtAGO6 and AtAGO9, respectively (Havecker et al., 2010). Previous studies have elucidated the structure of the MID domain of Argonaute proteins, highlighting its role in the selective recognition of the 5′ nucleotide (Frank et al., 2012). Interestingly, a coordination between the MID and PIWI domains was identified as the underlying cause of this specificity (W. Liu et al., 2022). In our research, we found that among the conserved sites involved in 5′ nucleotide anchoring, only one site varied between AtAGO4, HvAGO4A, and HvAGO4B, which was conserved between rice and barley (Fig. 11C). This site consists of four amino acids (QCxA), and the third amino acid could potentially be a determinant of specificity at the 5′ end of sRNAs. In the case of AGO6, this site remains unchanged between *Arabidopsis*, rice, and barley, retaining a QCIx sequence (Fig. 12); indeed, no differences in specificity are observed between *Arabidopsis* AGO6 and rice AGO16 (Havecker et al., 2010; Wu et al., 2010). *In silico* prediction of 3D protein structures has become an essential part of modern research, providing valuable insights into protein function and interactions even in the absence of experimentally determined structures. These computational approaches can generate reliable 3D models of proteins, which can then be used to study their physicochemical properties, identify binding sites, particularly for proteins that are difficult to crystallize, or analyze using experimental techniques (Abramson et al., 2024). AlphaFold3 predictions confirmed that the PIWI domain region, which is responsible for anchoring the 5' end of small RNAs, is the key determinant of binding specificity (Fig.

26). Interestingly, in the created model the sRNA automatically positioned itself with the 5'-end in direct contact with the MID-PIWI protein domains pocket, confirming the critical role of this protein pocket in recognizing and binding the 5'-end of the sRNA (X. Zhang et al., 2014). In addition, the predictions revealed that the sequence of the sRNA influences its precise positioning within the AGO4-sRNA complex. Focusing on the specific amino acid residues within the PIWI domain that interact with the 5' end of the sRNA, a single amino acid difference was identified among the three AGO4 proteins studied. This variable amino acid, located in the central region responsible for 5'-end recognition, and has an identity of a methionine (Met 609) in AtAGO4, a valine (Val 606) in HvAGO4A, and a leucine (Leu 603) in HvAGO4B. Although these amino acids share similar non-polar, aliphatic, and hydrophobic properties, differences in the side chain size and branching patterns may contribute to the altered binding affinities observed between the AGO4 proteins and their target small RNAs. The integration of experimental sRNA-IP sequencing data with high-resolution structural predictions using AlphaFold3 provides a powerful approach to unravel the molecular mechanisms underlying small RNA binding specificity. By pinpointing the critical amino acid residues responsible for 5'-end recognition and their impact on binding affinity, this study provides valuable insights into the evolution and diversification of AGO4 proteins in plants.

Therefore, in monocots, the duplication of the *AGO4* gene could lead to the specialization of AGO4A, resulting in hybrid characteristics between AtAGO4 and AtAGO6, since it mirrors the former in length selection and the latter in affinity for sRNAs starting with an A residue. Indeed, in *Arabidopsis*, AtAGO6 was found to be able to load 24 nt sRNAs as well as shorter sRNAs (19-22 nt), but with an exclusive preference for sRNAs with A residues at their 5′ ends. These length- and nucleotide-specific features ensure non-competition with other AGO proteins, in particular AGO1, which primarily binds miRNAs characterized by a length of 21 nucleotides and a starting base of U (W. Liu et al., 2022; McCue et al., 2015). This diversification may play a more specialized role in barley and tissues where both AGO4 proteins are expressed, possibly involving different localizations at the cell type level between the two.

In our research, we have outlined the unique binding capabilities of the barley HvAGO4 and HvAGO4B proteins to sRNAs in a broader context. Both barley AGO4 proteins bind to specific TE-derived sRNAs, but we found exclusive binding by HvAGO4A in certain cases, which was primarily determined by the first nucleotide of the sRNA sequence. Interestingly, our analysis revealed different patterns of active sRNA production within specific regions of *AtSN1*. The central regions mainly produce sRNAs starting with an A residue, whereas the two lateral regions, especially the 3′-end

terminal segment of the locus, produce sRNAs with G or U residues at the 5′ end (Fig. 25B). In this context, the different profiles of AGO4-bound, TE-derived sRNAs result in a significant difference in the cumulative sRNA abundance mapped to *AtSN1*, especially in the case of HvAGO4A compared to the other two proteins (Fig. 25A). A similar pattern is observed for another TE, *AT5TE27090*, which belongs to the *RathE3* family (Fig. 25C). Notably, the number of TE-derived sRNAs bound to HvAGO4A was almost negligible. This observation underscores the potential for HvAGO4 proteins to have, at least in part, distinct regulatory properties on TEs, especially on relatively short TEs that cannot produce sRNAs suitable for all AGO4-clade proteins. In contrast, the flexible binding properties of AtAGO4 and HvAGO4B may provide a functional advantage over those AGO proteins that bind sRNAs with absolute specificity regarding the 5′ terminal nucleotide of the sRNA.

8 CONCLUSIONS AND RECOMMENDATIONS

In conclusion, this study provides a comprehensive analysis of AGO proteins in barley and sheds light on their roles and mechanisms in small RNA binding and AGO1 homeostasis. Through genome-wide identification and expression analysis, we characterized the diverse set of AGO proteins in barley, highlighting their evolutionary significance and functional diversification. This analysis revealed the presence of multiple AGO proteins, each with unique structural features and regulatory roles, thereby expanding our understanding of the RNA silencing machinery in plants.

In silico analysis of the barley *AGO1* genes revealed their distinct expression patterns in different tissues and developmental stages. This detailed study of the *AGO1* genes family highlighted the conservation of the miR168 target site, underscoring the critical interplay between miR168 and AGO1 in maintaining homeostasis.

Investigation of the miR168 duplex structure revealed its critical role in determining loading efficiency into the AGO1-RISC, providing a dynamic regulatory mechanism that adjusts AGO1 protein levels in response to cellular stimuli. This competitive loading mechanism reveals a sophisticated level of control in miRNA pathway regulation and AGO1 homeostasis.

Further investigation of the barley AGO4 proteins demonstrated their overlapping functionality with distinct small RNA binding properties upon heterologous complementation, highlighting the evolutionary advantage provided by the *AGO4* gene duplication. These proteins exhibit unique binding capabilities, specificities, and involvement in TE regulation, underscoring their critical role in plant genetic regulatory mechanisms.

These findings highlight the complexity and precision of genetic regulatory mechanisms in plants and underscore the need for further research to fully understand these systems. Given the critical role of these regulatory mechanisms in plant growth, development and stress responses, these studies have the potential to contribute significantly to advances in crop improvement and plant biotechnology.

Future research should build on the findings of this study to further elucidate the roles and mechanisms of AGO proteins in barley. A major focus can be the further analysis of barley AGO1 proteins. Investigations could aim at unraveling of the specific functions of these proteins in barley, in particular their specificity for 5'-end nucleotides of small RNAs and tissue-specific functions. This could include high-throughput sequencing to map the small RNA populations associated with the individual AGO1 proteins. In addition, the generation of single and multiple *ago1* mutants using techniques such as

CRISPR/Cas9 would allow a more in-depth investigation of the individual and combined effects of these proteins on plant physiology and development. Comparative studies with similar mutants in rice and *Arabidopsis* could provide valuable insights into the conserved and species-specific roles of AGO1 proteins and enhance our understanding of their regulatory functions across plant species.

Another promising avenue for future research is to investigate how different regions of miRNA precursors influence loading efficiency into AGO1. This will involve dissecting different parts of the precursor structure, including stem-loop regions, flanking sequences and terminal loops, to identify which elements are critical for efficient loading. Understanding the molecular mechanisms behind this enhanced loading efficiency is also essential. Studies should explore the interactions between miRNA precursors and processing enzymes, such as Dicer-like proteins, as well as AGO1 itself.

Further investigation of the role of barley AGO4 proteins is also warranted. Future studies should generate single and double knockout of barley *AGO4* genes to assess their functional redundancy and specific roles in gene regulation. Phenotypic characterization of these mutants under different conditions will help determine the unique and overlapping functions of AGO4 proteins. Research should also focus on the regulatory roles of AGO4 in the regulation of TEs and other gene silencing pathways. This could include studying the small RNA populations associated with AGO4 proteins and identifying their target genes and TEs. In addition, it is important to understand the broader effects of *AGO4* mutations on plant growth, development, and stress responses. Transcriptome analysis of *ago4* mutants could reveal differentially expressed genes and shed light on the regulatory networks controlled by AGO4 proteins and their impact on plant biology.

These recommendations underscore the need for comprehensive functional analyses and mutant studies to deepen our understanding of RNA silencing mechanisms in barley. By addressing these areas, future research can uncover the intricate regulatory networks mediated by AGO proteins and miRNAs, ultimately contributing to advances in crop improvement and plant biotechnology.

9 NEW SCIENTIFIC RESULTS

- We identified and classified 21 putative candidate genes belonging to the Argonaute family in barley through whole genome analysis.

- *In silico* analysis of five barley *AGO1* genes revealed distinct tissue and developmental stage expression patterns and the conservation of the miR168 target site among these genes.

- We showed that modifying the miRNA duplex structure or expressing artificial precursors can alter the loading efficiency of miR168 into AGO1-RISC.

- We demonstrated that barley AGO4 proteins have overlapping functionality with distinct small RNA binding properties upon heterologous complementation.

- We observed distinct regulatory properties of barley AGO4 proteins on transposable elements, especially on relatively short TEs, affecting cumulative sRNA abundance.

10 SUMMARY

In this study we pursuit the in-depth examination of the role barley Argonaute (AGO) proteins and utilized the well-described model plant *Arabidopsis thaliana*, with a particular focus on miRNA loading efficiency in AGO1 and AGO4 functionality. The research explores the complex relationship between the loading efficiency of miR168 into the AGO1-RISC and the structural elements of the miRNA duplex. It reveals that the secondary structure of the *MIR168a* precursor not only affects miR168 processing, but also fine-tunes its loading into AGO1, thereby determining the biologically active subset of the miR168 pool. We have demonstrated the ability to alter the loading efficiency of miR168 into the AGO1-RISC by modifying the miRNA duplex structure or expressing artificial precursors. This reveals a dynamic regulatory mechanism and sheds light on a competitive loading mechanism model for miR168 action. In this model, excess miR168 acts as a buffer that continuously adjusts AGO1 protein levels in response to cellular stimuli. This provides a deeper understanding of miRNA pathway regulation and AGO1 homeostasis. The presence of an unbound cytoplasmic pool was demonstrated for many miRNAs, suggesting that the regulatory effect of competition-based loading efficiency may also apply to other miRNAs.

Through bioinformatic analysis, 21 putative AGO genes were initially identified in the barley genome and phylogenetic analysis placed these AGO proteins into three main clades: AGO1/5/10, AGO2/3/7, and AGO4/6/9, with an additional subclade AGO18 unique to grasses. Five different copies of the *AGO1* gene have been identified in barley, each with distinct expression patterns in different plant tissues and developmental stages. Four of these *AGO1* genes are primarily active in inflorescences, while the expression level of the *HvAGO1A* gene is absent. Comparison of the barley *AGO1* genes with *Arabidopsis AGO1* reveals structural similarity, particularly in the miR168 target site on the *AGO1* mRNAs, which shows conserved mismatches at specific positions. Notably, three barley *AGO1* genes align similarly with miR168, showing specific mismatches but no wobble base pairs, suggesting a conserved interaction pattern with miR168.

The study also investigated the role of AGO4 in *Arabidopsis* and its potential applications in monocots. Bioinformatic analysis of RNA sequencing data identified two actives putative *AGO4* genes in barley, *HvAGO4a* and *HvAGO4b*. These proteins function similarly to AtAGO4 in an *Arabidopsis* heterologous complementation system, primarily binding to 24-nucleotide long small RNAs (sRNAs) and inducing methylation at specific target loci. The diverse binding capacity of the

barley AGO4 proteins is reflected in the AGO4-associated, TE-derived sRNAs and their varying abundance. Both barley AGO4 proteins effectively restore extrachromosomal DNA levels and transcript abundance of the heat-activated *ONSEN* retrotransposon to those observed in wild-type *Arabidopsis* plants.

In conclusion, this work provides insight into the role and regulation of AGO proteins in RNA silencing pathways in barley. The results contribute to the understanding of RNA silencing mechanisms and epigenetic regulation in plants, particularly in barley, offering valuable contributions to the functional analysis of AGO proteins and their potential applications in crop improvement and plant biotechnology.

11 BIBLIOGRAPHY

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature 2024*, 1–3. https://doi.org/10.1038/s41586-024-07487-w

Achkar, N. P., Cambiagno, D. A., & Manavella, P. A. (2016). miRNA Biogenesis: A Dynamic Pathway. *Trends in Plant Science*, *21*(12), 1034–1044. https://doi.org/10.1016/j.tplants.2016.09.003

Afgan, E., Nekrutenko, A., Grüning, B. A., Blankenberg, D., Goecks, J., Schatz, M. C., Ostrovsky, A. E., Mahmoud, A., Lonie, A. J., Syme, A., Fouilloux, A., Bretaudeau, A., Nekrutenko, A., Kumar, A., Eschenlauer, A. C., Desanto, A. D., Guerler, A., Serrano-Solano, B., Batut, B., … Briggs, P. J. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Research*, *50*(W1), W345–W351. https://doi.org/10.1093/NAR/GKAC247

Allen, E., Xie, Z., Gustafson, A. M., & Carrington, J. C. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, *121*(2), 207–221. https://doi.org/10.1016/j.cell.2005.04.004

Arif, M. A., Frank, W., & Khraiwesh, B. (2013). Role of RNA Interference (RNAi) in the Moss Physcomitrella patens. *International Journal of Molecular Sciences 2013, Vol. 14, Pages 1516-1540*, *14*(1), 1516–1540. https://doi.org/10.3390/IJMS14011516

Aubert, J., Bellegarde, F., Oltehua-Lopez, O., Leblanc, O., Arteaga-Vazquez, M. A., Martienssen, R. A., & Grimanelli, D. (2022). AGO104 is a RdDM effector of paramutation at the maize b1 locus. *PLOS ONE*, *17*(8), e0273695. https://doi.org/10.1371/JOURNAL.PONE.0273695

Auweter, S. D., Oberstrass, F. C., & Allain, F. H. T. (2006). Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Research*, *34*(17), 4943. https://doi.org/10.1093/NAR/GKL620

Bajczyk, M., Bhat, S. S., Szewc, L., Szweykowska-Kulinska, Z., Jarmolowski, A., & Dolata, J. (2019). Novel nuclear functions of arabidopsis argonaute1: Beyond rna interference. *Plant Physiology*, *179*(3), 1030–1039. https://doi.org/10.1104/pp.18.01351

Bartel, D. P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. *Cell*, *136*(2), 215–233. https://doi.org/10.1016/J.CELL.2009.01.002

Baulcombe, D. (2004). RNA silencing in plants. *Nature 2004 431:7006*, *431*(7006), 356–363. https://doi.org/10.1038/nature02874

Baulcombe, D. (2015). RNA silencing in plants. *The Biochemist*, *37*(2), 10–13. https://doi.org/10.1042/BIO03702010

Baumberger, N., & Baulcombe, D. C. (2005). Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(33), 11928–11933. https://doi.org/10.1073/PNAS.0505461102

Blevins, T., Podicheti, R., Mishra, V., Marasco, M., Wang, J., Rusch, D., Tang, H., & Pikaard, C. S. (2015). Identification of pol IV and RDR2-dependent precursors of 24 nt siRNAs guiding de novo DNA methylation in arabidopsis. *ELife*, *4*(OCTOBER2015). https://doi.org/10.7554/ELIFE.09591

Böhmdorfer, G., Rowley, M. J., Kuciński, J., Zhu, Y., Amies, I., & Wierzbicki, A. T. (2014). RNA-directed DNA methylation requires stepwise binding of silencing factors to long non-coding RNA. *Plant Journal*, *79*(2), 181–191. https://doi.org/10.1111/tpj.12563

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., & Banning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. *The EMBO Journal*, *17*(1), 170–180. https://doi.org/10.1093/EMBOJ/17.1.170

Bologna, N. G., Iselin, R., Abriata, L. A., Sarazin, A., Pumplin, N., Jay, F., Grentzinger, T., Dal Peraro, M., & Voinnet, O. (2018). Nucleo-cytosolic Shuttling of ARGONAUTE1 Prompts a Revised Model of the Plant MicroRNA Pathway. *Molecular Cell*, *69*(4), 709-719.e5. https://doi.org/10.1016/j.molcel.2018.01.007

Bologna, N. G., Schapire, A. L., Zhai, J., Chorostecki, U., Boisbouvier, J., Meyers, B. C., & Palatnik, J. F. (2013). Multiple RNA recognition patterns during microRNA biogenesis in plants. *Genome Research*, *23*(10), 1675–1689. https://doi.org/10.1101/GR.153387.112

Bologna, N. G., & Voinnet, O. (2014). The diversity, biogenesis, and activities of endogenous silencing small RNAs in Arabidopsis. *Annual Review of Plant Biology*, *65*(Volume 65, 2014), 473–503. https://doi.org/10.1146/ANNUREV-ARPLANT-050213-035728

Borges, F., & Martienssen, R. A. (2015). The expanding world of small RNAs in plants. In *Nature Reviews Molecular Cell Biology* (Vol. 16, Issue 12, pp. 727–741). Nature Publishing Group. https://doi.org/10.1038/nrm4085

Casacuberta, E., & González, J. (2013). The impact of transposable elements in environmental adaptation. *Molecular Ecology*, *22*(6), 1503–1517. https://doi.org/10.1111/MEC.12170

Cavrak, V. V., Lettner, N., Jamge, S., Kosarewicz, A., Bayer, L. M., & Mittelsten Scheid, O. (2014). How a Retrotransposon Exploits the Plant's Heat Stress Response for Its Activation. *PLOS Genetics*, *10*(1), e1004115. https://doi.org/10.1371/JOURNAL.PGEN.1004115

Chan, S. W. L., Zilberman, D., Xie, Z., Johansen, L. K., Carrington, J. C., & Jacobsen, S. E. (2004). RNA Silencing Genes Control de Novo DNA Methylation. *Science*, *303*(5662), 1336. https://doi.org/10.1126/SCIENCE.1095989

Chang, Y., Zhu, C., Jiang, J., Zhang, H., Zhu, J., & Duan, C. (2020). Epigenetic regulation in plant abiotic stress responses. *Journal of Integrative Plant Biology*, *62*(5), 563–580. https://doi.org/10.1111/jipb.12901

Chano, V., Domínguez-Flores, T., Hidalgo-Galvez, M. D., Rodríguez-Calcerrada, J., & Pérez-Ramos, I. M. (2021). Epigenetic responses of hare barley (Hordeum murinum subsp. leporinum) to climate change: an experimental, trait-based approach. *Heredity 2021 126:5*, *126*(5), 748–762. https://doi.org/10.1038/s41437-021-00415-y

Chinnusamy, V., Dalal, M., & Zhu, J.-K. (2014). Epigenetic regulation of abiotic stress responses in plants. In *Plant Abiotic Stress* (pp. 203–229). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118764374.ch8

Cléry, A., Blatter, M., & Allain, F. H. T. (2008). RNA recognition motifs: boring? Not quite. *Current Opinion in Structural Biology*, *18*(3), 290–298. https://doi.org/10.1016/J.SBI.2008.04.002

Córdoba-Cañero, D., Cognat, V., Ariza, R. R., Roldán Arjona, T., & Molinier, J. (2017). Dual control of ROS1-mediated active DNA demethylation by DNA damage-binding protein 2 (DDB2). *The Plant Journal*, *92*(6), 1170–1181. https://doi.org/10.1111/TPJ.13753

Cuerda-Gil, D., & Slotkin, R. K. (2016). Non-canonical RNA-directed DNA methylation. *Nature Plants 2016 2:11*, *2*(11), 1–8. https://doi.org/10.1038/nplants.2016.163

Czotter, N., Molnár, J., Pesti, R., Demián, E., Baráth, D., Varga, T., & Várallyay, É. (2018). Use of siRNAs for diagnosis of viruses associated to woody plants in nurseries and stock collections. *Methods in Molecular Biology*, *1746*, 115–130. https://doi.org/10.1007/978-1-4939-7683-6_9

Dalmadi, Á., Gyula, P., Bálint, J., Szittya, G., & Havelda, Z. (2019). AGO-unbound cytosolic pool of mature miRNAs in plant cells reveals a novel regulatory step at AGO1 loading. *Nucleic Acids Research*, *47*(18), 9803–9817. https://doi.org/10.1093/NAR/GKZ690

Dalmadi, Á., Miloro, F., Bálint, J., Várallyay, É., & Havelda, Z. (2021). Controlled RISC loading efficiency of miR168 defined by miRNA duplex structure adjusts ARGONAUTE1 homeostasis. *Nucleic Acids Research*, *49*(22), 12912–12928. https://doi.org/10.1093/NAR/GKAB1138

Das, S., Swetha, C., Pachamuthu, K., Nair, A., & Shivaprasad, P. V. (2020). Loss of function of Oryza sativa Argonaute 18 induces male sterility and reduction in phased small RNAs. *Plant Reproduction*, *33*(1), 59–73. https://doi.org/10.1007/s00497-020-00386-w

Dasgupta, P., & Chaudhuri, S. (2019). Analysis of DNA Methylation Profile in Plants by Chop-PCR. *Methods in Molecular Biology (Clifton, N.J.)*, *1991*, 79–90. https://doi.org/10.1007/978-1-4939-9458-8_9

Dawson, I. K., Russell, J., Powell, W., Steffenson, B., Thomas, W. T. B., & Waugh, R. (2015). Barley: a translational model for adaptation to climate change. *New Phytologist*, *206*(3), 913–931. https://doi.org/10.1111/NPH.13266

Duan, C., Zhang, H., Tang, K., Zhu, X., Qian, W., Hou, Y., Wang, B., Lang, Z., Zhao, Y., Wang, X., Wang, P., Zhou, J., Liang, G., Liu, N., Wang, C., & Zhu, J. (2015). Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. *The EMBO Journal*, *34*(5), 581–592. https://doi.org/10.15252/embj.201489453

Endo, Y., Iwakawa, H. O., & Tomari, Y. (2013). Arabidopsis ARGONAUTE7 selects miR390 through multiple checkpoints during RISC assembly. *EMBO Reports*, *14*(7), 652–658. https://doi.org/10.1038/EMBOR.2013.73

Erdmann, R. M., & Picard, C. L. (2020). RNA-directed DNA Methylation. *PLOS Genetics*, *16*(10), e1009034. https://doi.org/10.1371/JOURNAL.PGEN.1009034

Eun, C., Lorkovic, Z. J., Sasaki, T., Naumann, U., Matzke, A. J. M., & Matzke, M. (2012). Use of forward genetic screens to identify genes required for RNA-directed DNA methylation in Arabidopsis thaliana. *Cold Spring Harbor Symposia on Quantitative Biology*, *77*, 195–204. https://doi.org/10.1101/SQB.2012.77.015099

Fang, X., & Qi, Y. (2016). RNAi in Plants: An Argonaute-Centered View. *The Plant Cell*, *28*(2), 272–285. https://doi.org/10.1105/TPC.15.00920

FAOSTAT. (2023). Retrieved July 26, 2023, from https://www.fao.org/faostat/en/#home

Frank, F., Hauver, J., Sonenberg, N., & Nagar, B. (2012). Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. *The EMBO Journal*, *31*(17), 3588–3595. https://doi.org/10.1038/EMBOJ.2012.204

Fukudome, A., Singh, J., Mishra, V., Reddem, E., Martinez-Marquez, F., Wenzel, S., Yan, R., Shiozaki, M., Yu, Z., Wang, J. C. Y., Takagi, Y., & Pikaard, C. S. (2021). Structure and RNA template requirements of Arabidopsis RNA-DEPENDENT RNA POLYMERASE 2. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(51), e2115899118. https://doi.org/10.1073/PNAS.2115899118

Gallego-Bartolomé, J. (2020). DNA methylation in plants: mechanisms and tools for targeted manipulation. *New Phytologist*, *227*(1), 38–44. https://doi.org/10.1111/NPH.16529

Ghildiyal, M., & Zamore, P. D. (2009). Small silencing RNAs: an expanding universe. *Nature Reviews Genetics 2009 10:2*, *10*(2), 94–108. https://doi.org/10.1038/nrg2504

Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS Letters*, *582*(14), 1977–1986. https://doi.org/10.1016/J.FEBSLET.2008.03.004

Gu, S., Jin, L., Huang, Y., Zhang, F., & Kay, M. A. (2012). Slicing-Independent RISC Activation Requires the Argonaute PAZ Domain. *Current Biology*, *22*(16), 1536–1542. https://doi.org/10.1016/J.CUB.2012.06.040

Guo, Q., Liu, Q., A. Smith, N., Liang, G., & Wang, M.-B. (2016). RNA Silencing in Plants: Mechanisms, Technologies and Applications in Horticultural Crops. *Current Genomics*, *17*(6), 476–489. https://doi.org/10.2174/1389202917666160520103117

Guo, W., Wang, D., & Lisch, D. (2021). RNA-directed DNA methylation prevents rapid and heritable reversal of transposon silencing under heat stress in Zea mays. *PLOS Genetics*, *17*(6), e1009326. https://doi.org/10.1371/JOURNAL.PGEN.1009326

Haag, J. R., & Pikaard, C. S. (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nature Reviews Molecular Cell Biology 2011 12:8*, *12*(8), 483–492. https://doi.org/10.1038/nrm3152

Haag, J. R., Ream, T. S., Marasco, M., Nicora, C. D., Norbeck, A. D., Pasa-Tolic, L., & Pikaard, C. S. (2012). *Molecular Cell In Vitro Transcription Activities of Pol IV, Pol V, and RDR2 Reveal Coupling of Pol IV and RDR2 for dsRNA Synthesis in Plant RNA Silencing*. https://doi.org/10.1016/j.molcel.2012.09.027

Hamar, E., Szaker, H. M., Kis, A., Dalmadi, A., Miloro, F., Szittya, G., Taller, J., Gyula, P., Csorba, T., & Havelda, Z. (2020). Genome-Wide Identification of RNA Silencing-Related Genes and Their Expressional Analysis in Response to Heat Stress in Barley (Hordeum vulgare L.). *Biomolecules 2020, Vol. 10, Page 929*, *10*(6), 929. https://doi.org/10.3390/BIOM10060929

Havecker, E. R., Wallbridge, L. M., Hardcastle, T. J., Bush, M. S., Kelly, K. A., Dunn, R. M., Schwach, F., Doonan, J. H., & Baulcombe, D. C. (2010). The Arabidopsis RNA-Directed DNA Methylation Argonautes Functionally Diverge Based on Their Expression and Interaction with Target Loci. *The Plant Cell*, *22*(2), 321. https://doi.org/10.1105/TPC.109.072199

Hayashi, Y., Takehira, K., Nozawa, K., Suzuki, T., Masuta, Y., Kato, A., & Ito, H. (2020). ONSEN shows different transposition activities in RdDM pathway mutants. *Genes & Genetic Systems*, *95*(4), 183–190. https://doi.org/10.1266/GGS.20-00019

He, X. J., Hsu, Y. F., Zhu, S., Wierzbicki, A. T., Pontes, O., Pikaard, C. S., Liu, H. L., Wang, C. S., Jin, H., & Zhu, J. K. (2009). An Effector of RNA-Directed DNA Methylation in Arabidopsis Is an ARGONAUTE 4- and RNA-Binding Protein. *Cell*, *137*(3), 498–508. https://doi.org/10.1016/J.CELL.2009.04.028

Huang, K., Wu, X. X., Fang, C. L., Xu, Z. G., Zhang, H. W., Gao, J., Zhou, C. M., You, L. L., Gu, Z. X., Mu, W. H., Feng, Y., Wang, J. W., & Zhang, Y. (2021). Pol IV and RDR2: A two-RNA-polymerase machine that produces double-stranded RNA. *Science*, *374*(6575), 1579–1586. https://doi.org/10.1126/SCIENCE.ABJ9184

Hung, Y. H., & Slotkin, R. K. (2021). The initiation of RNA interference (RNAi) in plants. *Current Opinion in Plant Biology*, *61*, 102014. https://doi.org/10.1016/J.PBI.2021.102014

Hutvagner, G., & Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. *Nature Reviews Molecular Cell Biology 2008 9:1*, *9*(1), 22–32. https://doi.org/10.1038/nrm2321

Iki, T. (2017). Messages on small RNA duplexes in plants. *Journal of Plant Research*, *130*(1), 7–16. https://doi.org/10.1007/S10265-016-0876-2

Iki, T., Clé Ry, A., Bologna, N. G., Sarazin, A., Brosnan, C. A., Pumplin, N., Dé, F., Allain, R. H. T., & Voinnet, O. (2018). Structural Flexibility Enables Alternative Maturation, ARGONAUTE Sorting and Activities of miR168, a Global Gene Silencing Regulator in Plants. *Molecular Plant*, *11*, 1008–1023. https://doi.org/10.1016/j.molp.2018.05.006

Iki, T., Cléry, A., Bologna, N. G., Sarazin, A., Brosnan, C. A., Pumplin, N., Allain, F. H. T., & Voinnet, O. (2018). Structural Flexibility Enables Alternative Maturation, ARGONAUTE Sorting and Activities of miR168, a Global Gene Silencing Regulator in Plants. *Molecular Plant*, *11*(8), 1008–1023. https://doi.org/10.1016/j.molp.2018.05.006

Ito, H. (2022). Environmental stress and transposons in plants. *Genes & Genetic Systems*, *97*(4), 169–175. https://doi.org/10.1266/GGS.22-00045

Ito, H., Gaubert, H., Bucher, E., Mirouze, M., Vaillant, I., & Paszkowski, J. (2011). An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. *Nature*, *472*(7341), 115–119. https://doi.org/10.1038/nature09861

Ito, H., Yoshida, T., Tsukahara, S., & Kawabe, A. (2013). Evolution of the ONSEN retrotransposon family activated upon heat stress in Brassicaceae. *Gene*, *518*(2), 256–261. https://doi.org/10.1016/J.GENE.2013.01.034

Izawa, T., & Shimamoto, K. (1996). Becoming a model plant: The importance of rice to plant science. *Trends in Plant Science*, *1*(3), 95–99. https://doi.org/10.1016/S1360-1385(96)80041-0

Ji, L., & Chen, X. (2012). Regulation of small RNA stability: methylation and beyond. *Cell Research 2012 22:4*, *22*(4), 624–636. https://doi.org/10.1038/cr.2012.36

Johnson, C., Kasprzewska, A., Tennessen, K., Fernandes, J., Nan, G. L., Walbot, V., Sundaresan, V., Vance, V., & Bowman, L. H. (2009). Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. *Genome Research*, *19*(8), 1429–1440. https://doi.org/10.1101/GR.089854.108

Jones, A. L., Thomas, C. L., & Maule, A. J. (1998). De novo methylation and co-suppression induced by a cytoplasmically replicating plant RNA virus. *The EMBO Journal*, *17*(21), 6385–6393. https://doi.org/10.1093/EMBOJ/17.21.6385

Kapazoglou, A., Drosou, V., Argiriou, A., & Tsaftaris, A. S. (2013). The study of a barley epigenetic regulator, HvDME, in seed development and under drought. *BMC Plant Biology*, *13*(1), 1–16. https://doi.org/10.1186/1471-2229-13-172

Kapoor, M., Arora, R., Lama, T., Nijhawan, A., Khurana, J. P., Tyagi, A. K., & Kapoor, S. (2008). Genome-wide identification, organization and phylogenetic analysis of Dicer-like, Argonaute and RNA-dependent RNA Polymerase gene families and their expression analysis during reproductive development and stress in rice. *BMC Genomics*, *9*(1), 1–17. https://doi.org/10.1186/1471-2164-9-451

Kazazian, H. H. (2004). Mobile Elements: Drivers of Genome Evolution. *Science*, *303*(5664), 1626–1632. https://doi.org/10.1126/SCIENCE.1089670

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nature Methods 2015 12:4*, *12*(4), 357–360. https://doi.org/10.1038/nmeth.3317

Kim, J. M., Sasaki, T., Ueda, M., Sako, K., & Seki, M. (2015). Chromatin changes in response to drought, salinity, heat, and cold stresses in plants. *Frontiers in Plant Science*, *6*(MAR), 124925. https://doi.org/10.3389/FPLS.2015.00114

Kis, A., Tholt, G., Ivanics, M., Várallyay, É., Jenes, B., & Havelda, Z. (2016). Polycistronic artificial miRNA-mediated resistance to Wheat dwarf virus in barley is highly efficient at low temperature. *Molecular Plant Pathology*, *17*(3), 427–437. https://doi.org/10.1111/MPP.12291

Kong, L., Liu, Y., Wang, X., & Chang, C. (2020). Insight into the Role of Epigenetic Processes in Abiotic and Biotic Stress Response in Wheat and Barley. *International Journal of Molecular Sciences 2020, Vol. 21, Page 1480*, *21*(4), 1480. https://doi.org/10.3390/IJMS21041480

Kryuvrysanaki, N., James, A., Tselika, M., Bardani, E., & Kalantidis, K. (2021). RNA silencing pathways in plant development and defense. *The International Journal of Developmental Biology*, *66*(1-2–3), 163–175. https://doi.org/10.1387/IJDB.210189KK

Lakatos, L., Szittya, G., Silhavy, D., & Burgyán, J. (2004). Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *EMBO Journal*, *23*(4), 876–884. https://doi.org/10.1038/SJ.EMBOJ.7600096

Lang-Mladek, C., Popova, O., Kiok, K., Berlinger, M., Rakic, B., Aufsatz, W., Jonak, C., Hauser, M. T., & Luschnig, C. (2010). Transgenerational Inheritance and Resetting of Stress-Induced Loss of Epigenetic Gene Silencing in Arabidopsis. *Molecular Plant*, *3*(3), 594–602. https://doi.org/10.1093/MP/SSQ014

Law, J. A., Du, J., Hale, C. J., Feng, S., Krajewski, K., Palanca, A. M. S., Strahl, B. D., Patel, D. J., & Jacobsen, S. E. (2013). Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature*, *498*(7454), 385–389. https://doi.org/10.1038/nature12178

Law, J. A., Vashisht, A. A., Wohlschlegel, J. A., & Jacobsen, S. E. (2011). SHH1, a Homeodomain Protein Required for DNA Methylation, As Well As RDR2, RDM4, and Chromatin Remodeling Factors, Associate with RNA Polymerase IV. *PLOS Genetics*, *7*(7), e1002195. https://doi.org/10.1371/JOURNAL.PGEN.1002195

Lee, T. F., Gurazada, S. G. R., Zhai, J., Li, S., Simon, S. A., Matzke, M. A., Chen, X., & Meyers, B. C. (2012). RNA polymerase V-dependent small RNAs in Arabidopsis originate from small, intergenic loci including most SINE repeats. *Epigenetics*, *7*(7), 781–795. https://doi.org/10.4161/EPI.20290

Lei, M., Zhang, H., Julian, R., Tang, K., Xie, S., & Zhu, J. K. (2015). Regulatory link between DNA methylation and active demethylation in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(11), 3553–3557. https://doi.org/10.1073/PNAS.1502279112

Li, S., Castillo-González, C., Yu, B., & Zhang, X. (2017). The functions of plant small RNAs in development and in stress responses. *The Plant Journal*, *90*(4), 654–670. https://doi.org/10.1111/TPJ.13444

Li, S., Vandivier, L. E., Tu, B., Gao, L., Won, S. Y., Li, S., Zheng, B., Gregory, B. D., & Chen, X. (2015). Detection of Pol IV/RDR2-dependent transcripts at the genomic scale in Arabidopsis reveals features and regulation of siRNA biogenesis. *Genome Research*, *25*(2), 235–245. https://doi.org/10.1101/GR.182238.114

Li, T., Li, Y., Shangguan, H., Bian, J., Luo, R., Tian, Y., Li, Z., Nie, X., & Cui, L. (2023). BarleyExpDB: an integrative gene expression database for barley. *BMC Plant Biology*, *23*(1), 1–10. https://doi.org/10.1186/S12870-023-04193-Z/FIGURES/3

Li, Y., Yang, Y., Liu, Y., Li, D., Zhao, Y., Li, Z., Liu, Y., Jiang, D., Li, J., Zhou, H., Chen, J., Zhuang, C., & Liu, Z. (2019). Overexpression of OsAGO1b Induces Adaxially Rolled Leaves by Affecting Leaf Abaxial Sclerenchymatous Cell Development in Rice. *Rice 2019 12:1*, *12*(1), 1–22. https://doi.org/10.1186/S12284-019-0323-9

Li, Z., Li, W., Guo, M., Liu, S., Liu, L., Yu, Y., Mo, B., Chen, X., & Gao, L. (2022). Origin, evolution and diversification of plant ARGONAUTE proteins. *The Plant Journal*, *109*(5), 1086–1097. https://doi.org/10.1111/TPJ.15615

Liu, B., & Zhao, M. (2023). How transposable elements are recognized and epigenetically silenced in plants? *Current Opinion in Plant Biology*, *75*, 102428. https://doi.org/10.1016/J.PBI.2023.102428

Liu, W., Shoji, K., Naganuma, M., Tomari, Y., & Iwakawa, H. O. (2022). The mechanisms of siRNA selection by plant Argonaute proteins triggering DNA methylation. *Nucleic Acids Research*, *50*(22), 12997–13010. https://doi.org/10.1093/NAR/GKAC1135

Liu, X., Lu, T., Dou, Y., Yu, B., & Zhang, C. (2014). Identification of RNA silencing components in soybean and sorghum. *BMC Bioinformatics*, *15*(1), 1–13. https://doi.org/10.1186/1471-2105-15-4

Liu, Y., Teng, C., Xia, R., & Meyers, B. C. (2020). PhasiRNAs in Plants: Their Biogenesis, Genic Sources, and Roles in Stress Responses, Development, and Reproduction. *The Plant Cell*, *32*(10), 3059–3080. https://doi.org/10.1105/TPC.20.00335

Loffer, A., Singh, J., Fukudome, A., Mishra, V., Wang, F., & Pikaard, C. S. (2022). A DCL3 dicing code within Pol IV-RDR2 transcripts diversifies the siRNA pool guiding RNA-directed DNA methylation. *ELife*, *11*. https://doi.org/10.7554/ELIFE.73260

Lü, B., Wu, J. J., & Fu, D. L. (2015). Constructing the barley model for genetic transformation in Triticeae. *Journal of Integrative Agriculture*, *14*(3), 453–468. https://doi.org/10.1016/S2095-3119(14)60935-7

Lunde, B. M., Moore, C., & Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nature Reviews Molecular Cell Biology 2007 8:6*, *8*(6), 479–490. https://doi.org/10.1038/nrm2178

Madsen, C. T., Stephens, J., Hornyik, C., Shaw, J., Collinge, D. B., Lacomme, C., & Albrechtsen, M. (2009). Identification and characterization of barley RNA-directed RNA polymerases. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1789*(5), 375–385. https://doi.org/10.1016/J.BBAGRM.2009.03.003

Makarevitch, I., Waters, A. J., West, P. T., Stitzer, M., Hirsch, C. N., Ross-Ibarra, J., & Springer, N. M. (2015). Transposable Elements Contribute to Activation of Maize Genes in Response to Abiotic Stress. *PLOS Genetics*, *11*(1), e1004915. https://doi.org/10.1371/JOURNAL.PGEN.1004915

Mallory, A. C., & Vaucheret, H. (2009). ARGONAUTE 1 homeostasis invokes the coordinate action of the microRNA and siRNA pathways. *EMBO Reports*, *10*(5), 521–526. https://doi.org/10.1038/EMBOR.2009.32

Mallory, A., & Vaucheret, H. (2012). Form, Function, and Regulation of ARGONAUTE Proteins. *The Plant Cell*, *22*(12), 3879–3889. https://doi.org/10.1105/TPC.110.080671

Manavella, P. A., Yang, S. W., & Palatnik, J. (2019). Keep calm and carry on: miRNA biogenesis under stress. *The Plant Journal*, *99*(5), 832–843. https://doi.org/10.1111/TPJ.14369

Marí-Ordóñez, A., Marchais, A., Etcheverry, M., Martin, A., Colot, V., & Voinnet, O. (2013). Reconstructing de novo silencing of an active plant retrotransposon. *Nature Genetics 2013 45:9*, *45*(9), 1029–1039. https://doi.org/10.1038/ng.2703

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., Radchuk, V., Dockter, C., Hedley, P. E., Russell, J., Bayer, M., Ramsay, L., Liu, H., Haberer, G., Zhang, X. Q., Zhang, Q., Barrero, R. A., Li, L., Taudien, S., … Stein, N. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature 2017 544:7651*, *544*(7651), 427–433. https://doi.org/10.1038/nature22043

Matzke, M. A., & Mosher, R. A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics 2014 15:6*, *15*(6), 394–408. https://doi.org/10.1038/nrg3683

McCue, A. D., Panda, K., Nuthikattu, S., Choudury, S. G., Thomas, E. N., & Slotkin, R. K. (2015). ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. *The EMBO Journal*, *34*(1), 20–35. https://doi.org/10.15252/EMBJ.201489499

Merai, Z., Kerenyi, Z., Molnar, A., Barta, E., Valoczi, A., Bisztray, G., Havelda, Z., Burgyan, J., & Silhavy, D. (2005). Aureusvirus P14 Is an Efficient RNA Silencing Suppressor That Binds Double-Stranded RNAs without Size Specificity. *Journal of Virology*, *79*(11), 7217–7226. https://doi.org/10.1128/jvi.79.11.7217-7226.2005

Mette, M. F., Aufsatz, W., Van der Winden, J., Matzke, M. A., & Matzke, A. J. M. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *The EMBO Journal*, *19*(19), 5194. https://doi.org/10.1093/EMBOJ/19.19.5194

Meyers, B. C., Simon, S. A., & Zhai, J. (2010). MicroRNA Processing: Battle of the Bulge. *Current Biology*, *20*(2), R68–R70. https://doi.org/10.1016/j.cub.2009.12.008

Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., Chen, S., Hannon, G. J., & Qi, Y. (2008). Sorting of Small RNAs into Arabidopsis Argonaute Complexes Is Directed by the 5′ Terminal Nucleotide. *Cell*, *133*(1), 116. https://doi.org/10.1016/J.CELL.2008.02.034

Molnar, A., Melnyk, C., & Baulcombe, D. C. (2011). Silencing signals in plants: A long journey for small RNAs. *Genome Biology*, *12*(1), 1–8. https://doi.org/10.1186/GB-2010-11-12-219

Montgomery, T. A., Seong, J. Y., Fahlgren, N., Gilbert, S. D., Howell, M. D., Sullivan, C. M., Alexander, A., Nguyen, G., Allen, E., Ji, H. A., & Carrington, J. C. (2008). AGO1-miR173 complex initiates phased siRNA formation in plants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(51), 20055–20062. https://doi.org/10.1073/pnas.0810241105

Morel, J. B., Godon, C., Mourrain, P., Béclin, C., Boutet, S., Feuerbach, F., Proux, F., & Vaucheret, H. (2002). Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *The Plant Cell*, *14*(3), 629–639. https://doi.org/10.1105/TPC.010358

Moro, B., Chorostecki, U., Arikit, S., Suarez, I. P., Hobartner, C., Rasia, R. M., Meyers, B. C., & Palatnik, J. F. (2018). Efficiency and precision of microRNA biogenesis modes in plants. *Nucleic Acids Research*, *46*(20), 10709–10723. https://doi.org/10.1093/NAR/GKY853

Mosher, R. A., Schwach, F., Studholme, D., & Baulcombe, D. C. (2008). PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(8), 3145–3150. https://doi.org/10.1073/PNAS.0709632105

Newton, A. C., Flavell, A. J., George, T. S., Leat, P., Mullholland, B., Ramsay, L., Revoredo-Giha, C., Russell, J., Steffenson, B. J., Swanston, J. S., Thomas, W. T. B., Waugh, R., White, P. J., & Bingham, I. J. (2011). Crops that feed the world 4. Barley: a resilient crop? Strengths and weaknesses in the context of food security. *Food Security 2011 3:2*, *3*(2), 141–178. https://doi.org/10.1007/S12571-011-0126-3

Niu, X., Chen, L., Kato, A., & Ito, H. (2022). Regulatory mechanism of a heat-activated retrotransposon by DDR complex in Arabidopsis thaliana. *Frontiers in Plant Science*, *13*, 1048957. https://doi.org/10.3389/FPLS.2022.1048957

Nonomura, K. I., Morohoshi, A., Nakano, M., Eiguchi, M., Miyao, A., Hirochika, H., & Kurataa, N. (2007). A Germ Cell–Specific Gene of the ARGONAUTE Family Is Essential for the Progression of Premeiotic Mitosis and Meiosis during Sporogenesis in Rice. *The Plant Cell*, *19*(8), 2583. https://doi.org/10.1105/TPC.107.053199

Nosaka, M., Itoh, J.-I., Nagato, Y., Ono, A., Ishiwata, A., & Sato, Y. (2012). Role of Transposon-Derived Small RNAs in the Interplay between Genomes and Parasitic DNA in Rice. *PLoS Genetics*, *8*(9), e1002953. https://doi.org/10.1371/journal.pgen.1002953

Nozawa, K., Masuda, S., Saze, H., Ikeda, Y., Suzuki, T., Takagi, H., Tanaka, K., Ohama, N., Niu, X., Kato, A., & Ito, H. (2022). Epigenetic regulation of ecotype-specific expression of the heat-activated transposon ONSEN. *Frontiers in Plant Science*, *13*, 899105. https://doi.org/10.3389/FPLS.2022.899105

Nuthikattu, S., McCue, A. D., Panda, K., Fultz, D., DeFraia, C., Thomas, E. N., & Keith Slotkin, R. (2013). The Initiation of Epigenetic Silencing of Active Transposable Elements Is Triggered by RDR6 and 21-22 Nucleotide Small Interfering RNAs. *Plant Physiology*, *162*(1), 116–131. https://doi.org/10.1104/PP.113.216481

Pall, G. S., & Hamilton, A. J. (2008). Improved northern blot method for enhanced detection of small RNA. *Nature Protocols*, *3*(6), 1077–1084. https://doi.org/10.1038/NPROT.2008.67

Panda, K., Ji, L., Neumann, D. A., Daron, J., Schmitz, R. J., & Slotkin, R. K. (2016). Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biology*, *17*(1), 1–19. https://doi.org/10.1186/S13059-016-1032-Y

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., Bileschi, M. L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Rivoire, C., … Bateman, A. (2023). InterPro in 2022. *Nucleic Acids Research*, *51*(D1), D418–D427. https://doi.org/10.1093/NAR/GKAC993

Pérez-Hormaeche, J., Potet, F., Beauclair, L., Le Masson, I., Courtial, B., Bouché, N., & Lucas, H. (2008). Invasion of the Arabidopsis Genome by the Tobacco Retrotransposon Tnt1 Is Controlled by Reversible Transcriptional Gene Silencing. *Plant Physiology*, *147*(3), 1264–1278. https://doi.org/10.1104/PP.108.117846

Pogorelcnik, R., Vaury, C., Pouchin, P., Jensen, S., & Brasset, E. (2018). SRNAPipe: A Galaxy-based pipeline for bioinformatic in-depth exploration of small RNAseq data. *Mobile DNA*, *9*(1), 1–6. https://doi.org/10.1186/S13100-018-0130-7

Popova, O. V., Dinh, H. Q., Aufsatz, W., & Jonak, C. (2013). The RdDM Pathway Is Required for Basal Heat Tolerance in Arabidopsis. *Molecular Plant*, *6*(2), 396–410. https://doi.org/10.1093/MP/SST023

Pourkheirandish, M., & Komatsuda, T. (2007). The Importance of Barley Genetics and Domestication in a Global Perspective. *Annals of Botany*, *100*(5), 999. https://doi.org/10.1093/AOB/MCM139

Qi, Y., Denli, A. M., & Hannon, G. J. (2005). Biochemical specialization within Arabidopsis RNA silencing pathways. *Molecular Cell*, *19*(3), 421–428. https://doi.org/10.1016/j.molcel.2005.06.014

Qi, Y., He, X., Wang, X. J., Kohany, O., Jurka, J., & Hannon, G. J. (2006). Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature 2006 443:7114*, *443*(7114), 1008–1012. https://doi.org/10.1038/nature05198

Qian, Y., Cheng, Y., Cheng, X., Jiang, H., Zhu, S., & Cheng, B. (2011). Identification and characterization of Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in maize. *Plant Cell Reports*, *30*(7), 1347–1363. https://doi.org/10.1007/S00299-011-1046-6

Quesneville, H. (2020). Twenty years of transposable element analysis in the Arabidopsis thaliana genome. *Mobile DNA*, *11*(1), 1–13. https://doi.org/10.1186/S13100-020-00223-X

Rajagopalan, R., Vaucheret, H., Trejo, J., & Bartel, D. P. (2006). A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes & Development*, *20*(24), 3407–3425. https://doi.org/10.1101/GAD.1476406

Rajendran, N. R., Qureshi, N., & Pourkheirandish, M. (2022). Genotyping by Sequencing Advancements in Barley. *Frontiers in Plant Science*, *13*, 931423. https://doi.org/10.3389/FPLS.2022.931423

Ramakrishnan, M., Satish, L., Kalendar, R., Narayanan, M., Kandasamy, S., Sharma, A., Emamverdian, A., Wei, Q., & Zhou, M. (2021). The dynamism of transposon methylation for plant development

and stress adaptation. *International Journal of Molecular Sciences*, *22*(21), 11387. https://doi.org/10.3390/IJMS222111387

Ramakrishnan, M., Zhang, Z., Mullasseri, S., Kalendar, R., Ahmad, Z., Sharma, A., Liu, G., Zhou, M., & Wei, Q. (2022). Epigenetic stress memory: A new approach to study cold and heat stress responses in plants. *Frontiers in Plant Science*, *13*, 1075279. https://doi.org/10.3389/FPLS.2022.1075279

Ray, D., Kazan, H., Chan, E. T., Castillo, L. P., Chaudhry, S., Talukder, S., Blencowe, B. J., Morris, Q., & Hughes, T. R. (2009). Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nature Biotechnology 2009 27:7*, *27*(7), 667–670. https://doi.org/10.1038/nbt.1550

Rensink, W. A., & Buell, C. R. (2004). Arabidopsis to Rice. Applying Knowledge from a Weed to Enhance Our Understanding of a Crop Species. *Plant Physiology*, *135*(2), 622–629. https://doi.org/10.1104/PP.104.040170

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), W320–W324. https://doi.org/10.1093/NAR/GKU316

Rogers, K., & Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. In *Plant Cell* (Vol. 25, Issue 7, pp. 2383–2399). https://doi.org/10.1105/tpc.113.113159

Ross, R. J., Weiner, M. M., & Lin, H. (2014). PIWI proteins and PIWI-interacting RNAs in the soma. *Nature 2014 505:7483*, *505*(7483), 353–359. https://doi.org/10.1038/nature12987

Rotasperti, L., Sansoni, F., Mizzotti, C., Tadini, L., & Pesaresi, P. (2020). Barley's Second Spring as a Model Organism for Chloroplast Research. *Plants*, *9*(7), 1–25. https://doi.org/10.3390/PLANTS9070803

Ruiz-Ferrer, V., & Voinnet, O. (2009). Roles of plant small RNAs in biotic stress responses. *Annual Review of Plant Biology*, *60*, 485–510. https://doi.org/10.1146/ANNUREV.ARPLANT.043008.092111

Saisho, D., & Takeda, K. (2011). Barley: Emergence as a New Research Material of Crop Science. *Plant and Cell Physiology*, *52*(5), 724–727. https://doi.org/10.1093/PCP/PCR049

Saitou, N., & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, *4*(4), 406–425. https://doi.org/10.1093/OXFORDJOURNALS.MOLBEV.A040454

Shimamoto, K., & Kyozuka, J. (2002). Rice as a model for comparative genomics of plants. *Annual Review of Plant Biology*, *53*, 399–419. https://doi.org/10.1146/ANNUREV.ARPLANT.53.092401.134447

Sigman, M. J., Panda, K., Kirchner, R., McLain, L. L., Payne, H., Peasari, J. R., Husbands, A. Y., Slotkin, R. K., & McCue, A. D. (2021). An siRNA-guided ARGONAUTE protein directs RNA polymerase V to initiate DNA methylation. *Nature Plants 2021 7:11*, *7*(11), 1461–1474. https://doi.org/10.1038/s41477-021-01008-7

Singh, J., Mishra, V., Wang, F., Huang, H. Y., & Pikaard, C. S. (2019). Reaction Mechanisms of Pol IV, RDR2, and DCL3 Drive RNA Channeling in the siRNA-Directed DNA Methylation Pathway. *Molecular Cell*, *75*(3), 576-589.e5. https://doi.org/10.1016/j.molcel.2019.07.008

Smith, L. M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F. K., Herr, A. J., Pikaard, C. S., & Baulcombe, D. C. (2007). An SNF2 Protein Associated with Nuclear RNA Silencing and the Spread of a Silencing Signal between Cells in Arabidopsis. *The Plant Cell*, *19*(5), 1507–1521. https://doi.org/10.1105/TPC.107.051540

Sun, L., Jing, Y., Liu, X., Li, Q., Xue, Z., Cheng, Z., Wang, D., He, H., & Qian, W. (2020). Heat stress-induced transposon activation correlates with 3D chromatin organization rearrangement in Arabidopsis. *Nature Communications 2020 11:1*, *11*(1), 1–13. https://doi.org/10.1038/s41467-020-15809-5

Szádeczky-Kardoss, I., Csorba, T., Auber, A., Schamberger, A., Nyikó, T., Taller, J., Orbán, T. I., Burgyán, J., & Silhavy, D. (2018). The nonstop decay and the RNA silencing systems operate cooperatively in plants. *Nucleic Acids Research*, *46*(9), 4632–4648. https://doi.org/10.1093/NAR/GKY279

Szádeczky-Kardoss, I., Szaker, H. M., Verma, R., Darkó, É., Pettkó-Szandtner, A., Silhavy, D., & Csorba, T. (2022). Elongation factor TFIIS is essential for heat stress adaptation in plants. *Nucleic Acids Research*, *50*(4), 1927–1950. https://doi.org/10.1093/NAR/GKAC020

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Tang, K., Lang, Z., Zhang, H., & Zhu, J. K. (2016). The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications. *Nature Plants*, *2*(11), 16169. https://doi.org/10.1038/NPLANTS.2016.169

Thiel, J., Koppolu, R., Trautewig, C., Hertig, C., Kale, S. M., Erbe, S., Mascher, M., Himmelbach, A., Rutten, T., Esteban, E., Pasha, A., Kumlehn, J., Provart, N. J., Vanderauwera, S., Frohberg, C., & Schnurbusch, T. (2021). Transcriptional landscapes of floral meristems in barley. *Science Advances*, *7*(18), 832–860. https://doi.org/10.1126/SCIADV.ABF0832

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/BIB/BBS017

Tiwari, R., & Rajam, M. V. (2022). RNA- and miRNA-interference to enhance abiotic stress tolerance in plants. *Journal of Plant Biochemistry and Biotechnology 2022 31:4*, *31*(4), 689–704. https://doi.org/10.1007/S13562-022-00770-9

Tolia, N. H., & Joshua-Tor, L. (2006). Slicer and the Argonautes. *Nature Chemical Biology 2007 3:1*, *3*(1), 36–43. https://doi.org/10.1038/nchembio848

Trujillo, J. T., Seetharam, A. S., Hufford, M. B., Beilstein, M. A., & Mosher, R. A. (2018). Evidence for a Unique DNA-Dependent RNA Polymerase in Cereal Crops. *Molecular Biology and Evolution*, *35*(10), 2454–2462. https://doi.org/10.1093/MOLBEV/MSY146

Várallyay, É., Válóczi, A., Ágyi, Á., Burgyán, J., & Havelda, Z. (2010). Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. *EMBO Journal*, *29*(20), 3507–3519. https://doi.org/10.1038/EMBOJ.2010.215

Vargason, J. M., Szittya, G., Burgyán, J., & Tanaka Hall, T. M. (2003). Size Selective Recognition of siRNA by an RNA Silencing Suppressor. *Cell*, *115*(7), 799–811. https://doi.org/10.1016/S0092-8674(03)00984-X

Vaucheret, H. (2008). Plant ARGONAUTES. *Trends in Plant Science*, *13*(7), 350–358. https://doi.org/10.1016/J.TPLANTS.2008.04.007

Vaucheret, H. (2009). AGO1 Homeostasis Involves Differential Production of 21-nt and 22-nt miR168 Species by MIR168a and MIR168b. *PLOS ONE*, *4*(7), e6442. https://doi.org/10.1371/JOURNAL.PONE.0006442

Vaucheret, H., Mallory, A. C., & Bartel, D. P. (2006). AGO1 Homeostasis Entails Coexpression of MIR168 and AGO1 and Preferential Stabilization of miR168 by AGO1. *Molecular Cell*, *22*(1), 129. https://doi.org/10.1016/J.MOLCEL.2006.03.011

Vaucheret, H., Vazquez, F., Crété, P., & Bartel, D. P. (2004). The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes & Development*, *18*(10), 1187–1197. https://doi.org/10.1101/GAD.1201404

Voinnet, O. (2005). Induction and suppression of RNA silencing: insights from viral infections. *Nature Reviews Genetics 2005 6:3*, *6*(3), 206–220. https://doi.org/10.1038/nrg1555

Wang, F., & Axtell, M. J. (2017). AGO4 is specifically required for heterochromatic siRNA accumulation at Pol V-dependent loci in Arabidopsis thaliana. *The Plant Journal*, *90*(1), 37–47. https://doi.org/10.1111/TPJ.13463

Wang, F., Huang, H. Y., Huang, J., Singh, J., & Pikaard, C. S. (2023). Enzymatic reactions of AGO4 in RNA-directed DNA methylation: siRNA duplex loading, passenger strand elimination, target RNA slicing, and sliced target retention. *Genes and Development*, *37*(3–4), 103–118. https://doi.org/10.1101/GAD.350240.122/-/DC1

Wang, J., Mei, J., & Ren, G. (2019). Plant microRNAs: Biogenesis, homeostasis, and degradation. In *Frontiers in Plant Science* (Vol. 10, p. 360). Frontiers Media S.A. https://doi.org/10.3389/fpls.2019.00360

Wang, P., Fang, X., Li, P., Li, M., Yang, Y., & Wang, C. (2023). Principles of amino-acid–ribonucleotide interaction revealed by binding affinities between homogeneous oligopeptides and single-stranded RNA molecules. *Nano Research*, *16*(12), 13294–13300. https://doi.org/10.1007/S12274-023-5971-9

Wang, Q., Xue, Y., Zhang, L., Zhong, Z., Feng, S., Wang, C., Xiao, L., Yang, Z., Jake Harris, C., Wu, Z., Zhai, J., Yang, M., Li, S., Jacobsen, S. E., & Du, J. (2021). Mechanism of siRNA production by a plant Dicer-RNA complex in dicing-competent conformation. *Science*, *374*(6571), 1152–1157. https://doi.org/10.1126/SCIENCE.ABL4546

Wang, Z., Ma, Z., Castillo-González, C., Sun, D., Li, Y., Yu, B., Zhao, B., Li, P., & Zhang, X. (2018). SWI2/SNF2 ATPase CHR2 remodels pri-miRNAs via Serrate to impede miRNA production. *Nature 2018 557:7706*, *557*(7706), 516–521. https://doi.org/10.1038/s41586-018-0135-x

Wassenegger, M., Heimes, S., Riedel, L., & Sänger, H. L. (1994). RNA-directed de novo methylation of genomic sequences in plants. *Cell*, *76*(3), 567–576. https://doi.org/10.1016/0092-8674(94)90119-8

Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J. L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., Paux, E., SanMiguel, P., & Schulman, A. H. (2007). A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics 2007 8:12*, *8*(12), 973–982. https://doi.org/10.1038/nrg2165

Wicker, T., Schulman, A. H., Tanskanen, J., Spannagl, M., Twardziok, S., Mascher, M., Springer, N. M., Li, Q., Waugh, R., Li, C., Zhang, G., Stein, N., Mayer, K. F. X., & Gundlach, H. (2017). The repetitive landscape of the 5100 Mbp barley genome. *Mobile DNA*, *8*(1), 22. https://doi.org/10.1186/s13100-017-0102-3

Wiegmann, M., Maurer, A., Pham, A., March, T. J., Al-Abdallat, A., Thomas, W. T. B., Bull, H. J., Shahid, M., Eglinton, J., Baum, M., Flavell, A. J., Tester, M., & Pillen, K. (2019). Barley yield formation under abiotic stress depends on the interplay between flowering time genes and environmental cues. *Scientific Reports 2019 9:1*, *9*(1), 1–16. https://doi.org/10.1038/s41598-019-42673-1

Wierzbicki, A. T., Cocklin, R., Mayampurath, A., Lister, R., Jordan Rowley, M., Gregory, B. D., Ecker, J. R., Tang, H., & Pikaard, C. S. (2012). Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the Arabidopsis epigenome. *Genes & Development*, *26*(16), 1825–1836. https://doi.org/10.1101/GAD.197772.112

Wierzbicki, A. T., Haag, J. R., & Pikaard, C. S. (2008). Noncoding Transcription by RNA Polymerase Pol IVb/Pol V Mediates Transcriptional Silencing of Overlapping and Adjacent Genes. *Cell*, *135*(4), 635–648. https://doi.org/10.1016/j.cell.2008.09.035

Wierzbicki, A. T., Ream, T. S., Haag, J. R., & Pikaard, C. S. (2009). RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genetics*, *41*(5), 630–634. https://doi.org/10.1038/NG.365

Wu, L., Zhang, Q., Zhou, H., Ni, F., Wu, X., & Qi, Y. (2009). Rice MicroRNA Effector Complexes and Targets. *The Plant Cell*, *21*(11), 3421. https://doi.org/10.1105/TPC.109.070938

Wu, L., Zhou, H., Zhang, Q., Zhang, J., Ni, F., Liu, C., & Qi, Y. (2010). DNA Methylation Mediated by a MicroRNA Pathway. *Molecular Cell*, *38*(3), 465–475. https://doi.org/10.1016/J.MOLCEL.2010.03.008

Xie, Z., Allen, E., Fahlgren, N., Calamar, A., Givan, S. A., & Carrington, J. C. (2005). Expression of Arabidopsis MIRNA Genes. *Plant Physiology*, *138*(4), 2145–2154. https://doi.org/10.1104/PP.105.062943

Yan, K. S., Yan, S., Farooq, A., Han, A., Zeng, L., & Zhou, M. M. (2003). Structure and conserved RNA binding of the PAZ domain. *Nature 2003 426:6965*, *426*(6965), 469–474. https://doi.org/10.1038/nature02129

Yang, Y., Zhong, J., Ouyang, Y. dan, & Yao, J. (2013). The integrative expression and co-expression analysis of the AGO gene family in rice. *Gene*, *528*(2), 221–235. https://doi.org/10.1016/J.GENE.2013.07.002

Yao, X., Wang, Y., Yao, Y., An, L., Bai, Y., Li, X., Wu, K., & Qiao, Y. (n.d.). *Use of gene family analysis to discover argonaut (AGO) genes for increasing the resistance of Tibetan hull-less barley to leaf stripe disease*. https://doi.org/10.17221/180/2020-PPS

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C. C. G., Tolia, N. H., Joshua-Tor, L., Mitani, S., Simard, M. J., & Mello, C. C. (2006). Analysis of the C. elegans Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. *Cell*, *127*(4), 747–757. https://doi.org/10.1016/J.CELL.2006.09.033

Zaheer, U., Munir, F., Salum, Y. M., & He, W. (2024). Function and regulation of plant ARGONAUTE proteins in response to environmental challenges: a review. *PeerJ*, *12*, e17115. https://doi.org/10.7717/PEERJ.17115

Zemach, A., Kim, M. Y., Hsieh, P. H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S. L., & Zilberman, D. (2013). The Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. *Cell*, *153*(1), 193–205. https://doi.org/10.1016/J.CELL.2013.02.033

Zhai, J., Bischof, S., Wang, H., Meyers, B. C., Ausin, I., & Jacobsen Correspondence, S. E. (2015). *A One Precursor One siRNA Model for Pol IV-Dependent siRNA Biogenesis*. https://doi.org/10.1016/j.cell.2015.09.032

Zhai, J., Zhang, H., Arikit, S., Huang, K., Nan, G. L., Walbot, V., & Meyers, B. C. (2015). Spatiotemporally dynamic, cell-type-dependent premeiotic and meiotic phasiRNAs in maize anthers. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(10), 3146–3151. https://doi.org/10.1073/PNAS.1418918112

Zhai, L., Sun, W., Zhang, K., Jia, H., Liu, L., Liu, Z., Teng, F., & Zhang, Z. (2014). Identification and characterization of Argonaute gene family and meiosis-enriched Argonaute during sporogenesis in maize. *Journal of Integrative Plant Biology*, *56*(11), 1042–1052. https://doi.org/10.1111/JIPB.12205

Zhai, L., Teng, F., Zheng, K., Xiao, J., Deng, W., & Sun, W. (2019). Expression analysis of Argonaute genes in maize (Zea mays L.) in response to abiotic stress. *Hereditas*, *156*, 27. https://doi.org/10.1186/S41065-019-0102-Z

Zhan, J., & Meyers, B. C. (2023). Plant Small RNAs: Their Biogenesis, Regulatory Roles, and Functions. *Annual Review of Plant Biology*, *74*(Volume 74, 2023), 21–51. https://doi.org/10.1146/ANNUREV-ARPLANT-070122-035226

Zhang, H., He, X., & Zhu, J. K. (2013). RNA-directed DNA methylation in plants: Where to start? *RNA Biology*, *10*(10), 1593. https://doi.org/10.4161/RNA.26312

Zhang, H., Lang, Z., & Zhu, J. K. (2018). Dynamics and function of DNA methylation in plants. In *Nature Reviews Molecular Cell Biology* (Vol. 19, Issue 8, pp. 489–506). Nature Publishing Group. https://doi.org/10.1038/s41580-018-0016-z

Zhang, H., Tang, K., Wang, B., Duan, C. G., Lang, Z., & Zhu, J. K. (2014). Protocol: A beginner's guide to the analysis of RNA-directed DNA methylation in plants. *Plant Methods*, *10*(1), 1–9. https://doi.org/10.1186/1746-4811-10-18

Zhang, H., Xia, R., Meyers, B. C., & Walbot, V. (2015). Evolution, functions, and mysteries of plant ARGONAUTE proteins. *Current Opinion in Plant Biology*, *27*, 84–90. https://doi.org/10.1016/J.PBI.2015.06.011

Zhang, X., Henderson, I. R., Lu, C., Green, P. J., & Jacobsen, S. E. (2007). Role of RNA polymerase IV in plant small RNA metabolism. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(11), 4536–4541. https://doi.org/10.1073/PNAS.0611456104

Zhang, X., Henriques, R., Lin, S. S., Niu, Q. W., & Chua, N. H. (2006). Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. *Nature Protocols 2006 1:2*, *1*(2), 641–646. https://doi.org/10.1038/nprot.2006.97

Zhang, X., Niu, D. D., Carbonell, A., Wang, A., Lee, A., Tun, V., Wang, Z., Carrington, J. C., Chang, C. E. A., & Jin, H. (2014). ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. *Nature Communications*, *5*, 5468. https://doi.org/10.1038/NCOMMS6468

Zhao, K., Zhao, H., Chen, Z., Feng, L., Ren, J., Cai, R., & Xiang, Y. (2015). The Dicer-like, Argonaute and RNA-dependent RNA polymerase gene families in Populus trichocarpa: gene structure, gene expression, phylogenetic analysis and evolution. *Journal of Genetics*, *94*(2), 317–321. https://doi.org/10.1007/S12041-015-0508-Y

Zhao, M., & Ma, J. (2013). Co-evolution of plant LTR-retrotransposons and their host genomes. *Protein Cell*, *4*(7), 493–501. https://doi.org/10.1007/s13238-013-3037-6

Zhao, T., Li, G., Mi, S., Li, S., Hannon, G. J., Wang, X. J., & Qi, Y. (2007). A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. *Genes & Development*, *21*(10), 1190–1203. https://doi.org/10.1101/GAD.1543507

Zheng, K., Wang, L., Zeng, L., Xu, D., Guo, Z., Gao, X., & Yang, D. L. (2021). The effect of RNA polymerase V on 24-nt siRNA accumulation depends on DNA methylation contexts and histone modifications in rice. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(30), e2100709118. https://doi.org/10.1073/PNAS.2100709118

Zheng, X., Zhu, J., Kapoor, A., & Zhu, J.-K. (2007). Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *The EMBO Journal*, *26*(6), 1691. https://doi.org/10.1038/SJ.EMBOJ.7601603

Zheng, Y., Cohen-Karni, D., Xu, D., Chin, H. G., Wilson, G., Pradhan, S., & Roberts, R. J. (2010). A unique family of Mrr-like modification-dependent restriction endonucleases. *Nucleic Acids Research*, *38*(16), 5527. https://doi.org/10.1093/NAR/GKQ327

Zhong, X., Du, J., Hale, C. J., Gallego-Bartolome, J., Feng, S., Vashisht, A. A., Chory, J., Wohlschlegel, J. A., Patel, D. J., & Jacobsen, S. E. (2014). Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell*, *157*(5), 1050–1060. https://doi.org/10.1016/j.cell.2014.03.056

Zhong, X., Hale, C. J., Law, J. A., Johnson, L. M., Feng, S., Tu, A., & Jacobsen, S. E. (2012). DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. *Nature Structural & Molecular Biology 2012 19:9*, *19*(9), 870–875. https://doi.org/10.1038/nsmb.2354

Zhou, M., Palanca, A. M. S., & Law, J. A. (2018). Locus-specific control of the de novo DNA methylation pathway in Arabidopsis by the CLASSY family. *Nature Genetics 2018 50:6*, *50*(6), 865–873. https://doi.org/10.1038/s41588-018-0115-y

Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S. H., Liou, L. W., Barefoot, A., Dickman, M., & Zhang, X. (2011). Arabidopsis Argonaute10 Specifically Sequesters miR166/165 to Regulate Shoot Apical Meristem Development. *Cell*, *145*(2), 242–256. https://doi.org/10.1016/J.CELL.2011.03.024

Zhu, J., Li, C., Peng, X., & Zhang, X. (2021). RNA architecture influences plant biology. *Journal of Experimental Botany*, *72*(11), 4144–4160. https://doi.org/10.1093/JXB/ERAB030

Zilberman, D., Cao, X., & Jacobsen, S. E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science*, *299*(5607), 716–719. https://doi.org/10.1126/SCIENCE.1079695

12 PUBLICATIONS LIST

Conference Presentations and Posters:

MBK Napok 30 (Gödöllő, 2020): presentation - "Role of *Arabidopsis* miRNA precursors in RISC-loading efficiency", Előadói Díj III..

XX. Genetikai Műhelyek Magyarországon Minikonferencia (Szeged, 2021): poster - "Functional analysis of barley *AGO4* genes in *Arabidopsis thaliana ago4-3* mutant background".

Microsymposium on RNA biology (Vienna, 2022): poster - "Genetic complementation analysis of barley *AGO4* genes in *Arabidopsis ago4* mutant".

FIKOK 2022 (Gödöllő, 2022): poster - "Identification of barley *AGO4* genes and complementation assay in *Arabidopsis*", abstract published on ISBN: 9789632699998.

GBI Nap (Gödöllő, 2022): presentation - "Genetic complementation analysis of barley *AGO4* genes in *Arabidopsis ago4* mutant", Előadói Díj I.

XXIX. Ifjúsági Tudományos Fórum (Keszthely, 2023): presentation - "Revealing the Diverse Functional Roles of Barley AGO4 Proteins in *Arabidopsis* through Heterologous Complementation Assay", conference paper published on ISBN: 9786156338082.

GBI Napok (Gödöllő, 2023): presentation – "Generation and characterization of barley *dcl3* RNAi mutants", Előadói Díj I.

Microsymposium on RNA biology (Vienna, 2024): poster - "Barley AGO4 proteins show overlapping functionality with distinct small RNA-binding properties in heterologous complementation".

Publications:

Hamar, E., Szaker, H. M., Kis, A., Dalmadi, A., **Miloro, F.**, Szittya, G., Taller, J., Gyula, P., Csorba, T., & Havelda, Z. (2020). Genome-Wide Identification of RNA Silencing-Related Genes and Their Expressional Analysis in Response to Heat Stress in Barley (Hordeum vulgare L.). Biomolecules 2020, Vol. 10, Page 929, 10(6), 929. https://doi.org/10.3390/BIOM10060929 (MDPI - Q2 - 2020).

Dalmadi, Á., **Miloro, F.**, Bálint, J., Várallyay, É., & Havelda, Z. (2021). Controlled RISC loading efficiency of miR168 defined by miRNA duplex structure adjusts ARGONAUTE1 homeostasis.

Nucleic Acids Research, 49(22), 12912–12928. https://doi.org/10.1093/NAR/GKAB1138 (Oxford Academic - D1 - 2021).

Contaldo, N., Zambon, Y., Galbacs, Z. N., **Miloro, F.**, Havelda, Z., Bertaccini, A., & Varallyay, E. (2023). Small RNA Profiling of Aster Yellows Phytoplasma-Infected Catharanthus roseus Plants Showing Different Symptoms. Genes, 14(5), 1114. https://doi.org/10.3390/GENES14051114 (MDPI - Q2 - 2023).

Miloro, F., Kis, A., Havelda, Z., & Dalmadi, Á. (2024). Barley AGO4 proteins show overlapping functionality with distinct small RNA-binding properties in heterologous complementation. Plant Cell Reports 2024 43:4, 43(4), 1–19. https://doi.org/10.1007/S00299-024-03177-Z (Springer Nature - D1 - 2024).

13 APPENDICES

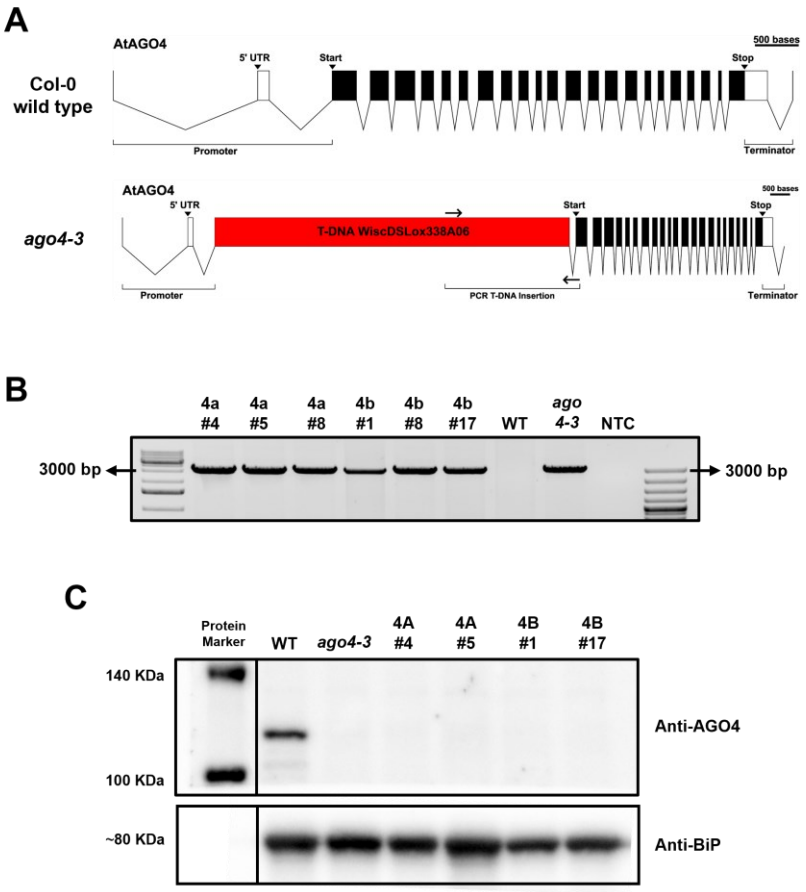

Figure S1. A. Visualization of the At*AGO4* gene structure in wild type (top) and in ago4-3 mutant (bottom). The promoter, 5' UTR, start and stop codons, and terminator are shown. T-DNA insertion between the 5' UTR and the start codon in ago4-3 is marked with red. The UTR regions are demonstrated with empty shapes. The primers for the amplification of the region containing the T-DNA are represented by arrows. The scale bars represent 500 bp. **B.** Agarose gel electrophoresis of PCR products from the region containing the T-DNA insertion in the ago4-3 mutant. ago4-3 and all lines derived from the ago4-3 background are positive for T-DNA insertion, while wild-type plants are negative. The forward primer is located on the T-DNA, specifically on the phosphinothricin resistance gene, while the reverse primer is located on the first coding exon, just after the start codon. The PCR product is 3405 bp and contains a fragment of the phosphinothricin resistance gene, its promoter, Ds transposon, LoxP site, T-DNA Left Border and a 170 bp promoter fragment before the start codon. The first lane is loaded with GeneRuler 1 kb Plus DNA Ladder and the last lane is loaded with GeneRuler 100 bp Plus DNA Ladder. NTC is for No Template Control. **C.** Western blot showing the endogenous protein level of AtAGO4 in wild type, ago4-3 mutant and the lines with the highest level of barley HA-tagged AGO4 protein. The first lane is loaded with ProSieve QuadColor Protein Marker (Lonza Bioscience), which gives an aspecific signal to the anti-AtAGO4 antibody. The membrane was cut into two, and lower part was used to detect BiP (lumenal-binding protein) as an internal control. The same membrane was divided in two and incubated each half separately with the two different antibodies.

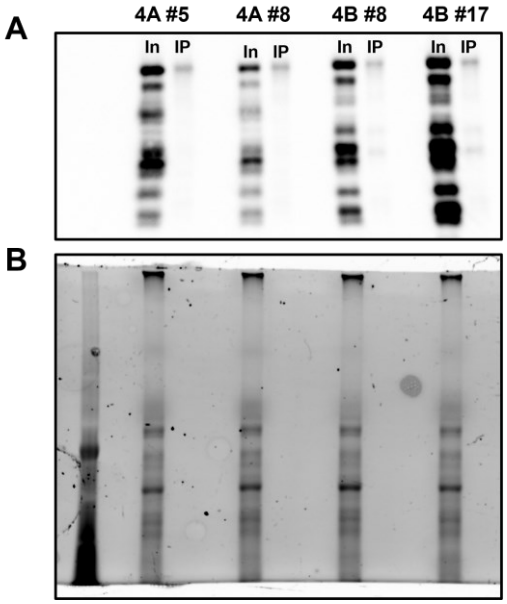

Figure S2. A. Western blot showing the protein expression of the barley HA-tagged proteins (HA-HvAGO4A and HA-HvAGO4B) used for small RNA-IP sequencing. The first lane is loaded with ProSieve^{TM} QuadColor^{TM} Protein Marker (Lonza Bioscience). The other lanes are loaded so that the left lane contains the input (In) from the protein extraction before immunoprecipitation and the right lane contains the sample after immunoprecipitation with anti-HA (IP). The membrane was exposed for 5 seconds. **B.** The gel on which the samples were run was a Mini-PROTEAN TGX Precast Protein Gel (BioRad) and shows the total amount of protein present in the different lanes after UV cross-linking.

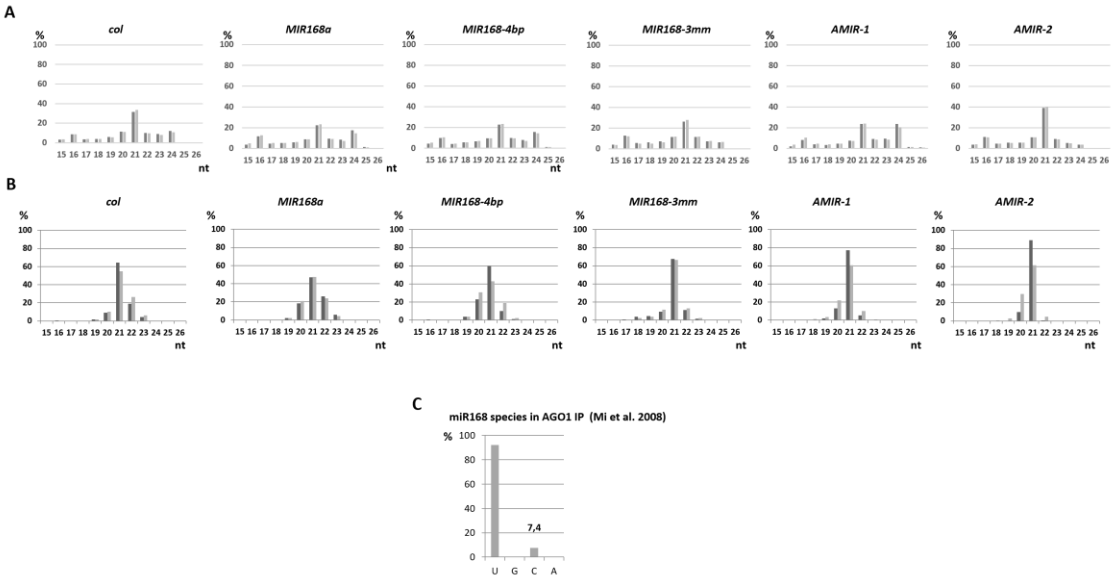

Figure S3. A. Size distribution of adaptor trimmed total reads, **B** and miR168 homologous sRNA sequences in HTS data of *A. thaliana* Col-0, wild type *MIR168a* precursor fragment, *MIR168-4bp* and *MIR168-3mm*, *AMIR-1* and *AMIR-2* overexpressing plants. Small RNAs derived from miR168 producing precursors were analyzed according to their size distribution. Data are presented as the percentage of the total reads or of *MIR168a* precursor homologous sequences, respectively. Colors of columns correspond to replica experiments. **C.** Distribution of AGO1-bound miR168 species according to their 5' end. AGO1 immuno-precipitated sample was sequenced on Illumina platform by Mi et al. (2008) and HTS data were deposited under the accession number GSE10036. Analysis of more than 90 000 miR168 homologous reads per million.

Table S1. Primers designed and used to perform PCR, qPCR and RT-qPCR in this study.

PCR/qPCR/RT-qPCR Primers					
Gene Name	**Gene Name**	**RT-qPCR primers**	**Sequence (5'-3')**	**Amplicon Length**	**Locus Name**
AtAGO4 (ago4-3 mutant)	Argonaute 4	*ago4-3* _F	CGGCGACGAGCCAGGGATAG	3405 bp	AT2G27040 (WiscDSLox338A06)
		ago4-3 _R	CTGGTGGAGGAGGTGGTAATG		
AtUBC9 (cDNA)	Ubiquitin-Conjugating enzyme 9	AtUBC9_qPCR_F	TCACAATTTCCAAGGTGCTGC	146 bp	AT4G27960
		AtUBC9_qPCR_R	ATACTTTTGGGTCCAGGTCCG		
AtUBC9 (gDNA)	Ubiquitin-Conjugating enzyme 9	AtUBC9_gDNA_F	GGTGCTGCTATCGATCTGTTC	69 bp	AT4G27960
		AtUBC9_gDNA_R	GCTATCTCAGGGACCAAAGGATC		
AtACT2	Actin 2	AtACT2_qPCR_F	GTGGTTCCATTCTTGCTTCCC	90 bp	AT3G18780
		AtACT2_qPCR_R	TGTGAACGATTCCTGGACCTG		
AtPP2AA3	Protein Phosphatase 2A subunit A3	AtPP2AA3_qPCR_F	TCCTCTGGCTAAGCGACTTTC	104 bp	AT1G13320
		AtPP2AA3_qPCR_R	TCTTTAGCACATCTGGGGCAC		
AtAGO4	Argonaute 4	AtAGO4_qPCR_F	CAAGTTCTTCCAGCCAACGTC	148 bp	AT2G27040
		AtAGO4_qPCR_R	CAGGACGTGGTAGTGAGTTGG		
AtROS1	DNA glycosylase/AP lyase	AtROS1_qPCR_F	TGGTGGGAGGAAGAACGTAATG	171 bp	AT2G36490
		AtROS1_qPCR_R	AGCCGAACTTGAGAGATGGTC		
AtSN1	SINE-like transposable element	AtSN1_qPCR_F	CTGGAAGTTCAGGCCCAAAG	113 bp	AT3TE63860
		AtSN1_qPCR_R	TAGAGGTGCTGGATTCGAGAC		
ONSEN (Ito et al., 2011)	Ty1/Copia-like retrotransposon	ONSEN_qPCR_F (COPIA78-4219F)	CCACAAGAGGAACCAACGAA	82 bp	AT1TE12295; AT1TE24850; AT1TE59755; AT1TE71045; AT3TE54550; AT3TE89830; AT3TE92525; AT5TE15240
		ONSEN_qPCR_R (COPIA78-4300R)	TTCGATCATGGAAGACCGG		
HvAGO4a	Argonaute 4a	HvAGO4a_qPCR_F	AGGACAGGAATCAGAGCACAC	151 bp	HORVU.MOREX.r2.3HG0213340
		HvAGO4a_qPCR_R	CATTACCCCTCCACCCAAGTC		
HvAGO4b	Argonaute 4b	HvAGO4b_qPCR_F	GAGCCCAAGTTCACAGTCATTG	104 bp	HORVU.MOREX.r2.1HG0078650
		HvAGO4b_qPCR_R	TGTTTATCCACCACCGTACCAG		

Table S2. Primers designed and used to perform the cloning of the different constructs used.

Cloning Primers					
Gene ID	**Gene Name**	**Cloning primers**	**Sequence (5'-3')**	**Amplicon Length**	**Locus Name**
MIR168A	*ath-MIR168A*	Pre168a_S	TTCCCGGGTCATATCCCTGTCTAAAGGG	158 bp	AT4G19395
		Pre168a_AS	TTGGATCCGCGTAGAAATCTTCCAGATC		
MIR168A_4bp	*ath-MIR168A*	miR168a_4bp_R	caGAATTCagtttttttaCACCTCGAGGATCCGATTCACTTGATGC	158 bp	AT4G19395
MIR168A_3mm	*ath-MIR168A*	miR168a_3mm_R	caGAATTCagtttttttaCACCTCGAGGATCCGATTCAGTTGAGGCGGGTCGGGATCC	158 bp	AT4G19395
AMIR1	*ath-MIR168A*	amiR168_1F	AACGGTCACCATTTCCCGACCAGCCCCAACCGAGACCACGCC	158 bp	AT4G19395
		amiR168_R	CACACGCGTGCATGAAAGAGCACCAATTCCCGACCTGCACCAAGCGAGAACCGCCG		
AMIR2	*ath-MIR168A*	amiR168_2F	AACGGTCACCATTTCCCGACCAGCTCCAATCGAGACCACGCC	158 bp	AT4G19395
		amiR168_R	CACACGCGTGCATGAAAGAGCACCAATTCCCGACCTGCACCAAGCGAGAACCGCCG		
AtAGO1	*Argonaute 1*	sensor168_F	TAGGATCCATGGTCACCACGCGTAGTAAAGGAGAAGAAC	558 bp	AT1G48410
		sensor168_R	GAGAATTCTTATTTGTATAGTTCATCCATG		
HvAGO4a	Argonaute 4a	Clon_HA-HvAGO4a_F	ACGTGCACGCGTATGTACCCATACGATGTTCCAGATTACGCTGAGTCGCATGTTGTTGAG	2820 bp	HORVU.MOREX.r2.3HG0213340
		Clon_HA-HvAGO4a_R	GCTACGTCTAGATTATCAGCAGAAGAACATGCTGCT		
HvAGO4b	Argonaute 4b	Clon_HA-HvAGO4b_F	AGTCATACGCGTATGTACCCATACGATGTTCCAGATTACGCTGACCCGCATGATGGAGAG	2811 bp	HORVU.MOREX.r2.1HG0078650
		Clon_HA-HvAGO4b_R	CGTCGGTCTAGATTATCAGCAGAAGAACATGGAGCT		
Promoter *AtAGO4*	Argonaute 4	Clon_pAGO4_F	TGCTCGGGTACCGTAAATTGTCCGATACTCTCATGTG	2536 bp	AT2G27040
		Clon_pAGO4_R	CGTTCGCTCGAGACGCGTCTCCTGCTCAAAGAAACCAAAC		
Terminator *AtAGO4*	Argonaute 4	Clon_tAGO4_F	ATTATAGCGGCCGCATGAGCAGCCCTACTTGGCT	576 bp	AT2G27040
		Clon_tAGO4_R	GTGTACGAGCTCACTTTGACTATGCTTATCTCGTTTG		

Table S3. Gene names and IDs of putative AGO proteins predicted in barley. The names and the clades are the same found in Fig. 9.

Clade	Protein Name	Gene ID (Morex V3)	Protein Length	Exon Number	cDNA Length
Clade 1/5/10	HvAGO1A	HORVU.MOREX.r3.6HG0603150.1	1086	23	3261
	HvAGO1B_1	HORVU.MOREX.r3.2HG0187760.1	1100	23	3728
	HvAGO1B_2	HORVU.MOREX.r3.7HG0639910.1	1216	23	4114
	HvAGO1C	HORVU.MOREX.r3.6HG0633820.3	1010	23	3282
	HvAGO1D	HORVU.MOREX.r3.7HG0750280.1	1022	21	3069
	HvAGO10	HORVU.MOREX.r3.7HG0714440.1	955	22	2870
	HvAGO10-like	HORVU.MOREX.r3.6HG0577170.1	949	21	2850
	HvMEL1	HORVU.MOREX.r3.5HG0520550.1	1042	22	3479
	HvMEL1-like	HORVU.MOREX.r3.5HG0520640.1/ HORVU.MOREX.r3.5HG0520650.1	1098	23	3294
	HvAGO11	HORVU.MOREX.r3.4HG0334500.1	1017	22	3054
	HvAGO12	HORVU.MOREX.r3.4HG0334490.1	978	22	2934
	HvAGO14	HORVU.MOREX.r3.2HG0125150.1	1063	22	3192
	HvAGO18	HORVU.MOREX.r3.3HG0238320.1	1090	21	3273
Clade 4/6/9	HvAGO4A	HORVU.MOREX.r3.3HG0256890.1	921	23	3210
	HvAGO4B	HORVU.MOREX.r3.1HG0095310.1	923	23	3198
	HvAGO15	HORVU.MOREX.r3.7HG0736980.1	980	22	2943
	HvAGO6	HORVU.MOREX.r3.5HG0468930.1	883	22	2893
Clade 2/3/7	HvAGO2	HORVU.MOREX.r3.2HG0190470.1	1047	3	3568
	HvAGO3_1	HORVU.MOREX.r3.2HG0190440.1	1027	3	3084
	HvAGO3_2	HORVU.MOREX.r3.2HG0190460.1	1020	3	3063
	HvAGO7	HORVU.MOREX.r3.2HG0189750.1	1047	3	3336

14 Acknowledgements

I would like to thank my two supervisors, Dr. Zoltán Havelda and Dr. Ágnes Dalmadi, for their moral and scientific support and for helping me to improve as a researcher. In addition, I would like to thank my lab members, Dr. András Kis and Erzsike Poldan, for their help in the lab and in plant management. Finally, I thank the people who have stood by my side during my PhD journey, supporting and encouraging me, with special thanks to my girlfriend Federica.