New Phytologist

Research

Phytophthora infestans Argonaute 1 binds microRNA and small RNAs from effector genes and transposable elements

Anna K. M. Åsman[1], Johan Fogelqvist[1], Ramesh R. Vetukuri[2] and Christina Dixelius[1]

[1]Department of Plant Biology, Swedish University of Agricultural Sciences, Uppsala BioCenter, Linnéan Center for Plant Biology, PO Box 7080, SE-75007 Uppsala, Sweden; [2]Department of Plant Protection Biology, Swedish University of Agricultural Sciences, Resistance Biology Unit, PO Box 102, SE-23053 Alnarp, Sweden

Author for correspondence:
Anna K. M. Åsman
Tel: +46 18673235
Email: anna.asman@slu.se

Received: *3 December 2015*
Accepted: *26 February 2016*

New Phytologist (2016) **211:** 993–1007
doi: 10.1111/nph.13946

Key words: Argonaute (Ago), Crinkler (CRN), effector, microRNA (miRNA), *Phytophthora infestans*, transposon.

Summary

• *Phytophthora* spp. encode large sets of effector proteins and distinct populations of small RNAs (sRNAs). Recent evidence has suggested that pathogen-derived sRNAs can modulate the expression of plant defense genes. Here, we studied the sRNA classes and functions associated with *Phytophthora infestans* Argonaute (Ago) proteins.

• sRNAs were co-immunoprecipitated with three PiAgo proteins and deep sequenced.

• Twenty- to twenty-two-nucleotide (nt) sRNAs were identified as the main interaction partners of PiAgo1 and high enrichment of 24–26-nt sRNAs was seen in the PiAgo4-bound sample. The frequencies and sizes of transposable element (TE)-derived sRNAs in the different PiAgo libraries suggested diversified roles of the PiAgo proteins in the control of different TE classes. We further provide evidence for the involvement of PiAgo1 in the *P. infestans* microRNA (miRNA) pathway. Protein-coding genes are probably regulated by the shared action of PiAgo1 and PiAgo5, as demonstrated by analysis of differential expression. An abundance of sRNAs from genes encoding host cell death-inducing Crinkler (CRN) effectors was bound to PiAgo1, implicating this protein in the regulation of the expanded *CRN* gene family.

• The data suggest that PiAgo1 plays an essential role in gene regulation and that at least two RNA silencing pathways regulate TEs in the plant-pathogenic oomycete *P. infestans*.

Introduction

RNA plays a central regulatory role in all cellular life forms. Not only is RNA the catalytic component of the ribosome, it also functions in defense against infectious agents and plasmids, mediates epigenetic control of gene expression and is an essential component of the pre-mRNA splicing machinery. Eukaryotic genomes contain large amounts of nonprotein-coding RNA (ncRNA) and it has been suggested that it is not the number of genes, but the proportion of ncRNA, that determines organismal complexity (Liu *et al.*, 2013; Morris & Mattick, 2014). Regulatory RNAs of < 200 nucleotides (nt) in length (small RNAs (sRNAs)) act as sequence-specific guide molecules to direct the silencing of complementary DNA and RNA. The main interactors of 20–30-nt sRNAs in eukaryotic cells are proteins of the Argonaute (Ago) family. These sRNA–Ago complexes mediate a multitude of functions, the most well-studied processes of which are gene expression regulation, antiviral defense and transposable element (TE) control (Mallory & Vaucheret, 2010; Clark *et al.*, 2013). Bacterial and archaeal genomes also encode Ago proteins, but prokaryotic Agos mediate DNA interference by binding either sRNA or small DNA guides in defense against invasion by plasmids, phages and transposons (Olovnikov *et al.*, 2013; Swarts *et al.*, 2015a,b).

The first RNA silencing process described in plants was double-stranded RNA-induced virus resistance, later termed RNA interference (RNAi; Waterhouse *et al.*, 1998). In addition to antiviral RNAi, additional sRNA-directed processes are now recognized as important regulators of plant defense and pathogen virulence. For instance, microRNAs (miRNAs) and phased small interfering RNAs (siRNAs) regulate the expression of resistance (*R*) genes in solanaceous and leguminous plants (Zhai *et al.*, 2011; Li *et al.*, 2012; Shivaprasad *et al.*, 2012). In the absence of pathogen infection, the *R* genes are silenced by sRNA-guided transcript cleavage, but decreased expression of specific plant miRNAs under pathogen attack leads to *R* gene activation. This mechanism enables pathogen-inducible upregulation of *R* gene expression, probably preventing autoimmunity and reducing fitness costs in the absence of pathogens.

In some species, sRNAs have the capacity to translocate between cells. In the worm *Caenorhabditis elegans*, sRNA uptake from the environment and transport within the organism require the transmembrane proteins SID-1 and SID-2 (Winston *et al.*, 2002, 2007). Plant sRNAs are mobile, both locally and systemically, and are probably transferred from infecting pathogens, but the mechanism of RNA uptake in plants is unknown (Knip *et al.*, 2014; Sarkies & Miska, 2014). As a result of their restricted sizes and universal target base-pairing rules, host cell targeting by sRNAs would be a potent mechanism for pathogens to manipulate host mRNA expression. Such

events have been demonstrated in the interaction between the fungal pathogen *Botrytis cinerea* and *Arabidopsis thaliana*, where fungal sRNAs target and downregulate host immunity genes (Weiberg *et al.*, 2013). This process requires fungal Dicer-like (Dcl) proteins and AtAgo1.

To establish disease on their hosts, plant pathogens produce effectors, which are secreted molecules that interfere with plant cellular processes to promote infection (Kamoun, 2007). The genomes of *Phytophthora* species code for numerous effectors, the majority of which are cytoplasmically translocated RxLR and Crinkler (CRN) proteins, defined by their distinctive amino acid motifs. The *Phytophthora infestans* reference genome (240 Mb) has been predicted to contain 563 *RxLR*s, 196 *CRN*s and 255 *CRN* pseudogenes (Haas *et al.*, 2009). The activities of most of these molecules are unknown, but, from the cases in which the function has been studied, it is clear that *Phytophthora* effectors target diverse plant processes (Fawke *et al.*, 2015). Two RxLRs from *Phytophthora sojae* have RNA silencing suppression activity (Qiao *et al.*, 2013). The effectors, named PSR1 and PSR2 (for *Phytophthora* suppressor of RNA silencing), repress host sRNA biogenesis. The interaction partner of PSR1 in *Arabidopsis* is a nuclear-localized helicase, whose silencing impairs AtDcl1 subnuclear localization (Qiao *et al.*, 2015). The fact that *P. sojae* has evolved the capacity to suppress plant RNA silencing points towards a critical role of sRNA-directed processes in defense against *Phytophthora* diseases.

Gene silencing in oomycetes is most well characterized in the potato late blight pathogen *P. infestans*. The *P. infestans* genome encodes the three core components involved in RNA silencing, comprising two *Dcl* genes (*PiDcl1*, *PiDcl2*), five *Ago* genes (*PiAgo1–5*) and one *RNA-dependent RNA polymerase* gene (*PiRdR1*; Vetukuri *et al.*, 2011a; Fahlgren *et al.*, 2013). Four distinct PiAgo proteins are produced, as *PiAgo1* and *PiAgo2* are nearly identical duplicated genes. Deep sequencing has shown that sRNAs in *Phytophthora* are centered on 21 and 25 nt (Vetukuri *et al.*, 2012; Fahlgren *et al.*, 2013). Repetitive sequences constitute *c.* 75% of the *P. infestans* genome (Haas *et al.*, 2009), and the majority of sRNAs derive from retrotransposon sequences. Specific sets of sRNAs, however, map to effector genes (Vetukuri *et al.*, 2012). The presence of 21-nt sRNAs from genes encoding CRNs suggests that these effectors are under sRNA-directed control. To deepen our understanding of the strategies used by *P. infestans* to evade detection by the host immune system and how the pathogen causes disease, further analysis of sRNAs and sRNA-interacting proteins in this organism is of critical importance.

In the present study, we aimed to identify the sRNA classes associating with *P. infestans* Ago proteins and to clarify whether the PiAgos have distinct or overlapping functions. We found that PiAgo1 and PiAgo4 bind distinct sets of sRNAs. PiAgo1-associated 20–22-nt sRNAs were enriched for protein-coding genes and *Gypsy* long terminal repeat (LTR) retrotransposons, whereas 24–26-nt PiAgo4-bound sRNAs derived mainly from *Helitron*, *Crypton*, *PiggyBac* and *Copia* transposons. sRNAs from *CRN* genes and the single verified miRNA in *P. infestans*, *pin*-miR8788, were identified as binding partners of PiAgo1.

Materials and Methods

Phytophthora infestans transformation, vector construction, 5′ rapid amplification of cDNA ends (RACE) and quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

Transformation of *Phytophthora infestans* (Mont.) de Bary and the maintenance of strains were performed as described by Vetukuri *et al.* (2012). The *green fluorescent protein* (*GFP*) gene was PCR amplified from pTOR-eGFP (Supporting Information Table S1) and ligated downstream of the *ham34* promoter in pTOR (GenBank EU257520). N-terminal GFP tagging employed *Spe*I and *Xba*I, creating pTOR-NGFP, whereas pTOR-GFPC was created using *Xba*I and *Sac*II. The *PiAgo* sequences were PCR amplified from genomic DNA of *P. infestans* 88069. Before ligation into pTOR-NGFP, the *PiAgo1* PCR product was digested with *Xba*I and *Sac*II, whereas the *PiAgo3/4/5* products were digested with *Xba*I and *Not*I. All *PiAgo* PCR products aimed for C-terminal GFP tagging were digested with *Spe*I and *Xba*I, and cloned into pTOR-GFPC. 5′ RACE was performed according to the manufacturer's instructions (Thermo Scientific 18374058, Waltham, MA, USA), using as input total RNA from isolate 88069. Total RNA extraction, cDNA synthesis and qRT-PCR were performed as described by Vetukuri *et al.* (2011a).

Protein and RNA co-immunoprecipitation

Immunoprecipitation (IP) of PiAgo-GFP and control GFP was performed according to the manufacturer's protocol (Chromotek gtm-100, Planegg-Martinsried, Germany), except for minor adjustments: dilution and wash buffers included 0.5% NP-40 and the wash buffer contained 500 mM NaCl. Protease inhibitor cocktail (Roche, Basel, Switzerland), 1 mM phenylmethylsulfonyl fluoride (PMSF) and 80 U ml^{-1} RiboLock RNase inhibitor (Thermo Scientific) were supplemented to all buffers. For Western blot, the immunocomplexes were eluted by boiling (95°C, Laemmli buffer, 8 M urea). Samples destined for RNA analysis were eluted by heating (70°C, 200 mM NaOAc, pH 5.5, 1 mM EDTA, 1% sodium dodecyl sulfate (SDS)) and extracted by phenol–chloroform.

Immunoblotting

Input and nonbound IP fractions were mixed 1 : 1 with 2 × Laemmli buffer. These, and the bound fraction, were separated by 12% SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membranes. Proteins were detected by the monoclonal GFP antibody JL-8 (1 : 1000, Clontech 632381, Mountain View, CA, USA). Blocking, secondary antibody incubation, development and detection were performed as described by Munch *et al.* (2015).

Denaturing PAGE and Northern hybridization

An aliquot of the co-immunoprecipitated RNA was 5′ dephosphorylated using CIAP (Life Technologies, Carlsbad, CA, USA), end-labeled with γ-^{32}P ATP and resolved on denaturing 15% polyacrylamide gels. Total RNA from isolate 88069, *PiAgo1-GFP* and *GFP* was extracted using TRIzol reagent (Life Technologies). *pin*-miR8788 was analyzed by Northern hybridization (Vetukuri *et al.*, 2012) using a γ-^{32}P ATP 5′ end-labeled DNA oligonucleotide. Sense and antisense riboprobes (Table S1) were generated as described by Vetukuri *et al.* (2011b).

sRNA sequencing and data analysis

RNA enriched for the low-molecular-weight fraction was extracted from the *PiAgo1-GFP* line using the mirVana™ miRNA isolation kit (Ambion, Waltham, MA, USA). These samples and co-IP RNA were sequenced on the Ion Proton platform using the Ion PI Sequencing 200 kit v3 (Life Technologies) at SciLifeLab (Uppsala, Sweden). The 18–30-nt adaptor-free reads (Table S2) were mapped to the *P. infestans* genome (http://www.broadinstitute.org) using Bowtie v.1.0.0 (Langmead *et al.*, 2009), allowing no mismatches and reporting all mapping locations. The sRNA reads were classified based on the mapping location. Annotations of protein-coding genes, transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) were downloaded from http://www.broadinstitute.org. The repeat library was based on annotations from Haas *et al.* (2009) with the inclusion of unannotated repeats identified using nucmer from the MUMer package v.3.23 with -maxmatch and -nosimplify (Kurtz *et al.*, 2004).

For the remaining analyses, reads mapping (Bowtie v.1.0.0, -v1) to annotated tRNA/rRNA genes from the *P. infestans* genome or from Rfam 10.0 (Gardner *et al.*, 2009) were discarded. Recording of sRNA length, number, starting base, orientation and the proportion of reads mapping to different classes was assessed using SAMtools v.0.1.19 (Li *et al.*, 2009), Bedtools v.2.23.0 (Quinlan & Hall, 2010) and custom Perl scripts (available at https://github.com/johanfogel/tools4sam). Reads mapping to multiple locations were scaled, so that, for each read, the sum of all mapping counts was equal to unity.

The numbers of reads mapping uniquely to transcripts were counted using HTSeq-count v.0.6.1.p1 (Anders *et al.*, 2015) and differential expression analysis was performed using DESeq2 v.1.8.1 (Anders & Huber, 2010). Genes were regarded as enriched for PiAgo-associated sRNAs if the $\log_2$ fold change to the GFP IP was at least two and the adjusted $P < 0.05$ (Benjamini & Hochberg, 1995). The $\log_2$ fold change was visualized in ternary plots using the R-package ggtern (Hamilton, 2015).

miRNAs were predicted using ShortStack v.2.1.0 (Axtell, 2013) and default plant miRNA settings, except for adjusting the sRNA size range to 18–26 nt. Strand-specific transcriptome sequences from *P. infestans* 88069 mycelium were mapped to the *P. infestans* genome using Tophat v.2.0.13 (Kim *et al.*, 2013) and default settings. After duplicated reads were marked using Picard v.1.137, GATK v.3.3.0 SplitNCigarReads and Indel Realignment (McKenna *et al.*, 2010) were applied, and variant discovery was performed using GATK HaplotypeCaller (DePristo *et al.*, 2011). Variants were filtered using GATK VariantFiltration. All GATK tools were applied using recommended RNA sequencing standard workflow settings (Van der Auwera *et al.*, 2013). A *de novo* transcriptome assembly was constructed using Trinity v.20140717 (Grabherr *et al.*, 2011) using jaccard clipping and default settings. Candidate coding regions were predicted using Transdecoder v.2.0.1 (Haas *et al.*, 2013).

Phylogenetic analysis

PIWI (P element-induced wimpy testis) domain sequences were obtained by Pfam analysis (http://pfam.xfam.org/) of full-length Ago protein sequences. All included oomycete proteins were predicted to have, at minimum, PAZ (Piwi Argonaute Zwille) and PIWI domains. The identified PIWI domain sequences were added to the multiple sequence alignment from Swarts *et al.* (2014) using Mafft v.7 (http://mafft.cbrc.jp/alignment/software/; Katoh & Frith, 2012) and default settings (Fig. S1; Table S3). Maximum likelihood phylogenetic reconstruction was performed in Mega5.2.2 (Tamura *et al.*, 2011), applying the Jones–Taylor–Thornton substitution model and 1000 bootstrap replications. Uniform rates of evolution among sites were assumed and gaps were partially deleted (cutoff 60% site coverage).

Immunostaining

Sporulating mycelia were fixed in 4% formaldehyde/Phosphate-Buffered Saline (PBS) for 20 min and washed for 3 × 5 min with PBS. Cell walls were partially digested by incubation for 20 min in 5 mg ml^{-1} lysing enzymes, 2 mg ml^{-1} cellulase (Sigma-Aldrich, St Louis, MO, USA) supplemented with protease inhibitor cocktail (Roche) and 1 mM PMSF. Samples were washed, permeabilized with PBS–0.2% Triton X-100 (PBST) for 10 min, blocked in 2% BSA/PBST for 45 min at room temperature (RT) and incubated with GFP antibody (1 : 250, Abcam ab6556, Cambridge, UK) overnight at 4°C. After washing, the samples were incubated with goat anti-rabbit IgG Alexa Fluor 555-conjugated secondary antibody (1 : 500, Abcam ab150078) for 2 h at RT and washed. Nuclei were counterstained with 0.4 μg ml^{-1} 4′,6-diamidino-2-phenylindole for 20 min. The samples were mounted in Vectashield (Vector Laboratories, Peterborough, UK). Image acquisition and analysis were performed as described previously (Munch *et al.*, 2015).

Accession numbers

The sRNA and mRNA transcriptome sequencing data are available through the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus, Series accession number GSE75282 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75282). The revised *PiAgo3* gene model has been deposited in GenBank, accession number KU168413.

Results

Oomycete Ago proteins form two clades

Eukaryotic Ago proteins contain four conserved functional domains. The N-terminal domain is important for target cleavage and sRNA duplex unwinding, whereas the PAZ and Mid domains bind the sRNA 3′ and 5′ ends, respectively. The PIWI domain adopts an RNase H fold and, similar to RNase H enzymes, cleavage-competent Ago proteins have four catalytic amino acids in their PIWI domain (Swarts *et al.*, 2014). This so-called catalytic tetrad (DEDH/D) forms the active site, where target RNA is cleaved via a two-metal-ion-dependent mechanism (Nakanishi *et al.*, 2012).

To clarify the relationship among the oomycete Ago proteins, we performed a phylogenetic analysis using the PIWI domains of selected bacterial, archaeal and eukaryotic Ago proteins. The eukaryotic sequences were chosen to represent the major eukaryotic supergroups, with a bias towards Stramenopiles, a lineage that encompasses oomycetes, brown algae and diatoms, among others (Adl *et al.*, 2012). The analysis placed the oomycete Ago proteins within the Ago subfamily (Shabalina & Koonin, 2008; Swarts *et al.*, 2014), although with low statistical support (Figs 1, S2). This result is in accordance with the overall higher sequence similarity of *P. infestans* Ago proteins to Ago-type than to Piwi-type proteins. The placement of *Ectocarpus siliculosus* Ago1 among the oomycete Ago proteins had low confidence, but suggests that oomycete Agos are more closely related to Agos in brown algae than in diatoms. An earlier study (De Riso *et al.*, 2009) also found that diatom Agos form a separate group.

The separation of the oomycete Agos into two well-defined clades was well supported in the phylogeny. As observed in a previous analysis of *P. sojae*, *Phytophthora ramorum* and *P. infestans* Ago sequences (Fahlgren *et al.*, 2013), one clade comprised proteins homologous to PiAgo3, PiAgo4 and PiAgo5, and the other clade included PiAgo1-like proteins (Fig. 1). The Ago4 and Ago5 proteins formed two separate subclades, but the PiAgo3-like proteins did not form a monophyletic group. Notably, we found that none of the Ago3/4/5 proteins had the amino acids DEDH/D at the positions corresponding to the catalytic tetrad. The majority, instead, had DDDH, suggesting that the Ago proteins in this clade mediate silencing through a cleavage-independent mechanism (Swarts *et al.*, 2014). By contrast, all oomycete Ago1 homologs, except one, had intact DEDH catalytic amino acids in their PIWI domains. The exception was one of the two *Saprolegnia diclina* Ago1 homologs (DDDK).

Establishment and validation of GFP-tagged Ago lines

To identify PiAgo-interacting sRNAs from *P. infestans*, we employed PiAgo-sRNA co-IP. Stable transgenic lines expressing GFP-tagged PiAgo proteins were generated by protoplast transformation. Both N- and C-terminal tagging was attempted for each protein, but only N-terminal GFP lines were recovered for *PiAgo1*, *PiAgo4* and *PiAgo5*, whereas the *PiAgo3* lines carried a C-terminal GFP construct. The most strongly expressing line of each construct (evaluated by fluorescence microscopy) was chosen for further studies. A line harboring *GFP* driven by *ham34* (Judelson *et al.*, 1992), the same constitutive promoter as applied in the *PiAgo-GFP* lines, was included in subsequent experiments (henceforth termed GFP control).

To assess the *PiAgo* transcript abundances in the selected *PiAgo-GFP* transformants, qRT-PCR analyses were performed, measuring the combined expression level of endogenous *PiAgo* and the introduced *PiAgo* genes. The *PiAgo1* level in the corresponding tagged line was about twice that of the wild-type (Fig. S3a). The *PiAgo3* expression level in *PiAgo3-GFP* was almost five-fold elevated and the *PiAgo5* transcript in *PiAgo5-GFP* accumulated to about the same level as in the wild-type (Fig. S3b). The highest overexpression was seen in the *PiAgo4-GFP* line, where the *PiAgo4* abundance was *c.* 200-fold elevated compared with the wild-type (Fig. S3a). *PiAgo1* and *PiAgo4* accumulated at comparable levels in their respective tagged lines, as endogenous PiAgo1 is expressed at high levels (Vetukuri *et al.*, 2011a). In the control *GFP* line, all four *PiAgo* transcripts were expressed at similar levels as in the wild-type (Fig. S3a,b).

Cytoplasmic localization of PiAgo1 and PiAgo4

To reveal to which subcellular compartment the PiAgo proteins localize, the GFP-tagged *P. infestans* lines were studied by confocal microscopy. Low GFP signal and high background fluorescence at the GFP emission wavelengths made it difficult to distinguish the PiAgo-GFP signal from autofluorescence. Instead, immunofluorescence with a GFP antibody and a fluorescent secondary antibody was employed, which increased the signal-to-noise ratio and allowed the detection of PiAgo1 and PiAgo4. Both proteins were observed primarily in the cytoplasm in *P. infestans* hyphae and sporangia (Fig. S4).

Ago-sRNA co-immunoprecipitation

Following co-IP, proteins and sRNAs were extracted and analyzed by Western blot and denaturing PAGE. GFP-tagged PiAgo1 (130 kDa) and PiAgo4 (124 kDa) were expressed at approximately equal levels, although a detectable level of free GFP was observed in the *PiAgo4-GFP* line (Fig. 2a). PiAgo3, with a predicted molecular mass of 160 kDa, was undetectable (Fig. S5a). PiAgo5-GFP (120 kDa) was expressed (Fig. S5a), but not detectable in the IP fraction. Western blotting of a protein dilution series showed that the amount of GFP in the control line was over 40-fold higher than the amounts of PiAgo1-GFP and PiAgo4-GFP in the tagged lines (Fig. S5b).

RNA was extracted from the co-IP samples from PiAgo1, PiAgo4, PiAgo5 and the GFP control, end-labeled with ^{32}P and separated by denaturing PAGE. Distinct sRNA size profiles were seen for the three PiAgo proteins. Notably, strong signals from *c.*

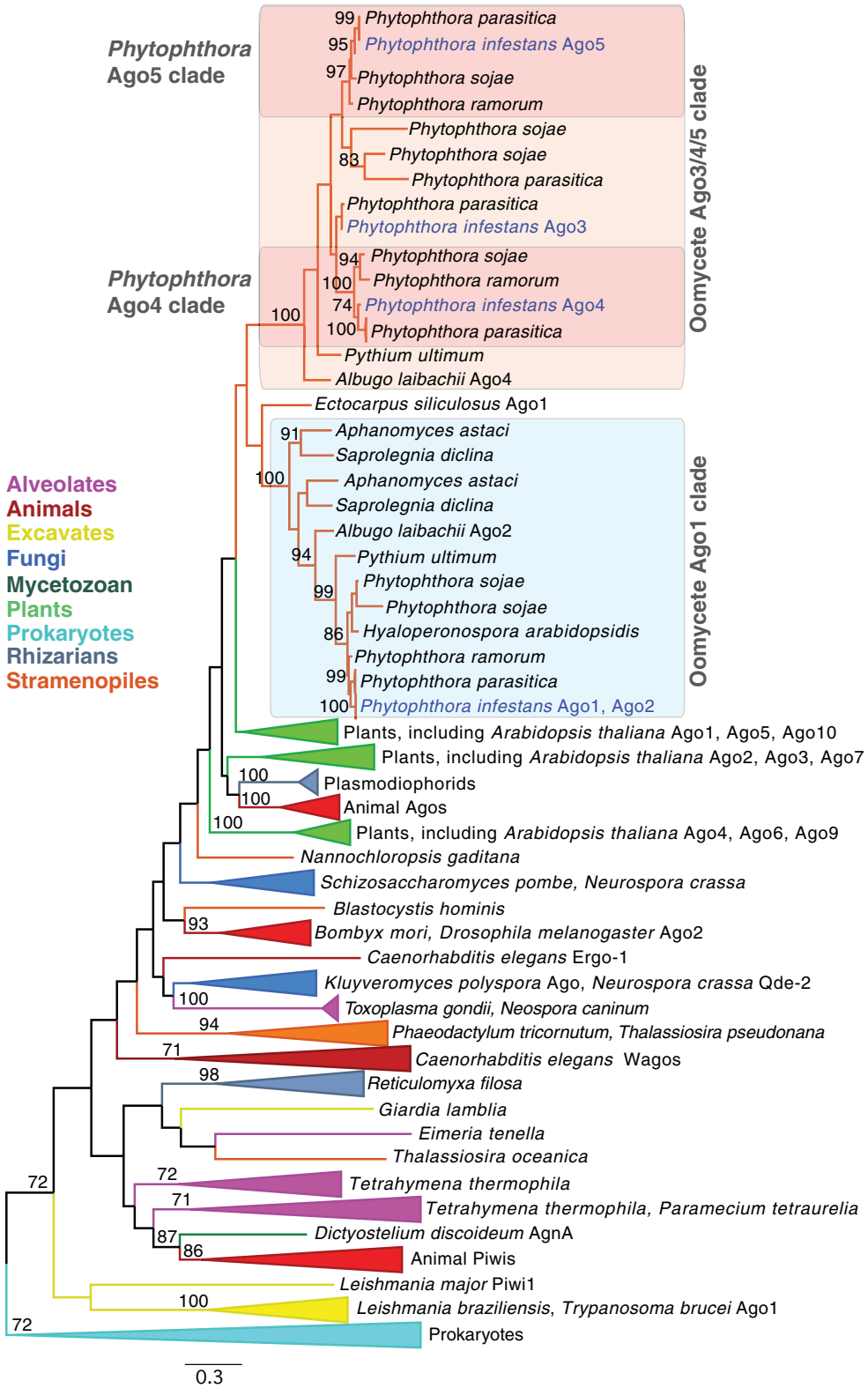

Fig. 1 *Phytophthora* Argonautes (Agos) form two distinct clades. Unrooted maximum likelihood phylogenetic reconstruction created using the PIWI (P element-induced wimpy testis) domains from selected Ago-like, Piwi-like and prokaryotic Agos. Numbers indicate interior branch bootstrap values > 70%. Clades were collapsed for clarity. Branches and clades are color-coded, as indicated to the left of the tree. Bar, number of amino acid substitutions per site.

21- and 25-nt sRNAs were observed in the samples from PiAgo1 and PiAgo4, respectively (Fig. 2b). These two sRNA fractions correspond to the major sRNA size classes previously identified in *P. sojae*, *P. infestans* and *P. ramorum* (Vetukuri *et al.*, 2012; Fahlgren *et al.*, 2013). The sRNA signal was weak in the PiAgo5 IP, but a band at *c.* 27 nt was visible.

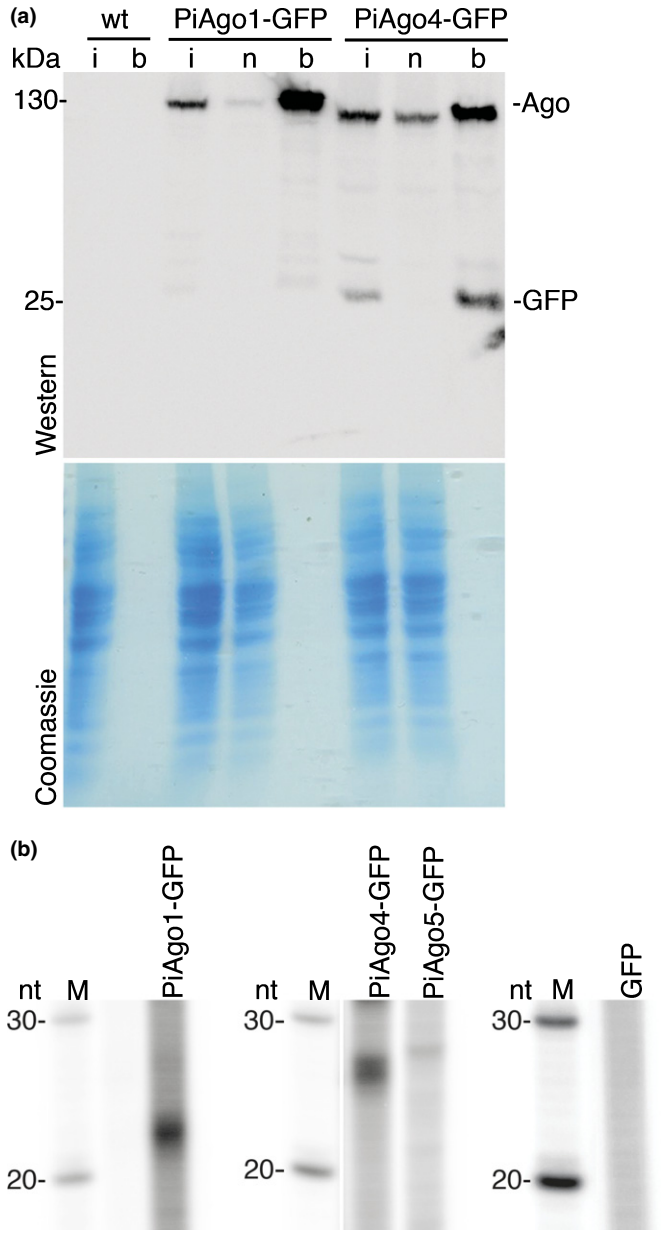

Fig. 2 Ago-sRNA co-immunoprecipitation (co-IP) from *Phytophthora infestans*. (a) Green fluorescent protein (GFP)-tagged PiAgo1 and PiAgo4 were detected with anti-GFP. i, input cell lysate; n, nonbound wash fraction; b, eluate after IP with anti-GFP antibody; wt, wild-type. The panel below the Western blot shows the Coomassie-stained gel as loading control. kDa, kilodaltons. (b) Denaturing polyacrylamide gel electrophoresis analysis of small RNAs (sRNAs) eluted from the PiAgo-GFP immunocomplexes and end-labeled with ^{32}P. M, RNA marker; nt, nucleotide.

Evidence for *PiAgo3* pseudogenization in isolate 88069

Detection of PiAgo3-GFP by Western blot was unsuccessful (Fig. S5a). On closer analysis of the *PiAgo3-GFP* construct, a single-nucleotide indel (a T deletion) was identified at position 1023 in *PiAgo3*, corresponding to the predicted third exon in the reference genome T30-4. This indel could also be found in a sequenced *PiAgo3* PCR product obtained with genomic DNA from isolate 88069 and in transcriptome sequence reads from the same isolate. Alignment of the transcriptome reads to the T30-4 *PiAgo3* genomic locus did not find support for the two predicted introns, suggesting that *PiAgo3* in 88069 differs from the T30-4 gene annotation.

5′ RACE confirmed the lack of introns in *PiAgo3* from 88069 and identified an alternative upstream start codon in the transcript (Fig. S6). The sequence between this start codon and the annotated stop codon was found to be highly similar to a full-length Ago-encoding open reading frame (ORF) in *Phytophthora parasitica*. The corresponding ORF in isolate 88069 was, however, disrupted by the earlier described indel, leading to *PiAgo3* pseudogenization (Fig. S7). The homologous relationship between the two *Ago3* loci is evidenced from both genomic synteny and phylogenetic co-clustering of the two PIWI domains (Fig. 1). Moreover, the 5′ terminal parts of both transcripts encode glycine–arginine-rich repeats, which, in the N-termini of *Trypanosoma brucei* and *Toxoplasma gondii* Ago proteins, are known to be important for Ago function (Shi *et al.*, 2004; Musiyenko *et al.*, 2012). Although the frequency of the *PiAgo3* indel is unknown among *P. infestans* isolates, pseudogenization would explain our failure to detect PiAgo3-GFP by Western blot. Consequently, PiAgo3 was excluded from further analyses.

Distinctive sRNA size preferences of PiAgo proteins

sRNAs from the PiAgo IP samples, the control GFP IP and non-IP sRNAs from the PiAgo1-GFP line ('Ago1 input') were subjected to high-throughput sRNA sequencing. In total, ten libraries were sequenced, generating 109 621 761 reads on an Ion Proton sequencing platform (Table S2). Genome-wide mapping of the sRNA sequence reads confirmed the size distribution observed by gel electrophoresis from PiAgo1 and PiAgo4 (Fig. 3). PiAgo1 was dominated by 20–22-nt sRNAs and clearly depleted for 24–26-nt species, whereas the majority of PiAgo4-associated sRNAs were 24–26 nt long. The size distributions in the PiAgo1 input and in GFP IP were bimodal and centered on 21 and 25 nt. The profile in the PiAgo5 IP sample differed from the control samples by being enriched for 21-nt sRNAs (Fig. 3). This argues for the presence of a low (Figs S3b, S5a), but significant, amount of PiAgo5-GFP fusion protein in this transgenic line. The 21-nt sRNAs were not observable by gel electrophoresis in the PiAgo5 sample (Fig. 2b), which was probably caused by the low PiAgo5 IP efficiency.

5′ nucleotide identities of co-purified sRNAs

The preferential association of PiAgo1 and PiAgo4 with distinctly sized sRNAs suggests that the sRNA length is an important determinant for Ago sorting in *P. infestans*. To further examine the sRNA binding properties of the PiAgo proteins, we analyzed the 5′ nt identities of the PiAgo-associated sRNAs. Recognition of the 5′ terminal nt has been shown to form the basis for Ago sRNA specificity in organisms such as *Arabidopsis*, *Drosophila* and humans (Mi *et al.*, 2008; Frank *et al.*, 2010; Cora *et al.*, 2014). A clear 5′ nt preference was observed in the PiAgo4

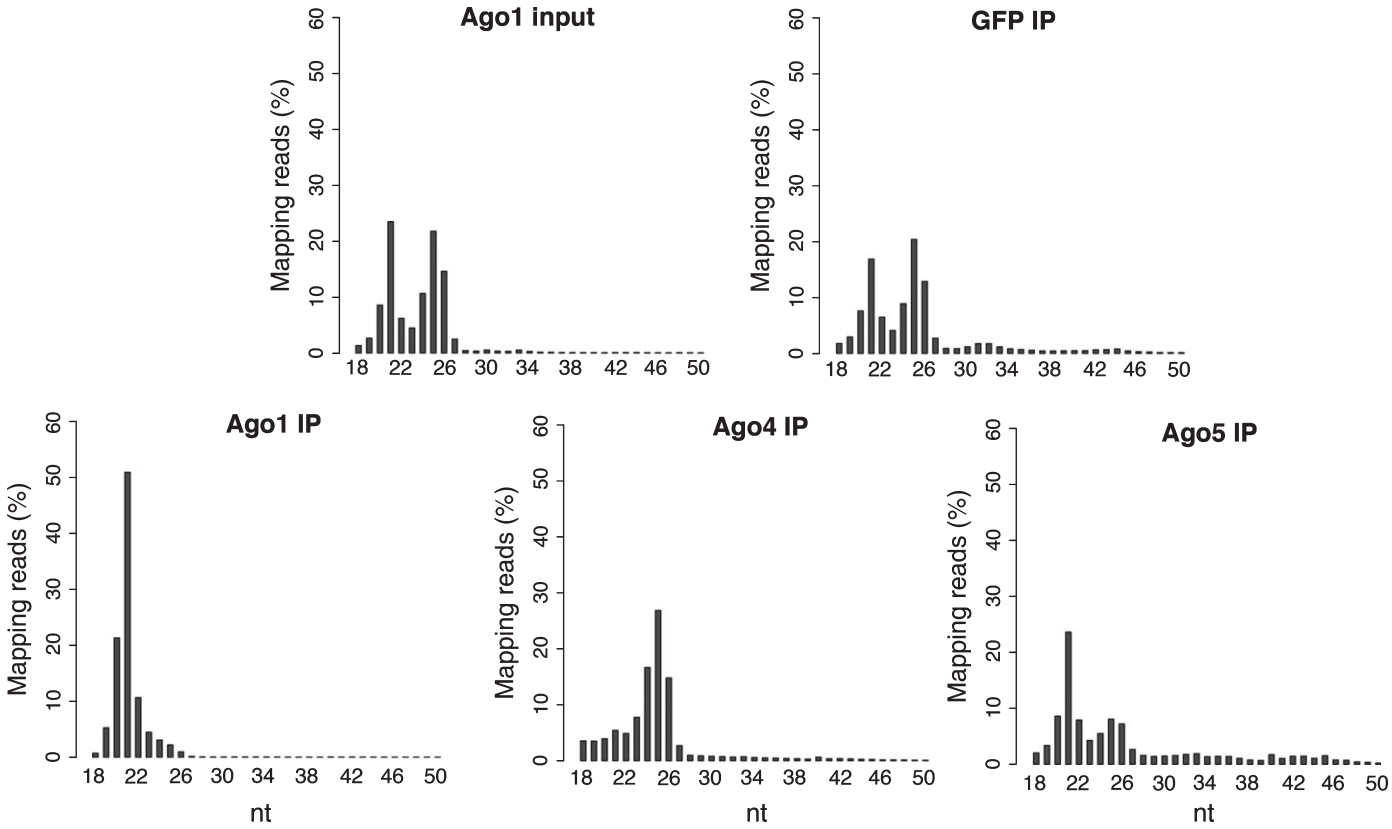

Fig. 3 Size distribution of small RNAs (sRNAs) co-eluted with *Phytophthora infestans* Argonaute (Ago) proteins. Combined sense and antisense sRNA reads mapped to the *P. infestans* genome. Distinct size profiles are evident in the PiAgo1, PiAgo4 and PiAgo5 immunoprecipitation (IP) samples. nt, nucleotide.

sample, where 78% of total sRNAs (18–30 nt) had 5′ uracil (U), compared with 54% in the GFP control. This bias was particularly pronounced for 24–26-nt PiAgo4-bound sRNAs, which had 5′ U in 91% of cases (Figs 4a, S8; Table S3). PiAgo1 showed enrichment of 5′ cytosine (C): 25% of the 18–30-nt sRNAs had this 5′ nt, compared with 13% in the PiAgo1 input and 11% in the GFP control. The enrichment of 5′ C in the PiAgo1 IP and 5′ U in the PiAgo4 IP agrees with the earlier observed high prevalence of 5′ C in 21-nt sRNAs and of 5′ U in 25/26-nt sRNAs in *P. infestans* (Vetukuri *et al.*, 2012).

Identity of PiAgo-enriched sRNAs

To investigate potential functional diversification between the PiAgo proteins, we analyzed the genomic origins of the sRNAs in the different PiAgo IP samples. The numbers of mapping sequence reads were counted in the following nine annotated feature sets from the *P. infestans* genome: intergenic regions, TEs and repeats, tRNA genes, rRNA genes, introns, total protein-coding genes (effector-encoding genes excluded) and effector genes (*RxLR*s, *CRN*s and *CRN* pseudogenes).

The majority of the sRNAs (18–30 nt) in the ten libraries derived from TEs (Fig. 4b; Table S3). A comparatively large proportion of sRNAs in the PiAgo1 IP mapped to protein-coding genes (10%), *CRN*s (1%) and pseudo-*CRN*s (2%). In the PiAgo4 IP, the fractions of sRNAs mapping to the different genomic locations were similar to those of the GFP control IP sample. sRNAs from introns and intergenic sequences were overrepresented in the PiAgo5 sample.

Repeat class-specific PiAgo association

We next analyzed the repetitive sequences in more detail to find repeat classes specifically associated with a particular PiAgo protein. The largest transposon class in the *P. infestans* genome consists of *Gypsy* LTR retroelements (Haas *et al.*, 2009), which account for approximately one-third of the genome. sRNAs from *Gypsy* LTRs were predominantly 20–22 nt in length and enriched in the PiAgo1 IP sample (Figs 5, S9; Table S3). Two classes of DNA transposons, *Mutator* and *helENtron*, were also biased towards 20–22-nt PiAgo1-associated sRNAs. Of the remaining transposon classes present in the *P. infestans* genome, *Copia* LTRs, and *Helitron*, *Crypton* and *PiggyBac* elements had more 24–26-nt than 20–22-nt sRNAs. A similar observation regarding *Helitron*, *Crypton* and *PiggyBac* repeats was found by Vetukuri *et al.* (2012). The sRNAs centered on 21 nt were associated with PiAgo1 and, to some extent, with PiAgo5, whereas 24–26-nt-long sRNAs dominated in the PiAgo4 IP (Fig. 5). To verify the co-IP of sRNAs from *Gypsy* and *helENtron* with PiAgo1, and of *Copia* and *Crypton* sRNAs with both proteins, Northern hybridizations were carried out. sRNAs from the analyzed TEs were undetectable in the IP samples and generated only weak signals in the input material (Fig. S10). The detected sRNAs were also longer than expected:

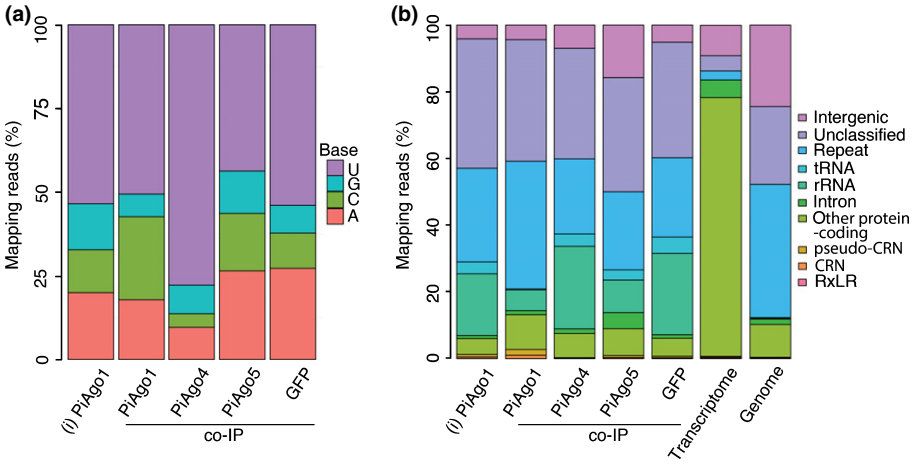

Fig. 4 Identity of 5′ nucleotide (nt) and genomic origin of small RNAs (sRNAs) (18–30 nt) associated with *Phytophthora infestans* Argonaute (Ago) proteins. (a) Relative frequency of each 5′ terminal nucleotide in total sRNAs (i, input) and in sRNAs co-purified with PiAgo proteins and the green fluorescent protein (GFP) control. Compared with GFP, the PiAgo4 sample shows enrichment for 5′ uracil (U) ($P < 0.05$). PiAgo1 is enriched for 5′ cytosine (C) ($P < 0.001$, relative to the PiAgo1 input; $P < 0.001$, relative to the GFP control). Two-tailed *t*-test with equal variance, arcsine-square root transformation (Snedecor & Cochran, 1980). The 5′ nt identity for each sRNA size class is presented in Supporting Information Fig. S8. (b) Proportion of sRNAs derived from different genomic features in PiAgo immunoprecipitation (IP) and control samples. The contributions of each genomic feature to the wild-type transcriptome (isolate 88069) and to the *P. infestans* reference genome (isolate T30-4) are included for comparison. i, input; Other protein-coding, all protein-coding genes except Crinkler (*CRN*), pseudo-*CRN* and *RxLR*; rRNA, ribosomal RNA; tRNA, transfer RNA; Unclassified, unclassified repeat.

25–28 nt from *Gypsy*, *Copia* and *Crypton* and > 40 nt from *Gypsy* and *helENtron*.

CRN-derived sRNAs associate with PiAgo1

sRNAs generated from *RxLR* genes were relatively low in number, but their size distribution was similar to that of *Copia*, *Helitron*, *Crypton* and *PiggyBac* transposons (Figs 5, 6). Conversely, *CRN*s and pseudo-*CRN*s were potent producers of PiAgo1-enriched 20–22-nt sRNAs (Fig. 6). This latter result is in agreement with previous studies, wherein the majority of reported *CRN* mapping sRNAs were 21 nt in length (Vetukuri *et al.*, 2012; Fahlgren *et al.*, 2013). The 21-nt sRNAs from *CRN*s and pseudo-*CRN*s were also found in the PiAgo5 IP, although to a lesser extent than in the PiAgo1 IP. The relative levels of 23–26-nt sRNAs from *CRN*s and pseudo-*CRN*s were higher in the PiAgo4 sample than in the GFP control (Fig. 6).

To examine the preferential association of sRNAs from *CRN*s, pseudo-*CRN*s and *RxLR*s with PiAgo1, PiAgo4 and PiAgo5 in comparison with GFP, a differential expression analysis was performed on *P. infestans* total protein-coding genes using DESeq2 (Anders & Huber, 2010). We focused on three sRNA size ranges: 18–30, 20–22 and 24–26 nt. *CRN*- and pseudo-*CRN*-derived 18–30-nt-long reads were present in all three PiAgo IPs, but significant sRNA enrichment was identified only in the PiAgo1 and PiAgo5 samples (Figs S11, S12; Table S4). Six *CRN*s were enriched for PiAgo1-associated 18–30-nt sRNAs. Among these, PITG_18133 was the most potent sRNA-producing *CRN* in terms of both sense and antisense sRNAs. This gene was significantly enriched for sense 18–30-nt sRNAs also in the PiAgo5 sample. PITG_18133, and one additional gene, PITG_04812, were found to be associated with PiAgo1 also in the analysis performed on 20–22-nt sRNAs (Fig. S11; Table S4). Five pseudo-*CRN*s showed significant sRNA expression differences in the PiAgo1 sample compared with the control. Among these loci, a strikingly large number of sRNAs mapped to PITG_22969 and PITG_22270 (Fig. S12; Table S4). No *CRN*s or pseudo-*CRN*s were significantly enriched for 24–26-nt sRNAs in any of the PiAgo IP samples.

The expanded *P. infestans CRN* gene family has 196 full-length members and 255 pseudogenes, and resides in several large clusters in the genome (Haas *et al.*, 2009). The CRN proteins have conserved N-terminal parts, containing LFLAK and DWL domains, and highly diversified C-termini. In total, 35 C-terminal domain architectures can be identified among *P. infestans* CRNs (Stam *et al.*, 2013). Taking non-uniquely mapping reads into account, a number of hotspots for PiAgo1-associated *CRN* and pseudo-*CRN* sRNAs were identified. Among these, only DN5 and DC domain-encoding genes were represented (21 and 23 genes). The majority of the DC-type *CRN*s were clustered on supercontig 6, whereas most DN5-type *CRN*s were found in two clusters on supercontigs 16 and 63.

As a result of the small numbers of sRNA reads from *RxLR*s, no member of this gene family was identified as significantly enriched for PiAgo-associated sRNAs (Fig. S13; Table S5). A noticeable depletion of antisense *RxLR*-derived sRNAs was observed in the PiAgo1 sample.

sRNAs from other protein-coding genes

To further characterize the sRNA binding properties of the PiAgo proteins, we employed the same enrichment analysis on *P. infestans* total non-effector protein-coding genes. In a previous sRNA deep sequencing study from *P. infestans*, an abundance of 18–23-nt sRNAs was observed from inverted repeats and protein-coding genes, whereas 24–30-nt sRNAs mainly derived from transposons (Fahlgren *et al.*, 2013). The pattern of PiAgo-associated sRNA reads mapping to protein-coding genes was in line with these

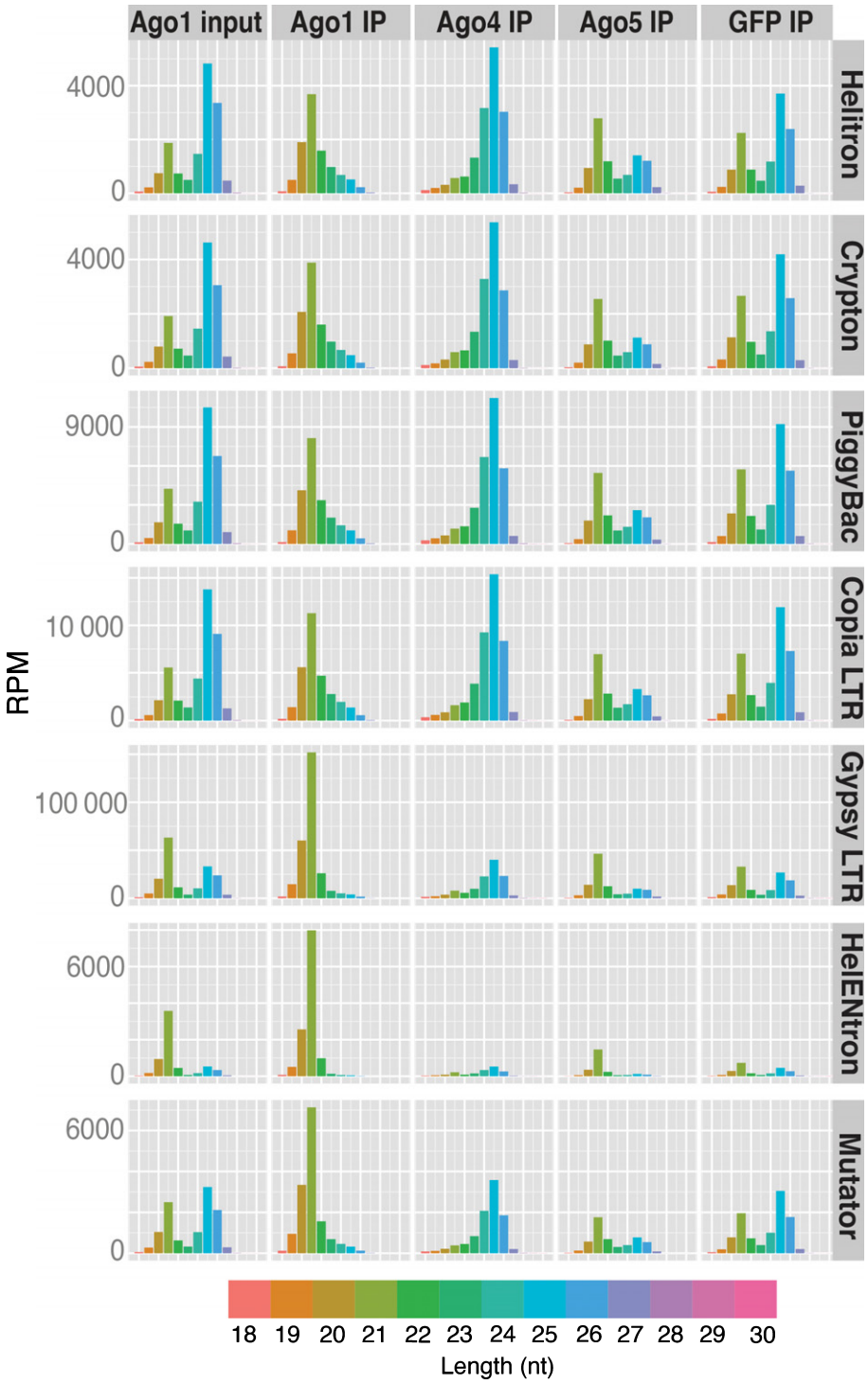

Fig. 5 Size distribution of small RNAs (sRNAs) mapping to different transposon classes in *Phytophthora infestans*. sRNAs from *Helitron*, *Crypton*, *PiggyBac* and *Copia* long terminal repeat (LTR) co-purify with both PiAgo1 and PiAgo4, whereas sRNAs from *Gypsy* LTR, *HelENtron* and *Mutator* associate preferentially with PiAgo1. GFP, green fluorescent protein; nt, nucleotide; RPM, reads per million.

findings: the majority were 20–22 nt, rather than 24–26 nt (Fig. 6). The sRNAs that showed significant differential accumulation were, to the largest extent, 20–22 nt in length and associated with PiAgo1 (Fig. S14; Table S5). This is consistent with the total sRNA size distribution and the enrichment of sRNAs from protein-coding genes in the PiAgo1 sample (Figs 3, 4b).

Among the genes significantly enriched for sRNAs in the PiAgo4 IP, *PiAgo4* itself was represented. The sRNAs from *PiAgo4* were, however, not of the PiAgo4-preferred 24–26-nt size class. Instead, the read distribution covered the whole 18–30-nt size span. In terms of non-uniquely mapping reads, *PiAgo1* and its gene copy *PiAgo2* were among the top 25 contributors of 18–30-nt sRNAs in the PiAgo1 IP sample (*CRN* genes included). As uniquely mapping reads were used in the differential expression analysis, very few *PiAgo1/2*-derived sRNAs were considered. (The two coding sequences differ at only two nucleotide positions.) Nonetheless, the few unique 18–30- and 20–22-nt sRNAs that were detected from *PiAgo1* were significantly enriched in the

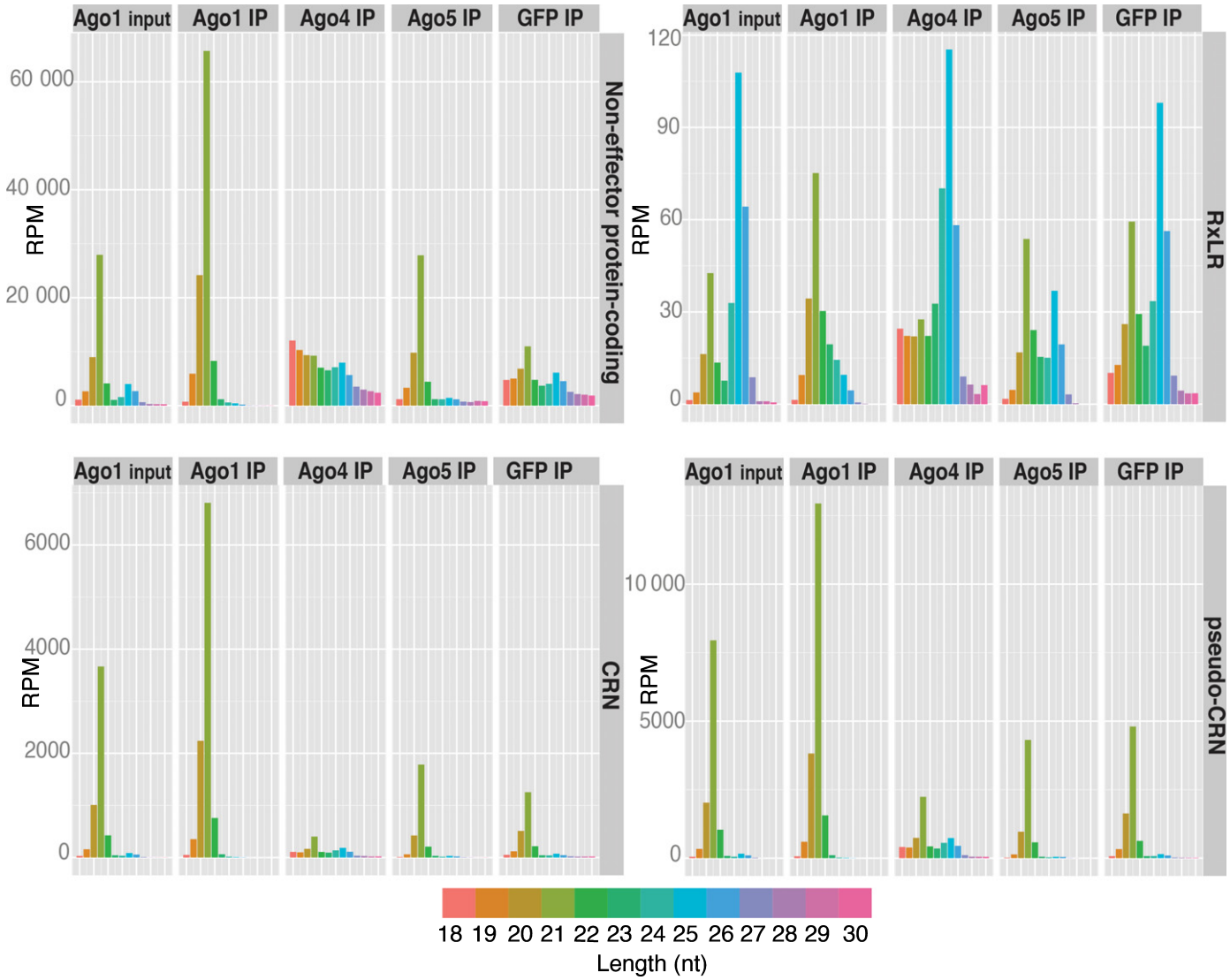

Fig. 6 Size distribution of small RNAs (sRNAs) from *Phytophthora infestans* protein-coding genes. sRNAs were mapped to individual datasets consisting of all protein-coding genes (effectors excluded), effector-coding genes (*RxLR*, *CRN* (*Crinkler*)) and *CRN* pseudogenes. GFP, green fluorescent protein; nt, nucleotide; RPM, reads per million.

PiAgo1 IP. Unique *PiAgo2*-derived sRNAs were not significantly enriched in any of the PiAgo IPs.

Counting the total *PiAgo1/2* mapping reads in the PiAgo1 input and IP samples, a roughly equal abundance of sense and antisense sRNAs was seen overlapping the two genomic loci, suggesting the involvement of PiRdR1 in sRNA biogenesis. With few exceptions, all *PiAgo1/2* sRNA reads in these samples were 19–22 nt in length, indicating specific PiDcl-mediated cleavage (Fig. S15). By contrast, the sRNA read coverage over *PiAgo1/2* was low in the PiAgo4, PiAgo5 and GFP IPs, where the majority of reads mapped in the sense direction and did not show any sRNA size class preference.

To examine the sRNA coverage over *PiAgo1* and *PiAgo2* in the wild-type isolate from which the GFP-tagged PiAgo lines were generated (88069), Illumina sRNA sequencing data (Åsman *et al.*, 2014) were analyzed. Few sRNA reads overlapped the two genes in 88069. The scarcity of sRNAs in the wild-type and in the PiAgo4, PiAgo5 and GFP IP samples indicates that the boost in sRNA production was restricted to the *PiAgo1-GFP* transformant and therefore was an effect of *PiAgo1* transgene expression. In line with this interpretation, large numbers of sense and antisense sRNAs mapping to *GFP* were observed specifically in the PiAgo1 input and IP samples (Fig. S15).

The genes identified as significantly enriched for sRNAs in the PiAgo5 IP overlapped to a large extent with the PiAgo1-enriched genes (Fig. S14; Table S5). For the majority of the shared genes, the sRNAs co-purified to a larger extent with PiAgo1 than with PiAgo5. Nevertheless, a significant expression difference was shown in the PiAgo5 IP compared with the GFP control. The gene most highly enriched for 18–30- and 20–22-nt sRNAs in the PiAgo5 IP was *PiAgo3*. We were unable to detect any sRNAs from *PiAgo3* by Northern hybridization with PiAgo co-IP RNA. Nonetheless, this observation suggests control of the *PiAgo3* pseudogene transcript by PiAgo5 and 20–22-nt sRNAs.

PiAgo1 associates with miRNA

A single experimentally verified miRNA, miR8788, is reported from *Phytophthora*. miR8788 is conserved in at least three species (*P. infestans*, *P. ramorum* and *P. sojae*) and is 21–22 nt in length (Fahlgren *et al.*, 2013). To search for additional miRNAs in *P. infestans*, we analyzed the IP sRNA sequence libraries using the miRNA annotation program ShortStack (Axtell, 2013), but no new candidate could be identified. *pin*-miR8788 was predicted in the PiAgo1 input sample and in the IP samples from PiAgo1 and PiAgo5. Reads from the miRNA and miRNA* sequences had uniform 5′ ends, but more variable 3′ ends and 2-nt 3′ overhangs at both sides of the predicted duplex (Table S6). The support for miR8788 was rather weak (Kozomara & Griffiths-Jones, 2014; http://www.mirbase.org/), as less than four miRNA* reads were detectable in the different samples.

The vast majority of sRNA reads from the *pin*-miR8788 locus were 19, 20 and 21 nt in length. Notably, over five times as many reads were 19 nt as 21 nt in PiAgo1 IP (Table S6). Similar to our observations, the *P. sojae* homolog *psj*-miR8788a is 22 nt in length, but sRNA sequencing identified 19 nt as the most abundant length (Fahlgren *et al.*, 2013). The association of PiAgo1 with *pin*-miR8788 was confirmed by Northern hybridization with co-IP sRNAs (Fig. 7). In addition to the PiAgo1 IP sample, the miRNA was detectable in total RNA samples from 88069, *PiAgo1-GFP* and *GFP*. In agreement with sRNA sequencing, three isoforms of *pin*-miR8788 were seen, but the sizes apparent on the blot were 20–22 nt rather than 19–21 nt. Analyzing the reads aligning to the *pin*-miR8788 locus in further detail, we identified 21-nt *pin*-miR8788 sequences carrying one to five non-templated uridines at their 3′ ends (total length, 22–26 nt). Such extra added uridines were seen in all samples that contained reads from *pin*-miR8788 (Table S6). Uridylation of plant sRNAs promotes 3′–5′ exoribonuclease degradation (Ruegger & Grosshans, 2012), suggesting that the 19-nt reads represent degradation intermediates.

Discussion

A previous study predicted a small number of candidate miRNAs in *P. infestans* (Vetukuri *et al.*, 2012), but miR8788 is, to date, the only verified miRNA among three analyzed *Phytophthora* species (Fahlgren *et al.*, 2013). This is in stark contrast with the numerous miRNAs found in plants (Nozawa *et al.*, 2012) and their well-studied biogenesis pathways (Rogers & Chen, 2013). Fungi do not have miRNAs, but instead produce a class of sRNAs, termed miRNA-like RNAs (Lee *et al.*, 2010). These have silencing capacity, but do not match the criteria of current miRNA prediction algorithms. Compared with miRNAs, the miRNA-like precursor stemloop structures are more flexible in length, and their biogenesis pathways are not as well defined as those of miRNAs. miRNA-like sRNAs have also been described in diatoms (Rogato *et al.*, 2014), a sister lineage to oomycetes within the Stramenopiles. Two miRNA candidates were predicted in *Phaeodactylum tricornutum*, but these sRNAs were designated as miRNA-like based on the presence of a 2-nt 3′ overhang only on one side of the miRNA/miRNA* duplex. As *P. infestans* has the capacity to generate one miRNA and has active sRNA silencing pathways, it is likely that atypical miRNAs also exist in this oomycete. The conserved vertebrate miR-451 is generated via an unusual Dcl-independent pathway that does not proceed via a miRNA/miRNA* duplex, but instead requires the slicer activity of Ago2 and exonuclease trimming (Cheloufi *et al.*, 2010). This example shows that failure to detect a miRNA* sequence does not strictly imply that miRNAs are absent.

The detection of *pin*-miR8788 in the PiAgo1 IP by both sRNA sequencing and Northern hybridization suggests that PiAgo1 is the major miRNA-binding partner in *P. infestans*. The observation of 20–22-nt-long *pin*-miR8788 molecules by Northern hybridization indicates that the abundant 19-nt-long sequence reads represent degradation intermediates. In line with this interpretation, 21-nt perfectly matching *pin*-miR8788 reads carrying additional nontemplated U nucleotides at the 3′ end were identified. Whether 3′ uridylation is a mark of sRNA degradation in *P. infestans*, similar to *Arabidopsis* and *Chlamydomonas reinhardtii* (Ruegger & Grosshans, 2012), remains to be established. Notably, *P. infestans* lacks a homolog of the methyltransferase *Hen1* (Vetukuri *et al.*, 2012), which 2′-*O*-methylates sRNAs at the 3′ end as a protection against uridylation, trimming and degradation (Ameres & Zamore, 2013). In comparison with the large number of reads from the miRNA strand in the PiAgo1 and PiAgo5 IP samples, very few miRNA* reads were detected, most probably reflecting the incorporation of only the guide strand in the *P. infestans* RNA silencing complex.

There was a size discrepancy between the PiAgo5-bound sRNAs detected by ^{32}P labeling (27 nt) and by sRNA sequencing (21 nt). As the latter technique is the more sensitive of the two, the true size of PiAgo5-bound sRNAs is probably 21 nt. The greater sensitivity of sRNA sequencing is evidenced by the

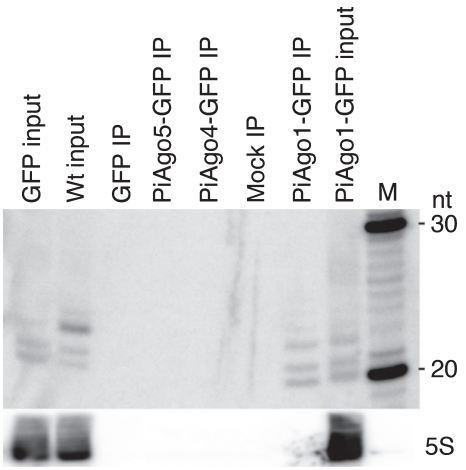

Fig. 7 Association of *pin*-miR8788 with *Phytophthora infestans* Ago1. *pin*-miR8788 is detectable by Northern hybridization in total small RNA (sRNA) samples (input) and in the PiAgo1-bound fraction using a probe antisense to *pin*-miR8788 (Fahlgren *et al.*, 2013). Shown below the blot is the membrane reprobed for 5S rRNA (Åsman *et al.*, 2014) as loading control. Mock IP, IP in a strain that does not express a PiAgo-GFP protein; M, RNA marker.

generation of large numbers of sRNA reads and a distinct sRNA size profile in the GFP IP (Fig. 3; Table S2), despite the lack of visible sRNAs on the PAGE gel in this sample (Fig. 2b).

Our phylogenetic analysis showed that oomycete Ago1 and Ago3/4/5 homologs form two separate clades. Most likely, functional diversification has occurred between members of the two clades, which is perhaps in part reflected by the presence of conserved catalytic tetrad amino acids only in Ago1-type proteins. The IP data indicate that PiAgo1, apart from having a gene regulatory function, is also involved in the silencing of a number of TEs. This is illustrated by the identification of PiAgo1-bound 20–22-nt sRNAs from the greatly expanded class of *Gypsy* LTR transposons, and from *Mutator* and *helENtron* elements. By contrast, the shared action of PiAgo1 and PiAgo4 seems to silence *Copia* LTRs and *Helitron*, *Crypton* and *PiggyBac* transposons. Both 20–22-nt and 24–26-nt sRNAs were produced from these TEs, with the latter size class in the majority and co-purified with PiAgo4. The TE mode of transposition is probably not a determinant for sRNA-PiAgo sorting, as retrotransposons and DNA transposons were found to be similarly distributed between PiAgo1 and PiAgo4.

Clear functional subdivision between different Dcl/Ago pairs is found in *Drosophila*, where DmDcr1 generates miRNAs that bind DmAgo1, and DmDcr2 produces siRNAs that load into DmAgo2 (Tomari *et al.*, 2007). Four Dcl proteins in *Arabidopsis* generate different sRNA classes, which are channeled into ten different AtAgos based on sRNA size and 5′ nt identity (Bologna & Voinnet, 2014). *Phytophthora infestans* has two Dcl proteins (Vetukuri *et al.*, 2011a; Fahlgren *et al.*, 2013), and knockdown of *PiDcl1* disrupts the accumulation of 21-nt sRNAs (Vetukuri *et al.*, 2012). This observation, together with the results from the present study, suggests a tight linkage between PiDcl1, PiAgo1 and 21-nt sRNAs.

Compared with DNA mutations, silencing at the transcriptional level might allow a pathogen to re-employ an *Avr* gene and regain virulence in a scenario in which the cognate *R* gene-expressing host plant is not present (Gijzen *et al.*, 2014). sRNA-guided silencing of *PsAvr3a* in *P. sojae* leads to virulence on *Rps*3a-expressing plants (Qutob *et al.*, 2013). Virulence and *PsAvr3a* silencing are inherited in the offspring in particular crosses between virulent and avirulent isolates. The size profile of the sRNAs identified from the silenced *PsAvr3a* locus is very similar to the population of PiAgo4-bound sRNAs in our study, suggesting that transgenerational gene silencing in *P. sojae* might be mediated by this organism's Ago4 homolog.

Feedback regulation of Ago protein levels is known to occur in plants and animals. *Arabidopsis Ago1* homeostasis is regulated by both siRNAs and miRNAs; miR168 guides AtAgo1 to cleave the *AtAgo1* mRNA, which induces siRNA production and reinforces *AtAgo1* self-silencing (Mallory & Vaucheret, 2009). Moreover, both *Arabidopsis* and animal Ago proteins are regulated by autophagy. Accumulation of mammalian Ago2 and *Drosophila* Ago1 is dependent on miRNA availability, in such a way that unloaded Ago is unstable (Smibert *et al.*, 2013; Derrien & Genschik, 2014). Whether or not such feedback regulatory loops control the abundance of *P. infestans* Agos remains to be investigated. Highly abundant and uniquely sized 20–22-nt sRNAs from the duplicated genes *PiAgo1* and *PiAgo2* and from *GFP* were, however, observed in the PiAgo1 IP. The presence of low levels of sRNAs from *PiAgo1*, *PiAgo2* and *GFP* in the IP samples from PiAgo4 and PiAgo5 indicates that this sRNA accumulation is a result of the presence of the *PiAgo1-GFP* construct. This is similar to sense cosuppression in plants: silencing of an endogenous gene by the introduction of a homologous transgene (Jorgensen, 1995). Large numbers of sRNAs were also identified from *PiAgo4* in the highly overexpressing PiAgo4-GFP line. These sRNAs, however, covered the whole 18–30-nt size span, suggesting non-specific binding to PiAgo4. Possibly, both PiAgo1-GFP and PiAgo4-GFP are silenced post-transcriptionally by 21-nt PiAgo1-bound sRNAs.

This study provides new insights into the PiAgo-mediated sRNA silencing pathways in *P. infestans*. The results indicate that PiAgo1 is a multifunctional protein involved in the regulation of *CRN* effector-encoding genes, in TE control and in *pin*-miR8788-directed gene silencing. The identification of 24–26-nt transposon-derived sRNAs as the main binding partners of PiAgo4 implicates this protein in the repression of TE activity. Endogenous gene regulation is apparently mediated by PiAgo1, PiAgo5 and 20–22-nt sRNAs. Our findings provide an important platform for upcoming functional and mechanistic studies of RNA silencing processes in *P. infestans* and of their impact on host responses to this important plant pathogen.

Acknowledgements

We thank Sarosh Bejai for assistance with confocal microscopy, Steve Whisson for providing the GFP transformant and plasmid, and the National Genomics Infrastructure (NGI)/UPPMAX for assistance in massive parallel sequencing. This work was supported by Formas, the Swedish Research Council (VR), Helge Ax:son-Johnson's Foundation and the Swedish University of Agricultural Sciences. NGI was funded by the Council for Research Infrastructures/VR and the Science for Life Laboratory.

Author contributions

A.K.M.Å. designed and performed the experiments, analyzed and interpreted the data and wrote the manuscript. R.R.V. designed the experiments and contributed to the interpretation of the data. J.F. performed the bioinformatics data analyses. C.D. guided all the experiments and wrote the manuscript.

References

Adl SM, Simpson AG, Lane CE, Lukes J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V *et al.* 2012. The revised classification of eukaryotes. *Journal of Eukaryotic Microbiology* **59**: 429–493.

Ameres SL, Zamore PD. 2013. Diversifying microRNA sequence and function. *Nature Reviews: Molecular Cell Biology* **14**: 475–488.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. *Genome Biology* **11**: R106.

Anders S, Pyl PT, Huber W. 2015. HTSeq – a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.

Åsman AK, Vetukuri RR, Jahan SN, Fogelqvist J, Corcoran P, Avrova AO, Whisson SC, Dixelius C. 2014. Fragmentation of tRNA in *Phytophthora infestans* asexual life cycle stages and during host plant infection. *BMC Microbiology* **14**: 308.

Axtell MJ. 2013. ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA* **19**: 740–751.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate – a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* **57**: 289–300.

Bologna NG, Voinnet O. 2014. The diversity, biogenesis, and activities of endogenous silencing small RNAs in *Arabidopsis*. *Annual Review of Plant Biology* **65**: 473–503.

Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ. 2010. A Dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**: 584–589.

Clark MB, Choudhary A, Smith MA, Taft RJ, Mattick JS. 2013. The dark matter rises: the expanding world of regulatory RNAs. *Essays in Biochemistry* **54**: 1–16.

Cora E, Pandey RR, Xiol J, Taylor J, Sachidanandam R, McCarthy AA, Pillai RS. 2014. The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs. *RNA* **20**: 773–781.

De Riso V, Raniello R, Maumus F, Rogato A, Bowler C, Falciatore A. 2009. Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Research* **37**: e96.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M *et al.* 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**: 491–498.

Derrien B, Genschik P. 2014. When RNA and protein degradation pathways meet. *Frontiers in Plant Science* **5**: 161.

Fahlgren N, Bollmann SR, Kasschau KD, Cuperus JT, Press CM, Sullivan CM, Chapman EJ, Hoyer JS, Gilbert KB, Grunwald NJ *et al.* 2013. *Phytophthora* have distinct endogenous small RNA populations that include short interfering and microRNAs. *PLoS ONE* **8**: e77181.

Fawke S, Doumane M, Schornack S. 2015. Oomycete interactions with plants: infection strategies and resistance principles. *Microbiology and Molecular Biology Reviews* **79**: 263–280.

Frank F, Sonenberg N, Nagar B. 2010. Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **465**: 818–822.

Gardner PP, Daub J, Tate JG, Nawrocki EP, Kolbe DL, Lindgreen S, Wilkinson AC, Finn RD, Griffiths-Jones S, Eddy SR *et al.* 2009. Rfam: updates to the RNA families database. *Nucleic Acids Research* **37**: D136–D140.

Gijzen M, Ishmael C, Shrestha SD. 2014. Epigenetic control of effectors in plant pathogens. *Frontiers in Plant Science* **5**: 638.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**: 644–652.

Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T *et al.* 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* **461**: 393–398.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M *et al.* 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* **8**: 1494–1512.

Hamilton N. 2015. *ggtern: an extension to 'ggplot2', for the creation of ternary diagrams*. R package v.1.0.5.0. [WWW document] URL http://CRAN.R-project.org/package=ggtern. [accessed 25 May 2015].

Jorgensen RA. 1995. Cosuppression, flower color patterns, and metastable gene expression states. *Science* **268**: 686–691.

Judelson HS, Tyler BM, Michelmore RW. 1992. Regulatory sequences for expressing genes in oomycete fungi. *Molecular and General Genetics* **234**: 138–146.

Kamoun S. 2007. Groovy times: filamentous pathogen effectors revealed. *Current Opinion in Plant Biology* **10**: 358–365.

Katoh K, Frith MC. 2012. Adding unaligned sequences into an existing alignment using MAFFT and LAST. *Bioinformatics* **28**: 3144–3146.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* **14**: R36.

Knip M, Constantin ME, Thordal-Christensen H. 2014. Trans-kingdom cross-talk: small RNAs on the move. *PLoS Genetics* **10**: e1004602.

Kozomara A, Griffiths-Jones S. 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research* **42**: D68–D73.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biology* **5**: R12.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**: R25.

Lee HC, Li L, Gu W, Xue Z, Crosthwaite SK, Pertsemlidis A, Lewis ZA, Freitag M, Selker EU, Mello CC *et al.* 2010. Diverse pathways generate microRNA-like RNAs and Dicer-independent small interfering RNAs in fungi. *Molecular Cell* **38**: 803–814.

Li F, Pignatta D, Bendix C, Brunkard JO, Cohn MM, Tung J, Sun H, Kumar P, Baker B. 2012. MicroRNA regulation of plant innate immune receptors. *Proceedings of the National Academy of Sciences, USA* **109**: 1790–1795.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Liu G, Mattick JS, Taft RJ. 2013. A meta-analysis of the genomic and transcriptomic composition of complex life. *Cell Cycle* **12**: 2061–2072.

Mallory A, Vaucheret H. 2009. ARGONAUTE 1 homeostasis invokes the coordinate action of the microRNA and siRNA pathways. *EMBO Reports* **10**: 521–526.

Mallory A, Vaucheret H. 2010. Form, function, and regulation of ARGONAUTE proteins. *Plant Cell* **22**: 3879–3889.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M *et al.* 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**: 1297–1303.

Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, Wu L, Li S, Zhou H, Long C *et al.* 2008. Sorting of small RNAs into *Arabidopsis* Argonaute complexes is directed by the 5′ terminal nucleotide. *Cell* **133**: 116–127.

Morris KV, Mattick JS. 2014. The rise of regulatory RNA. *Nature Reviews Genetics* **15**: 423–437.

Munch D, Teh OK, Malinovsky FG, Liu Q, Vetukuri RR, El Kasmi F, Brodersen P, Hara-Nishimura I, Dangl JL, Petersen M *et al.* 2015. Retromer contributes to immunity-associated cell death in Arabidopsis. *Plant Cell* **27**: 463–479.

Musiyenko A, Majumdar T, Andrews J, Adams B, Barik S. 2012. PRMT1 methylates the single Argonaute of *Toxoplasma gondii* and is important for the recruitment of Tudor nuclease for target RNA cleavage by antisense guide RNA. *Cellular Microbiology* **14**: 882–901.

Nakanishi K, Weinberg DE, Bartel DP, Patel DJ. 2012. Structure of yeast Argonaute with guide RNA. *Nature* **486**: 368–374.

Nozawa M, Miura S, Nei M. 2012. Origins and evolution of microRNA genes in plant species. *Genome Biology and Evolution* **4**: 230–239.

Olovnikov I, Chan K, Sachidanandam R, Newman DK, Aravin AA. 2013. Bacterial Argonaute samples the transcriptome to identify foreign DNA. *Molecular Cell* **51**: 594–605.

Qiao Y, Liu L, Xiong Q, Flores C, Wong J, Shi J, Wang X, Liu X, Xiang Q, Jiang S *et al.* 2013. Oomycete pathogens encode RNA silencing suppressors. *Nature Genetics* **45**: 330–333.

Qiao Y, Shi J, Zhai Y, Hou Y, Ma W. 2015. *Phytophthora* effector targets a novel component of small RNA pathway in plants to promote infection. *Proceedings of the National Academy of Sciences, USA* **112**: 5850–5855.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841–842.

Qutob D, Chapman BP, Gijzen M. 2013. Transgenerational gene silencing causes gain of virulence in a plant pathogen. *Nature Communications* **4**: 1349.

Rogato A, Richard H, Sarazin A, Voss B, Cheminant Navarro S, Champeimont R, Navarro L, Carbone A, Hess WR, Falciatore A. 2014. The diversity of small non-coding RNAs in the diatom *Phaeodactylum tricornutum*. *BMC Genomics* **15**: 698.

Rogers K, Chen X. 2013. Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell* **25**: 2383–2399.

Ruegger S, Grosshans H. 2012. MicroRNA turnover: when, how, and why. *Trends in Biochemical Sciences* **37**: 436–446.

Sarkies P, Miska EA. 2014. Small RNAs break out: the molecular cell biology of mobile small RNAs. *Nature Reviews: Molecular Cell Biology* **15**: 525–535.

Shabalina SA, Koonin EV. 2008. Origins and evolution of eukaryotic RNA interference. *Trends in Ecology & Evolution* **23**: 578–587.

Shi H, Ullu E, Tschudi C. 2004. Function of the Trypanosome Argonaute 1 protein in RNA interference requires the N-terminal RGG domain and arginine 735 in the Piwi domain. *Journal of Biological Chemistry* **279**: 49889–49893.

Shivaprasad PV, Chen HM, Patel K, Bond DM, Santos BA, Baulcombe DC. 2012. A microRNA superfamily regulates nucleotide binding site-leucine-rich repeats and other mRNAs. *Plant Cell* **24**: 859–874.

Smibert P, Yang JS, Azzam G, Liu JL, Lai EC. 2013. Homeostatic control of Argonaute stability by microRNA availability. *Nature Structural & Molecular Biology* **20**: 789–795.

Snedecor GW, Cochran WG. 1980. *Statistical methods*. Ames, IA, USA: Iowa State University Press.

Stam R, Jupe J, Howden AJ, Morris JA, Boevink PC, Hedley PE, Huitema E. 2013. Identification and characterisation of CRN effectors in *Phytophthora capsici* shows modularity and functional diversity. *PLoS ONE* **8**: e59517.

Swarts DC, Hegge JW, Hinojo I, Shiimori M, Ellis MA, Dumrongkulraksa J, Terns RM, Terns MP, van der Oost J. 2015a. Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Research* **43**: 5120–5129.

Swarts DC, Koehorst JJ, Westra ER, Schaap PJ, van der Oost J. 2015b. Effects of Argonaute on gene expression in *Thermus thermophilus*. *PLoS ONE* **10**: e0124880.

Swarts DC, Makarova K, Wang Y, Nakanishi K, Ketting RF, Koonin EV, Patel DJ, van der Oost J. 2014. The evolutionary journey of Argonaute proteins. *Nature Structural & Molecular Biology* **21**: 743–753.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**: 2731–2739.

Tomari Y, Du T, Zamore PD. 2007. Sorting of *Drosophila* small silencing RNAs. *Cell* **130**: 299–308.

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J *et al.* 2013. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics* **11**: 11.10.1–11.10.33.

Vetukuri RR, Åsman AK, Tellgren-Roth C, Jahan SN, Reimegard J, Fogelqvist J, Savenkov E, Soderbom F, Avrova AO, Whisson SC *et al.* 2012. Evidence for small RNAs homologous to effector-encoding genes and transposable elements in the oomycete *Phytophthora infestans*. *PLoS ONE* **7**: e51399.

Vetukuri RR, Avrova AO, Grenville-Briggs LJ, Van West P, Soderbom F, Savenkov EI, Whisson SC, Dixelius C. 2011a. Evidence for involvement of Dicer-like, Argonaute and histone deacetylase proteins in gene silencing in *Phytophthora infestans*. *Molecular Plant Pathology* **12**: 772–785.

Vetukuri RR, Tian Z, Avrova AO, Savenkov EI, Dixelius C, Whisson SC. 2011b. Silencing of the *PiAvr3a* effector-encoding gene from *Phytophthora infestans* by transcriptional fusion to a short interspersed element. *Fungal Biology* **115**: 1225–1233.

Waterhouse PM, Graham MW, Wang MB. 1998. Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proceedings of the National Academy of Sciences, USA* **95**: 13959–13964.

Weiberg A, Wang M, Lin FM, Zhao H, Zhang Z, Kaloshian I, Huang HD, Jin H. 2013. Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* **342**: 118–123.

Winston WM, Molodowitch C, Hunter CP. 2002. Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. *Science* **295**: 2456–2459.

Winston WM, Sutherlin M, Wright AJ, Feinberg EH, Hunter CP. 2007. *Caenorhabditis elegans* SID-2 is required for environmental RNA interference. *Proceedings of the National Academy of Sciences, USA* **104**: 10565–10570.

Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, Gonzalez AJ, Yan Z, Kitto SL, Grusak MA *et al.* 2011. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes & Development* **25**: 2540–2553.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1 Multiple amino acid sequence alignment.

Fig. S2 Complete phylogenetic tree of Argonaute (Ago) proteins.

Fig. S3 Transcript levels of *Phytophthora infestans Argonaute* (*Ago*) genes in *PiAgo-GFP* transgenic lines.

Fig. S4 Cytoplasmic localization of *Phytophthora infestans* Argonaute1 (Ago1) and Ago4.

Fig. S5 Analysis of protein expression with anti-green fluorescent protein (anti-GFP).

Fig. S6 Multiple nucleotide sequence alignment.

Fig. S7 Schematic representation of the old and the revised gene model of *Phytophthora infestans* Argonaute3 (Ago3).

Fig. S8 Relative frequency of each 5′ terminal nucleotide for different small RNA (sRNA) size classes in *Phytophthora infestans*.

Fig. S9 Genomic origins of *Phytophthora infestans* small RNAs (sRNAs) including transposon class information.

Fig. S10 Detection of low levels of small RNAs (sRNAs) from transposable elements.

Fig. S11 Small RNAs (sRNAs) from particular *Crinkler* (*CRN*) genes associate preferentially with PiAgo1 and PiAgo5.

Fig. S12 Small RNAs (sRNAs) from particular *Crinkler* (*CRN*) pseudogenes associate preferentially with PiAgo1.

Fig. S13 Small RNAs (sRNAs) from *RxLR* genes do not show any PiAgo preference.

Fig. S14 Small RNAs (sRNAs) from protein-coding genes associate preferentially with PiAgo1 and PiAgo5.

Fig. S15 Size distribution of small RNAs (sRNAs) mapping to *Phytophthora infestans* Argonaute1 (Ago1) and to green fluorescent protein (GFP).

Table S1 Sequences of oligonucleotides used for PiAgo-green fluorescent protein (GFP) construct cloning and production of riboprobes

Table S2 Read mapping data in small RNA (sRNA) sequence samples from *Phytophthora infestans*

Table S3 Accession numbers of proteins used for phylogenetic reconstruction, identity of small RNA (sRNA) 5′ terminal nucleotide, proportions of reads mapping to different genomic features and sRNA read directionality

Table S4 Fold-change small RNA (sRNA) expression per *Crinkler* (*CRN*)- and pseudo-*CRN*-encoding gene in the PiAgo1, PiAgo4 and PiAgo5 immunoprecipitation (IP) samples relative to the green fluorescent protein (GFP) IP

Table S5 Fold-change small RNA (sRNA) expression per *RxLR*-encoding gene and non-effector protein-coding gene in the PiAgo1, PiAgo4 and PiAgo5 immunoprecipitation (IP) samples relative to the green fluorescent protein (GFP) IP

Table S6 Summary of *pin*-miR8788 mapping data in the different small RNA (sRNA) sequencing samples

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.