# The stem rust fungus *Puccinia graminis* f. sp. *tritici* induces waves of small RNAs with opposing profiles during wheat infection

Jana Sperschneider[1,2#], Silke Jacques[3], Bo Xu[4], Narayana M. Upadhyaya[2], Rohit Mago[2], Karam B. Singh[3,5], Eric A. Stone[1], Ming-Bo Wang[2], [Peter N. Dodds[2#], Jennifer M. Taylor[2]][joint]

[1]Centre for Genomics, Metabolomics and Bioinformatics, The Australian National University, Canberra, ACT 2601, Australia

[2]Black Mountain Science and Innovation Park, CSIRO Agriculture and Food, Canberra, ACT 2601, Australia

[3]Centre for Environment and Life Sciences, CSIRO Agriculture and Food, Perth, WA 6014, Australia

[4]Thermo Fisher Scientific, 5 Caribbean Drive, Scoresby, VIC 3179, Australia

[5]Centre for Crop and Disease Management, Department of Environment and Agriculture, Curtin University, Bentley, WA 6102, Australia

Jana Sperschneider: jana.sperschneider@anu.edu.au
Silke Jacques: silke.jacques@csiro.au
Bo Xu: xbxubo@icloud.com
Narayana M. Upadhyaya: narayana.upadhyaya@csiro.au
Rohit Mago: rohit.mago@csiro.au
Karam B. Singh: karam.singh@csiro.au
Eric A. Stone: eric.stone@anu.edu.au
Ming-Bo Wang: ming-bo.wang@csiro.au
Peter N. Dodds: peter.dodds@csiro.au
Jennifer M. Taylor: jen.taylor@csiro.au

# corresponding authors:

Jana Sperschneider, jana.sperschneider@anu.edu.au
Peter N. Dodds: peter.dodds@csiro.au

## Abstract

### Background

The fungus *Puccinia graminis* f. sp. *tritici* (*Pgt*) causes devastating stem rust disease on wheat. Infection occurs when rust spores germinate on the leaf surface and subsequently, specialized infection structures called haustoria form inside host cells. Small RNAs (sRNAs) are critical regulators of gene expression and can play a significant role in plant-pathogen crosstalk. How *Pgt* utilizes sRNAs during infection is thus far unknown.

### Results

Small RNA and transcriptome sequencing during *Pgt*-wheat infection reveals that the *Pgt* RNA interference (RNAi) machinery has functionally diversified. A number of *Pgt* RNAi genes are strongly up-regulated during late infection and in haustorial tissue and this coincides with the production of distinct *Pgt* sRNA profiles during infection. Strikingly, over 90% of *Pgt* sRNAs are differentially expressed during infection, compared to only 2% of wheat sRNAs. An early wave of *Pgt* sRNAs expressed during infection constitutes of 21 nt sRNAs with a 5‘ uracil derived from genes, whereas the late wave of *Pgt* sRNAs encompasses 22 nt sRNAs with a 5‘ adenine derived from repetitive regions. The late wave of *Pgt* sRNAs appears to target largely transposable elements and genes located in close proximity to TEs targeted by sRNAs exhibit reduced average expression, resembling the epigenetic silencing mechanism in plants.

### Conclusion

We conclude that the *Pgt* RNAi machinery has functionally diversified and is highly regulated, resulting in differential accumulation of *Pgt* sRNAs in germinated spores and specialized-infection tissues. Such tight temporal control of the RNAi machinery has thus far not been observed in fungi.

**Keywords:** small RNA, miRNA, stem rust, *Puccinia graminis* f. sp. *tritici*, fungal pathogen, wheat, plant defence

## Background

The basidiomycete fungus *Puccinia graminis* f. sp. *tritici* (*Pgt*) is a plant pathogen that causes wheat stem rust, resulting in devastating crop losses [1]. During the asexual infection phase on a cereal host, *Pgt* produces single-celled dikaryotic urediniospores that germinate on the leaf surface [2, 3]. Subsequently, appressoria form and penetration through stomata occurs with development of specialized infection structures called haustoria by around 2 days. Haustoria enable nutrient uptake as well as the delivery of secreted pathogen proteins called effectors into the host plant cell [4]. The start of urediniospore production occurs approximately at 6-7 days and urediniospore pustules typically erupt through the leaf or stem surface (sporulation) after 8–10 days [2]. Whilst genomic and transcriptomic resources are available for *Pgt* [5-8], how it utilizes its RNAi machinery during infection is thus far unknown.

Small RNAs (sRNAs) are 20-30 nucleotide (nt) short regulatory non-coding RNAs that function in transcriptional or posttranscriptional gene silencing through sequence complementarity. In plants, sRNAs are predominantly in the size range of 20-24 nt and can be divided into two classes: small interfering RNAs (siRNAs) processed by Dicer proteins from long double-stranded RNA (dsRNA) and microRNAs (miRNAs) processed from stem-loop regions of single-stranded primary RNAs. Endogenous dsRNA is generated by RNA-dependent RNA polymerases (RdRPs) from single-stranded RNA (ssRNA), whereas self-folding miRNA precursors are transcribed by RNA Polymerase II from MIRNA genes. Both miRNAs and siRNAs are bound to argonaute (AGO) proteins to induce silencing of targets by base-pairing interactions and complementarity. Post-transcriptional gene silencing (PTGS) is induced by sRNAs in the cytoplasm, which target complementary mRNAs for degradation or translational repression. In plants, successful identification of a target by miRNA/AGO complexes and some siRNA/AGO complexes induces PTGS and can trigger secondary siRNA production. In contrast, transcriptional gene silencing (TGS) is induced by nucleus-localized sRNAs through epigenetic modifications, such as DNA cytosine methylation and histone methylation, to homologous regions of the genome [9]. In plants, these nucleus-localized heterochromatic siRNAs (hc-siRNAs) are the most abundant sRNAs. They are predominantly 24 nts in length, derived from intergenic or repetitive regions and are associated with the AGO4 clade to regulate epigenetic silencing through RNA-directed DNA methylation (RdDM). Adenine is the most common 5′ base of AGO4-bound 24 nt siRNAs in *Arabidopsis* [10]. In plants, RdDM is implicated in maintaining genome stability through transposon control, pathogen defence and stress responses, intercellular communication and germ cell specification [9].

In the diverse fungal kingdom, the RNAi machinery of the fission yeast *Schizosaccharomyces pombe* and the ascomycete fungus *Neurospora crassa* are thus far the best-studied [11]. Quelling is an RNAi-related gene-silencing mechanism in *Neurospora* that is induced by repetitive transgenic sequences and occurs in the vegetative growth stage to control transposons [12]. In *Neurospora*, meiotic silencing by unpaired DNA (MSUD) occurs during meiosis [11]. In *Schizosaccharomyces pombe*, RNAi components are required for heterochromatin formation [13]. The roles of sRNAs in eukaryotic plant pathogens have thus far not been extensively characterized [14]. In *Phytophthora* spp., sRNAs are putatively involved in effector gene and transposable element (TE) regulation and are predominantly of the size classes of 21 nt, 25-26 nt and 32 nt [15]. Many *Phytophthora* sRNAs of all size classes map to TEs, particularly to long terminal repeat (LTR) retrotransposons. Another class of sRNAs mapped to Crinkler effector genes and were predominantly of the 21 nt size class. In *Magnaporthe oryzae*, 18-23 nt sRNAs are produced from repetitive elements and are implicated in TE regulation in vegetative tissue, whereas 28-35 nt sRNAs mapping to transfer RNA (tRNA) loci are enriched in the appressoria [16]. Several cross-kingdom RNAi events between fungal pathogens and plants have been uncovered. Some *Botrytis cinerea* sRNAs silence *Arabidopsis* and tomato genes involved in plant immunity and are mainly derived from LTR retrotransposons and are 21 nt in size with a 5’ uracil [17], while *Arabidopsis* cells secrete exosome-like extracellular vesicles to deliver sRNAs into the fungal pathogen *Botrytis cinerea* to silence pathogenicity genes [18]. The wheat stripe rust fungus *Puccinia striiformis* f. sp. *tritici* (*Pst*) produces a large number of 20-22 nt sRNAs and expresses RNAi genes during infection [19]. One 20 nt sRNA appears to target the wheat defence pathogenesis-related 2 (*PR2*) gene [20]. The production of sRNAs and their potential roles in *Pgt* development and pathogenicity has thus far not been investigated.

## Results

### The distinct expression profiles of *Pgt* RNAi genes suggest their functional diversification

RNAi machinery genes (two argonautes, five dicers and five RdRPs) were previously identified in the reference genome *Pgt* p7a [7, 21] and orthologs of these genes are present in *Pgt* 21-0 (Table 1). We assessed the expression of these RNAi genes during a time course of *Pgt* 21-0 infecting wheat from 2-7 days post infection (dpi) [6] and in germinated spores and haustorial tissue [8]. Clustering of their expression profiles shows two main groups: one set of RNAi genes that appear to be constitutively highly expressed during infection and another set of RNAi genes that appear highly expressed only during the later stages of infection (Figure 1A). Argonaute A, dicer C and RdRP C in

particular are highly expressed at 5-7 dpi (Figure 1A), indicating a potential role in developmental or infection-related processes.

We then assessed the up-regulation of *Pgt* RNAi genes, comparing the infection stages to germinated spores (Figure 1B). RdRPs A and C, argonautes A and B as well as dicer C are up-regulated (*log2* fold change >= 1) at the later stages of infection (5-7 dpi). Particularly argonaute A and RdRP A are highly up-regulated in haustorial tissue (*log2* fold change > 3). Overall, absolute expression of argonaute B appears higher than for argonaute A. Argonaute B has high expression in germinated spores and during all stages of infection, whereas argonaute A is expressed at lower levels during early infection and in germinated spores (Figure 1A). The expression patterns of the two *Pgt* argonautes are strikingly different and hint towards their functional diversification. The phylogenetic tree for argonaute A, which is expressed at higher levels during late infection and in haustoria reveals significant sequence similarity to the argonautes of pathogenic oomycetes (Supplementary Figure S1). This was not observed for argonaute B (Supplementary Figure S2). Taken together, the gene expression and phylogenetic analysis suggests that *Pgt* uses its RNAi machinery to regulate stage-specific infection processes, which could encompass haustorial development or the formation of new urediniospores during late infection.

**Table 1:** RNAi genes in *Pgt*.

| RNAi genes | *Pgt* p7a | *Pgt* 21-0 | Identifier |
|---|---|---|---|
| **Argonautes** | PGTG_08429 | HSGS210\|asmbl_13027\|m.9230 | Argonaute A |
| | PGTG_11327 | HSGS210\|asmbl_19441\|m.33522 | Argonaute B |
| **Dicer** | PGTG_19535 | HSGS210\|asmbl_38227\|m.101748 | Dicer A |
| | PGTG_19538 | HSGS210\|asmbl_38232\|m.101783 | Dicer B |
| | PGTG_12289 | HSGS210\|asmbl_22132\|m.43860 | Dicer C |
| | PGTG_13081 | HSGS210\|asmbl_22868\|m.46534 | Dicer D |
| | PGTG_13088 | HSGS210\|asmbl_22874\|m.46597 | Dicer E |
| **RdRPs** | PGTG_20838 | HSGS210\|comp16198_c0_seq2\|m.183295 | RdRP A |
| | PGTG_17766 | HSGS210\|asmbl_34494\|m.89212 | RdRP B |
| | PGTG_02834 | HSGS210\|asmbl_36553\|m.96284 | RdRP C |
| | PGTG_05092 | HSGS210\|asmbl_1600\|m.20443 | RdRP D |
| | PGTG_09533 | HSGS210\|asmbl_16348\|m.22057 | RdRP E |

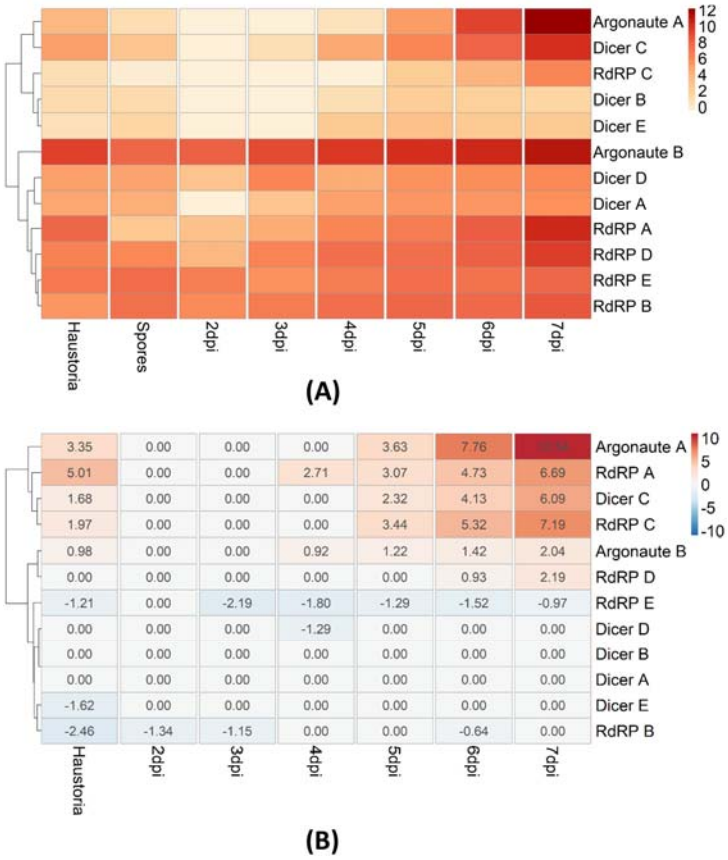


**Figure 1: *Pgt* 21-0 RNAi gene expression. (A)** Hierarchical clustering of expression levels of *Pgt* RNAi genes in *log* of counts per million (*logCPM*, red color intensity relates to high expression strength). **(B)** Up- or down-regulation of RNAi genes in *log2* fold changes (FC) with FDR < 0.05 relative to germinated spores are shown. A value of zero indicates that no significant up- or down-regulation relative to germinated spores could be observed. In particular, RdRP A, RdRP C, argonaute A, argonaute B and dicer C are up-regulated in haustoria (FC >= 2) and at the later stages of infection (5-7 dpi).

### *Pgt* and wheat produce distinct sRNA profiles during infection

To assess the role of the RNAi machinery during rust infection, we performed small RNA-sequencing on germinated spores, uninfected wheat and infected wheat at 3 dpi, 5 dpi and 7 dpi. Adapter-trimmed and tRNA/rRNA-filtered clean sRNA reads were first mapped to the wheat genome (IWGSC RefSeq v1.0) and the *Pgt* 21-0 genome. Strikingly, the read alignment rates show a strong presence of *Pgt* sRNAs in the 7 dpi sample (Table 2). The mapping rates to rust at 3 dpi and 5 dpi are low at 3.5% and 5.2%, respectively, but increase drastically to 45.4% at 7 dpi. In contrast, 64.1% of sRNA sequences map to the wheat genome in the uninfected wheat samples and sRNA mapping rates to wheat decrease to 28.2% at 7 dpi. The rust-derived sRNAs mostly map to unique positions (~70% for germinated spores), with fewer reads mapping to multiple positions (~30% multi-mapped for germinated spores). On the other hand, most wheat-derived sRNAs map to multiple positions (eg. ~75% for uninfected wheat), which likely relates to the polyploidy in wheat as well as potentially the mapping preferably to repetitive regions.

**Table 2:** Small RNA read mapping rates to the wheat and rust genomes.

| Sample | % mapped to *Pgt* | % uniquely mapped to *Pgt* | % multi-mapped to *Pgt* | % mapped to wheat | % uniquely mapped to wheat | % multi-mapped to wheat |
|---|---|---|---|---|---|---|
| **Germinated spores** | 51.5% | 35.7% | 15.8% | - | - | - |
| **Uninfected wheat** | - | - | - | 64.1% | 16.3% | 47.8% |
| **Infected wheat 3dpi** | 3.5% | 2.9% | 0.6% | 63.5% | 13% | 50.4% |
| **Infected wheat 5dpi** | 5.2% | 4% | 1.3% | 61.4% | 17.1% | 44.3% |
| **Infected wheat 7dpi** | 45.4% | 25% | 20.4% | 28.2% | 7.7% | 20.6% |

To annotate high-confidence *Pgt* and wheat sRNAs from the sequencing data, we used the ShortStack software [22]. ShortStack predicts and quantifies sRNA-producing loci in a genome based on clusters of sRNA reads and miRNA-producing loci according to a series of tests, such as strandedness of the locus and the predicted precursor secondary structure. A total of 6,362 *Pgt* sRNA clusters (6,352 siRNAs and 10 miRNAs) and 1,352 wheat sRNA clusters (1,290 siRNAs and 62 miRNAs) were predicted using ShortStack (Table 3, Supplementary Files S1-S4). The read length distributions of rust and wheat sRNAs show characteristic peaks and deviate from a random distribution (Figure 2). As expected for plant sRNAs, there are two peaks at 21 nt and 24 nt for the wheat sRNAs. However, the *Pgt* sRNAs show a strong peak in sequence length distribution at 20-22 nts, with very low relative abundance of 24 nt sRNAs. Overall, 63.5% of *Pgt* sRNAs have a preference for a 5' uracil, compared to only 28% of the wheat sRNAs. In particular, the 24 nt wheat sRNAs have a preference for a 5' adenine (45.2%), therefore likely representing siRNAs involved in RdDM [10]. The majority of *Pgt* sRNA clusters are produced from non-stranded loci (90.9%). Only 10 of the stranded sRNA loci pass the ShortStack criteria for miRNA loci that are optimized for properties of plant miRNAs, and eight of the predicted *Pgt* miRNAs (80%) are 22 nt in length. This indicates that *Pgt* produces some miRNA-like sequences that are predominantly 22 nt in length. Similar observations have also been made in the stripe rust fungus *Pst* [19]. Of the ten *Pgt* miRNAs, 50% have a 5' adenine and 50% have a 5' uracil. None of the predicted *Pgt* miRNA-like sequences have an exact match to a known miRNA in the RNACentral database [23], compared to 37 of the 62 predicted wheat miRNAs (59.7%). The remaining stranded sRNA loci in *Pgt* (57.7%) are not annotated as miRNAs mainly because the predicted precursor secondary structures do not form a stable stem-loops, with more than 5 unpaired bases in the sRNA region.

**Table 3:** Predicted sRNAs in *Pgt* and wheat and their properties.

| | *Pgt* siRNAs | *Pgt* miRNAs | Wheat siRNAs | Wheat miRNAs |
|---|---|---|---|---|
| **# of sRNAs** | 6,352 | 10 | 1,290 | 62 |
| **Average length of sRNA loci (nts)** | 867 | 357 | 293 | 144 |
| **5' adenine** | 31.5% | 50% | 37.7% | 11.3% |
| **5' cytosine** | 4.1% | 0.0% | 17.6% | 4.8% |
| **5' guanine** | 0.9% | 0.0% | 18.9% | 9.7% |
| **5' uracil** | 63.5% | 50% | 25.8% | 74.2% |

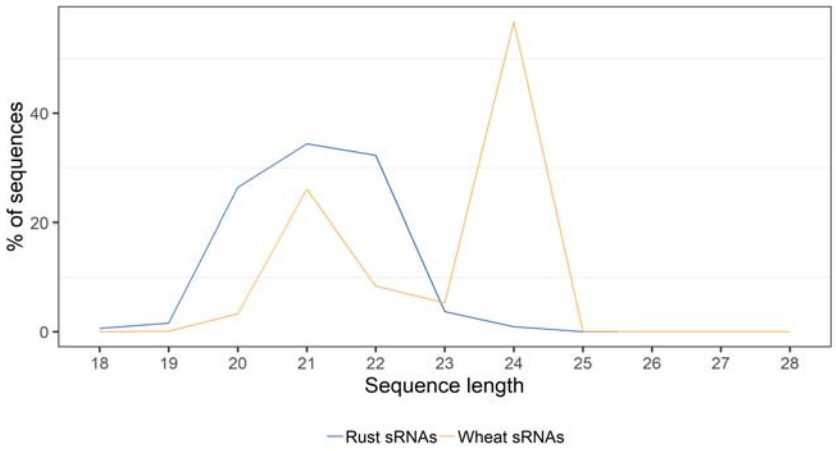



**Figure 2:** Sequence length distributions of predicted sRNAs in rust and wheat. The plant sRNAs show strong peaks at 21 nt and 24 nt, whereas the rust sRNAs are predominantly 20-22 nt in length.

### *Pgt* induces early and late waves of sRNAs during infection, whereas most wheat sRNAs are not differentially expressed

We assessed the differential expression of *Pgt* sRNA clusters in germinated spores, during early infection when haustoria are present inside the plant cells (3 dpi and 5 dpi) and during late infection when haustoria are present and sporulation starts (7 dpi) using the sRNA cluster read counts and edgeR [24]. We detected no differential expression of *Pgt* sRNAs between 3 dpi and 5 dpi and thus these two time points were together considered as early infection (Figure 3A). Strikingly, 5,868 of the 6,362 *Pgt* sRNA clusters are predicted as differentially expressed (92.2%), with 2,424 up-regulated in germinated spores (38.1%), 497 up-regulated during early infection (7.8%) and 3,553 up-regulated during late infection (55.9%) (Supplementary Files S5-S8). sRNAs with the following two expression patterns are over-represented, namely up-regulation in late infection compared to germinated spores and early infection ($n$ = 3,191) and down-regulation in late infection compared to germinated spores ($n$ = 1,737). Nine of the ten predicted *Pgt* miRNAs are down-regulated during late infection and up-regulated in germinated spores. Of these nine *Pgt* miRNAs, three are up-regulated during early infection, four are down-regulated during early infection and two are not differentially expressed in early infection. This indicates that a major switch in *Pgt* sRNA expression occurs between the germinated spore stage and late infection, coinciding with the differential expression of several RNAi genes (Figure 1). In contrast to *Pgt*, which exhibits prominent waves of sRNA expression during infection, only 26 of the 1,352 wheat sRNAs (1.9%) are predicted to be differentially expressed. None of the differentially expressed wheat sRNAs have an exact match to a known miRNA in the RNACentral database.

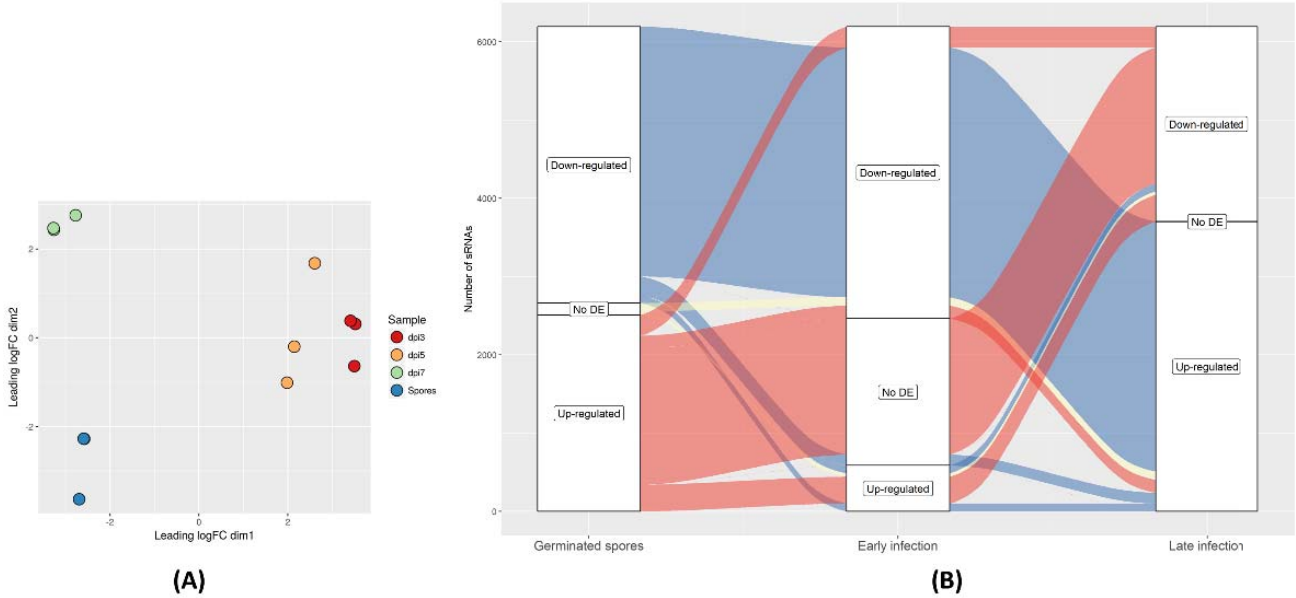


**Figure 3:** *Pgt* sRNAs expression patterns in germinated spores, early infection and late infection. **(A)** A multi-dimensional scaling plot of the sRNA samples in edgeR represents the relationship between groups of samples and shows a clear separation between the different infection timepoints. **(B)** Two major waves of sRNA expression occur in *Pgt* and these predominantly involve sRNAs differentially expressed when comparing germinated spores to late infection.

### The late wave of *Pgt* sRNAs expressed during infection is predominantly 22 nt in length with a 5' adenine

We assessed the length distributions and 5' nucleotide preferences of differentially expressed *Pgt* sRNAs (Figure 4A). *Pgt* sRNAs up-regulated during early infection or in germinated spores show a preference for sRNAs of 21 nt in length (41.2% and 45.8%, respectively), whereas sRNAs up-regulated during late infection are predominantly 22 nt in length (41%). *Pgt* sRNAs with no detected differential expression follow a similar size distribution pattern to those that are up-regulated in germinated spores, with 21 nt sRNAs being the most prevalent class (41.1%, Figure 4A). The majority of the 20-22 nt sRNAs up-regulated in germinated spores, in early infection and those with no differential expression contain a 5' uracil (Figure 4B). In contrast, 22 nt sRNAs that are up-regulated during late infection have a strong preference for 5' adenines (68.2%). In plants, the 5' nucleotide of sRNA has a strong effect on preferential loading into different argonautes. In *Arabidopsis thaliana*, AGO1 and AGO10 bind preferentially small RNAs with a 5′ uracil, whereas AGO2, AGO4, AGO6, AGO7 and AGO9 prefer sRNAs with 5‘ adenines and AGO5 5‘ cytosines [25]. The high expression of argonaute B and relatively low expression of argonaute A in germinated spores and during early infection suggests that argonaute B preferentially loads sRNAs with a 5‘ uracil. In contrast, argonaute A is up-regulated during late infection and so may be involved in loading of 22 nt sRNAs with a 5‘ adenine.

**Figure 4: (A)** Sequence lengths and **(B)** 5‘ nucleotide distribution of up-regulated *Pgt* sRNAs and of *Pgt* sRNAs with no detected differential expression.

### Waves of early and late infection *Pgt* sRNAs exhibit opposing profiles in their genomic origin

We assessed the mapping of the predicted *Pgt* sRNA clusters to the genome and found that the waves of early and late infection *Pgt* sRNAs exhibit opposing profiles in their genomic origin (Table 4). *Pgt* sRNAs up-regulated in germinated spores predominantly map to genes (67.2%), compared to only 5.1% of sRNAs up-regulated during late infection. *Pgt* sRNAs induced during late infection are largely generated from repetitive elements (72.2%, Table 4). Repetitive elements associated with sRNAs are predominantly LTR retrotransposons, DNA transposons and unknown repeats. Notably, over a third of *Pgt* sRNAs up-regulated during late infection map to DNA transposons, compared to only a fifth of *Pgt* sRNAs up-regulated in germinated spores (Table 4).

ShortStack also conducts phasing analysis for unstranded loci where > 80% of reads are either 21 nt or 24 nt in length. We observed that 14.1% of *Pgt* sRNA clusters induced during early infection and 5.7% of *Pgt* sRNA clusters induced in germinated spores exhibit a high phase score and thus resemble phasiRNA-like sRNAs. In plants, phasiRNAs are produced in a one-hit or two-hit model, where mRNA targets are cleaved by a 22 nt miRNA [25]. This subsequently triggers dsRNA production and processing up- or down-stream of the cleaved target site into phased 21-mers by dicer activity. PhasiRNA-like sRNAs have thus far not been described in fungi.

**Table 4: Genomic origins of *Pgt* sRNAs**.

| | **Up-regulated in germinated spores** | **Up-regulated during early infection** | **Up-regulated during late infection** | **No differential expression** |
|---|---|---|---|---|
| # of sRNAs | 2,424 | 497 | 3,553 | 494 |
| Average length of sRNA loci (nts) | 918 | 1,090 | 916 | 588 |
| Strand call for loci | | | | |
| + genomic strand | 5.2% | 7.4% | 2.8% | 12.2% |
| - genomic strand | 6.4% | 6% | 2.4% | 10.7% |
| no strand | 88.4% | 86.5% | 94.9% | 77.1% |
| % phased loci (phase score > 30.0) | 5.7% | 14.1% | 0.9% | 3.4% |
| Mapping to repeats | 25.7% | 24.7% | 72.2% | 56.1% |
| Mapping to genes | 67.2% | 59.8% | 5.1% | 40.7% |
| **Genomic origin of sRNAs from repetitive regions** | | | | |
| LTR retrotransposons | 37.8% | 31.4% | 37.9% | 33.5% |
| DNA transposons | 20.6% | 24.3% | 36% | 32.1% |
| Unknown repeat | 21.4% | 25.4% | 21.8% | 23.7% |
| LINEs | 3.6% | 6.5% | 1.5% | 5.3% |

A GO term analysis of the 1,651 genes that produce *Pgt* sRNAs up-regulated in germinated spores reveals an enrichment in proteins with serine/threonine kinase and endopeptidase activity as well as proteins with ATP and chromatin binding. Genes that produce *Pgt* sRNAs up-regulated during early infection are also enriched in proteins with a serine/threonine kinase annotation. In particular, we observed several members of the calcium/calmodulin-dependent kinase (CaMK) family, the fungal specific FunK1 family, Cell division cycle 7 (CDC7) protein kinase family and AGC and STE protein kinases. The abundance of proteins related to kinase activity might indicate a potential role of the early wave of *Pgt* sRNAs in regulation of signal transduction and virulence during early infection. This is not observed in genes that produce *Pgt* sRNAs with no differential expression, which are enriched for proteins involved in zinc ion binding, mismatched DNA binding and serine O-acetyltransferase activity. Taken together, the distinct genomic origins of differentially expressed *Pgt* sRNA clusters further underlines their involvement in stage-specific regulatory processes.

**Table 5:** *Pgt* genes that produce sRNAs up-regulated in germinated spores and their functional GO term enrichment (molecular function) with FDR < 0.001 compared to all *Pgt* genes.

| **Enriched GO term category** | **False discovery rate (FDR)** | **# of proteins** |
|---|---|---|
| **Genes that produce smRNAs up-regulated in germinated spores** | | |
| Protein serine/threonine kinase activity | $1.6 \times 10^{-13}$ | 59 |
| ATP binding | $4.9 \times 10^{-10}$ | 151 |
| Aspartic-type endopeptidase activity | $1.5 \times 10^{-8}$ | 27 |
| Chromatin binding | $9.8 \times 10^{-5}$ | 23 |
| DNA-directed DNA polymerase activity | $2 \times 10^{-4}$ | 13 |
| Single-stranded DNA binding | $3.3 \times 10^{-4}$ | 20 |
| **Genes that produce smRNAs up-regulated during early infection** | | |
| Protein serine/threonine kinase activity | $5.6 \times 10^{-4}$ | 19 |
| **Genes that produce smRNAs with no differential expression** | | |
| Zinc ion binding | $8.5 \times 10^{-8}$ | 37 |
| Mismatched DNA binding | $1.2 \times 10^{-5}$ | 7 |
| Serine O-acetyltransferase activity | $2.2 \times 10^{-4}$ | 3 |

### TEs targeted by sRNAs are associated with reduced expression of genes in close proximity

To assess the correlation between sRNAs, their mapping to TEs and the expression of nearby genes, we mapped *Pgt* sRNAs that are up-regulated during late infection without mismatches to the *Pgt* 21-0 genome. We labelled a TE as sRNA+TE if a sRNA mapped to it and as sRNA-TE if no sRNA mapped to it. To investigate the relationship between

TE proximity and gene expression, we measured the distance from a gene to its nearest neighboring TE, including both upstream and downstream TEs. This included both gene models that overlap with a TE as well as gene models that are either up- or downstream of a TE. We then separated genes into two groups, one group contained genes with the closest TE being a sRNA+TE and the other group contained genes with the closest TE being a sRNA-TE. We then assessed the average gene expression levels at late infection (7 dpi) and compared the different groups (Figure 5). When a TE was overlapping with or in close proximity to a gene (< 200 bp), then genes with a sRNA+TE were expressed at significantly lower levels, on average, than genes with a sRNA-TE. Genes that were more than 200 bp away from a sRNA+TE were, on average, not expressed at significantly lower levels than genes more than 200 bp away from a sRNA-TE. We then repeated this experiment with *Pgt* sRNAs that are up-regulated in germinated spores and assessed the average gene expression levels in germinated spores. We observed a significant effect on the expression of nearby genes only up to a distance of ∼100 bp (Figure 5B). Overall, we observed lower expression of the genes that are likely to be targeted by sRNAs during late infection than for the genes targeted by sRNAs in germinated spores. This suggests that TE silencing and the reduced expression of nearby genes during late infection is much more pronounced during late infection than in germinated spores.

During late infection, the proximity to a sRNA+TE or a sRNA−TE has a significant effect on the expression of nearby genes but only up to a distance of ∼200 bp. The most abundant repeat type amongst the sRNA+TEs is the family of Gypsy LTR retrotransposons (44.4%). We then investigated all genes that have sRNA+TE as their closest TE to see if they are enriched for effectors. However, only 6.5% of genes with a sRNA+TE as their closest TE encode secreted proteins and the AvrSr50 effector has a sRNA-TE in close proximity (< 100 bp). Therefore, regulation of TEs and their neighbouring genes is likely unrelated to effector expression.

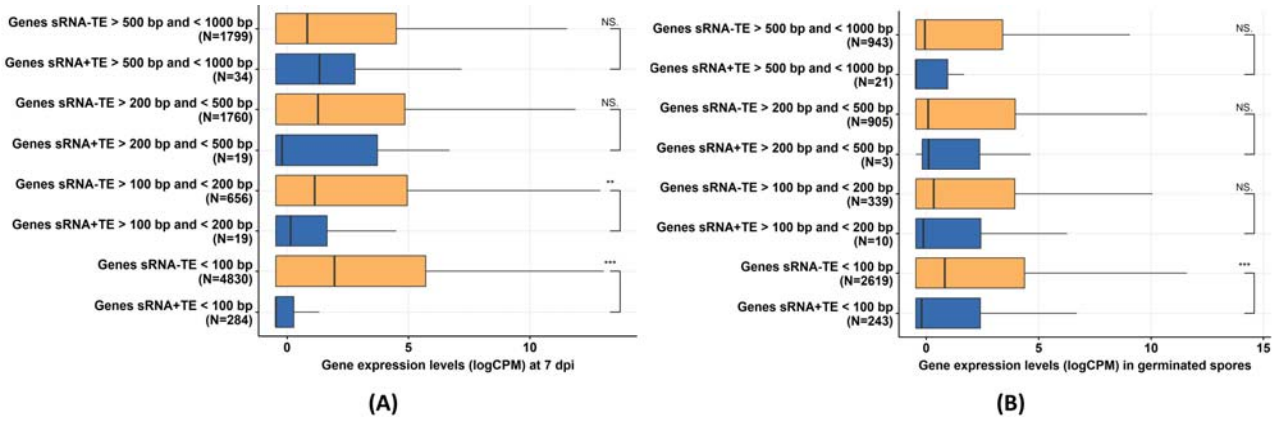



**Figure 5:** Expression levels (*log*-normalized read counts) for genes in each bin of increasing distance from the nearest TE that is predicted to be either targeted by a *Pgt* sRNA **(A)** induced at 7dpi (sRNA+TE) or not (sRNA-TE) or **(B)** induced in germinated spores. A distance of zero indicates genes that overlap with TEs. The total number of observations in each group (N) is shown on the y-axis.

### Predicted targets of *Pgt* sRNAs include secreted proteins and proteins involved in pathogenicity

We predicted putative target genes of differentially expressed *Pgt* sRNA candidates of length 20-22 nts with a 5' uracil in both wheat and rust using psRNATarget [26]. To reduce the high false positive rate of sRNA target prediction, we assessed the gene expression of the predicted targets in both *Pgt* and wheat using RNA-sequencing data from the same infection time course [6] as well as from haustorial tissue and germinated spores [8]. We kept a predicted target gene if it is down-regulated in the same conditions where the corresponding sRNA is up-regulated. For example, for *Pgt* sRNAs that are up-regulated in germinated spores and in early infection compared to late infection, we kept predicted rust target genes if they are down-regulated in germinated spores and in early infection compared to late infection and we kept predicted wheat targets if they are down-regulated at 2dpi (as the germinated spore stage roughly overlaps with 1-2dpi) and in early infection compared to late infection. We dismissed sRNAs that have predicted down-regulated targets in both rust and wheat as ambiguous. Whilst this is a stringent filtering procedure that might miss true targets, it will reduce false positive target predictions.

We predicted a total of 542 unique rust targets with a down-regulated expression profile (Supplementary Files S9-S14). 24.8% of rust genes predicted to be silenced by sRNAs in germinated spores encode secreted proteins (Table 6). Whilst these genes might encode for haustorially-delivered rust effectors, we did not observe a significant enrichment for predicted effector candidates using EffectorP 2.0 [27]. However, we observed putative sRNA-targeting of the rust effector AvrSr35 [5], which is not predicted as an effector by EffectorP possibly due to its larger size. Additional *Pgt*

effectors might be regulated by sRNAs, but are not predicted as effectors by EffectorP, or are regulated by sRNAs with relatively low abundance in our sequencing data. Alternatively, *Pgt* sRNAs up-regulated in germinated spores might predominantly silence rust genes encoding secreted proteins that are involved in other haustorially-related functions or signalling. The highest proportion of secreted proteins (38.3%) were found amongst the 47 rust targets predicted to be down-regulated in late infection versus early infection. We found that these 18 secreted proteins are mostly without functional annotation, except for two glycoside hydrolase family proteins, a mannan endo-1,6-alpha-mannosidase, a thioredoxin and a copper/zinc superoxide dismutase protein. 27.8% of these secreted proteins are predicted as effectors by EffectorP 2.0, suggesting that a subset might function in virulence.

**Table 6: Predicted rust target genes of *Pgt* sRNAs with a 5' uracil**.

| ***Pgt* sRNA expression** | **# of predicted down-regulated *Pgt* target genes** | **# of secreted proteins** | **# secreted proteins that are predicted effectors** |
|---|---|---|---|
| Up-regulated in spores vs. early infection (3 or 5 dpi) | 209 | 54 (25.8%) | 7 (13%) |
| Up-regulated in spores vs. late infection (7 dpi) | 331 | 80 (24.2%) | 11 (13.8%) |
| Up-regulated in early infection vs. spores | 75 | 6 (8%) | 0 (0%) |
| Up-regulated in early infection vs. late infection | 101 | 20 (19.8%) | 3 (15%) |
| Up-regulated in late infection vs. spores | 118 | 22 (18.6%) | 5 (22.7%) |
| Up-regulated in late infection vs. early infection | 47 | **18 (38.3%)** | **5 (27.8%)** |

We further examined the annotations of rust genes that are targeted by differentially expressed *Pgt* sRNAs with high expression levels (Figure 6). The most abundant *Pgt* sRNA (22 nts) up-regulated in germinated spores is predicted to down-regulate a homolog of the pheromone response factor Prf1 (HSGS210|asmbl_9894|m.145435). The pheromone response factor Prf1 coordinates filamentous growth and pathogenicity in *Ustilago maydis* [28]. Another target of highly abundant *Pgt* sRNAs is a bud-site selection protein (HSGS210|asmbl_2532|m.55818). There is evidence that some fungal pathogens use sRNAs to suppress plant defence responses during infection [17]. We therefore looked for *Pgt* sRNAs that may have a down-regulated target in wheat, but not in rust. Again, we only considered 20-22 nt sRNAs with a 5' uracil. We predicted a total of 3,884 targets in wheat with a down-regulated expression profile (Supplementary Files S9-S14). The most abundant *Pgt* sRNA with a predicted wheat target (TraesCS2D02G574400, putative disease resistance RPP13-like protein 3) is up-regulated in germinated spores and 20 nt in length. The most abundant *Pgt* sRNA up-regulated during early infection with a predicted wheat target is 21 nt in length, maps to a *Pgt* intergenic region and appears to target a receptor-like protein kinase HSL1 (TraesCS1B02G355700) and a homolog of the BAG family molecular chaperone regulator 6 (BAG6, TraesCS4D02G357900). In *Arabidopsis thaliana*, BAG6 is processed *in planta* upon fungal recognition and its cleavage confers resistance via the induction of autophagy-mediated cell death [29].

**Figure 6: Predicted targets of the most abundant *Pgt* sRNAs in (A) rust and in (B) wheat.** Gene expression levels of *Pgt* RNAi genes targeted by the 20 most abundant *Pgt* sRNAs are shown in *log* of counts per million (*logCPM*, red color intensity relates to high expression strength and blue color intensity to low expression strength).

## Discussion

Small RNAs play a vital role in regulation of gene expression and in plant-pathogen crosstalk [14]. Previous studies on small RNA characterization in fungal plant pathogens mostly rely on sequencing of one time point of infection only. This obscures the expression profiles of pathogen sRNAs over a time course of infection. For example, a previous study in the wheat stripe rust fungus (*Puccinia striiformis* f.sp. *tritici*) sequenced sRNAs at 4 dpi and found that the majority (75 %) of the predicted 20–22 nt *Pst* sRNAs began with uracil [19]. The presence of distinct sRNA profiles in mycelia and appressoria tissues was suggested in the rice blast fungal pathogen, *Magnaporthe oryzae* [16]. However, prominent waves of sRNA expression profiles during fungal infection of plants have thus far not been reported. Through small RNA sequencing over a time course of *Pgt*-wheat infection, we uncovered that *Pgt* produces two distinct sRNA profiles during infection. Using the ShortStack software which uses criteria tailored to plant miRNA properties, we predicted 6,352 *Pgt* sRNAs and only 10 *Pgt* miRNAs. However, it is possible that *Pgt* produces a larger contingent of miRNA-like RNAs that follow currently unknown fungal-specific rules. For example, plant and animal miRNAs are different in many respects such as in their degree of complementarity to their target mRNA [30]. Loci with some, but insufficient, evidence for miRNA biogenesis (such as strandedness) using ShortStack might be worth exploring as miRNA-like candidates in the future [31].

*Pgt* sRNA expression appears to be under tight temporal control, with over 90% of *Pgt* sRNAs induced over the time course, compared to only 2% of wheat sRNAs. In germinated spores and during early infection, *Pgt* sRNAs predominantly overlap with gene models and are 21 nts in length with a 5’ uracil. A switch to 22 nt *Pgt* sRNAs with a 5’ adenine occurs during late infection, which coincides with formation of new urediniospores, and these 22 nt sRNAs are mostly produced from repetitive elements. The presence of two distinct sRNA profiles and their differential expression during rust developmental stages has thus far not been observed and indicates functional diversification of the RNAi machinery, with a strong role in the infection process. Many 22 nt *Pgt* sRNAs with a 5’ adenine appear to silence TEs, suggesting that their primary role is in maintaining genome stability. The high level accumulation of 22 nt sRNAs occurs at the later stage of infection, suggesting that the 22 nt sRNA function to silence TEs during formation of new urediniospores. The up-regulation of argonaute A, one dicer and two RdRPs at the late stage of infection implies their involvement in such a functionally diversified TE silencing pathway. In plants, TEs are silenced mainly by transcriptional gene silencing (TGS) via 24 nt small RNA-directed DNA methylation (RdDM) [25]. 24 nt sRNAs are most abundant during seed development in plants, presumably to ensure stable inheritance of the genome. However, when TEs are activated, 21 nt sRNAs can be generated to direct post-transcriptional gene silencing (PTGS) of TEs in plants, and conversely, when TGS occurs 21 nt sRNAs disappear and 24nt sRNAs appear. So TGS and PTGS appear to be antagonistic to each other although both function to inhibit TE activity. One can speculate that the majority of 22 nt *Pgt* sRNAs are responsible for TGS and the majority of 21 nt *Pgt* sRNAs for PTGS.

In rust one can speculate that during germination and establishment of infection, TEs are less regulated and so have an opportunity to cause mutation in effector genes that would allow infection of a resistant host. Having established a successful infection leading to pustules and urediniospore production, the genome has proved itself to be compatible so it might reduce TE element activity to maintain this successful combination. The high activity of 22 nt sRNAs in the later stages of infection might ensure that the genome is passed on stably to subsequent generations. Rust genomes are amongst the largest sequenced fungal genomes and feature expanded gene families and massive amplification of TEs, mobile genetic elements that shape the evolution of genomes [32]. Over 45% of the assembled *Pgt* reference genome consists of TEs, with 12.4% and 9.8% of class I LTR retrotransposons and class II TIR elements, respectively [7]. Genome sequencing has revealed that some plant pathogenic fungi harbour their effector genes preferentially in TE-rich genomic compartments. In particular, highly expressed effector genes can associate with TE-rich genomic regions [33]. However, a correlation between TE-rich genomic regions and enrichment in effector genes has not been reported for rust genomes thus far [34, 35]. The TE silencing function can be hijacked by some genes for regulation and we showed that this occurs in ~300 *Pgt* genes up to ~200 bps of sRNA-targeted TEs. In plants, insertion of TE near genes can provide cis-elements for stress responsive or tissue-specific expression, and the expression level can be modulated by DNA methylation and/or histone modification at the TEs. It is likely that a similar DNA methylation or histone modification mechanism exists in *Pgt*. In the future, the availability of a long-read *Pgt* genome with less fragmentation will enable more accurate mapping of sRNA reads to repetitive regions and further empower the investigation of TE-associated genes.

By correlating target predictions with gene expression data, we were able to associate down-regulated targets to induced sRNA levels. In germinated spores, *Pgt* appears to down-regulate some secreted proteins with high expression in haustoria, low expression in spores and high expression during infection, including the effector AvrSr35. Future experiments are needed to confirm the role of sRNAs in direct regulation of effector expression. Overall, our work provides strong evidence for the involvement of sRNAs in *Pgt* gene and TE regulation as well as in host-pathogen interactions. The down-regulation of wheat defence-related genes in combination with highly abundant *Pgt* sRNAs that are predicted to target them, but that have no confidently predicted target in rust, highlights the possibility that *Pgt* may use sRNAs to silence host genes. In the future, small RNA-sequencing specifically of haustorial tissues can elucidate if haustoria are the sites of sRNA transfer between host and pathogen [36].

## Conclusions

In summary, we demonstrated that the stem rust fungus *Puccinia graminis* f. sp. *tritici* induces two distinct waves of small RNAs during infection of wheat. Such tight temporal control of the RNAi machinery, resulting in differential accumulation of *Pgt* sRNAs in germinated spores and specialized-infection tissues, has thus far not been observed in fungi. Our work has uncovered fundamental characteristics of the stem rust RNAi machinery. Future research can use this knowledge to optimize methods of host-induced gene silencing where small RNAs from the plant operate via the fungus's own RNAi machinery to silence genes important for causing disease.

## Methods

### Small RNA sequencing, read processing, filtering and alignment and gene expression analysis of wheat and *Pgt*

For rust infection, host plants (cv. Sonora) were grown at high density (~25 seeds per 12cm pot with compost as growth media) to the two leaf stage (~7 days) in a growth cabinet set at 18-23°C temperature and 16 h light. Spores (-80°C stock) were first thawed and heated to 42°C for 3 minutes, mixed with talcum powder and dusted over the plants. Pots were placed in a moist chamber for 24 hours and then transferred back to the growth cabinet. Leaf samples were harvested at specified days after inoculation, snap frozen and stored at -80°C until use. 100 mg freshly collected spores were germinated in 800 ml sterile $ROH_2O$ (four 15 cm petri dishes, 200ml water in each) at 9 degree overnight. All spores were harvested via filtering through nylon mesh 15 µm. Small RNAs were extracted from the germinated spores and infected leaf samples with the Purelink microRNA Isolation Kit from Invitrogen. We sequenced sRNAs (50 bp reads) from the following five conditions (3 replicates each) on the Illumina HiSeq: germinated spores, uninfected wheat and infected wheat at 3 dpi, 5 dpi and 7 dpi. Adapters were trimmed using cutadapt (-m18 –M28 -q30 –trim-n –discard-untrimmed) [37]. Untrimmed reads, reads shorter than 18 nts or reads larger than 28 nts were discarded and flanking N bases were removed from each read [37]. FASTQC was run on the resulting reads (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). *Triticum aestivum* and *Puccinia* tRNAs, rRNAs and spliceosomal RNAs were collected from the RNACentral database [23] as well as the tRNA and rRNA RFAM families RF00001, RF00002, RF00005, RF01852, RF01960 and RF02543 [38], snoRNAs from dbsnOPY, 5S and 23S ribosomal RNAs from the European Nucleotide Archive (ENA) and the tRNA/rRNA file from the sRNA workbench [39]. Contaminants were de-duplicated using bbmap and its tool dedupe.sh (sourceforge.net/projects/bbmap/). Reads that mapped to the contaminant set were removed using bowtie 1.1.2 [40]. To assess read length distributions across the different samples, clean small RNA reads were mapped to the wheat genome IWGSC RefSeq v1.0 [41] and *PGT* 21-0 genome using bowtie (alignment settings: one mismatch allowed –v0; report all alignments: -a –best –strata; suppress all alignments with more than 50 reportable alignments: -m50).

From the same infected leaf samples, previously published RNA-sequencing data (0 dpi, 2 dpi, 3 dpi, 4 dpi, 5 dpi, 6 dpi, 7dpi) was used for the gene expression analysis [6]. This was complemented with previously published RNA-sequencing data of *Pgt* 21-0 germinated spores and haustorial tissue [42]. The program STAR [43] was used with default parameters to align the reads against the *Triticum aestivum* (IWGSC 1.0) [41] and *Pgt 21-0* [8] genomes. We performed differential expression analysis of the *Pgt* 21-0 and IWGSC genes (high confidence models v1.1) with edgeR [24]. FeatureCounts [44] was used to obtain a read count matrix, counting multi-mapped reads (-M) for wheat but not for *Pgt*. The *Pgt* 21-0 gff file was manually corrected for the AvrSr35 [5] and argonaute gene models. Differentially expressed genes were annotated with the B2GO software [45]. We downloaded the top 100 Blastp hits [46] for the *Pgt 21-0* argonaute proteins and used the ete3 toolkit [47] to construct phylogenetic trees with the RAxML standard workflow. This used Clustal Omega as the aligner [48] and RAxML for phylogenetic tree prediction [49].

### *Pgt* sRNA prediction, differential expression analysis and sRNA target prediction

To annotate and quantify high-confidence *Pgt* and wheat small RNAs from the sequencing data, we used the ShortStack 3.8.5 software [22] on the clean sRNA reads (setting for rust: --foldsize 1000, --pad 25, --mincov 30; setting for wheat: --foldsize 300, --mincov 30). We further filtered the predicted sRNA clusters to include only those

where >= 80% of reads are within 20-24 nts of length. This is the recommended procedure in ShortStack to avoid degradation products. We also included only sRNA clusters that produce a major RNA with at least 5 reads in total and where the sRNA cluster has at least 5 RPM. The ShortStack software outputs sRNA cluster properties such as the most abundant sRNA (termed sRNA candidate) in the cluster, strandedness of the locus, miRNA annotation and phasing [22]. Strandedness of sRNA loci is determined by forcing the bowtie aligner to select one strand or the other with a probability that is proportional to the number of best sites on the strand. Stranded loci are typical of miRNA production in plants and are a requirement for annotation of a locus as a miRNA by ShortStack. We used the read counts returned by ShortStack for all predicted sRNA clusters and used edgeR [24] to assess which are differentially expressed at any of the infection stages versus germinated spores (FDR < 0.05, fold change > 2).

We used psRNATarget 2017 to predict targets against *Pgt* and wheat transcripts with expectation score <= 5.0 [26]. If a sRNA has a predicted target in rust with perfect complementarity (expectation score = 0.0), we only kept this transcript as the putative target and dismissed those transcript with less complementarity. Secreted proteins were predicted using using SignalP 3 [50], TMHMM [51] and TargetP [52] as described in [53]. Effectors were predicted using EffectorP 2.0 [27]

All plots were produced using Ggplot2 (Wickham, 2009) and statistical significance was assessed with *t*□tests using the ggsignif package (https://cran.r-project.org/web/packages/ggsignif/index.html). Significance thresholds according to *t*□test are: NS, not significant; *, < 0.05; **, < 0.01; ***, < 0.001.

## Declarations

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Availability of data and material

Sequence data for the *Pgt* infection RNAseq is available at the National Center for Biotechnology Information Sequencing Read Archive under bioproject PRJNA415866. Sequence data for the *Pgt* small RNAseq is available at CSIRO Data Access Portal under https://doi.org/10.25919/5bd93643b1e41.

### Competing interests

Not applicable.

### Funding

JS was supported by a CSIRO Office of the Chief Executive (OCE) Postdoctoral Fellowship.

### Authors' contributions

JS, JMT and PND conceived and designed the study. JS analysed and interpreted the data with input and feedback from all authors. SJ, BX, NMU and RM performed experiments. All authors were involved in drafting the manuscript and revised, read and approved the final manuscript.

### Acknowledgements

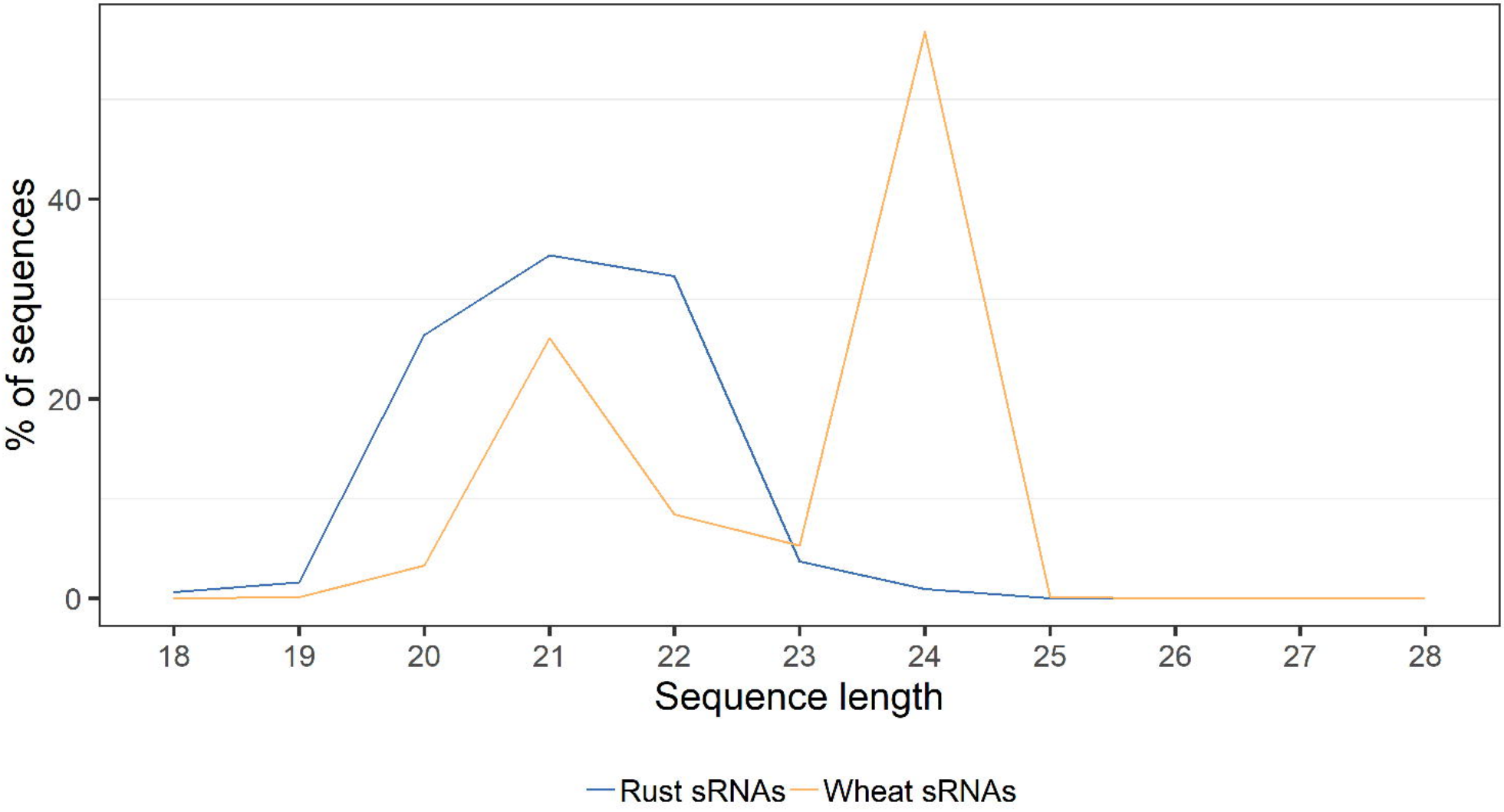
We thank Xiaodi Xia for excellent technical assistance.

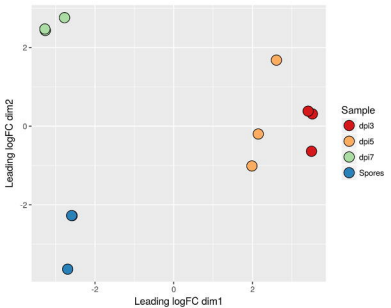
## References

1. Dubin HJB, J.P.: **Combating stem and leaf rust of wheat: Historical perspective, impacts, and lessons learned.** *International Food Policy Research Institute* 2009.
2. Figueroa M, Upadhyaya NM, Sperschneider J, Park RF, Szabo LJ, Steffenson B, Ellis JG, Dodds PN: **Changing the Game: Using Integrative Genomics to Probe Virulence Mechanisms of the Stem Rust Pathogen Puccinia graminis f. sp. tritici.** *Front Plant Sci* 2016, **7:**205.
3. Leonard KJ, Szabo LJ: **Stem rust of small grains and grasses caused by *Puccinia graminis.*** *Mol Plant Pathol* 2005, **6:**99-111.
4. Garnica DP, Nemri A, Upadhyaya NM, Rathjen JP, Dodds PN: **The ins and outs of rust haustoria.** *PLoS Pathog* 2014, **10:**e1004329.

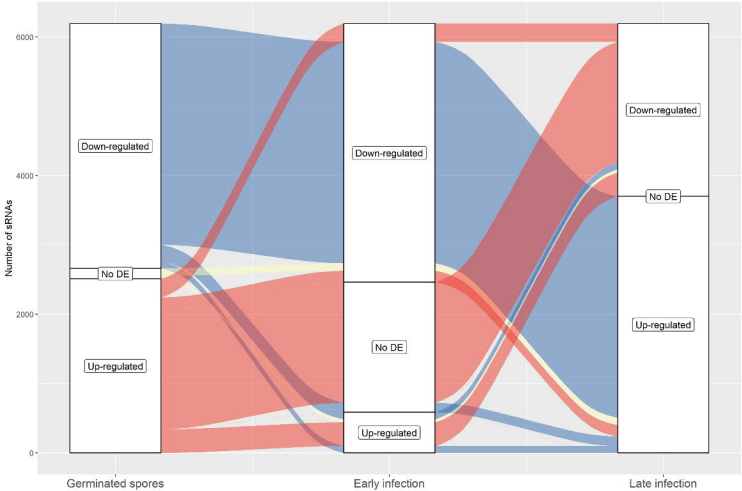
5. Salcedo A, Rutter W, Wang S, Akhunova A, Bolus S, Chao S, Anderson N, De Soto MF, Rouse M, Szabo L, et al: **Variation in the AvrSr35 gene determines Sr35 resistance against wheat stem rust race Ug99.** *Science* 2017, **358:**1604-1606.
6. Chen J, Upadhyaya NM, Ortiz D, Sperschneider J, Li F, Bouton C, Breen S, Dong C, Xu B, Zhang X, et al: **Loss of AvrSr50 by somatic exchange in stem rust leads to virulence for Sr50 resistance in wheat.** *Science* 2017, **358:**1607-1610.
7. Duplessis S, Cuomo CA, Lin YC, Aerts A, Tisserant E, Veneault-Fourrey C, Joly DL, Hacquard S, Amselem J, Cantarel BL, et al: **Obligate biotrophy features unraveled by the genomic analysis of rust fungi.** *Proc Natl Acad Sci U S A* 2011, **108:**9166-9171.
8. Upadhyaya NM, Garnica DP, Karaoglu H, Sperschneider J, Nemri A, Xu B, Mago R, Cuomo CA, Rathjen JP, Park RF, et al: **Comparative genomics of Australian isolates of the wheat stem rust pathogen Puccinia graminis f. sp. tritici reveals extensive polymorphism in candidate effector genes.** *Front Plant Sci* 2015, **5:**759.
9. Matzke MA, Mosher RA: **RNA-directed DNA methylation: an epigenetic pathway of increasing complexity.** *Nat Rev Genet* 2014, **15:**394-408.
10. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR: **Highly integrated single-base resolution maps of the epigenome in Arabidopsis.** *Cell* 2008, **133:**523-536.
11. Chang SS, Zhang Z, Liu Y: **RNA interference pathways in fungi: mechanisms and functions.** *Annu Rev Microbiol* 2012, **66:**305-323.
12. Romano N, Macino G: **Quelling: transient inactivation of gene expression in Neurospora crassa by transformation with homologous sequences.** *Mol Microbiol* 1992, **6:**3343-3353.
13. Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA: **Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi.** *Science* 2002, **297:**1833-1837.
14. Weiberg A, Wang M, Bellinger M, Jin H: **Small RNAs: a new paradigm in plant-microbe interactions.** *Annu Rev Phytopathol* 2014, **52:**495-516.
15. Vetukuri RR, Asman AK, Tellgren-Roth C, Jahan SN, Reimegard J, Fogelqvist J, Savenkov E, Soderbom F, Avrova AO, Whisson SC, Dixelius C: **Evidence for small RNAs homologous to effector-encoding genes and transposable elements in the oomycete Phytophthora infestans.** *PLoS One* 2012, **7:**e51399.
16. Nunes CC, Gowda M, Sailsbery J, Xue M, Chen F, Brown DE, Oh Y, Mitchell TK, Dean RA: **Diverse and tissue-enriched small RNAs in the plant pathogenic fungus, Magnaporthe oryzae.** *BMC Genomics* 2011, **12:**288.
17. Weiberg A, Wang M, Lin FM, Zhao H, Zhang Z, Kaloshian I, Huang HD, Jin H: **Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways.** *Science* 2013, **342:**118-123.
18. Cai Q, Qiao L, Wang M, He B, Lin FM, Palmquist J, Huang HD, Jin H: **Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes.** *Science* 2018.
19. Mueth NA, Ramachandran SR, Hulbert SH: **Small RNAs from the wheat stripe rust fungus (Puccinia striiformis f.sp. tritici).** *BMC Genomics* 2015, **16:**718.
20. Wang B, Sun Y, Song N, Zhao M, Liu R, Feng H, Wang X, Kang Z: **Puccinia striiformis f. sp. tritici microRNA-like RNA 1 (Pst-milR1), an important pathogenicity factor of Pst, impairs wheat resistance to Pst by suppressing the wheat pathogenesis-related 2 gene.** *New Phytol* 2017, **215:**338-350.
21. Choi J, Kim KT, Jeon J, Wu J, Song H, Asiegbu FO, Lee YH: **funRNA: a fungi-centered genomics platform for genes encoding key components of RNAi.** *BMC Genomics* 2014, **15 Suppl 9:**S14.
22. Axtell MJ: **ShortStack: comprehensive annotation and quantification of small RNA genes.** *RNA* 2013, **19:**740-751.
23. Consortium TR: **RNAcentral: a comprehensive database of non-coding RNA sequences.** *Nucleic Acids Res* 2016.
24. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26:**139-140.
25. Borges F, Martienssen RA: **The expanding world of small RNAs in plants.** *Nat Rev Mol Cell Biol* 2015, **16:**727-741.
26. Dai X, Zhuang Z, Zhao PX: **psRNATarget: a plant small RNA target analysis server (2017 release).** *Nucleic Acids Res* 2018.
27. Sperschneider J, Dodds PN, Gardiner DM, Singh KB, Taylor JM: **Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0.** *Mol Plant Pathol* 2018.
28. Hartmann HA, Kahmann R, Bolker M: **The pheromone response factor coordinates filamentous growth and pathogenicity in Ustilago maydis.** *EMBO J* 1996, **15:**1632-1641.
29. Li Y, Kabbage M, Liu W, Dickman MB: **Aspartyl Protease-Mediated Cleavage of BAG6 Is Necessary for Autophagy and Fungal Resistance in Plants.** *Plant Cell* 2016, **28:**233-247.

30. Millar AA, Waterhouse PM: **Plant and animal microRNAs: similarities and differences.** *Funct Integr Genomics* 2005, **5:**129-135.
31. Axtell MJ, Westholm JO, Lai EC: **Vive la difference: biogenesis and evolution of microRNAs in plants and animals.** *Genome Biol* 2011, **12:**221.
32. Raffaele S, Kamoun S: **Genome evolution in filamentous plant pathogens: why bigger can be better.** *Nat Rev Microbiol* 2012, **10:**417-430.
33. Moller M, Stukenbrock EH: **Evolution and genome architecture in fungal plant pathogens.** *Nat Rev Microbiol* 2017.
34. Schwessinger B, Sperschneider J, Cuddy WS, Garnica DP, Miller ME, Taylor JM, Dodds PN, Figueroa M, Park RF, Rathjen JP: **A near-complete haplotype-phased genome of the dikaryotic wheat stripe rust fungus Puccinia striiformis f. sp. tritici reveals high interhaplotype diversity.** *mBio* 2018, **9**.
35. Miller ME, Zhang Y, Omidvar V, Sperschneider J, Schwessinger B, Raley C, Palmer JM, Garnica D, Upadhyaya NM, Rathjen JP, et al: **De novo assembly and phasing of dikaryotic genomes from two isolates of Puccinia coronata f. sp. avenae, the causal agent of oat crown rust.** *9* 2018, **e01650-17**.
36. Shahid S, Kim G, Johnson NR, Wafula E, Wang F, Coruh C, Bernal-Galeano V, Phifer T, dePamphilis CW, Westwood JH, Axtell MJ: **MicroRNAs from the parasitic plant Cuscuta campestris target host messenger RNAs.** *Nature* 2018, **553:**82-85.
37. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** *2011* 2011, **17**.
38. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate J, Finn RD: **Rfam 12.0: updates to the RNA families database.** *Nucleic Acids Res* 2015, **43:**D130-137.
39. Stocks MB, Moxon S, Mapleson D, Woolfenden HC, Mohorianu I, Folkes L, Schwach F, Dalmay T, Moulton V: **The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets.** *Bioinformatics* 2012, **28:**2059-2061.
40. Langmead B, Trapnell C, Pop M, Salzberg SL: **Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.** *Genome Biol* 2009, **10:**R25.
41. International Wheat Genome Sequencing C, investigators IRp, Appels R, Eversole K, Feuillet C, Keller B, Rogers J, Stein N, investigators Iw-gap, Pozniak CJ, et al: **Shifting the limits in wheat research and breeding using a fully annotated reference genome.** *Science* 2018, **361**.
42. Upadhyaya NM, Garnica DP, Karaoglu H, Sperschneider J, Nemri A, Xu B, Mago R, Cuomo CA, Rathjen JP, Park RF, et al: **Comparative genomics of Australian isolates of the wheat stem rust pathogen Puccinia graminis f. sp. tritici reveals extensive polymorphism in candidate effector genes.** *Front Plant Sci* 2014, **5:**759.
43. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics* 2013, **29:**15-21.
44. Liao Y, Smyth GK, Shi W: **featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.** *Bioinformatics* 2014, **30:**923-930.
45. Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A: **High-throughput functional annotation and data mining with the Blast2GO suite.** *Nucleic Acids Res* 2008, **36:**3420-3435.
46. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215:**403-410.
47. Huerta-Cepas J, Serra F, Bork P: **ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data.** *Mol Biol Evol* 2016, **33:**1635-1638.
48. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, et al: **Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.** *Mol Syst Biol* 2011, **7:**539.
49. Stamatakis A: **RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.** *Bioinformatics* 2014, **30:**1312-1313.
50. Bendtsen JD, Nielsen H, von Heijne G, Brunak S: **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol* 2004, **340:**783-795.
51. Krogh A, Larsson B, von Heijne G, Sonnhammer EL: **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J Mol Biol* 2001, **305:**567-580.
52. Emanuelsson O, Nielsen H, Brunak S, von Heijne G: **Predicting subcellular localization of proteins based on their N-terminal amino acid sequence.** *J Mol Biol* 2000, **300:**1005-1016.
53. Sperschneider J, Williams AH, Hane JK, Singh KB, Taylor JM: **Evaluation of Secretion Prediction Highlights Differing Approaches Needed for Oomycete and Fungal Effectors.** *Front Plant Sci* 2015, **6:**1168.

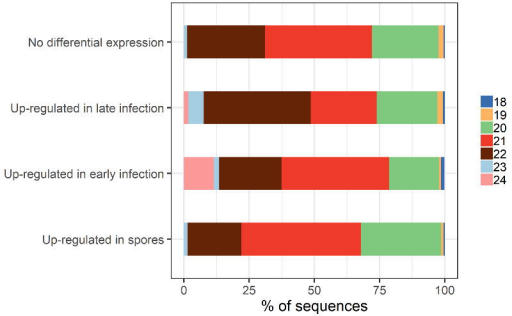

% of sequences
40
20
0
18
19
20
21
22
23
24
25
26
27
28
Sequence length
Rust sRNAs
Wheat sRNAs


Leading logFC dim2
Leading logFC dim1
2
0
-2
-2
0
2
Sample
dpi3
dpi5
dpi7
Spores

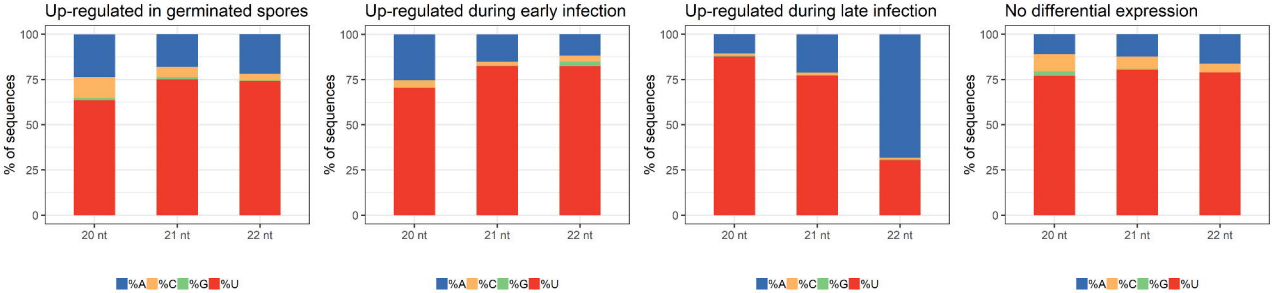
(A)


Number of sRNAs
6000
4000
2000
0
Down-regulated
No DE
Up-regulated
Down-regulated
No DE
Up-regulated
Down-regulated
No DE
Up-regulated
Germinated spores
Early infection
Late infection

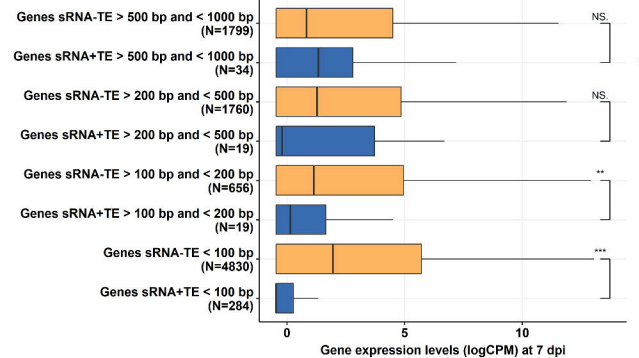
(B)


No differential expression
Up-regulated in late infection
Up-regulated in early infection
Up-regulated in spores
0
25
50
75
100
% of sequences
18
19
20
21
22
23
24

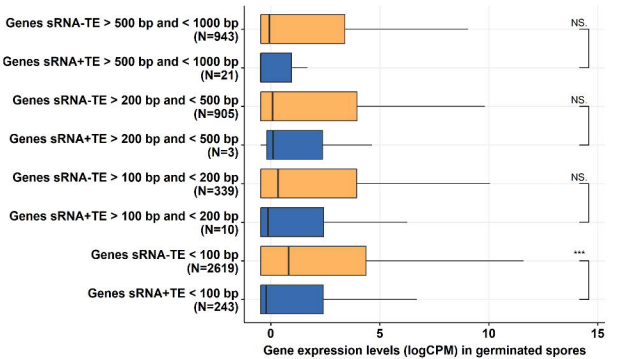
**(A)**


Up-regulated in germinated spores
Up-regulated during early infection
Up-regulated during late infection
No differential expression
% of sequences
0
25
50
75
100
20 nt
21 nt
22 nt
%A
%C
%G
%U

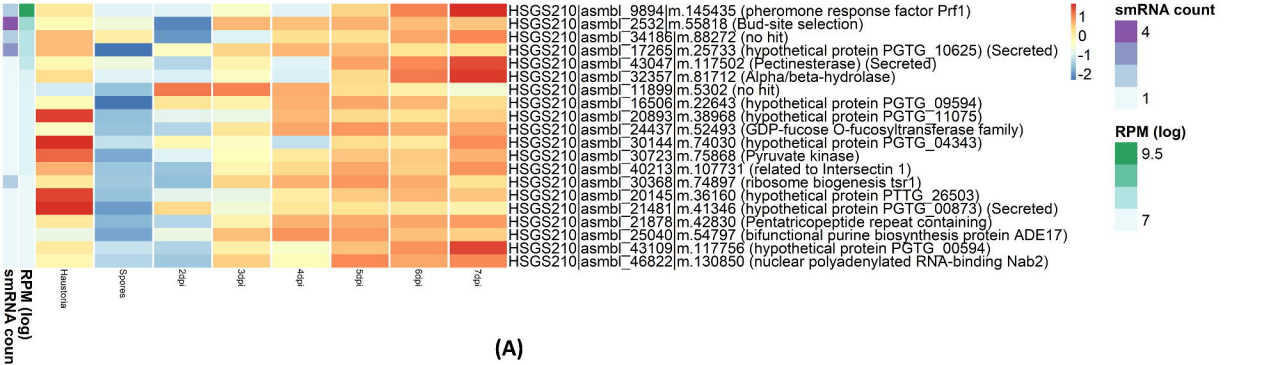
**(B)**


Genes sRNA-TE > 500 bp and < 1000 bp (N=1799)
Genes sRNA+TE > 500 bp and < 1000 bp (N=34)
Genes sRNA-TE > 200 bp and < 500 bp (N=1760)
Genes sRNA+TE > 200 bp and < 500 bp (N=19)
Genes sRNA-TE > 100 bp and < 200 bp (N=656)
Genes sRNA+TE > 100 bp and < 200 bp (N=19)
Genes sRNA-TE < 100 bp (N=4830)
Genes sRNA+TE < 100 bp (N=284)
NS.
NS.
**
***
0
5
10
Gene expression levels (logCPM) at 7 dpi

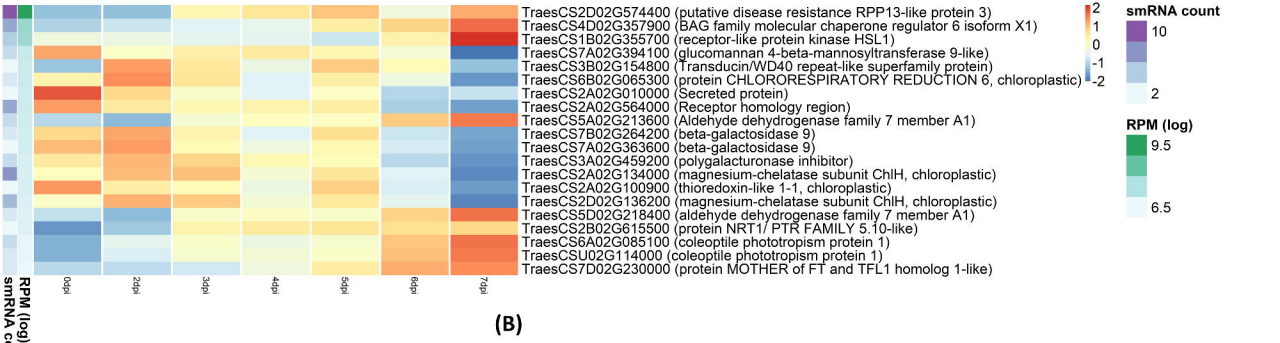
(A)


Genes sRNA-TE > 500 bp and < 1000 bp (N=943)
Genes sRNA+TE > 500 bp and < 1000 bp (N=21)
Genes sRNA-TE > 200 bp and < 500 bp (N=905)
Genes sRNA+TE > 200 bp and < 500 bp (N=3)
Genes sRNA-TE > 100 bp and < 200 bp (N=339)
Genes sRNA+TE > 100 bp and < 200 bp (N=10)
Genes sRNA-TE < 100 bp (N=2619)
Genes sRNA+TE < 100 bp (N=243)
NS.
NS.
NS.
***
0
5
10
15
Gene expression levels (logCPM) in germinated spores

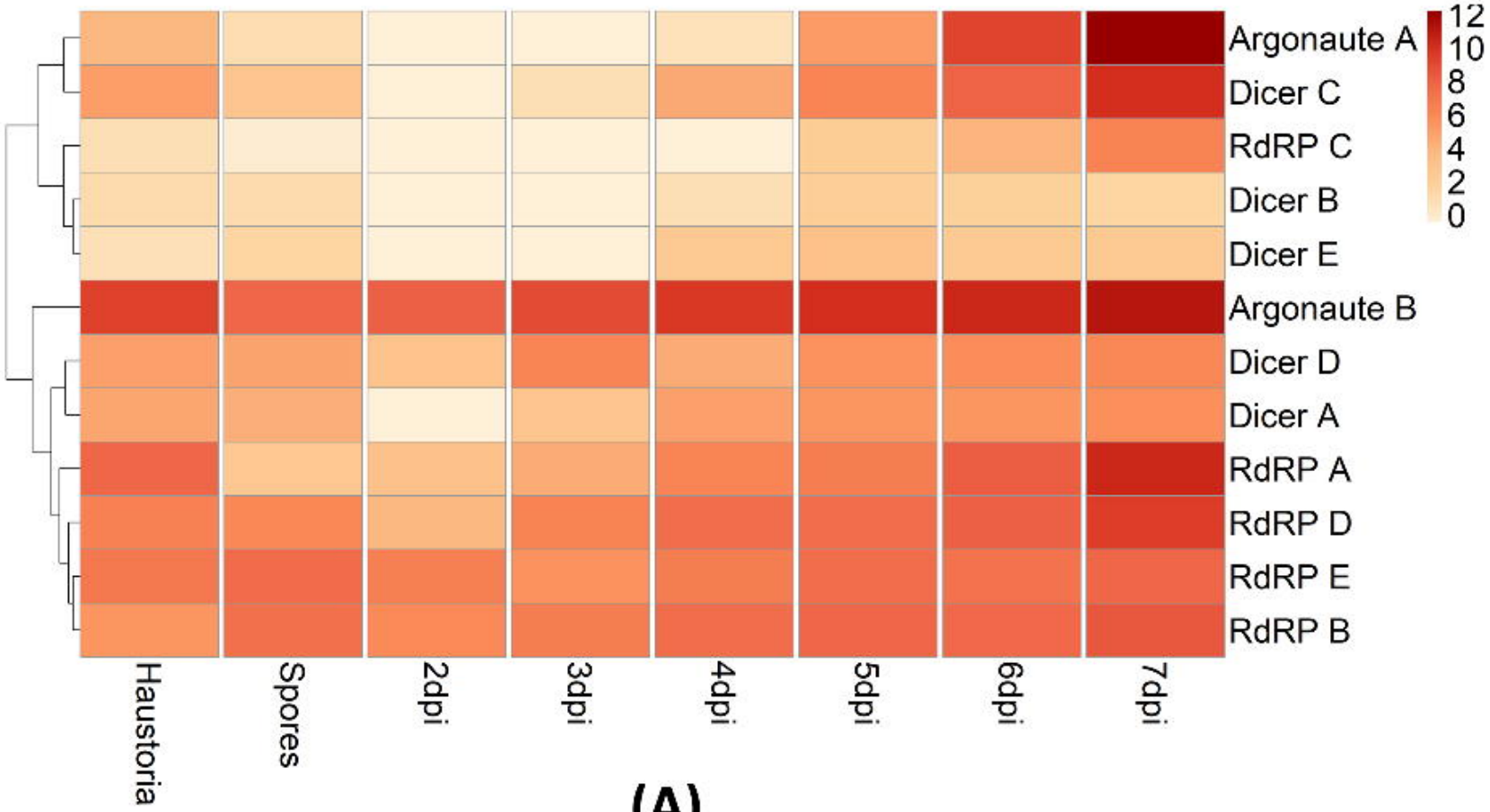
(B)


HSGS210|asmbl_9894|m.145435 (pheromone response factor Prf1)
HSGS210|asmbl_2532|m.55818 (Bud-site selection)
HSGS210|asmbl_34186|m.88272 (no hit)
HSGS210|asmbl_17265|m.25733 (hypothetical protein PGTG_10625) (Secreted)
HSGS210|asmbl_43047|m.117502 (Pectinesterase) (Secreted)
HSGS210|asmbl_32357|m.81712 (Alpha/beta-hydrolase)
HSGS210|asmbl_11899|m.5302 (no hit)
HSGS210|asmbl_16506|m.22643 (hypothetical protein PGTG_09594)
HSGS210|asmbl_20893|m.38968 (hypothetical protein PGTG_11075)
HSGS210|asmbl_24437|m.52493 (GDP-fucose O-fucosyltransferase family)
HSGS210|asmbl_30144|m.74030 (hypothetical protein PGTG_04343)
HSGS210|asmbl_30723|m.75868 (Pyruvate kinase)
HSGS210|asmbl_40213|m.107731 (related to Intersectin 1)
HSGS210|asmbl_30368|m.74897 (ribosome biogenesis tsr1)
HSGS210|asmbl_20145|m.36160 (hypothetical protein PTTG_26503)
HSGS210|asmbl_21481|m.41346 (hypothetical protein PGTG_00873) (Secreted)
HSGS210|asmbl_21878|m.42830 (Pentatricopeptide repeat containing)
HSGS210|asmbl_25040|m.54797 (bifunctional purine biosynthesis protein ADE17)
HSGS210|asmbl_43109|m.117756 (hypothetical protein PGTG_00594)
HSGS210|asmbl_46822|m.130850 (nuclear polyadenylated RNA-binding Nab2)
Haustoria
Spores
2dpi
3dpi
4dpi
5dpi
6dpi
7dpi
RPM (log)
smRNA count
1
0
-1
-2
smRNA count
4
1
RPM (log)
9.5
7

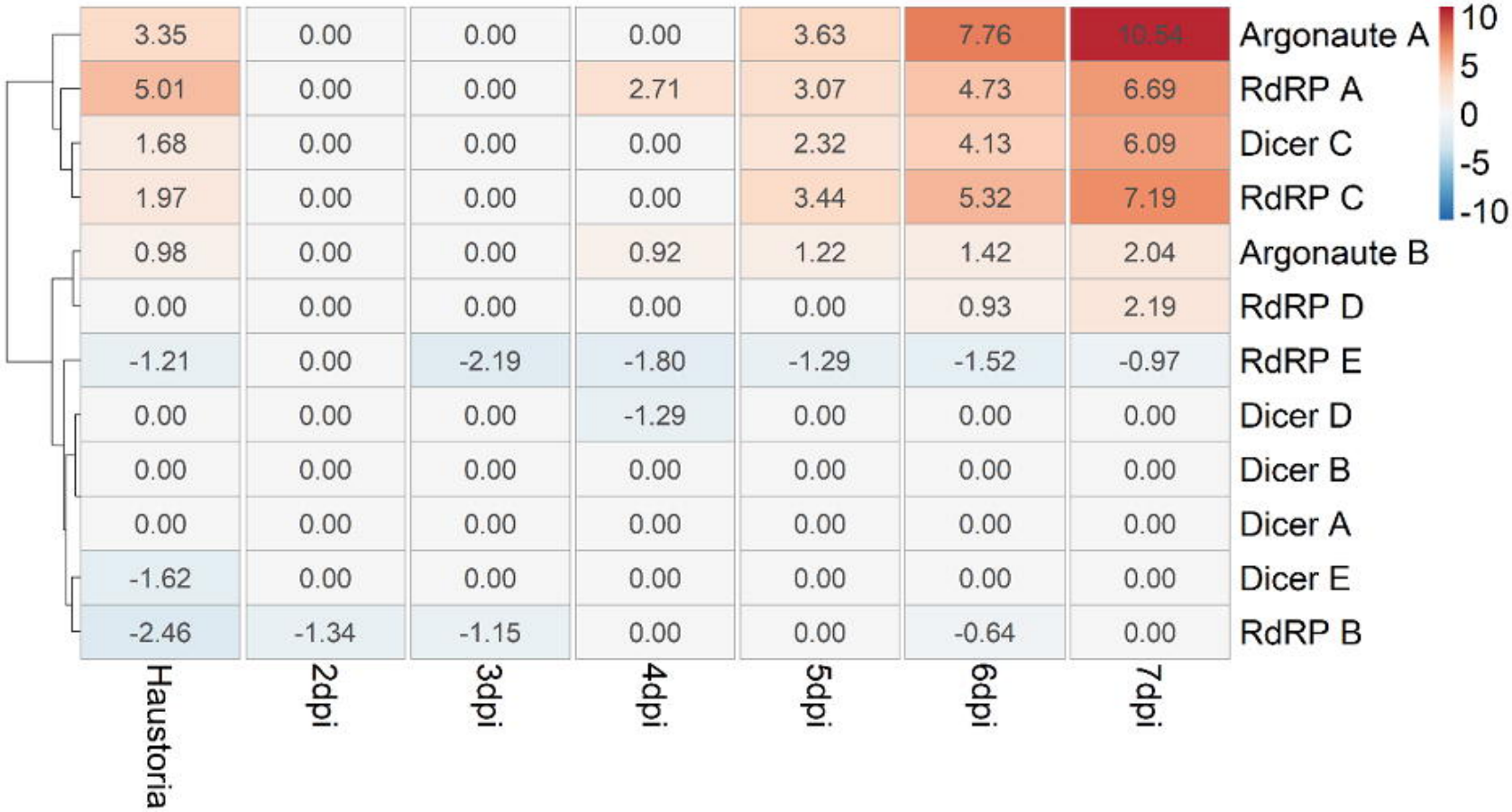
(A)


TraesCS2D02G574400 (putative disease resistance RPP13-like protein 3)
TraesCS4D02G357900 (BAG family molecular chaperone regulator 6 isoform X1)
TraesCS1B02G355700 (receptor-like protein kinase HSL1)
TraesCS7A02G394100 (glucomannan 4-beta-mannosyltransferase 9-like)
TraesCS3B02G154800 (Transducin/WD40 repeat-like superfamily protein)
TraesCS6B02G065300 (protein CHLORORESPIRATORY REDUCTION 6, chloroplastic)
TraesCS2A02G010000 (Secreted protein)
TraesCS2A02G564000 (Receptor homology region)
TraesCS5A02G213600 (Aldehyde dehydrogenase family 7 member A1)
TraesCS7B02G264200 (beta-galactosidase 9)
TraesCS7A02G363600 (beta-galactosidase 9)
TraesCS3A02G459200 (polygalacturonase inhibitor)
TraesCS2A02G134000 (magnesium-chelatase subunit ChlH, chloroplastic)
TraesCS2A02G100900 (thioredoxin-like 1-1, chloroplastic)
TraesCS2D02G136200 (magnesium-chelatase subunit ChlH, chloroplastic)
TraesCS5D02G218400 (aldehyde dehydrogenase family 7 member A1)
TraesCS2B02G615500 (protein NRT1/ PTR FAMILY 5.10-like)
TraesCS6A02G085100 (coleoptile phototropism protein 1)
TraesCSU02G114000 (coleoptile phototropism protein 1)
TraesCS7D02G230000 (protein MOTHER of FT and TFL1 homolog 1-like)
0dpi
2dpi
3dpi
4dpi
5dpi
6dpi
7dpi
RPM (log)
smRNA count
2
1
0
-1
-2
smRNA count
10
2
RPM (log)
9.5
6.5

(B)


Argonaute A
Dicer C
RdRP C
Dicer B
Dicer E
Argonaute B
Dicer D
Dicer A
RdRP A
RdRP D
RdRP E
RdRP B
12
10
8
6
4
2
0
Haustoria
Spores
2dpi
3dpi
4dpi
5dpi
6dpi
7dpi

**(A)**


3.35 0.00 0.00 0.00 3.63 7.76 10.54 Argonaute A
5.01 0.00 0.00 2.71 3.07 4.73 6.69 RdRP A
1.68 0.00 0.00 0.00 2.32 4.13 6.09 Dicer C
1.97 0.00 0.00 0.00 3.44 5.32 7.19 RdRP C
0.98 0.00 0.00 0.92 1.22 1.42 2.04 Argonaute B
0.00 0.00 0.00 0.00 0.00 0.93 2.19 RdRP D
-1.21 0.00 -2.19 -1.80 -1.29 -1.52 -0.97 RdRP E
0.00 0.00 0.00 -1.29 0.00 0.00 0.00 Dicer D
0.00 0.00 0.00 0.00 0.00 0.00 0.00 Dicer B
0.00 0.00 0.00 0.00 0.00 0.00 0.00 Dicer A
-1.62 0.00 0.00 0.00 0.00 0.00 0.00 Dicer E
-2.46 -1.34 -1.15 0.00 0.00 -0.64 0.00 RdRP B
10
5
0
-5
-10
Haustoria
2dpi
3dpi
4dpi
5dpi
6dpi
7dpi

**(B)**