



**HAL**  
open science

## **Dominance modifiers at the Arabidopsis self-incompatibility locus retain proto-miRNA features and act through non-canonical pathways**

Rita A Batista, Eléonore Durand, Monika Mörchen, Jacinthe Azevedo-Favory, Samson Simon, Manu Dubin, Vinod Kumar, Eléonore Lacoste, Corinne Cruaud, Christelle Blassiau, et al.

### ► To cite this version:

Rita A Batista, Eléonore Durand, Monika Mörchen, Jacinthe Azevedo-Favory, Samson Simon, et al.. Dominance modifiers at the Arabidopsis self-incompatibility locus retain proto-miRNA features and act through non-canonical pathways. 2024. hal-04805332

**HAL Id: hal-04805332**

**<https://hal.science/hal-04805332v1>**

Preprint submitted on 26 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Dominance modifiers at the *Arabidopsis* self-incompatibility locus retain proto-miRNA features and act through non-canonical pathways

Rita A. Batista<sup>1,a</sup>, Eléonore Durand<sup>1</sup>, Monika Mörchen<sup>1</sup>, Jacinthe Azevedo-Favory<sup>2</sup>, Samson Simon<sup>1</sup>, Manu Dubin<sup>1</sup>, Vinod Kumar<sup>3</sup>, Eléonore Lacoste<sup>4</sup>, Corinne Cruaud<sup>5</sup>, Christelle Blassiau<sup>1</sup>, Matteo Barois<sup>1</sup>, Anne-Catherine Holl<sup>1</sup>, Chloé Ponitzki<sup>1</sup>, Nathalie Faure<sup>1</sup>, William Marande<sup>6</sup>, Sonia Vautrin<sup>6</sup>, Isabelle Fobis-Loisy<sup>7</sup>, Jean-Marc Aury<sup>4</sup>, Sylvain Legrand<sup>1</sup>, Ute Krämer<sup>3</sup>, Thierry Lagrange<sup>2</sup>, Xavier Vekemans<sup>1</sup>, Vincent Castric<sup>1</sup>

1. Univ. Lille, CNRS, UMR 8198 – Evo-Eco-Paleo, F-59000 Lille, France
2. CNRS, Laboratoire Génome et Développement des Plantes, UMR 5096, 66860 Perpignan, France, and Université Perpignan *Via Domitia*, Laboratoire Génome et Développement des Plantes, UMR 5096, F-66860 Perpignan, France
3. Department of Molecular Genetics and Physiology of Plants, Faculty of Biology and Biotechnology, Ruhr University Bochum, D-44801 Bochum, Germany
4. Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France
5. Genoscope, Institut François Jacob, CEA, Université Paris-Saclay, Evry, 91057, France
6. INRAE - CNRGV French Plant Genomic Resource Center - Castanet Tolosan, France
7. Laboratoire Reproduction et Développement des Plantes, Université de Lyon, ENS de Lyon, UCB Lyon 1, CNRS, INRAE, F-69342 Lyon, France
- a. Current address: Max Planck Institute for Biology, Department of Algal Development and Evolution, 72076 Tübingen, Germany

**Corresponding author:** Vincent Castric, [vincent.castric@univ-lille.fr](mailto:vincent.castric@univ-lille.fr)

**Author Contributions:** Designed research – RAB, ED, MM, JA-F, MD, SS, VK, IF-L, J-MA, SL, UK, TL, XV, VC; Performed research – RAB, ED, MM, JA-F, MD, SS, VK, EL, CC, CB, A-CH, CP, NF, IF-L; Contributed new reagents or analytic tools – WM, SV, IF-L, J-MA, UK, TL, XV, VC; Analyzed data – RAB, ED, MM, JA-F, MD, SS, VK, EL, J-MA, SL, XV, VC; Wrote the paper – RAB, VC. All authors reviewed and edited the manuscript.

**Classification:** Biological Sciences, Plant Biology

**Keywords:** self-incompatibility; dominance; miRNA; gene silencing

**This PDF file includes:**

Main Text

Figures 1 to 5, with corresponding supplementary figures

## 1 **Abstract**

2 Self-incompatibility in flowering plants is a common mechanism that prevents self-fertilization and  
3 promotes outcrossing. In Brassicaceae, there is high genetic diversity at the locus controlling self-  
4 incompatibility, and dozens of distinct alleles are organized in a complex dominance hierarchy: the gene  
5 controlling self-incompatibility specificity in pollen shows monoallelic expression in heterozygote  
6 individuals. This is achieved through the action of sRNA precursors that resemble miRNAs, although  
7 the underlying molecular mechanisms remain elusive. Here, we engineered *Arabidopsis thaliana* lines  
8 expressing components of the *Arabidopsis halleri* self-incompatibility system, and used a reverse  
9 genetics approach to pinpoint the pathways underlying the function of these sRNA precursors. We  
10 showed that they trigger a robust decrease in transcript abundance of the recessive pollen self-  
11 incompatibility genes, but not through the canonical transcriptional or post-transcriptional gene silencing  
12 pathways. Furthermore, we observed that single sRNA precursors are typically processed into  
13 hundreds of sRNA molecules of distinct sizes, abundance levels and ARGONAUTE loading  
14 preferences. This heterogeneity closely resembles that of proto-miRNAs, the evolutionary ancestors of  
15 miRNAs. Our results suggest that these apparently arbitrary features, which are often associated with  
16 lack of effects on gene expression, are crucial in the context of the self-incompatibility dominance  
17 hierarchy since they allow for one sRNA precursor of a given allele to repress multiple other recessive  
18 alleles. This study not only provides an in-depth characterization of the molecular features underlying  
19 complex dominance interactions, but also constitutes a unique example of how specific evolutionary  
20 constraints shape the progression of sRNA precursors along the proto-miRNA - miRNA evolutionary  
21 continuum.

## 22 **Significance statement**

23 miRNAs are genetically encoded small non-coding RNA molecules that control gene expression. Here,  
24 we characterized a set of regulatory loci present in the self-incompatibility locus of *Arabidopsis halleri*  
25 which resemble proto-miRNAs, the evolutionary ancestors of miRNAs. We report that despite the  
26 absence of classical miRNA features thought to be essential for gene expression regulation, proto-  
27 miRNAs generated by these loci robustly repress the activity of self-incompatibility genes through non-  
28 canonical silencing pathways, and thus function to establish a complex dominance hierarchy between  
29 self-incompatibility alleles. We show that the evolutionary forces promoting diversity of self-  
30 incompatibility alleles have likely favored the maintenance of proto-miRNA features in order to mediate  
31 dominance interactions between numerous alleles, thus preventing the transition of self-incompatibility  
32 proto-miRNAs into canonical miRNAs.

## 33 Main text

### 34 Introduction

35 Genetic dominance is the widespread property by which one allele can mask the phenotypic effect  
36 of another in a heterozygous genotype **(1)**. Dominance introduces non-linearity in the genotype-to-  
37 phenotype map and influences the exposure of mutations to natural selection, thus altering the  
38 probability and dynamics of their fixation, elimination, or maintenance in natural populations **(2)**. The  
39 evolution of dominance has been subject to a fierce debate in the field of genetics throughout the 20<sup>th</sup>  
40 century **(3)**, specifically around the existence of genetic elements dedicated to altering the dominance  
41 of other alleles in *trans* **(4)**, and the possibility for natural selection to act effectively on these elements  
42 **(5)**. The existence of these dominance modifiers has remained elusive for many years; nevertheless,  
43 theoretical studies have predicted that dominance modifiers could exist in cases where heterozygosity  
44 is high, such as in Batesian and Müllerian mimicry in butterflies, or at the self-incompatibility locus of  
45 Brassicaceae plants **(3)**. More recently, dominance modifiers similar to those envisioned by Fisher **(4)**  
46 were empirically identified at the self-incompatibility locus of Brassica and Arabidopsis **(6–8)**. However,  
47 their mode of action and evolutionary properties remain elusive.

48 Self-incompatibility is a widespread self-recognition mechanism in hermaphroditic flowering plants  
49 that prevents selfing and enforces outcrossing. In Brassicaceae, self-pollen recognition is controlled at  
50 the sporophytic level by a single non-recombining self-incompatibility locus (S-locus) containing two  
51 protein-coding genes: one encoding a trans-membrane receptor (SRK) expressed in papillae cells as  
52 the female component, and the male component consisting of an SCR peptide produced in the tapetum  
53 of the anther and deposited on the surfaces of the pollen grains. The SRK and SCR proteins function  
54 as a receptor-ligand system, such that pollen germination is stopped when a pair of cognate SRK and  
55 SCR proteins interact upon pollen deposition on the pistil, ultimately preventing fertilization **(9–13)**. The  
56 S-locus is subject to an intense long-term balancing selection, which favors the maintenance of a high  
57 number of distinct S-alleles **(14–17)**. Although heterozygosity is typically very high at the S-locus,  
58 monoallelic expression of the male SCR gene is the rule, following a transitive and mostly linear  
59 dominance hierarchy among S-alleles **(18–20)**. These dominance interactions maximize reproductive  
60 fitness, since repressing one of the two S-alleles carried by a diploid heterozygous male individual  
61 avoids rejection by pistils expressing the same allelic specificity, and thus maximizes the number of  
62 available female partners **(21)**. In *Arabidopsis lyrata* and *A. halleri*, respectively, 56 and 43 S-alleles  
63 have been documented so far **(15, 22)**, and are divided in four main dominance classes, with class IV  
64 being the most dominant and class I the most recessive **(23, 24)**.

65 Underlying these dominance interactions are S-locus dominance modifiers, which take the form of  
66 sRNA generating loci. Similarly to miRNAs genes, these loci are transcribed into an RNA molecule  
67 that folds into a hairpin structure, which is subsequently cleaved into sRNAs. In *A. halleri*, eight distinct  
68 families of sRNA precursors exist, and they are predicted to control dominance interactions between  
69 and within the four S-allele dominance classes **(6)**. Interestingly, in Brassica where only two  
70 dominance classes exist, only two sRNA precursors genes have been identified **(7, 8, 20)**, suggesting  
71 that there is an association between the complexity of the dominance network and the total number

72 of dominance modifiers. The two sRNA precursors of Brassica (*Smi* and *Smi2*) have each been  
73 suggested to form a hairpin and produce a 24-nt sRNA with sequence similarity to the 5' region of  
74 *SCR* (referred to *SP11* in the Brassica nomenclature). The presence of this sRNA is associated with  
75 decreased expression of the recessive *SCR* allele as well as deposition of DNA methylation at the  
76 targeted region in the tapetum cells (7, 8, 20). This is reminiscent of what is observed in several plant  
77 species where 21/24-nt sRNAs derived from miRNAs genes or inverted repeats can be loaded onto  
78 proteins of the ARGONAUTE4 (AGO4) clade and subsequently elicit *de novo* DNA methylation at  
79 target loci, such as transposable elements or protein-coding genes, through the RNA-directed DNA  
80 Methylation (RdDM) pathway (25–29). Similar mechanisms could be utilized by *Smi* and *Smi2*;  
81 nevertheless, empirical tests to validate this hypothesis are missing.

82 In Brassica, the dominance hierarchy of *S*-alleles is proposed to rely on the combinatorial effect of  
83 single nucleotide variants at the *Smi* and *Smi2* sequences and at their respective *SCR* target sites (8).  
84 In Arabidopsis, where more dominance classes exist, the dominance network is hypothesized to rely  
85 on two properties: i) recessive alleles carry more target sites and these sites are more generalist (*i.e.*  
86 they are targeted by more sRNA precursors than dominant alleles); and ii) sRNA precursors of dominant  
87 alleles are more generalist (*i.e.* they target a higher number of alleles than sRNA precursors in less  
88 dominant alleles) (6). Despite this observation, the molecular features allowing sRNA precursors to  
89 target multiple alleles remain to be identified. In addition, among the eight families of sRNA precursors  
90 previously predicted to control the hierarchy in Arabidopsis, only one has been subject to formal  
91 experimental validation.

92 In this study, we aimed at identifying the molecular pathway underlying the sRNA-mediated  
93 dominance interactions between *S*-alleles in Arabidopsis. To achieve this, we engineered and  
94 characterized a series of *A. thaliana* lines recapitulating the self-incompatibility phenotype, and used  
95 these lines to show that two different sRNA precursors from dominant *A. halleri* *S*-loci trigger a decrease  
96 in abundance of recessive *SCR* transcripts. We thus provide a proof of concept that this heterologous  
97 approach can be used to validate sRNA precursor function, both at the phenotypic and at the molecular  
98 level. This system also allowed us to investigate if key components of the canonical RdDM pathway,  
99 such as the ARGONAUTE proteins AGO4 and AGO6, and subunits of the DNA-dependent RNA  
100 polymerases POLIV and POLV, are required for the repression of *SCR* transcript by *S*-locus sRNA  
101 precursors, and our analyses suggest that the two selected sRNA precursors do not act through the  
102 canonical RdDM pathway.

103 Additionally, and in order to better understand the role of sRNA precursors genes in the extensive  
104 and complex *S*-locus dominance network, we characterized their molecular features, revealing that  
105 despite their structural similarity to miRNAs genes, they rather resemble their evolutionary precursors:  
106 proto-miRNAs (30, 31). Proto-miRNAs loci are inverted repeats present in several plant genomes that  
107 have acquired transcriptional competence, forming long hairpins which can be diced by multiple Dicer-  
108 like proteins (DCLs), leading to a complex population of sRNAs. There is little evidence that such a  
109 sRNA population could cause substantial effects in gene expression, at least through the canonical  
110 PTGS pathway (32–36); therefore, these structures have been mostly regarded as substrates for  
111 selection to act on, rather than having immediate biological significance. Notwithstanding, we show that

112 S-locus proto-miRNA genes are able to influence the transcript abundance of *SCR* alleles, affect the  
113 pollen germination rate and, consequently, mediate S-allele dominance interactions. S-locus proto-  
114 miRNAs produce a large population of sRNAs with different sizes and sequences, which are important  
115 to maximize the number of targeted alleles, thus allowing one allele to be dominant over many others.  
116 Overall, our results provide a unique window into the function and evolution of proto-miRNAs,  
117 constituting the first example of a biological system that depends on their action.

## 118 Results

119 *The sRNA precursors genes Ah04mir1887 and Ah20mirS3 control dominance interactions at the S-*  
120 *locus by decreasing the transcript level of the recessive SCR01 allele*

121 To study in detail the molecular mechanisms controlling the activity of dominance modifiers in the  
122 self-incompatibility system of *A. halleri*, we first recreated the self-incompatibility reaction of *A. halleri* in  
123 the selfing species *A. thaliana*, since the latter is a more tractable study model. To do this, we expressed  
124 the *SRK01* receptor and its cognate *SCR01* ligand in female and male parents, respectively (6, 37, 38).  
125 Manual crosses between these plants revealed that the germination of pollen from a fully fertile *SCR01*-  
126 expressing male was abolished when deposited on pistils of an *SRK01*-expressing female, the hallmark  
127 of a successful self-incompatible reaction (Fig. 1C and D, wt panels). Given the challenges faced in  
128 prior studies to reconstruct this reaction in the Col-0 ecotype, all *A. thaliana* plants used in this study  
129 are of the C24 ecotype, where the self-incompatibility can be faithfully recreated as shown both in this,  
130 and other studies (37, 39, 40).

131 Among the eight known families of *A. halleri* sRNA precursors loci, only one - mirS3 - has been  
132 experimentally validated to regulate dominance between two distinct S-alleles: specifically, the  
133 dominance of allele Ah20 over Ah01, mediated by Ah20mirS3 sRNAs, which target the *SCR01* intronic  
134 region (Fig. 1B) (6). In this study, we set out to examine if the yet uncharacterized mir1887 family could  
135 have a similar function. For this, we focused on testing the role of the sRNA precursor Ah04mir1887 in  
136 controlling the dominance of allele Ah04 over Ah01, whose sRNAs are predicted to target the 5' region  
137 of *SCR01* (Fig. 1A). We thus generated transgenic lines expressing Ah04mir1887 and introgressed this  
138 transgene into an *SCR01*-expressing male, and used these plants to pollinate *SRK01*-expressing  
139 females, testing whether the presence of the Ah04mir1887 sRNA precursor locus could transform the  
140 previously incompatible interaction between *SCR01* and *SRK01* into a compatible one (Fig. 1C). In the  
141 particular case of Ah20mirS3, the transgenic line previously generated by Durand et al. (6) and used  
142 here expresses both Ah20mirS3 and Ah20mir1887 because of their close physical proximity within the  
143 Ah20 allele (Fig. 1B). To specifically isolate the impact of Ah20mirS3 in the dominance interaction of  
144 Ah20 over Ah01, we developed the *SCR01*<sup>mir1887</sup> transgenic line, which contains five point mutations  
145 designed to fully disrupt the slight homology observed between some Ah20mir1887 sRNAs and *SCR01*  
146 (Fig. 1B; Fig. 1 S1). We then used this modified *SCR01* line for all subsequent tests on the functional  
147 role of Ah20mirS3.

148 We observed that combining either Ah04mir1887 (Interaction A), or Ah20mirS3 (Interaction B) with  
149 *SCR01* results in a switch from an incompatible to a compatible pollen germination phenotype, which  
150 is accompanied by a dramatic decrease in the abundance of the *SCR01* transcript, as measured by

151 RT-qPCR (**Fig. 1C and D, wt panels**). This shows that both Ah04mir1887 and Ah20mirS3 act as  
152 dominance modifiers in their respective allelic interactions, and that they accomplish this by decreasing  
153 *SCR01* transcript abundance, thus preventing SRK-SCR cognate recognition and allowing pollen  
154 germination to occur. Moreover, these results demonstrate that in Arabidopsis, *SCR* transcript  
155 abundance can be modulated by multiple sRNA precursor families that use distinct target sites in the  
156 *SCR* locus.

#### 157 *Decrease in SCR01 transcript abundance does not depend on RdDM*

158 Next, we aimed at elucidating the exact molecular mechanism responsible for the reduction of  
159 *SCR01* transcript abundance triggered by the two *S*-locus sRNA precursors. Given their resemblance  
160 to miRNAs genes, one potential scenario is that sRNAs produced by Ah04mir1887 and Ah20mirS3  
161 could trigger post-transcriptional gene silencing (PTGS) of the *SCR01* transcript. However, since these  
162 sRNAs are predicted to target the 5' region and intronic region of *SCR01*, and since the substrate of  
163 PTGS is generally thought to be the mature mRNA (**41**), this does not support PTGS as the pathway  
164 responsible for the decrease in *SCR01* transcript abundance. An alternative scenario is that  
165 Ah04mir1887 and Ah20mirS3 sRNAs could trigger the RNA-directed DNA methylation pathway  
166 (RdDM), resulting in the deposition of DNA methylation in their respective *SCR01* target sites, and  
167 reduced *SCR01* transcription. Interestingly, previous studies have established that the silencing of  
168 recessive alleles in the homologous self-incompatibility system of *Brassica* relies on sRNA-mediated  
169 DNA methylation, although the specific pathway underlying this process remains elusive (**7, 8, 20**).  
170 Thus, we focused our attention on the RdDM pathway, and set out to test if mutating key components  
171 of this pathway would prevent Ah04mir1887 and Ah20mirS3 from reducing *SCR01* transcript  
172 abundance. To circumvent the absence of RdDM mutant alleles in the C24 ecotype, we generated and  
173 validated C24 *ago4 ago6* CRISPR/Cas9 mutant alleles (**Fig. 1 S2**), and used the previously published  
174 C24 *polIV* and *polV* knock-out alleles (**42**).

175 To our surprise, none of the RdDM mutants impaired the activity of the tested sRNA precursors. In  
176 the case of Ah04mir1887, *SCR01* transcript abundance is reduced in all mutant backgrounds to a  
177 similar level as in wild-type, which is reflected by a compatible pollen germination phenotype (**Fig 1C**).  
178 Similarly, the presence of Ah20mirS3 is associated with a decrease in *SCR01* transcript in all mutant  
179 backgrounds, showing that this sRNA precursor is still functional (**Fig. 1D**). Remarkably, this is not  
180 reflected in a compatible pollen germination phenotype since the basal *SCR01* transcript level is  
181 increased in both *polIV* and *polV* mutant backgrounds when compared with wild-type (**Fig. 1D, polIV**  
182 **and polV panels vs. wt panel**), pointing to an effect of these mutants on the expression of the *SCR01*  
183 transgene. This suggests that the *SCR01<sup>\*mir1887</sup>* construct might be intrinsically targeted for repression  
184 by RdDM, a phenomenon often observed in transgenic lines (**43**) (note that the *SCR01* and  
185 *SCR01<sup>\*mir1887</sup>* transgenes used in Interaction A and B, respectively, correspond to two independent  
186 transgenic lines with different insertion sites; **see Methods section**). As such, and even though the  
187 addition of Ah20mirS3 in these mutant backgrounds strongly reduces *SCR01* transcript abundance, it  
188 does not do so sufficiently to compromise the self-incompatibility reaction. Indeed, our data shows that

189 a strong reduction of *SCR01* transcript abundance is required for a compatible pollen germination  
190 phenotype, since even very low levels of transcript can trigger an incompatibility reaction (**Fig. 1 S3**).

191 To further assess if the decrease in *SCR01* transcript abundance triggered by sRNA precursors is  
192 associated with deposition of DNA methylation at *SCR01* target sites, we used the Bisulfite Amplicon  
193 Sequencing technique - BSAS (**28, 44**), which relies on PCR enrichment of regions of interest using  
194 bisulfite converted DNA as a template. *SCR01* expression is restricted to the tapetum cells of immature  
195 floral buds, which are challenging to manually isolate in *A. thaliana* due to their small size and  
196 embedment within other floral tissues. To overcome this limitation, we performed BSAS on 8-10  
197 biological replicates of *SCR01* transgenic lines both in the presence and absence of the sRNA precursor  
198 transgenes (**Fig. 1 S4**), and generated an average coverage of 1400x for each cytosine located in the  
199 *SCR01* target sites and in the 800 bp around them (**Table S3**). This allows us to detect minute DNA  
200 methylation differences between samples, thus compensating for the difficulty in enriching for tapetum  
201 tissue. We observed strong DNA methylation levels at the *SCR01* target sites (**Fig. 1 S4**); however, no  
202 significant changes were detected in the presence of Ah04mir1887 or Ah20mirS3, suggesting that the  
203 decrease in *SCR01* transcript abundance is not mediated by changes in DNA methylation at the target  
204 sites (**Fig. 1 S4**), which is in line with the absence of effects in the tested RdDM mutants. The observed  
205 high methylation levels at *SCR01* target sites, coupled with the specific *SCR01* expression in tapetum  
206 cells, led us to hypothesize that these methylation levels might represent the methylation status of  
207 surrounding floral tissues rather than specifically reflecting tapetum methylation. To investigate this  
208 further, we categorized BSAS reads based on the fraction of methylated cytosines within a single read.  
209 Our analysis revealed that a significant proportion of reads exhibited complete methylation, with a  
210 notable fraction showing little to no methylation (**Fig. 1 S5**). We postulate that this subset of  
211 unmethylated reads may originate from tapetum cell DNA, where *SCR01* is actively expressed.  
212 Notwithstanding, no significant changes in unmethylated read abundance were detected in the  
213 presence of Ah04mir1887 or Ah20mirS3, again suggesting these sRNA precursors have no impact on  
214 DNA methylation.

215 Together, these results indicate that neither PTGS nor RdDM, at least in their canonical forms,  
216 appear to be responsible for the reduction in *SCR01* transcript abundance triggered by S-locus sRNA  
217 precursors. Despite this, both Ah04mir1887 and Ah20mirS3 strongly and robustly decrease *SCR01*  
218 transcript levels, which we show is necessary to abolish the self-incompatibility reaction.

#### 219 *Ah04mir1887* and *Ah20mirS3* loci strongly resemble proto-miRNAs genes

220 S-locus sRNA precursors, including those of the mir1887 and mirS3 families studied here, consist  
221 of inverted repeats that are transcribed and predicted to form a RNA hairpin structures containing a 5'  
222 arm, a terminal loop and a complementary 3' arm (**Fig. 2A**) (**6**), akin to the structures formed by  
223 precursor miRNA (pri-miRNA) molecules. To delve deeper into the molecular properties of  
224 Ah04mir1887 and Ah20mirS3, and to explore potential non-canonical pathways through which these  
225 precursors may act, we conducted sRNA sequencing on *A. thaliana* plants harboring the Ah04mir1887  
226 and Ah20mirS3 transgenes. We observed that dicing of S-locus sRNA precursor hairpins leads to  
227 numerous sRNAs – at least 106 in the case of Ah04mir867, and 224 in the case of Ah20mirS3 (**Fig.**



228 **2B-D)**. These sRNAs range from 18 to 25 nt in length, with no clear predominance of a specific length  
229 **(Fig. 2 S1 A-B)**, and their abundances vary by several orders of magnitude **(Fig. 2B, Fig. 2 S1 C-D)**.

230 Given the large population of sRNAs produced by a single precursor, we aimed at evaluating the  
231 functional relevance of each sRNA in controlling the abundance of the *SCR01* transcript. For this, we  
232 used a previously published algorithm **(6)**, which attributes a targeting score based on the homology  
233 between a single sRNA and the *SCR* genomic region. A previous study has demonstrated that in  
234 multiple *A. halleri* *S*-locus heterozygotes, a targeting score equal to or above 18 is associated with  
235 phenotypic dominance and a sharp decrease of the recessive *SCR* transcript level **(18)**. Using this  
236 algorithm, we found that 12 out of the total 106 sRNAs produced by Ah04mir1887, and 33 out of the  
237 total 224 sRNAs produced by Ah20mirS3, have a base-pairing score of  $\geq 18$  and are thus predicted to  
238 target *SCR01*. This subset of putatively functional sRNAs displays a variety of sizes and expression  
239 levels **(Fig. 2B, see grey box)**.

240 We additionally performed ARGONAUTE immunoprecipitation (IP) assays, followed by sRNA  
241 sequencing in the Ah04mir1887 and Ah20mirS3 *A. thaliana* transgenic lines, which showed that  
242 functional sRNAs produced from either precursor can be loaded into AGO1 (characteristic of the PTGS  
243 pathway) and AGO4 (characteristic of the RdDM pathway) **(Fig. 2 C-D)**. Notwithstanding this, a large  
244 part of these functional sRNAs (75% in Ah04mir1887 and 48% in Ah20mirS3) are not found to be loaded  
245 in either of these two AGO proteins, again consistent with our observations that both the canonical  
246 PTGS and RdDM pathways are unlikely to underlie the activity of these sRNA precursors. We also  
247 noted that the majority of sRNAs loaded in AGO1 have a 5' terminal uracil, reflecting the known AGO1  
248 preference for binding these sRNAs **(45)**. However, sRNAs loaded into AGO4, or with an unknown  
249 loading pattern, do not show a significant bias towards a specific 5' nucleotide **(Fig. 2 S1 E-F)**. Overall,  
250 these results show that the predicted functionality of a given sRNA does not seem to be tied to a specific  
251 set of molecular features.

252 Despite the structural similarity between *S*-locus sRNA precursors and miRNAs, and the existence  
253 of shared features, such as the presence of mature 21-nt sRNAs and loading of sRNAs into AGO1, it  
254 is evident that *S*-locus sRNA precursors deviate from the typical characteristics associated with  
255 canonical miRNAs, and cannot be characterized as such. Instead, our data show that Ah04mir1887  
256 and Ah20mirS3 highly resemble proto-miRNAs, the evolutionary predecessors of miRNAs: they have  
257 an unusually staggered hairpin processing pattern with numerous sRNAs derived from both the 5' and  
258 3' arms, and differ in size, abundance and association with AGO proteins. Proto-miRNAs are thought  
259 to have no target genes and/or to cause no significant effect in transcript abundance **(30, 31, 33–36)**.  
260 This is not the case for Ah04mir1887 and Ah20mirS3 since their effect on *SCR01* transcript levels is  
261 evident and necessary to modulate the self-incompatibility phenotype **(Fig. 1C and D, wt panels)**.

#### 262 *proto-miRNA features allow S-locus sRNA precursors to target multiple S alleles*

263 We hypothesized that this heterogeneity in hairpin processing pattern could play an important role  
264 in the context of the complex *S*-locus dominance hierarchy by allowing a single sRNA precursor to exert  
265 dominance over multiple other alleles. To test this, we expanded our analysis of the function of

266 Ah04mir1887 and Ah20mirS3 beyond the interaction with allele Ah01, and looked at their role within  
267 the *A. halleri* dominance network (**Fig. 3**).

268 We compiled previously published phenotypic data obtained by controlled reciprocal crosses  
269 between different *A. halleri* *S*-locus heterozygotes to determine the position of allele Ah20 and Ah04 in  
270 the dominance hierarchy of a total of 15 *S*-alleles (**Fig. 3**) (**6, 18, 46**). From this, we could infer that  
271 Ah20 is dominant over 13 other *S*-alleles, with 10 of these interactions relying on the action of  
272 Ah20mirS3 (**Fig. 3**), as determined by the sRNA-*SCR* targeting algorithm described above. Thus,  
273 Ah20mirS3 is not only able to mediate the repression of *SCR01*, but also the repression of at least nine  
274 other distinct *SCR* alleles. Similarly, Ah04mir1887 mediates the repression of at least five different *SCR*  
275 alleles that are recessive to Ah04 (**Fig. 3**).

276 To test if the large population of sRNAs produced by these *S*-locus precursors is important in  
277 repressing multiple target alleles, we examined the role of each individual sRNA produced by both  
278 Ah04mir1887 and Ah20mirS3 (**Fig. 4**): 11% of Ah04mir1887 sRNAs (**Fig. 4A, cluster 1**), and 26% of  
279 Ah20mirS3 sRNAs (**Fig. 4B, cluster 1**) are predicted to target at least one *SCR* allele within the inferred  
280 dominance network presented in **Figure 3**. As is the case in Interaction A and B, there does not seem  
281 to be a clear correlation between abundance and sRNA functionality in any of the analyzed interactions  
282 (**Fig. 2B and Fig. 4A-B, see cluster 1 vs. cluster 2**). Nevertheless, sRNAs that are predicted to be  
283 functional have a larger size than those that are not (median of 23 nt vs. 21 nt), and in the case of  
284 Ah20mirS3, this is also associated with a larger number of sRNAs being loaded in AGO proteins (**Fig.**  
285 **4A-B; see cluster 1 vs. cluster 2**). Additionally, we can observe that most *S*-alleles are redundantly  
286 targeted by multiple sRNAs (**Fig. 4**).

287 When investigating in more detail the positional correspondence between sRNAs and the  
288 Ah04mir1887/Ah20mirS3 precursor hairpin structure we noticed that sRNAs are produced in similar  
289 amounts from both the 5' and the 3' arms of the hairpin (**Fig. 5**). Each hairpin arm produces numerous  
290 sRNAs that show a staggered pattern, suggesting imprecise dicing and resulting in a population of  
291 sRNAs that show small base-pair differences between them. In the case of Ah04mir1887, only the  
292 sRNAs produced from the 5' hairpin arm are predicted to be functional against the *S*-alleles tested here  
293 (**Fig. 5A**). Remarkably, in the case of Ah20mirS3, each hairpin arm shows functional specialization  
294 since sRNAs from the 5' arm target all allele classes, except Class IV; while sRNAs from the 3' arm  
295 exclusively target Class IV and Class III alleles. Together, these observations suggest that the high  
296 number and heterogeneous processing of sRNAs is a conserved and important feature of *S*-locus sRNA  
297 precursors which allows targeting of multiple alleles and thus underlies their generalism.

## 298 Discussion

299 *S*-locus sRNA precursors act through non-canonical gene silencing pathways

300 Our characterization of self-incompatibility transgenic lines of *A. thaliana* contributes additional  
301 evidence to the work of Durand et al. (**6**): collectively, these studies establish *A. thaliana* as a tractable  
302 system to study dominance interactions among *A. halleri* *S*-alleles and validate the role of two distinct  
303 sRNA precursor families in these interactions, yet the precise mechanisms underlying their action  
304 remain elusive. Our study reveals that some of the sRNAs generated in the processing of the *S*-locus

305 sRNA precursors have a length of 24 nt and are loaded into AGO4, potentially indicating a role for  
306 RdDM. Nevertheless, *ago4 ago6*, *polIV* and *polV* mutations did not significantly affect the action of *A.*  
307 *halleri* S-locus sRNA precursors, suggesting that canonical RdDM is not the primary mechanism  
308 through which Ah20mirS3 and Ah04mir1887 mediate the reduction in *SCR01* transcript levels.  
309 Considering their hairpin structure and their ability to also generate AGO1-loaded 21-nt sRNAs, it could  
310 also be envisioned that these precursors might act through PTGS. Still, the location of target sites in  
311 the 5' region and intronic regions of *SCR01* poses a conundrum, since PTGS is reported to occur in the  
312 cytoplasm, using mature mRNA as a cleavage substrate (41). Notably, in animals, several lines of  
313 evidence point to miRNAs acting both in the cytoplasm and in the nucleus, with effects at the  
314 transcriptional and post-transcriptional levels (47–49). In plants, it has been observed that an inverted  
315 repeat targeting the intron of the soybean gene *FAD12-1A* triggers cleavage of its precursor mRNA  
316 (pre-mRNA) in the nucleus, which is accompanied by accumulation of siRNAs of the cleaved pre-mRNA  
317 (50). Additionally, all main miRNA processing components, including AGO1, are present in the nucleus  
318 (51–53). Thus, given the independent slicer activity of AGO1 (54), it could be hypothesized that this  
319 ARGONAUTE uses sRNAs to target the *SCR01* pre-mRNA in the nucleus, promoting its cleavage co-  
320 transcriptionally. Alternatively, the interaction between AGO1 and the *SCR* pre-mRNA could promote  
321 disassembly of the Poll II complex leading to the arrest of transcription, as has been suggested to occur  
322 in some miRNA genes upon salt stress (55). Even though we could not detect any siRNAs  
323 corresponding to the cleavage of the *SCR01* pre-mRNA in any of our sRNA sequencing experiments  
324 (data not shown), future experiments should investigate the possibility that the sRNA precursors at the  
325 S-locus could act through nuclear PTGS or through transcriptional arrest.

326 Our results that Ah20mirS3 and Ah04mir1887 work through an RdDM-independent pathway contrast  
327 with previous observations in Brassica: in this latter system, the sRNA precursors *Smi* and *Smi2* are  
328 associated with deposition of DNA methylation at *SCR* target sites and a decrease in *SCR* expression  
329 (7, 8, 20). DNA methylation may result directly from the incorporation of *Smi* and *Smi2* sRNAs into  
330 AGO4-clade proteins, activating RdDM. Alternatively, as suggested by Finnegan et al. (56), it could  
331 depend on *Smi*-induced cleavage of an antisense transcript at the *SCR* locus, generating sRNAs that  
332 could subsequently be co-opted into the RdDM pathway. Even though empirical investigations are still  
333 necessary to test these hypotheses, the cumulative evidence from these and our study suggests that  
334 different sRNA precursors within the Brassicaceae self-incompatibility system can elicit distinct effector  
335 pathways, that nevertheless converge into the same molecular phenotype: reduction of *SCR* transcript  
336 abundance. This would imply that these gene regulatory elements have an outstanding molecular  
337 flexibility, despite functioning within the specific evolutionary constraints of the self-incompatibility  
338 system. Thus, it cannot be excluded that different *A. halleri* sRNA precursors, as well as precursors  
339 from other self-incompatibility species such as *A. lyrata* and *Capsella grandiflora* could co-opt different  
340 gene regulatory pathways. Future investigations in these different species, using heterologous  
341 expression systems as done in this study, or in the native context whenever genetic transformation is  
342 feasible, will hopefully answer this outstanding question.

343 *The complex S-allele dominance network relies on the unique features of S-locus sRNA precursors*

344 Durand et al. (6) previously showed that the S-locus sRNA precursors of *A. halleri* have high  
345 generalism, and suggested that dominant S-alleles in particular, rely on this ability to target multiple  
346 recessive alleles, rather than on carrying an increased number of sRNA precursors. Our study uncovers  
347 the molecular basis for this generalism, and shows that the apparent haphazardous processing of S-  
348 locus sRNA precursors is required to generate numerous functional sRNAs that target multiple alleles  
349 cooperatively: remarkably, all SCR alleles analyzed in this study are targeted by multiple sRNAs of a  
350 single sRNA precursor, and/or are targeted by distinct sRNA precursors. This contrasts with the less  
351 complex dominance hierarchy of Brassica, where fewer allelic interactions are present, and a single  
352 functional sRNA molecule *per* precursor has been suggested to mediate all dominance relationships  
353 (7, 8). It is important to note that while we investigated a total of 15 S-alleles in this study, at least 43  
354 are known to exist in *A. halleri* (15), and > 80 in the closely related *Capsella grandiflora* (22). Therefore,  
355 in this study we most likely underestimated the fraction of functional sRNAs in each precursor, since  
356 those that are non-functional in the allelic interactions explored here could potentially have a role in  
357 interactions that have not yet been characterized.

358 Overall, the evidence suggests that collaborative targeting, either by multiple sRNAs of a single  
359 precursor, or by sRNAs from distinct sRNA precursors, is an essential feature of this system: firstly, it  
360 is essential for granting an S-allele its complete targeting spectrum across the many S-alleles  
361 segregating in the population; and secondly, the use of multiple sRNAs with distinct molecular features  
362 could trigger diverse effector pathways acting on the same target allele, which could perhaps underlie  
363 the robust reduction of recessive SCR transcripts observed in all allelic interactions studied thus far  
364 (this study and (18)).

365 Intriguingly, this mirrors the collaborative non-self recognition model observed in Solanaceae, where  
366 each self-incompatibility allele contains multiple SLF genes. SLF proteins play a crucial role in  
367 recognizing and removing the toxic non-self female S-RNase, which otherwise inhibits pollen  
368 germination. In this system, the presence of multiple SLF genes within a single S-locus allele, each  
369 capable of recognizing a distinct S-RNase allele, coupled with the ability of some SLF genes to  
370 recognize and detoxify multiple S-RNases (generalism), is crucial. This collaborative functionality  
371 among SLF genes ensures a comprehensive recognition spectrum for a given self-incompatibility allele,  
372 thereby preserving the intricate inter-allelic interaction network within Solanaceae (57). While in  
373 Solanaceae, the generalism of SLF proteins likely depends on their ability to establish protein-protein  
374 interactions with distinct S-RNases (58), in *A. halleri*, sRNA precursor generalism relies on their unique  
375 molecular features, specifically their ability to produce numerous sRNAs with distinct targeting spectra.

376 *S-locus sRNA precursors provide a window into the miRNA evolution continuum*

377 Although sharing some structural similarities with miRNAs, the features of S-locus sRNA precursors  
378 more closely resemble those of proto-miRNAs. We hypothesize that the specific selective constraints  
379 of the S-locus dominance network have likely contributed to the persistence of proto-miRNA features  
380 over extended evolutionary times, and prevented their progression towards canonical miRNA features,  
381 as this would likely render them unable to effectively target multiple S-alleles. It would be interesting to

382 test if apparently stabilized proto-miRNAs are also used in other biological systems where the  
383 expression of multiple genes/alleles needs to be coordinated by a single regulator and under similar  
384 evolutionary constraints as those acting on the *S*-locus. The mimicry of wing patterns in *Heliconius*  
385 butterflies would be an interesting model system to test this, since balancing selection has favored the  
386 appearance of numerous alleles of the supergene controlling wing pattern, and dominance relationships  
387 between them are prevalent (59, 60).

388 To our knowledge, the *A. halleri* *S*-locus sRNA precursors represent the first instance where proto-  
389 miRNA-like features are biologically relevant, and actively selected for. And despite not conforming to  
390 canonical gene regulatory pathways, our data shows that *S*-locus proto-miRNAs exert a considerable  
391 impact on gene expression. Prior investigations of proto-miRNA function have primarily evaluated the  
392 functionality of these structures in the light of canonical miRNA pathways, often concluding that their  
393 functional impact is negligible (32–36). Based on our findings, we propose that further exploration is  
394 necessary to determine whether proto-miRNAs can employ non-canonical pathways to regulate gene  
395 expression in diverse biological contexts beyond the *S*-locus. Investigating the function and mode of  
396 action of proto-miRNAs in different organisms will contribute not only to better understand their potential  
397 roles, but also to uncover the key steps in the emergence of proto-miRNAs and their transition into  
398 canonical miRNAs.

## 399 **Materials and Methods**

### 400 *Plant material & growing conditions*

401 Wild-type (wt) and transgenic *Arabidopsis thaliana* plants used in this study were grown in a  
402 greenhouse, using standard growth conditions (16 h light/8 h dark; 110  $\mu\text{mol/s/m}^2$ ; 21°C; 70%  
403 humidity). Seeds were germinated directly on soil, or in MS-medium (0.43% MS-salts, 0.8% Bacto agar,  
404 0.19% MES hydrate, and 1% sucrose) when antibiotic selection was necessary for transgene selection.

405 Because the self-incompatibility reaction cannot be faithfully reconstructed in the *A. thaliana* Col-0  
406 ecotype (37–40) all wild-type and mutant plants used in this study were of the C24 ecotype. *polIV* and  
407 *polV* correspond to the previously published and validated *rdm5* (*nrdp1a*) and *rdm6* (*nrdp1b*) mutant  
408 alleles (42). These two mutants were generated in the *ros1* mutant background, which was removed in  
409 our study by backcrossing to wt C24 plants. The *ago4 ago6* double mutant was generated in the C24  
410 background using CRISPR/Cas9 (see section **Generation and validation of CRISPR/Cas9 mutants**  
411 **for more details**).

412 To obtain the *S*-allele sequences of Ah65, Ah60, Ah03, Ah33 and Ah19, five *Arabidopsis halleri*  
413 individuals containing a combination of these alleles were collected from diverse natural European  
414 populations (61), and grown under standard greenhouse conditions in order to obtain plant material for  
415 DNA extraction and sequencing (see the section **DNA extraction, Nanopore sequencing and**  
416 **assembly of S-alleles for more details**). Following Goubet et al. (23), we further obtained the *S*-allele  
417 sequence of Ah01, Ah02 and Ah25 by constructing and screening BAC libraries from fresh leaf samples,  
418 and fully sequencing the positive clones using PACBIO.

## 419 Cloning

420 Several transgenic lines were used in this study: lines carrying SRK01, SCR01, and the sRNA  
421 precursor Ah20mirS3 were cloned as detailed previously (6).

422 The precursors Ah20mir1887 and Ah20mirS3 are in very close physical proximity within the Ah20  
423 allele, and the previously published Ah20mirS3 line contains both sRNA precursors (6). While  
424 performing sRNA sequencing of AGO-IP experiments (see section **ARGONAUTE**  
425 **immunoprecipitation and sRNA sequencing**), we noticed that some Ah20mir1887 sRNAs show  
426 homology to the 5' region of SCR01. The base-pairing score of these predicted interactions varies  
427 between 12.5 and 18.5, depending on the sRNAs (Fig. 1 S1, see the section **Target site inference**  
428 **for Ah04mir1887 and Ah20mirS3 for more details on these scores**). To confidently isolate the effect  
429 of Ah20mirS3 in the dominance interaction of Ah20 over Ah01 (Interaction B), we introduced five point  
430 mutations in the SCR01<sup>\*mir1887</sup> sequence that disrupt the homology with the Ah20mir1887 sRNAs,  
431 rendering the Ah20mir1887 target site non-functional. The SCR01<sup>\*mir1887</sup> transgene was generated by  
432 performing site-directed mutagenesis on the already published SCR01 clone (6), using the primers  
433 detailed in table S1. The resulting amplicon was then recombined into the entry vector pDONR-Zeo,  
434 and subsequently recombined into the pB7m34GW destination vector using the 3-fragment Gateway  
435 Cloning System (Invitrogen), including 5' and 3' mock sequences, as detailed previously (6).

436 Ah04mir1887 was cloned by amplifying a region spanning this sRNAs precursor and containing 2  
437 kb of upstream and downstream sequences from a BAC clone carrying allele Ah04, using the primers  
438 detailed in table S1. The resulting amplicon was then recombined into the entry vector pDONR-Zeo,  
439 and subsequently recombined into the pH7m34GW destination vector using the 3-fragment Gateway  
440 Cloning System (Invitrogen), including 5' and 3' mock sequences, as detailed previously (6).

441 Arabidopsis thaliana C24 plants were transformed using the floral dip method (62), and transgenic  
442 lines were selected using the appropriate antibiotics.

## 443 Generation and validation of CRISPR/Cas9 mutants

444 Given the aforementioned limitation that the self-incompatibility reaction of *A. halleri* can only be  
445 reconstituted in the *A. thaliana* C24 ecotype (37–40), we created the *ago4 ago6* RdDM mutants in this  
446 specific ecotype using CRISPR/Cas9. The expression cassette of pCBC-DT1T2 (63) was amplified  
447 using two overlapping forward and reverse primers that contained two gene-specific sgRNA sequences  
448 (Fw1: 5'-ATATATGGTCTCGATTGNNNNNNNNNNNNNNNNNNNNNGTT-3'; Fw2: 5'-  
449 TGNNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGC-3'; Rv1: 5'-  
450 ATTATTGGTCTCGAAACNNNNNNNNNNNNNNNNNNNNNCAA-3'; Rv2: 5'-  
451 AACNNNNNNNNNNNNNNNNNNNNNCAATCTCTTAGTCTACTAC-3', where N corresponds to the  
452 gene-specific gRNAs (table S1), and underlined regions correspond to BsaI restriction sites). The  
453 amplicons obtained from this PCR reaction were gel-isolated and inserted into pHEE401E (64) using  
454 BsaI and T4 ligase (Thermo Fisher Scientific). The pHEE401E vectors containing gRNAs against AGO4  
455 and AGO6 were transformed into the *Agrobacterium tumefaciens* strain GV3101. The floral-dip method  
456 (62) was then used to transform wt C24 *A. thaliana* plants.

457 To screen for T1 mutant plants, the regions targeted by the gRNAs were Sanger sequenced to  
458 identify *ago4 ago6* mutants with premature stop codons in both genes (**Fig. 1 S2**). In the T2 generation,  
459 double homozygous mutants that did not carry the pHEE401E were selected by genotyping (**see table**  
460 **1 for primers**).

461 To confirm that *ago4 ago6* mutations impair DNA methylation, a Chop-PCR assay was performed  
462 (**65**). Briefly, this assay uses genomic DNA treated with the methylation-sensitive enzyme HaeIII as a  
463 template for PCR. Primers targeting regions known to have AGO4/AGO6-dependent DNA methylation,  
464 as well as control unmethylated regions are amplified (**see table S1 for primers**). Using this method,  
465 we could determine that the *ago4 ago6* mutant generated in this study shows reduced DNA methylation  
466 levels, as has been previously published for other mutant alleles of these genes (**Fig. 1 S2 C**) (**66**).

#### 467 *Pollen germination assays*

468 All lines used for the reconstruction of the self-incompatibility reaction in *A. thaliana* were used in  
469 hemizygous state to mimic the abundance of sRNA precursors, and *SCR/SRK* in *A. halleri* individuals  
470 heterozygous at the S-locus. Genotyping of each transgenic line/mutant was done by classical PCR, or  
471 using the KASPar assay (**67**) (**table S1**). Paternal plants hemizygous for *SCR01* and sRNA precursors  
472 were obtained by crossing *SCR01* to Ah04mir1887 homozygous lines or by crossing *SCR01<sup>\*mir1887</sup>* to  
473 Ah20mirS3 homozygous lines.

474 To assess the self-incompatibility reaction at the phenotypic level, pollen germination assays were  
475 performed on manually crossed plants. Maternal plants hemizygous for *SRK01* were emasculated one  
476 day before anthesis, and manually pollinated 24h later with paternal plants hemizygous for *SCR01*,  
477 *SCR01<sup>\*mir1887</sup>*, *SCR01* Ah04mir1887 or *SCR01<sup>\*mir1887</sup>* Ah20mirS3. Pollinated pistils were transferred to  
478 fixing solution 6h after pollination and stained with aniline blue, as described before (**6**). These pistils  
479 were then mounted on microscope slides and imaged using a Zeiss AX10 fluorescence microscope,  
480 where we counted the number of germinated pollen grains present in each pistil.

481 To evaluate the effect of RdDM mutants on the activity of S-locus sRNA precursors we crossed the  
482 RdDM mutants *ago4 ago6*, *polIV* and *polV* with the previously described paternal plants carrying *SCR01*  
483 and sRNA precursors (*SCR01*, *SCR01<sup>\*mir1887</sup>*, *SCR01* Ah04mir1887 and *SCR01<sup>\*mir1887</sup>* Ah20mirS3). We  
484 then obtained plants hemizygous for the *SCR01* and sRNA precursor transgenes and homozygous for  
485 the mutations of interest and used these as paternal plants in crosses with maternal hemizygous *SRK01*  
486 plants, as described above.

487 As a control, all maternal and paternal lines used in this study were crossed to wt C24. Germination  
488 and seed set for all these genotypes were comparable to that of wt x wt crosses (**data not shown**),  
489 confirming that the transgenes and mutations had no detectable effect on plant fertility.

#### 490 *RT-qPCR*

491 To measure the abundance of *SCR01* transcripts, we isolated immature buds of plants hemizygous  
492 for *SCR01*, *SCR01<sup>\*mir1887</sup>*, *SCR01* Ah04mir1887 or *SCR01<sup>\*mir1887</sup>* Ah20mirS3. RNA was extracted using  
493 the NucleoSpin RNA Plus kit (Macherey-Nagel) and cDNA was synthesized with the RevertAid First  
494 Strand cDNA Synthesis Kit (Thermo Scientific). qPCR was performed using the iTaq Universal SYBR

495 Green Supermix (BioRad) in a Lightcycler 480 instrument (Roche), with the primers found in **table S1**.  
496 *ACT8* was used as a reference gene. Transcript abundance was quantified using the Pfaffl method  
497 **(68)**.

#### 498 *ARGONAUTE immunoprecipitation and sRNA sequencing*

499 AGO1 and AGO4 immunoprecipitation (IP) assays were performed on immature buds of plants  
500 carrying a homozygous Ah04mir1887 or Ah20mirS3 transgenes, using a previously published protocol  
501 **(69)**, with the following modifications: IPs were performed on 1.4 g of finely ground powder of frozen  
502 immature buds. The anti-AGO1 (AS09 527, Agrisera) and anti-AGO4 (AS09 617, Agrisera) antibodies  
503 were added at a 1/500 dilution for the incubation step and their immobilization was performed using  
504 Dynabeads Protein G (Invitrogen). The IP products were washed six times with PBS prior to sRNA  
505 extraction using the Trizol reagent (Ambion). In parallel, an aliquot of each input extract was used for  
506 total sRNA isolation using the Trizol LS reagent (Invitrogen), according to the supplier instructions. For  
507 each genotype three sRNA fractions were obtained: AGO1-IP, AGO4-IP and input.

508 These three sRNA fractions were subjected to an acrylamide gel-based size selection for sRNAs.  
509 These sRNAs were then used for TruSeq Small RNA (Illumina) library preparation and sequenced on  
510 a NextSeq 500 platform (Illumina), using 75 bp single-reads. Reads were trimmed, quality filtered, and  
511 size-selected using TrimGalore **(70)**. We followed a two-step mapping procedure with ShortStack  
512 (allowing for one mismatch) **(71)**, where we mapped sRNAs from each sample to the *A. thaliana* TAIR10  
513 genome (masking the *S*-locus region), and to the sequence of the corresponding *A. halleri* *S*-allele  
514 (Ah04, for samples carrying Ah04mir1887 and Ah20 for samples carrying Ah20mirS3). Only reads that  
515 mapped exclusively to the *S*-locus were kept for further analysis. sRNA read counts were normalized  
516 to rpm to allow comparison of abundance levels across samples.

517 AGO loading was inferred based on the ratio of reads found in the AGO1-IP fraction and the AGO4-  
518 IP fraction: a sRNA was considered to be loaded in AGO1 or AGO4 when more than 50% of its total  
519 reads were derived from either the AGO1-IP fraction or the AGO4-IP fraction, respectively. A given  
520 sRNA was classified as having an unknown loading when no reads were found in both AGO-IP  
521 fractions, or when the amount of reads found in the input fraction was superior to that found in any of  
522 the two AGO-IP fractions.

#### 523 *Target site inference for Ah04mir1887 and Ah20mirS3*

524 To identify sRNA target sites on *SCR* genomic regions, we first generated a database containing all  
525 sRNAs sequenced in the AGO-IP experiments described above. In parallel, a database containing *SCR*  
526 genomic regions of different *A. halleri* *S*-loci was compiled. This included alleles from all four dominance  
527 classes, and comprised previously published sequencing data **(see Data Availability section for a**  
528 **complete list of data sources)**. In addition to this, long read Nanopore sequences newly obtained  
529 from individuals carrying alleles Ah65, Ah60, Ah03, Ah33 and Ah19 were also added to this database  
530 **(see section Nanopore sequencing and assembly of S-alleles for more details)**.

531 Using the databases of sRNA sequences and *SCR* alleles, we applied a previously published  
532 homology detection algorithm **(6)** to find target sites. Functional sRNA are defined as those with a



533 homology score of  $\geq 18$ , since this score has been consistently shown to be associated with a decrease  
534 in *SCR* transcript abundance of the targeted allele (6, 18). The identified sRNA - *SCR* interactions were  
535 then represented in a heatmap (**Fig. 4**) using ComplexHeatmap (**72**), and in a circos plot (**Fig. 5**) using  
536 circlize (**73**).

#### 537 *DNA extraction, Nanopore sequencing and assembly of S-alleles*

538 To obtain the *S*-locus sequences of alleles Ah65, Ah60, Ah03, Ah33 and Ah19, 2 g of fresh leaves  
539 were collected from five different individuals carrying these alleles and flash-frozen. High molecular  
540 weight genomic DNA was extracted as described before (**74**). Long read sequences were obtained by  
541 Oxford Nanopore Technologies (ONT) flowcells v9. To polish the contigs obtained by long read  
542 sequencing Illumina data were also generated (with no step of PCR amplification to minimize  
543 sequencing bias).

544 For Nanopore library preparation, the smallest genomic DNA fragments were first eliminated using  
545 the Short Read Eliminator Kit (Pacific Biosciences). Libraries were then prepared according to the  
546 protocol 1D Native barcoding genomic DNA (with EXP-NBD104 and SQK-LSK109) provided by Oxford  
547 Nanopore Technologies. Depending on how many samples were pooled, 250 ng (pool of 9 samples) to  
548 1  $\mu\text{g}$  (pool of 4 samples) of genomic DNA fragments were repaired and end-prepped with the NEBNext  
549 FFPE DNA Repair Mix and the NEBNext Ultra II End Repair/dA-Tailing Module (New England Biolabs).  
550 Barcodes provided by ONT were then ligated using the Blunt/TA Ligase Master Mix (NEB). Barcoded  
551 fragments were purified with AMPure XP beads (Beckmann Coulter), then pooled and ONT adapters  
552 were added using the NEBNext Quick Ligation Module (NEB). After purification with AMPure XP beads  
553 (Beckmann Coulter), each library was mixed with the sequencing buffer (ONT) and the loading beads  
554 (ONT) and loaded on a PromethION R9.4.1 flow cell. In order to maintain the translocation speed, flow  
555 cells were refueled with 250  $\mu\text{l}$  Flush Buffer when necessary. Reads were basecalled using Guppy  
556 version 3.2.10, 4.0.11, 5.0.12, 5.0.13, 5.0.16 or 5.1.12. The Nanopore long reads were not cleaned and  
557 raw reads were used for genome assembly.

558 For Illumina PCR-free library preparation, 1.5  $\mu\text{g}$  of genomic DNA was sonicated to a 100–1500-bp  
559 size range using a Covaris E220 sonicator (Covaris). The fragments (1 $\mu\text{g}$ ) were end-prepped, and  
560 Illumina adapters (NEXTFLEX Unique Dual Index Barcodes, Perkin Elmer) were added using the Kapa  
561 Hyper Prep Kit (Roche). The ligation products were purified twice with 1X AMPure XP beads (Beckman  
562 Coulter). The libraries were then quantified by qPCR using the KAPA Library Quantification Kit for  
563 Illumina Libraries (Roche), and their profiles were assessed on an Agilent Bioanalyzer (Agilent  
564 Technologies). The libraries were sequenced on an Illumina NovaSeq 6000 instrument (Illumina) using  
565 150 base-length read chemistry in a paired-end mode.

566 Nanopore sequencing data were assembled using Necat (**75**) with a genome size of 240 Mbp and  
567 the remaining parameters left as default. Contigs produced by Necat were polished one time using  
568 Racon (**76**) with Nanopore reads, then one time with Medaka  
569 (<https://github.com/nanoporetech/medaka>, model r941\_prom\_hac\_g507) and Nanopore reads, and two  
570 times with Hapo-G v1.3.4 (**77**) and Illumina short reads. However, the cumulative sizes of our long-read  
571 assemblies were larger than the estimated 240 Mb, suggesting that the assembly size was currently

572 inflated by the presence of allelic duplications. Thus, HaploMerger2 (**78**) was run (Batch A twice to  
573 remove major misjoin and one Batch B) to generate a haploid version of each assembly.

574 To detect the *S*-locus region in the generated assemblies, blast searches were performed using the  
575 genes flanking the *S*-locus (*ARK3* and *U-box* (**23**)) as queries. This allowed us to precisely narrow down  
576 the genomic regions corresponding to the full Ah65, Ah60, Ah03, Ah33 and Ah19 *S*-alleles. We  
577 annotated *SRK* using blast searches, and following Goubet et al. (**23**), we used Fgenesh+ (**79**) to predict  
578 the genomic sequence of *SCR* based on a database of *SCR* protein sequences from already known *S*-  
579 alleles.

#### 580 *Phenotypic determination of the S-loci dominance hierarchy*

581 The *S*-locus dominance hierarchy in **Fig. 3** is based on a compilation of controlled reciprocal crosses  
582 between different *A. halleri* individuals (**6, 18, 46, 80**). In cases where interactions between alleles of  
583 two different dominance classes had not been tested phenotypically, dominance was inferred based on  
584 the phylogenetic class of each allele, since interclass dominance interactions are predictable (**15, 24**).  
585 In the case of intraclass relationships, such as for alleles Ah65, Ah60 and Ah33 (all belonging to class  
586 III), phenotypic dominance has not been determined yet, and thus the ordering of these alleles along  
587 the hierarchy remains arbitrary at this stage.

#### 588 *BSAS sequencing*

589 To measure DNA methylation in *SCR01* target sites we used the Bisulfite Amplicon Sequencing  
590 Technique - BSAS (**28, 44**) which relies on PCR enrichment of regions of interest using bisulfite  
591 converted DNA as a template. To assess if the presence of *S*-locus sRNA precursors is associated with  
592 DNA methylation changes in the respective *SCR01* target regions, we isolated immature buds of  
593 *SCR01*, *SCR01*<sup>\*mir1887</sup>, *SCR01* Ah04mir1887 and *SCR01*<sup>\*mir1887</sup> Ah20mirS3 plants. Genomic DNA was  
594 extracted using the NucleoSpin Plant II kit (Macherey-Nagel), and converted with the EZ DNA  
595 Methylation Kit (Zymo Research), including the modifications A and B, as described by the supplier, to  
596 optimize the conversion efficiency. Eight to ten biological replicates were extracted and converted for  
597 each genotype. Converted DNA was then amplified using primers flanking the Ah04mir1887 or  
598 Ah20mirS3 target sites in the intron and 5' region regions of *SCR01*, respectively (**table S1**). For each  
599 sRNA precursor, seven amplicons were generated, which amplify the top and bottom strands of the  
600 ~800 bp that flank the target site. As a negative control, we also amplified *AT2G20610*, which has been  
601 previously described to be an unmethylated negative control in *A. thaliana* Col-0 plants (**28**).  
602 Amplification of bisulfite converted DNA was performed using the TaKaRa EpiTaq™ HS (Takara), and  
603 the resulting PCR reactions were cleaned-up using the NucleoSpin Gel and PCR Clean-up XS kit  
604 (Macherey-Nagel). All amplicons obtained from a single biological replicate were pooled in equimolar  
605 amounts and a total of 100 ng were used for library preparation following the modified Illumina Nextera  
606 Flex library preparation protocol described before (**28, 81**). Libraries were sequenced in a MiSeq  
607 (Illumina) platform using 250 bp paired-end reads. Reads were trimmed, quality-filtered, and mapped  
608 to the Ah01 allele using methlypy (**82**), following the paired-end pipeline and specifying *AT2G20610* as  
609 the unmethylated control. Conversion efficiencies for each sample were estimated based on the

610 methylation levels detected in the *AT2G20610* locus, and were between 97% and 83% (**table S2**). The  
611 number of methylated reads covering each *SCR01* cytosine position, as outputted by methylpy, can be  
612 found in **table S3**. Average DNA methylation levels within the top and bottom strands of the ~800 bp  
613 regions flanking Ah04mir1887 and Ah20mirS3 target sites were summarized in R (**83**) and plotted with  
614 ggplot2 (**84**).

615 To identify DNA methylation heterogeneity at the single-molecule level, the methylation rate over  
616 each single read aligned to the *SCR01* sequence in three different cytosine contexts (CG, CHG and  
617 CHH) was computed using the Bismark methylation call output (**85**). For each sample, reads were  
618 classified in three distinct categories: highly methylated reads with 90% or more methylated cytosines,  
619 weakly methylated reads with a methylation rate inferior to 10%, and everything in between. Statistical  
620 differences in number of highly and weakly methylated reads between the *SCR01* and *SCR01*  
621 Ah04mir1887 or *SCR01*<sup>\*mir1887</sup> and *SCR01*<sup>\*mir1887</sup> Ah20mirS3 genotypes were assessed using a two-  
622 tailed Mann-Whitney test with continuity correction.

#### 623 *Data availability*

624 The AGO-IP and sRNA sequencing data is deposited in the Gene Expression Omnibus database,  
625 under the SubSeries reference GSE249507.

626 The Illumina and Oxford Nanopore sequencing data of the five *A. halleri* individuals carrying alleles  
627 Ah65, Ah60, Ah03, Ah33 and Ah19 are available in the European Nucleotide Archive, under the project  
628 number PRJEB70880.

629 The S-locus sequences used in this study are a combination of previously published data (6, 23)  
630 and data generated in this study. All allele sequences can be found in GenBank, under the following  
631 references: Ah01 - this study, PP396894; Ah02 - this study, PP396895; Ah03 - this study, PP537897;  
632 Ah04 - KJ461484; Ah10 - KM592810 and KM592817; Ah12 - KJ772373 and KJ772377; Ah13 -  
633 KJ461479 and KJ461483; Ah19 - this study, PP537898; Ah20 - FO203486; Ah25 - this study,  
634 PP396896; Ah28 - KJ461478; Ah29 - KM592803; Ah33 - this study, PP537899; Ah60 - this study,  
635 PP537900; Ah65 - this study, PP537901.

#### 636 **Acknowledgments**

637 This work was supported by a European Molecular Biology Organization (EMBO) Postdoctoral  
638 Fellowship attributed to RAB, by the European Research Council (NOVEL project, grant #648321) and  
639 by the French National Research Agency (TE-MoMa project, grant ANR-18-CE02-0020-01). This work  
640 was performed using the infrastructure and technical support of the Plateforme Serre, Cultures et  
641 Terrains Expérimentaux - Université de Lille for the greenhouse/field facilities. Research at the  
642 Lagrange laboratory was supported by the Centre National de la Recherche Scientifique (CNRS) and  
643 Université de Perpignan Via Domitia (UPVD). This study is set within the framework of the "Laboratoires  
644 d'Excellences (LABEX)" TULIP (ANR-10-LABX-41) and the "Ecole Universitaire de Recherche" (EUR)  
645 TULIP-GS (ANR-18-EURE-0019). Research at the Krämer laboratory was supported by the Deutsche  
646 Forschungsgemeinschaft DFG in SPP 1529 (ADAPTOMICS, Kr1967/10-1,-2), and by ERC-AdG LEAP-

647 EXTREME (788380). We thank Hervé Vaucheret, Vincent Colot, Claudia Köhler and Nicolas Butel for  
648 helpful discussions, and Hervé Vaucheret for sharing seeds.

## 649 References

- 650 1. Mendel, G., “Experiments on Plant Hybrids (Sherwood, E.R., trans.)” in *The Origin of Genetics: A*  
651 *Mendel Source Book* (Stern, C. and Sherwood, E.R., Eds), (W.H. Freeman, 1865), pp. 1–48.
- 652 2. S. Billiard, V. Castric, V. Llaurens, The integrative biology of genetic dominance. *Biol. Rev.* **96**,  
653 2925–2942 (2021).
- 654 3. S. Billiard, V. Castric, Evidence for Fisher’s dominance theory: How many “special cases”? *Trends*  
655 *Genet.* **27**, 441–445 (2011).
- 656 4. R. A. Fisher, The Possible Modification of the Response of the Wild Type to Recurrent Mutations.  
657 *Am. Nat.* **62**, 115–126 (1928).
- 658 5. S. Wright, Physiological and Evolutionary Theories of Dominance. *Am. Nat.* **68**, 24–53 (1934).
- 659 6. E. Durand, *et al.*, Dominance hierarchy arising from the evolution of a complex small RNA  
660 regulatory network. *Science* **346**, 1200–1205 (2014).
- 661 7. Y. Tarutani, *et al.*, Trans-acting small RNA determines dominance relationships in Brassica self-  
662 incompatibility. *Nature* **466**, 983–986 (2010).
- 663 8. S. Yasuda, *et al.*, A complex dominance hierarchy is controlled by polymorphism of small RNAs  
664 and their targets. *Nat. Plants* **3**, 1–5 (2016).
- 665 9. J. Huang, *et al.*, Stigma receptors control intraspecies and interspecies barriers in Brassicaceae.  
666 *Nature* **614**, 303–308 (2023).
- 667 10. C. Liu, *et al.*, Pollen PCP-B peptides unlock a stigma peptide–receptor kinase gating mechanism  
668 for pollination. *Science* **372**, 171–175 (2021).
- 669 11. J. B. Nasrallah, Stop and go signals at the stigma–pollen interface of the Brassicaceae. *Plant*  
670 *Physiol.* kiad301 (2023). <https://doi.org/10.1093/plphys/kiad301>.
- 671 12. J. B. Nasrallah, “Self-incompatibility in the Brassicaceae: Regulation and mechanism of self-  
672 recognition” in *Current Topics in Developmental Biology*, (Academic Press, 2019), pp. 435–452.
- 673 13. L. Zhang, *et al.*, FERONIA receptor kinase-regulated reactive oxygen species mediate self-  
674 incompatibility in Brassica rapa. *Curr. Biol.* **31**, 3004–3016.e4 (2021).
- 675 14. V. Castric, J. Bechsgaard, M. H. Schierup, X. Vekemans, Repeated adaptive introgression at a  
676 gene under multiallelic balancing selection. *PLoS Genet.* **4** (2008).
- 677 15. M. Genete, V. Castric, X. Vekemans, Genotyping and De Novo Discovery of Allelic Variants at  
678 the Brassicaceae Self-Incompatibility Locus from Short-Read Sequencing Data. *Mol. Biol. Evol.*  
679 **37**, 1193–1201 (2020).
- 680 16. M. J. Lawrence, Population Genetics of the Homomorphic Self-incompatibility Polymorphisms in  
681 Flowering Plants. *Ann. Bot.* **85**, 221–226 (2000).
- 682 17. K. Sato, *et al.*, Coevolution of the S-Locus Genes SRK, SLG and SP11/SCR in Brassica oleracea  
683 and B. rapa. *Genetics* **162**, 931–940 (2002).
- 684 18. N. Burghgraeve, *et al.*, Base-Pairing Requirements for Small RNA-Mediated Gene Silencing of  
685 Recessive Self-Incompatibility Alleles in Arabidopsis halleri. *Genetics* **215**, genetics.303351.2020

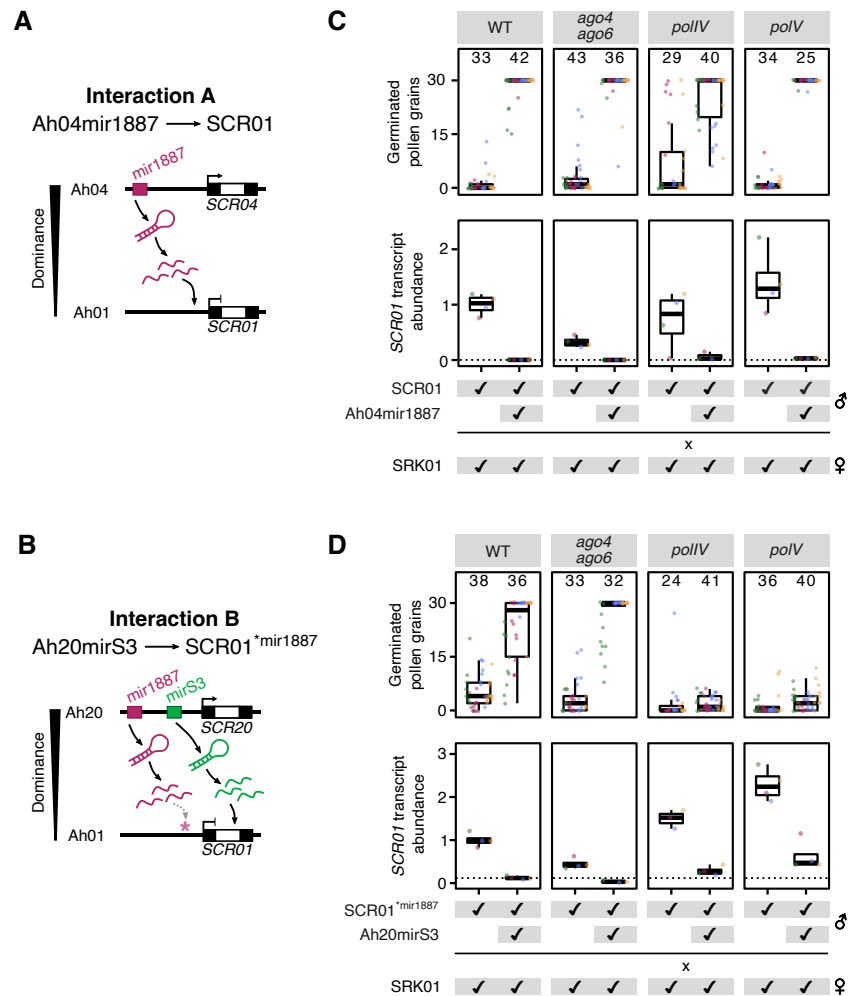
- 686 (2020).
- 687 19. T. Kakizaki, *et al.*, Linear Dominance Relationship among Four Class-II S Haplotypes in Pollen is  
688 Determined by the Expression of SP11 in Brassica Self-Incompatibility. *Plant Cell Physiol.* **44**,  
689 70–75 (2003).
- 690 20. H. Shiba, *et al.*, The Dominance of Alleles Controlling Self-Incompatibility in Brassica Pollen Is  
691 Regulated at the RNA Level. *Plant Cell* **14**, 491–504 (2002).
- 692 21. V. Llaurens, S. Billiard, V. Castric, X. Vekemans, Evolution of dominance in sporophytic self-  
693 incompatibility systems: I. genetic load and coevolution of levels of dominance in pollen and pistil.  
694 *Evolution* **63**, 2427–2437 (2009).
- 695 22. T. Duan, *et al.*, Dominance between self-incompatibility alleles determines the mating system of  
696 *Capsella* allopolyploids. [Preprint] (2023). Available at:  
697 <https://www.biorxiv.org/content/10.1101/2023.04.17.537155v1> [Accessed 19 September 2023].
- 698 23. P. M. Goubet, *et al.*, Contrasted Patterns of Molecular Evolution in Dominant and Recessive Self-  
699 Incompatibility Haplotypes in Arabidopsis. *PLoS Genet.* **8**, e1002495 (2012).
- 700 24. N. L. Prigoda, A. Nassuth, B. K. Mable, Phenotypic and genotypic expression of self-  
701 incompatibility haplotypes in Arabidopsis lyrata suggests unique origin of alleles in different  
702 dominance classes. *Mol. Biol. Evol.* **22**, 1609–1620 (2005).
- 703 25. P. Chellappan, *et al.*, siRNAs from miRNA sites mediate DNA methylation of target genes. *Nucleic*  
704 *Acids Res.* **38**, 6883–6894 (2010).
- 705 26. B. Khraiwesh, *et al.*, Transcriptional Control of Gene Expression by MicroRNAs. *Cell* **140**, 111–  
706 122 (2010).
- 707 27. K. Panda, *et al.*, Full-length autonomous transposable elements are preferentially targeted by  
708 expression-dependent forms of RNA-directed DNA methylation. *Genome Biol.* **17**, 170 (2016).
- 709 28. M. J. Sigman, *et al.*, An siRNA-guided ARGONAUTE protein directs RNA polymerase V to initiate  
710 DNA methylation. *Nat. Plants* **7**, 1461–1474 (2021).
- 711 29. L. Wu, *et al.*, DNA Methylation Mediated by a MicroRNA Pathway. *Mol. Cell* **38**, 465–475 (2010).
- 712 30. E. Allen, *et al.*, Evolution of microRNA genes by inverted duplication of target gene sequences in  
713 Arabidopsis thaliana. *Nat. Genet.* **36**, 1282–1290 (2004).
- 714 31. M. J. Axtell, J. O. Westholm, E. C. Lai, Vive la différence: biogenesis and evolution of microRNAs  
715 in plants and animals. *Genome Biol.* **12**, 221 (2011).
- 716 32. J. Cui, C. You, X. Chen, The evolution of microRNAs in plants. *Curr. Opin. Plant Biol.* **35**, 61–67  
717 (2017).
- 718 33. J. T. Cuperus, N. Fahlgren, J. C. Carrington, Evolution and Functional Diversification of MIRNA  
719 Genes. *Plant Cell* **23**, 431–442 (2011).
- 720 34. N. Fahlgren, *et al.*, MicroRNA gene evolution in Arabidopsis lyrata and Arabidopsis thaliana. *Plant*  
721 *Cell* **22**, 1074–1089 (2010).
- 722 35. N. Fahlgren, *et al.*, High-throughput sequencing of Arabidopsis microRNAs: Evidence for frequent  
723 birth and death of MIRNA genes. *PLoS ONE* **2**, e219 (2007).
- 724 36. R. Rajagopalan, H. Vaucheret, J. Trejo, D. P. Bartel, A diverse and evolutionarily fluid set of  
725 microRNAs in Arabidopsis thaliana. *Genes Dev.* **20**, 3407–3425 (2006).

- 726 37. M. E. Nasrallah, P. Liu, S. Sherman-Broyles, N. A. Boggs, J. B. Nasrallah, Natural variation in  
727 expression of self-incompatibility in *Arabidopsis thaliana*: Implications for the evolution of selfing.  
728 *Proc. Natl. Acad. Sci.* **101**, 16070–16074 (2004).
- 729 38. M. E. Nasrallah, P. Liu, J. B. Nasrallah, Generation of Self-Incompatible *Arabidopsis thaliana* by  
730 Transfer of Two S Locus Genes from *A. lyrata*. *Science* **297**, 247–249 (2002).
- 731 39. N. A. Boggs, J. B. Nasrallah, M. E. Nasrallah, Independent S-Locus Mutations Caused Self-  
732 Fertility in *Arabidopsis thaliana*. *PLoS Genet.* **5**, e1000426 (2009).
- 733 40. S. Fujii, *et al.*, Parallel evolution of dominant pistil-side self-incompatibility suppressors in  
734 *Arabidopsis*. *Nat. Commun.* **2020 111 11**, 1–9 (2020).
- 735 41. O. Voinnet, Origin, Biogenesis, and Activity of Plant MicroRNAs. *Cell* **136**, 669–687 (2009).
- 736 42. X.-J. He, *et al.*, NRPD4, a protein related to the RPB4 subunit of RNA polymerase II, is a  
737 component of RNA polymerases IV and V and is required for RNA-directed DNA methylation.  
738 *Genes Dev.* **23**, 318–330 (2009).
- 739 43. S. W.-L. Chan, *et al.*, RNA Silencing Genes Control de Novo DNA Methylation. *Science* **303**,  
740 1336–1336 (2004).
- 741 44. D. R. Masser, A. S. Berg, W. M. Freeman, Focused, high accuracy 5-methylcytosine quantitation  
742 with base resolution by benchtop next-generation sequencing. *Epigenetics Chromatin* **6**, 33  
743 (2013).
- 744 45. S. Mi, *et al.*, Sorting of Small RNAs into *Arabidopsis* Argonaute Complexes Is Directed by the 5'  
745 Terminal Nucleotide. *Cell* **133**, 116–127 (2008).
- 746 46. A. Le Vève, “Balancing selection, genetic load and dominance between self-incompatibility alleles  
747 in *Arabidopsis*: an empirical and theoretical study of this ménage à trois,” Université de Lille  
748 (2022-....). (2022).
- 749 47. H. Liu, *et al.*, Nuclear functions of mammalian MicroRNAs in gene regulation, immunity and  
750 cancer. *Mol. Cancer* **17**, 64 (2018).
- 751 48. J. O'Brien, H. Hayder, Y. Zayed, C. Peng, Overview of MicroRNA Biogenesis, Mechanisms of  
752 Actions, and Circulation. *Front. Endocrinol.* **9** (2018).
- 753 49. T. C. Roberts, The MicroRNA Biology of the Mammalian Nucleus. *Mol. Ther. - Nucleic Acids* **3**  
754 (2014).
- 755 50. P. Hoffer, *et al.*, Posttranscriptional gene silencing in nuclei. *Proc. Natl. Acad. Sci.* **108**, 409–414  
756 (2011).
- 757 51. N. G. Bologna, *et al.*, Nucleo-cytosolic Shuttling of ARGONAUTE1 Prompts a Revised Model of  
758 the Plant MicroRNA Pathway. *Mol. Cell* **69**, 709-719.e5 (2018).
- 759 52. Y. Fang, D. L. Spector, Identification of Nuclear Dicing Bodies Containing Proteins for MicroRNA  
760 Biogenesis in Living *Arabidopsis* Plants. *Curr. Biol.* **17**, 818–823 (2007).
- 761 53. W. Wang, *et al.*, An Importin  $\beta$  Protein Negatively Regulates MicroRNA Activity in *Arabidopsis*[W].  
762 *Plant Cell* **23**, 3565–3576 (2011).
- 763 54. N. Baumberger, D. C. Baulcombe, *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively  
764 recruits microRNAs and short interfering RNAs. *Proc. Natl. Acad. Sci.* **102**, 11928–11933 (2005).
- 765 55. J. Dolata, *et al.*, Salt Stress Reveals a New Role for ARGONAUTE1 in miRNA Biogenesis at the

- 766           Transcriptional and Posttranscriptional Levels. *Plant Physiol.* **172**, 297–312 (2016).
- 767 56. E. J. Finnegan, D. Liang, M.-B. Wang, Self-incompatibility: Smi silences through a novel sRNA  
768 pathway. *Trends Plant Sci.* **16**, 238–241 (2011).
- 769 57. S. Fujii, K. Kubo, S. Takayama, Non-self- and self-recognition models in plant self-incompatibility.  
770 *Nat. Plants* **2**, 1–9 (2016).
- 771 58. H. Zhao, *et al.*, Origin, loss, and regain of self-incompatibility in angiosperms. *Plant Cell* **34**, 579–  
772 596 (2022).
- 773 59. H. Bastide, S. V. Saenko, M. Chouteau, M. Joron, V. Llaurens, Dominance mechanisms in  
774 supergene alleles controlling butterfly wing pattern variation: insights from gene expression in  
775 *Heliconius numata*. *Heredity* **130**, 92–98 (2023).
- 776 60. M. Joron, C. D. Jiggins, A. Papanicolaou, W. O. McMillan, *Heliconius* wing patterns: an evo-devo  
777 model for understanding phenotypic diversity. *Heredity* **97**, 157–167 (2006).
- 778 61. R. J. Stein, *et al.*, Relationships between soil and leaf mineral composition are element-specific,  
779 environment-dependent and geographically structured in the emerging model *Arabidopsis halleri*.  
780 *New Phytol.* **213**, 1274–1286 (2017).
- 781 62. S. J. Clough, A. F. Bent, Floral dip: A simplified method for *Agrobacterium*-mediated  
782 transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
- 783 63. H.-L. Xing, *et al.*, A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.*  
784 **14**, 327 (2014).
- 785 64. Z.-P. Wang, *et al.*, Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates  
786 homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome*  
787 *Biol.* **16**, 144 (2015).
- 788 65. S. Lahmy, *et al.*, Evidence for ARGONAUTE4–DNA interactions in RNA-directed DNA  
789 methylation in plants. *Genes Dev.* **30**, 2565–2570 (2016).
- 790 66. C.-G. Duan, *et al.*, Specific but interdependent functions for *Arabidopsis* AGO4 and AGO6 in RNA-  
791 directed DNA methylation. *EMBO J.* **34**, 581–592 (2015).
- 792 67. S. M. Smith, P. J. Maughan, SNP genotyping using KASPar assays. *Methods Mol. Biol. Clifton*  
793 *NJ* **1245**, 243–256 (2015).
- 794 68. M. W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic*  
795 *Acids Res.* **29**, 45e–445 (2001).
- 796 69. C. Barre-Villeneuve, *et al.*, The unique dual targeting of AGO1 by two types of PRMT enzymes  
797 promotes phasiRNA loading in *Arabidopsis thaliana*. *Nucleic Acids Res.* gkae045 (2024).  
798 <https://doi.org/10.1093/nar/gkae045>.
- 799 70. F. Krueger, TrimGalore. (2016). Deposited 2016.
- 800 71. M. J. Axtell, ShortStack: Comprehensive annotation and quantification of small RNA genes. *RNA*  
801 **19**, 740 (2013).
- 802 72. Z. Gu, Complex heatmap visualization. *iMeta* **1**, e43 (2022).
- 803 73. Z. Gu, L. Gu, R. Eils, M. Schlesner, B. Brors, circlize implements and enhances circular  
804 visualization in R. *Bioinformatics* **30**, 2811–2812 (2014).
- 805 74. C. Belser, *et al.*, Chromosome-scale assemblies of plant genomes using nanopore long reads

- 806 and optical maps. *Nat. Plants* **4**, 879–887 (2018).
- 807 75. Y. Chen, *et al.*, Efficient assembly of nanopore reads via highly accurate and intact error  
808 correction. *Nat. Commun.* **12**, 60 (2021).
- 809 76. R. Vaser, I. Sović, N. Nagarajan, M. Šikić, Fast and accurate de novo genome assembly from  
810 long uncorrected reads. *Genome Res.* **27**, 737–746 (2017).
- 811 77. J.-M. Aury, B. Istace, Hapo-G, haplotype-aware polishing of genome assemblies with accurate  
812 reads. *NAR Genomics Bioinforma.* **3**, lqab034 (2021).
- 813 78. S. Huang, M. Kang, A. Xu, HaploMerger2: rebuilding both haploid sub-assemblies from high-  
814 heterozygosity diploid genome assembly. *Bioinformatics* **33**, 2577–2579 (2017).
- 815 79. V. Solovyev, P. Kosarev, I. Seledsov, D. Vorobyev, Automatic annotation of eukaryotic genes,  
816 pseudogenes and promoters. *Genome Biol.* **7**, S10 (2006).
- 817 80. V. Llaurens, *et al.*, Does frequency-dependent selection with complex dominance interactions  
818 accurately predict allelic frequencies at the self-incompatibility locus in *Arabidopsis halleri*?  
819 *Evolution* **62**, 2545–2557 (2008).
- 820 81. D. Gaio, *et al.*, Hackflex: low cost Illumina Nextera Flex sequencing library construction. *bioRxiv*  
821 779215 (2021). <https://doi.org/10.1101/779215>.
- 822 82. M. D. Schultz, *et al.*, Human body epigenome maps reveal noncanonical DNA methylation  
823 variation. *Nature* **523**, 212–216 (2015).
- 824 83. R Core Team, R: A Language and Environment for Statistical Computing | BibSonomy. (2017).  
825 Available at: <https://www.R-project.org/>.
- 826 84. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2009).
- 827 85. F. Krueger, S. R. Andrews, Bismark: a flexible aligner and methylation caller for Bisulfite-Seq  
828 applications. *Bioinformatics* **27**, 1571–1572 (2011).

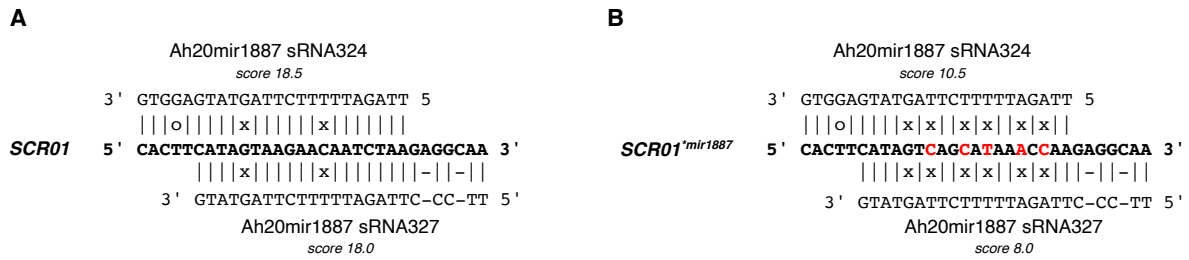




829 **Figure 1. The RdDM pathway is not required for the activity of Ah04mir1887 and Ah20mirS3.**

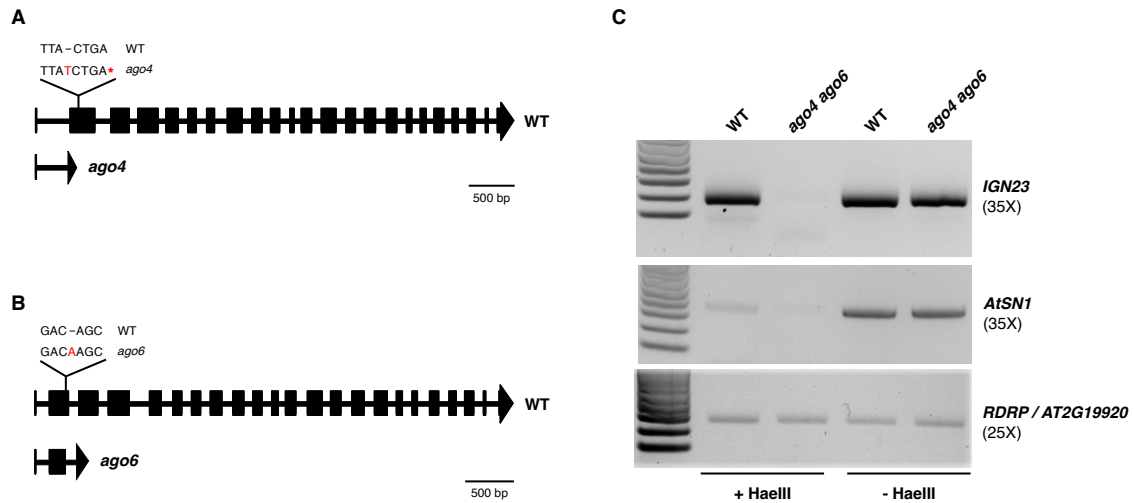
830 **(A,B)** Dominance interactions studied in this paper. The black and white regions within *SCR* genes  
 831 correspond to exons and introns, respectively. **(A)** Interaction A focuses on the role of the sRNA  
 832 precursor Ah04mir1887 in mediating the dominance of Ah04 over Ah01. **(B)** Interaction B focuses on  
 833 the role of Ah20mirS3 in mediating the dominance of Ah20 over Ah01. S-locus sRNA precursors consist  
 834 of inverted repeats that form hairpins which are processed into multiple sRNAs showing sequence  
 835 homology to specific *SCR* regions: Ah04mir1887 sRNAs show homology to the 5' region region of  
 836 *SCR01*, while Ah20mirS3 sRNAs show homology to the intronic region (white) of *SCR01*. Because  
 837 sRNAs produced from the Ah20mir1887 precursor show homology to the wt *SCR01* 5' region region,  
 838 and to study the specific effect of Ah20mirS3 in Interaction B, the *SCR01*\*mir1887 line containing five point  
 839 mutations in the *SCR01* 5' region was created, which disrupts the homology with Ah20mir1887 sRNAs.  
 840 **(C,D)** Phenotypic and expression assay evaluating the role of key RdDM components on the action of  
 841 Ah04mir1887 **(C)**, and Ah20mirS3 **(D)**. Crosses between *SRK01*-expressing mothers and *SCR01*-  
 842 expressing fathers were performed in wt and RdDM mutant backgrounds in the absence or presence  
 843 of the respective sRNA precursor. The number of germinated pollen grains of the indicated genotype  
 844 was counted on pistils of *SRK01* mothers. *SCR01* expression in immature buds of the indicated  
 845 genotype was measured by RT-qPCR and normalized to the wt level. Five biological replicates

846 (represented by the differently colored dots) were analyzed for pollen germination and *SCR01*  
847 expression. Numbers on top of columns indicate the total number of siliques sampled for pollen  
848 germination assays.



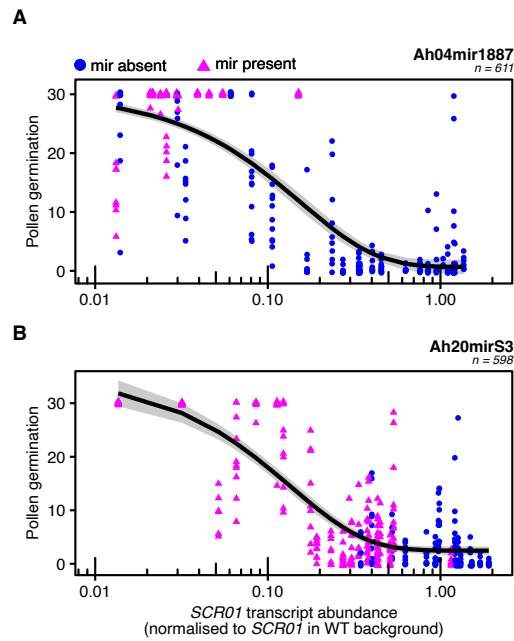
**849 Figure 1 S1. SCR01<sup>\*mir1887</sup> transgenic line**

**850 (A)** Alignment between the native mir1887 target site on *SCR01* and Ah20mir1887 sRNAs with a  
**851** targeting score  $\geq 18$ . **(B)** Alignment between the same sRNAs and the mutated mir1887 target site of  
**852** the transgenic line *SCR01<sup>\*mir1887</sup>*. Note that the mutations decrease the homology score below 18,  
**853** suggesting that these sRNAs are not able to target the mutated site. Mutated sites are in red.



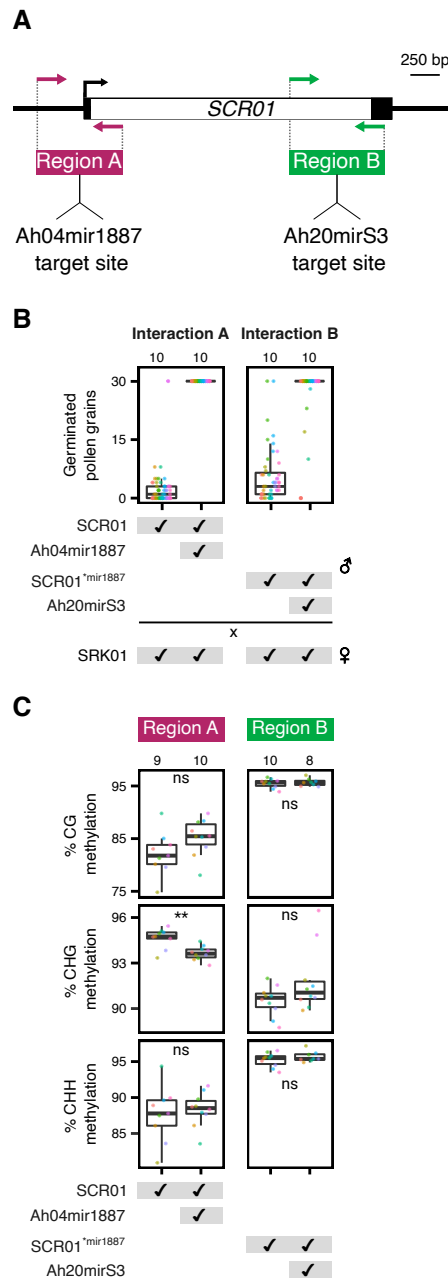
854 **Figure 1 S2. Validation of the RdDM CRISPR/Cas9 mutants generated in this study.**

855 **(A-B)** Schematic representation of CRISPR/Cas9 induced mutations in the *AGO4* and *AGO6* genes.  
856 Sequence changes are shown in red, premature stop codons are represented by the asterisk. **(C)** Chop-  
857 PCR assay on wt C24 plants and *ago4 ago6* double mutant plants. *IGN23* and *AtSN1* correspond to  
858 RdDM-dependent methylated regions, while *RDRP* corresponds to a non-methylated region.



859 **Figure 1 S3. A compatible pollen germination phenotype requires a low *SCR01* transcript**  
860 **abundance.**

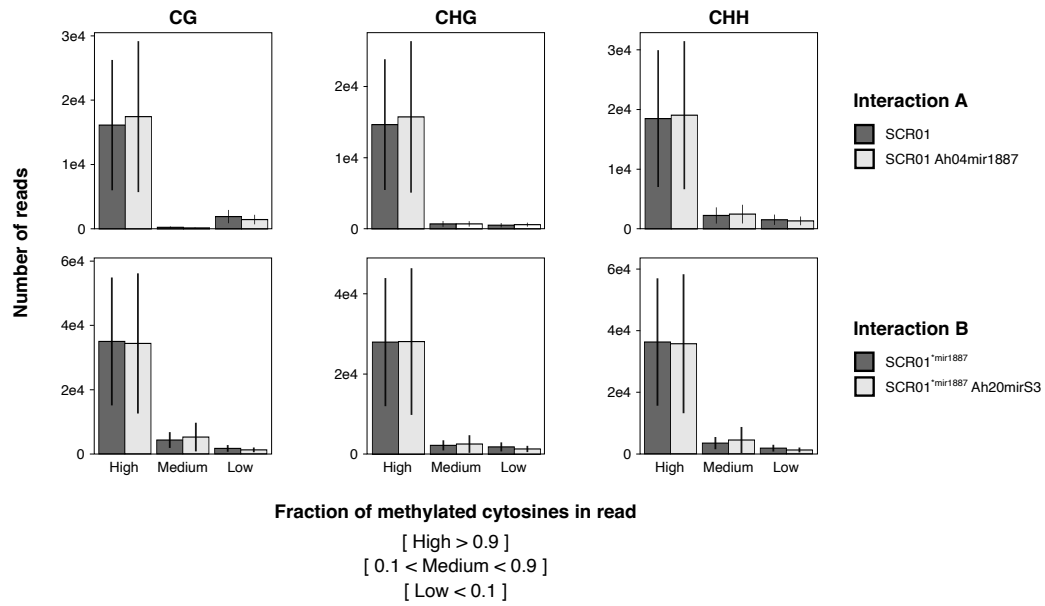
861 Non-linear regression analysis modeling the relationship between *SCR01* transcript abundance levels  
862 and pollen germination. For both Ah04mir1887 (A) and Ah20mirS3 (B) sRNA precursors a compatible  
863 pollen germination phenotype (~20-30 pollen grains) requires a 10-fold decrease in *SCR01* transcript  
864 abundance (compared to the level observed in the wt). Analysis was performed using the phenotypic  
865 and expression data presented in Fig. 1 C-D. The black line represents the predicted model and the  
866 gray area corresponds to the 95% confidence interval.



867 **Figure 1 S4. Ah04mir1887 and Ah20mirS3 are not associated with detectable DNA methylation**  
 868 **changes at SCR01 target sites.**

869 **(A)** Schematic representation of the BSAS assay performed to measure DNA methylation at the  
 870 Ah04mir1887 target site (*SCR01* 5' region - Region A), and at the Ah20mirS3 target site (*SCR01* intron  
 871 - Region B). Region A and Region B correspond to fragments of 794 and 871 bp, respectively, centered  
 872 in each sRNA precursor's target site. **(B)** Phenotypic validation of lines used for the BSAS assay.  
 873 Colored dots correspond to values obtained for each biological replicate. Number of biological replicates  
 874 is indicated at the top of the panel. For each biological replicate pollen germination on a total of 5 *SRK01*  
 875 pistils was counted. **(C)** Average DNA methylation levels on Region A (Interaction A) and Region B  
 876 (Interaction B) of *SCR01*. DNA methylation levels correspond to the average of methylation in all

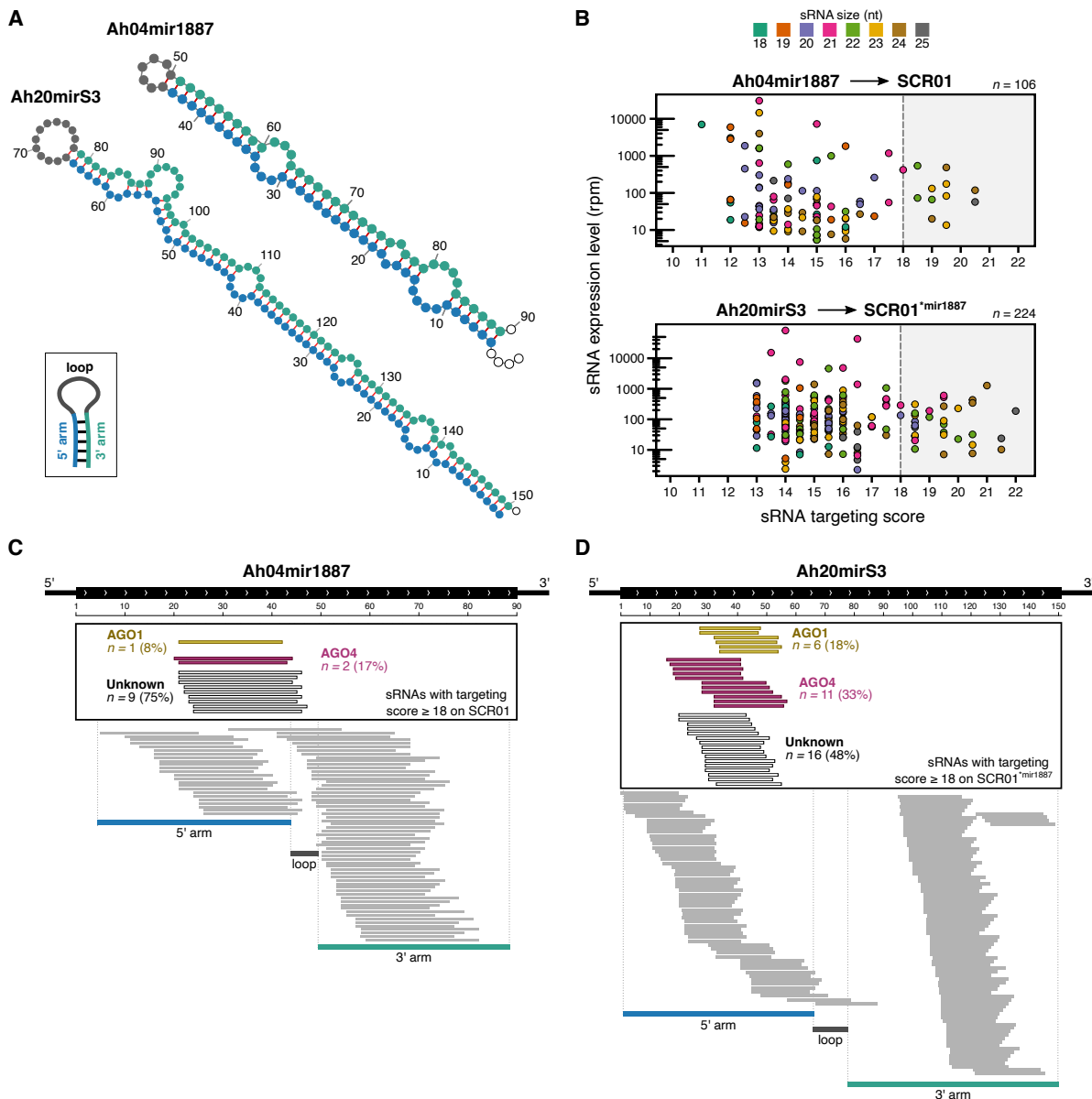
877 cytosines in the top and bottom DNA strands of sampled regions. Number of biological replicates is  
878 indicated at the top of the panel. Statistical differences were assessed using a two-tailed Mann-Whitney  
879 test with continuity correction (ns - non significant, p-value > 0.05; \*\* - p-value ≤ 0.01).



**880 Figure 1 S5. Ah04mir1887 and Ah20mirS3 target sites are heavily methylated.**

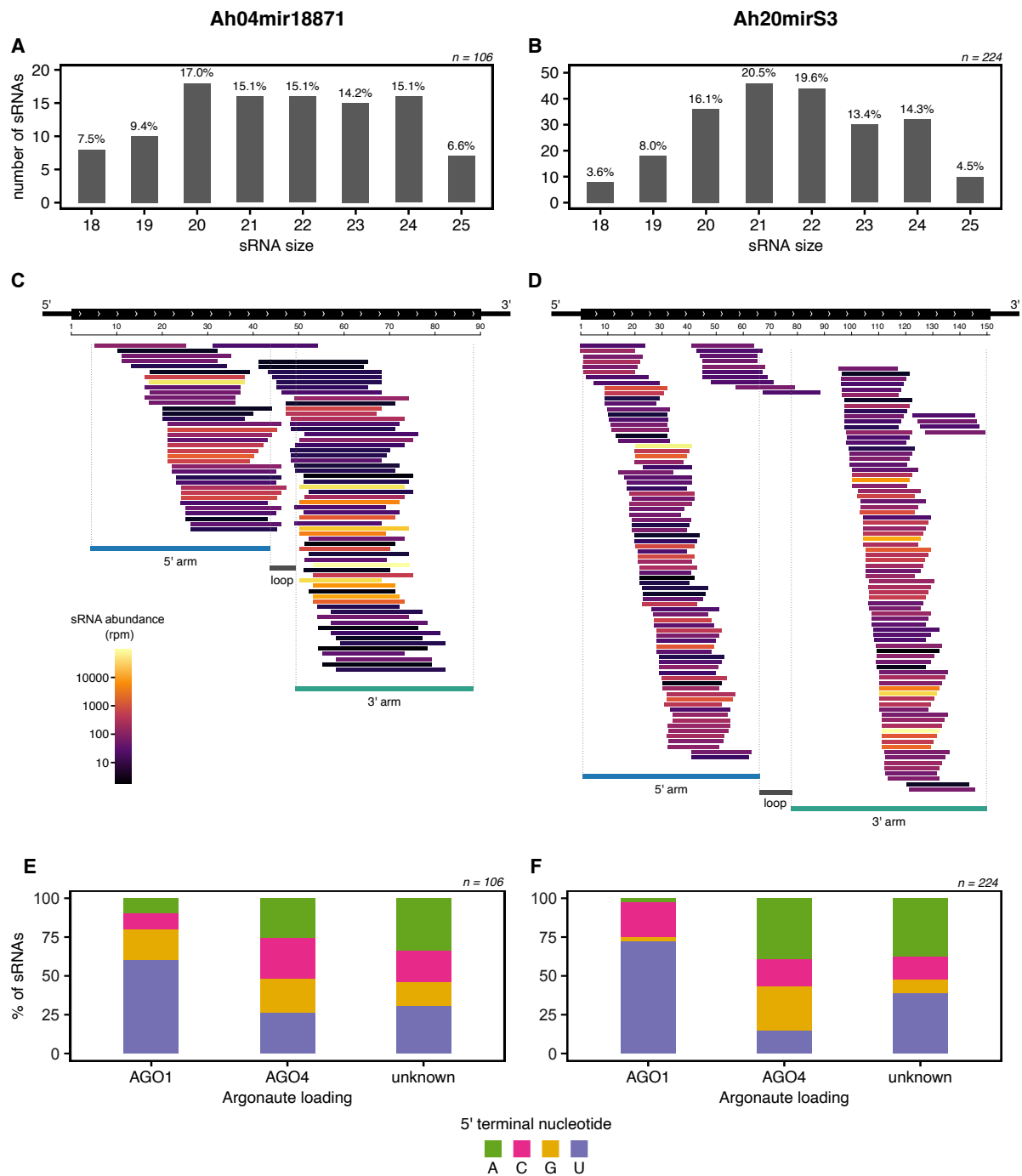
**881** BSAS reads were grouped according to the fraction of methylated cytosines within each read. High:  
**882** reads where nearly all cytosine positions are methylated (> 90% cytosines); Low: reads where very few  
**883** cytosine positions are methylated (< 10% cytosines); Medium: all reads with more than 10% and less  
**884** than 90% cytosines methylated. Error bars represent the standard deviation between replicates. No  
**885** statistically significant differences were detected in the number of highly and lowly methylated reads  
**886** between the *SCR01* and *SCR01* Ah04mir1887 or *SCR01*<sup>mir1887</sup> and *SCR01*<sup>mir1887</sup> Ah20mirS3  
**887** genotypes (p-value > 0.05, two-tailed Mann-Whitney test with continuity correction).





888 **Figure 2. Ah04mir1887 and Ah20mirS3 hairpins are processed into numerous sRNAs.**

889 **(A)** Predicted RNA hairpin structures of the *S*-locus sRNA precursors Ah04mir1887 and Ah20mirS3.  
 890 **(B)** Expression of Ah04mir1887 and Ah20mirS3 sRNAs as a function of their targeting score against  
 891 *SCR01*. The color of the dots corresponds to the size of the sRNA. The grey box highlights sRNAs  
 892 predicted to induce a reduction in *SCR01* transcript abundance (score ≥ 18). *n* corresponds to the total  
 893 number of unique sRNAs identified upon sequencing for each sRNA precursor. **(C-D)** Genomic  
 894 representation of the Ah04mir1887 and Ah20mirS3 loci, and their respective sRNAs. The black  
 895 rectangle highlights sRNAs predicted to induce a reduction in *SCR01* transcript abundance (score ≥  
 896 18). AGO loading information for each functional sRNA is indicated by the sRNA color (yellow – sRNA  
 897 loaded predominantly in AGO1; magenta – sRNA loaded predominantly in AGO4; white – sRNAs with  
 898 unknown loading profile), see section **ARGONAUTE immunoprecipitation and sRNA sequencing**  
 899 for more details.



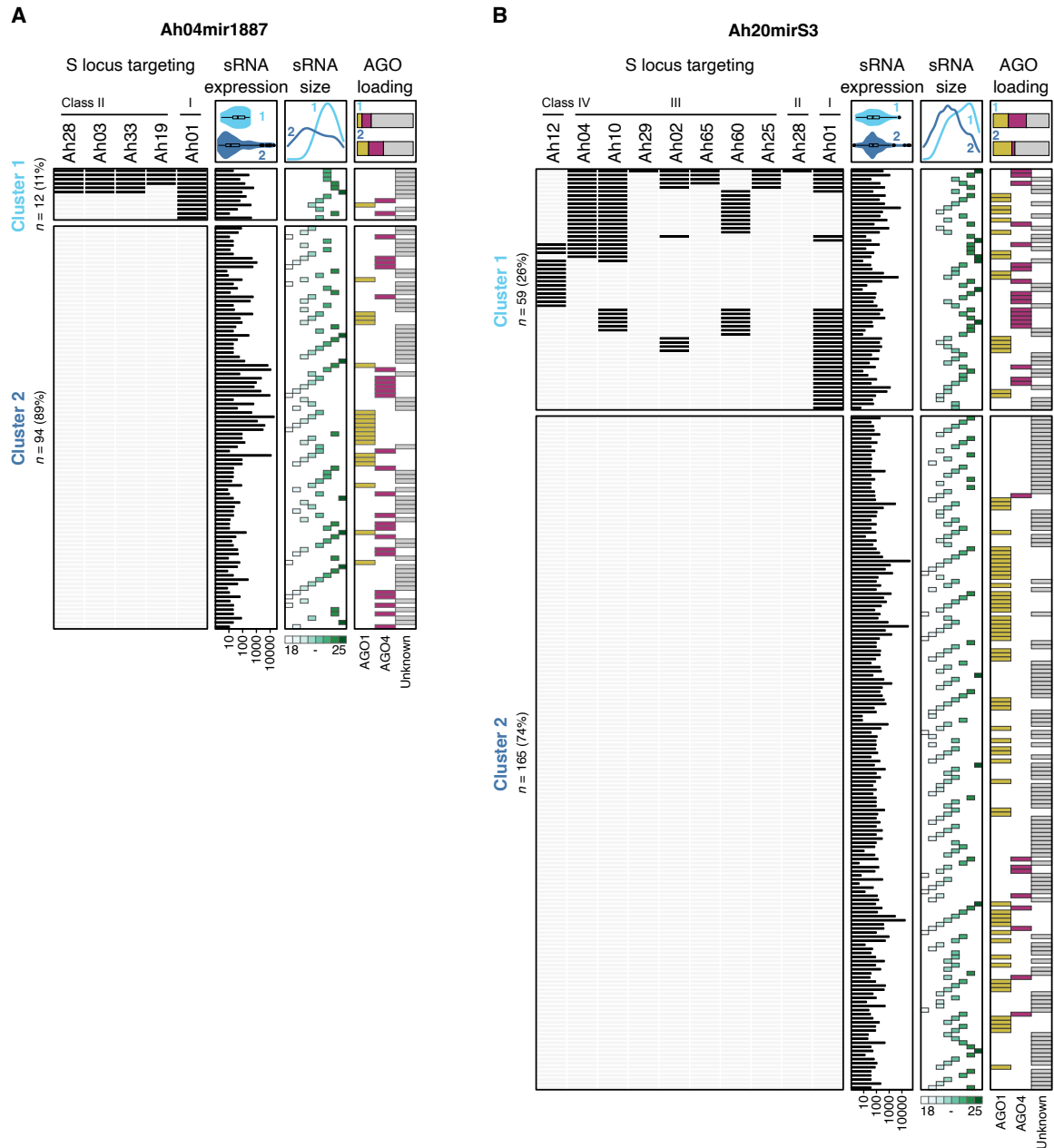
900 **Figure 2 S1. sRNAs produced from Ah04mir1887 and Ah20mirS3 do not show a specific size**  
 901 **bias but vary in their abundance levels.**

902 Size distribution of sRNAs derived from Ah04mir1887 **(A)** and Ah20mirS3 **(B)**. **(C-D)** Genomic  
 903 representation of the Ah04mir1887 and Ah20mirS3 loci, and their respective sRNAs. sRNAs are colored  
 904 according to their abundance level. **(E-F)** 5' terminal nucleotide frequencies in Ah04mir1887 and  
 905 Ah20mirS3 sRNAs, related to their inferred AGO loading pattern.

sRNA source		Targeted allele														
		IV			III						II				I	
		Ah20	Ah13	Ah12	Ah04	Ah10	Ah29	Ah02	Ah65	Ah60	Ah25	Ah28	Ah03	Ah33	Ah19	Ah01
Ah20	Dominance relationship	■	CD	D	D	D	D	D*	D*	D*	D	D	D	D	D	D
	mirS3			■	■	■	■	■	■	■	■					■ Int. B
	other mirs			■	■		■	■	■	■	■	■	■	■	■	■
Ah04	Dominance relationship				■	D	D	D	na	na	na	D	D	D*	D*	D
	mir1887											■	■	■	■	■ Int. A
	other mirs					■	■	■	■	■	■	■	■	■	■	■

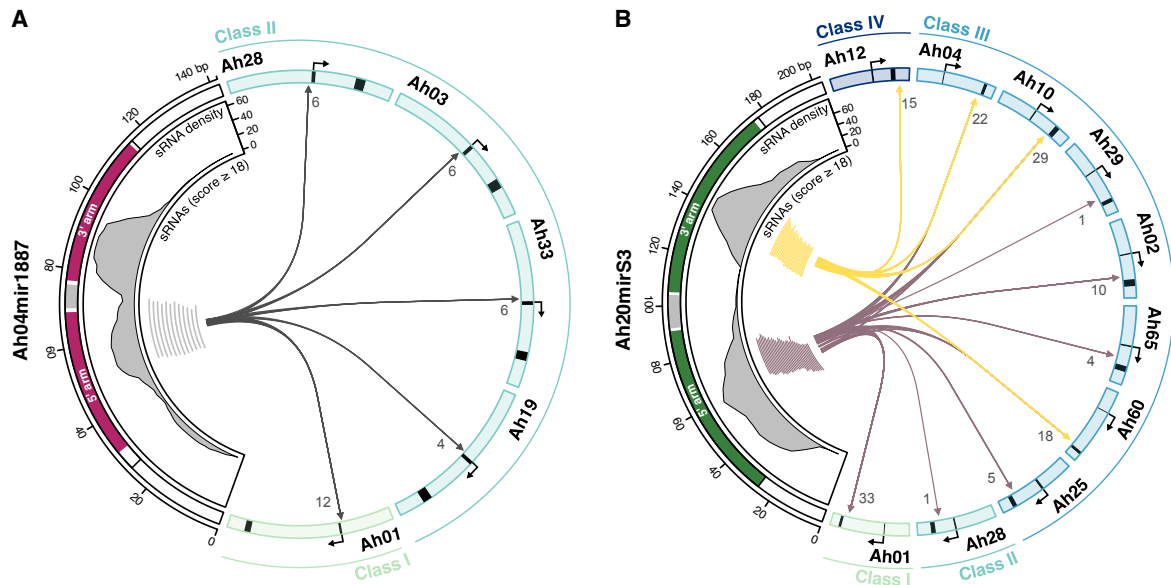
906 **Figure 3. Ah04mir1887 and Ah20mirS3 underlie numerous S-allele dominance interactions.**

907 Schematic representation of the dominance relationship between alleles Ah20, Ah04 and other S-  
 908 alleles in the current *A. halleri* dominance network. The dominance phenotypes observed are  
 909 represented by D (dominant, as determined by phenotypic assays), CD (codominant, as determined by  
 910 phenotypic assays), and na (no phenotypic data available). D\* refers to a predicted dominance  
 911 relationship, based on the difference between the dominance class of the sRNA source and the targeted  
 912 allele (15, 24). The sRNA precursors underlying each dominance relationship are highlighted in green  
 913 (mirS3), magenta (mir1887), or grey (other mirs). The specific interactions studied earlier in this  
 914 manuscript – Interaction A and Interaction B - are highlighted by a dashed box.



915 **Figure 4. Ah04mir1887 and Ah20mirS3 produce multiple sRNAs targeting one or more S-alleles.**

916 Heatmap representing the S-loci targeted by Ah04mir1887-derived sRNAs **(A)** and Ah20mirS3-derived  
 917 sRNAs **(B)**. A black rectangle indicates that a specific sRNA has a targeting score of  $\geq 18$  against the  
 918 corresponding S-allele. Cluster 1 groups sRNAs that target at least one S-allele, while Cluster 2 groups  
 919 sRNAs that have no targets in the represented S-alleles. sRNA expression levels, sRNA size and AGO  
 920 loading patterns are indicated for each sRNA. Summary plots comparing these features in Cluster 1  
 921 and Cluster 2 are shown in the top panel.  $n$  corresponds to the number of sRNAs produced by each  
 922 sRNA precursor.



923 **Figure 5. The targeting repertoire of S-locus sRNA precursors is enhanced by their**  
924 **heterogeneous processing pattern.**

925 Circos plot representing the positional relationship between sRNAs produced by Ah04mir1887 (A),  
926 Ah20mirS3 (B) and the corresponding targeted S-alleles. The genomic regions harboring Ah04mir1887  
927 and Ah20mirS3 are represented on the left side of the plots. The predicted 5' and 3' hairpin arms  
928 are represented in magenta (Ah04mir1887) and green (Ah20mirS3), and the sRNA density (number of  
929 sRNAs) is plotted along the sRNA precursor region. Individual sRNAs with a score of  $\geq 18$  are  
930 represented by the grey, yellow or dark purple segments. Arrowed segments represent the sRNA  
931 source - target relationship. The numbers at the end of these arrowed segments indicate the total  
932 number of sRNAs targeting that respective allele. The SCR genomic region of targeted S-alleles is  
933 represented on the right side and includes 1500 bp upstream of the TSS and 500 bp downstream of  
934 the TES. SCR exons 1 and 2 are colored in black.