

Base-pairing requirements for small RNA-mediated gene silencing of recessive self-incompatibility alleles in Arabidopsis halleri

Nicolas Burghgraeve, Samson Simon, Simon Barral, Isabelle Fobis-Loisy, Anne-Catherine Holl, Chloé Poniztki, Eric Schmitt, Xavier Vekemans,

Vincent Castric

▶ To cite this version:

Nicolas Burghgraeve, Samson Simon, Simon Barral, Isabelle Fobis-Loisy, Anne-Catherine Holl, et al.. Base-pairing requirements for small RNA-mediated gene silencing of recessive self-incompatibility alleles in Arabidopsis halleri. Genetics, 2020, 215 (3), pp.653-664. 10.1534/genetics.120.303351. hal-02992033

HAL Id: hal-02992033 https://hal.science/hal-02992033

Submitted on 6 Nov 2020

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.

1 Base-pairing requirements for small RNA-mediated gene

2 silencing of recessive self-incompatibility alleles in Arabidopsis

- 3 *halleri*.
- 4
- 5 N. Burghgraeve¹, S. Simon¹, S. Barral¹, I. Fobis-Loisy², A-C Holl¹, C. Poniztki¹, E.
- 6 Schmitt¹, X. Vekemans¹, V. Castric¹

7

8 Affiliations

- 9 1. CNRS, Univ. Lille, UMR 8198 Evo-Eco-Paleo, F-59000 Lille, France
- 10 2. Laboratoire Reproduction et Développement des Plantes, Univ de Lyon, ENS de
- 11 Lyon, UCB Lyon1, CNRS, INRA, F-69342, Lyon, France

12

13 FigShare accession number :

- 14 Supplementary data, figures and tables are available at:
- 15 https://doi.org/10.6084/m9.figshare.c.4877223

17 Running title

18 SI dominance and sRNA silencing

19 Key Words

- 20 Dominance/recessivity. Sporophytic self-incompatibility. RT-qPCR. Allele-specific
- 21 expression assay. Arabidopsis halleri.

22 Corresponding author

- 23 Vincent Castric
- 24 CNRS, Univ. Lille, UMR 8198 Evo-Eco-Paleo, F-59000 Lille, France
- 25 Batiment SN2, bureau 207
- 26 59655 Villeneuve d'Ascq FRANCE
- 27 Tel: 33 (0)3-20-33-63-03
- 28 Mail: <u>Vincent.Castric@univ-lille1.fr</u>

29

30

32

Abstract

33 Small non-coding RNAs are central regulators of genome activity and stability. Their 34 regulatory function typically involves sequence similarity with their target sites, but 35 understanding the criteria by which they specifically recognize and regulate their targets 36 across the genome remains a major challenge in the field, especially in the face of the 37 diversity of silencing pathways involved. The dominance hierarchy among self-38 incompatibility alleles in Brassicaceae is controlled by interactions between a highly 39 diversified set of small non-coding RNAs produced by dominant S-alleles and their 40 corresponding target sites on recessive S-alleles. By controlled crosses, we created 41 numerous heterozygous combinations of S-alleles in Arabidopsis halleri and developed an 42 RT-qPCR assay to compare allele-specific transcript levels for the pollen determinant of 43 self-incompatibility (SCR). This provides the unique opportunity to evaluate the precise 44 base-pairing requirements for effective transcriptional regulation of this target gene. We 45 found strong transcriptional silencing of recessive SCR alleles in all heterozygote 46 combinations examined. A simple threshold model of base-pairing for the sRNA-target 47 interaction captures most of the variation in SCR transcript levels. For a subset of S-alleles, 48 we also measured allele-specific transcript levels of the determinant of pistil specificity 49 (SRK) and found sharply distinct expression dynamics throughout flower development 50 between SCR and SRK. In contrast to SCR, both SRK alleles were expressed at similar 51 levels in the heterozygote genotypes examined, suggesting no transcriptional control of 52 dominance for this gene. We discuss the implications for the evolutionary processes 53 associated with the origin and maintenance of the dominance hierarchy among self-54 incompatibility alleles.

55

Introduction

56	Small non-coding RNAs are short RNA molecules (20-25nt) with a range of regulatory						
57	functions (Vazquez et al., 2010; Aalto & Pasquinelli, 2012). The best-known members of						
58	this class of molecules are microRNAs, which are typically involved in post-transcriptional						
59	gene silencing and regulate the activity of their target gene in <i>trans</i> by either mRNA						
60	cleavage (quickly followed by degradation) or by blocking translation (Li et al., 2014). In						
61	some cases, the action of microRNAs leads to the production of secondary phased short						
62	interfering RNAs (pha-siRNAs) by their target coding or non-coding sequence, which in						
63	turn can regulate other downstream targets (Fei et al., 2013). Another major set of small						
64	RNAs is heterochromatic short interfering RNAs (hc-siRNAs) which are mediating						
65	transcriptional silencing of repeat sequences in the genome through epigenetic						
66	modification by the RNA-dependent DNA methylation pathway (RdDM, Matzke et al.,						
67	2009).						
68	Both microRNAs and siRNAs guide their effector molecules (members of the						
69	ARGONAUTE gene family: AGO1 and AGO4, respectively) to their target sites by						
70	sequence similarity through base-pairing. For plant microRNAs, sequence similarity with						
71	the target sequence is typically very high and appears to be a shared feature of all						
72							
	functionally verified interactions (Wang et al., 2015). High base-pairing complementarity,						
73	functionally verified interactions (Wang <i>et al.</i> , 2015). High base-pairing complementarity, however, is not the sole determinant of target specificity, and the position of the						
73 74	functionally verified interactions (Wang <i>et al.</i> , 2015). High base-pairing complementarity, however, is not the sole determinant of target specificity, and the position of the mismatches along the microRNA:target duplex is also important. Indeed, expression assays						
73 74 75	functionally verified interactions (Wang <i>et al.</i> , 2015). High base-pairing complementarity, however, is not the sole determinant of target specificity, and the position of the mismatches along the microRNA:target duplex is also important. Indeed, expression assays showed that while individual mismatches typically have limited functional consequences,						
73 74 75 76	functionally verified interactions (Wang <i>et al.</i> , 2015). High base-pairing complementarity, however, is not the sole determinant of target specificity, and the position of the mismatches along the microRNA:target duplex is also important. Indeed, expression assays showed that while individual mismatches typically have limited functional consequences, they can also entirely inactivate the interaction when they hit specific positions such as, for						

78	et al., 2006). Furthermore, the position of mismatches along the microRNA:target duplex
79	also seems to be crucial, with a greater tolerance in the 3' than the 5' region of the
80	microRNA (up to four mismatches generally have limited functional consequences in the
81	3' region, while only two mismatches in the 5' region seem sufficient to abolish the target
82	recognition capability; Liu et al., 2014, Mallory et al., 2004; Parizotto et al., 2004; Schwab
83	et al., 2005). These observations have led to the formulation of general "rules" for
84	microRNA targeting (Axtell & Meyers, 2018), but at the same time they also revealed a
85	large number of exceptions. As a result, in silico prediction of microRNA target sites
86	currently remains a difficult challenge (Ding et al., 2012; Axtell & Meyers, 2018). For
87	other types of small RNAs (pha-siRNAs and hc-siRNAs), even less is known about the
88	base-pairing requirements for targeting, mostly because of the absence of experimentally
89	confirmed examples of discrete, single siRNA target sites either in cis or in trans (Wang et
90	al., 2015).

91 In this context, the recent discovery by Tarutani et al. (2010), Durand et al. (2014) and 92 Yasuda et al., (2016) of a highly diversified set of small non-coding RNAs at the gene 93 cluster controlling self-incompatibility (SI) in Brassicaceae, provides an experimentally 94 tractable model to evaluate the base-pairing requirements for silencing by a set of sRNAs 95 that are regulating expression of a single gene. Sporophytic SI is a genetic system that 96 evolved in several hermaphroditic plant lineages to enforce outcrossing by preventing self-97 fertilization, hence avoiding inbreeding depression (De Nettancourt, 2001). In the 98 Brassicaceae family, SI is controlled by a single genomic region called the "S-locus", 99 which contains two tightly linked genes, namely SCR and SRK, that encode the pollen S-100 locus cysteine-rich and the stigma S-locus receptor kinase recognition proteins,

101 respectively. This system involves a polymorphism in which multiple deeply diverged 102 allelic lines are maintained, and accordingly a large number of S-alleles is typically found 103 in natural populations of self-incompatible species (Castric & Vekemans, 2004). With such 104 a large allelic diversity and the very process of self-rejection, most individual plants are 105 heterozygotes at the S-locus. Yet in most cases, only one of the two S-alleles in a 106 heterozygous genotype is expressed at the phenotypic level in either pollen or pistil, as can 107 be revealed by controlled pollination assays on pollen or pistil tester lines (Llaurens *et al.*, 108 2008; Durand *et al.*, 2014). Which of the two alleles is expressed is determined by their 109 relative position along a dominance hierarchy, whose molecular basis for the pollen 110 phenotype has been initially studied in the genus Brassica. In this genus, dominance is 111 controlled at the transcriptional level in pollen (Schopfer 1999, Kakizaki et al. 2003). 112 Transcriptional silencing of recessive alleles by dominant alleles is caused by 24nt-long 113 trans-acting small RNAs produced by dominant S-alleles and capable of targeting a DNA 114 sequence in the promoter sequence of the SCR gene of recessive S-alleles, provoking DNA 115 methylation (Shiba et al. 2006). Details of how these sRNAs achieve their silencing 116 function remain incompletely understood (Finnegan *et al.*, 2011), but it is clear that their 117 biogenesis is similar to that of microRNAs (*i.e.*, they are produced by a short hairpin 118 structure), while their mode of action is rather reminiscent of that of siRNAs (*i.e.*, the 119 transcriptional gene silencing functions through recruitment of the methylation machinery). 120 Strikingly, the full dominance hierarchy in the Brassica genus seems to be controlled by 121 just two small RNAs called *Smi* and *Smi2* (Tarutani et al., 2010, Yasuda et al. 2016). *Smi* 122 and Smi2 target distinct DNA sequences, but both are located in the promoter region of 123 SCR, and both seem to involve DNA methylation and 24-nt active RNA molecules.

124 The dominance hierarchy in Brassica is, however, peculiar in that only two ancestral allelic 125 lineages segregate in that genus (the class I and class II alleles referred to above, see *e.g.* 126 Leducq *et al.*, 2014), whereas self-incompatible species in Brassicaceae typically retain 127 dozens of highly divergent ancestral allelic lineages (Castric & Vekemans, 2004). A recent 128 study showed that in Arabidopsis halleri, a Brassicaceae species with multiple allelic 129 lineages at the S-locus, the dominance hierarchy among S-alleles in pollen is controlled by 130 not just two but as many as eight different sRNA precursor families and their target sites, 131 whose interactions collectively determine the position of the alleles along the hierarchy 132 (Durand *et al.*, 2014). In that genus, much less is known about the mechanisms by which 133 the predicted sRNA-target interactions translate into the dominance phenotypes. First, the 134 expression dynamics of the SCR gene across flower development stages is poorly known. 135 Indeed, Kusaba et al. (2002) measured expression of SCR alleles in A. lyrata, but focused 136 on only two S-alleles (SCRa and SCRb, also known as AlSCR13 and AlSCR20, 137 respectively, in Mable *et al.* 2003) and showed striking differences in their expression 138 dynamics in anthers. Hence, the developmental stage at which the transcriptional control of 139 dominance in pollen should be tested is not precisely known. Second, while they did 140 confirm monoallelic expression, consistent with the observed dominance relationship 141 between the two alleles (SCRb > SCRa, Kusaba *et al.* 2002), the fact that only a single 142 heterozygote combination was measured among the myriad possible combinations given 143 the large number of S-alleles segregating in that species (at least 43 S-alleles: Genete *et al.*, 144 2020) prevents generalization at this step. Hence, a proper experimental validation of the 145 transcriptional control of dominance among S-alleles in the Arabidopsis genus is still 146 lacking. Third, Durand et al., (2014) observed rare sRNA-target interaction predictions that

147 did not agree with the observed dominance phenotype. In particular, they identified pairs 148 of S-alleles where no sRNA observed as being produced by the dominant allele was 149 predicted to target the SCR gene of the recessive one, while the dominance phenotype had 150 been well established phenotypically by controlled crosses (e.g. Ah04>Ah03) suggesting 151 the possibility that mechanisms other than transcriptional control may be acting. 152 Conversely, in other rare cases, sRNAs produced by a recessive S-allele were predicted to 153 target the SCR gene of a more dominant allele, suggesting exceptions to the set of base-154 pairing rules used to predict target sites. Fourth, the target sites for the two sRNAs in 155 Brassica were both located in the promoter sequence (Tarutani et al., 2010, Yasuda et al. 156 2016), and can thus reasonably be expected to prevent transcriptional initiation through 157 local modification of the chromatin structure associated with DNA methylation. Many of 158 the predicted sRNA target sites in A. halleri, however, are rather mapped to the SCR intron 159 or the intron-exon boundary (beside some in the promoter as well, Durand et al. 2014), 160 which suggests that distinct silencing pathways might be acting (Cuerda-Gil & Slotkin, 161 2016). It thus remains to be determined whether transcriptional control is also valid when 162 the targets are at other locations along the SCR gene structure. Finally, the dominance 163 hierarchy at the female determinant SRK differs from that at SCR, co-dominance being 164 more frequent than on the pollen side both in Brassica (Hatakeyama et al., 2001) and in A. 165 halleri (Llaurens et al., 2008). Limited transcriptional analysis in Brassica and Arabidopsis 166 suggests that dominance in pistils is not associated with SRK expression differences, but 167 again the number of allelic pairs tested has remained limited (Suzuki et al. 1999; Kusaba et 168 al. 2002).

169	Here, we take advantage of the fact that dominance interactions in Arabidopsis SI are
170	controlled in pollen by a diversity of sRNAs and the diversity of their target sites to
171	determine the base-pairing requirements for successful small-RNA mediated
172	transcriptional silencing of recessive SCR alleles. We first used controlled crosses to obtain
173	a large collection of A. halleri plants in which S-alleles were placed in various
174	homozygous and heterozygote combinations for which pairwise dominance interactions
175	had been determined. We then developed and validated a qPCR protocol for allele-specific
176	expression of a set of nine SCR and five SRK alleles in A. halleri. This enabled us to
177	analyse the expression dynamics across four flower developmental stages of each of these
178	alleles and test the transcriptional control of dominance for both genes in many
179	heterozygote combinations. We quantified the strength of silencing of recessive SCR
180	alleles and propose a quantitative threshold model for how sequence identity between the
181	small non-coding RNAs and their target sites results in silencing. We discuss the
182	implications of this model on the evolutionary processes associated with the origin and
183	maintenance of the S-locus dominance hierarchy in Brassicaceae.
184	

185

Material & Methods

186 Plant material

187 We used controlled crosses to create a collection of 88 A. halleri plants containing nine

188 different S-alleles (S1, S2, S3, S4, S10, S12, S13, S20, and S29) in a total of 37 of all 45

189 possible homozygous and heterozygous combinations. Some S-locus genotypes were

190 obtained independently by different controlled crosses and were considered below as

191	"biological replicates" (different genetic backgrounds, on average $n=2.05$ biological
192	replicates per S-locus genotype, Table S1 & S2). Three plants were cloned by cuttings and
193	considered as "clone replicates" (identical genetic background, Table S1) that we used to
194	evaluate the expression variance associated with different genetic backgrounds.
195	Each plant was genotyped at the S-locus using the PCR-based protocol described in
196	Llaurens et al. (2008). Pairwise dominance interactions between S-alleles of the
197	heterozygote combinations were either taken from Llaurens et al. (2008); Durand et al.
198	(2014); Leducq et al. (2014) or were newly determined by controlled pollination assays
199	following the protocol of Durand et al., (2014). In a few instances, relative dominance
200	status of the two alleles had not been resolved phenotypically and were inferred from the
201	phylogeny of SRK alleles, which is largely consistent with the dominance hierarchy
202	(Durand et al. 2014). The pairwise dominance interactions between these alleles as
203	determined by pollen and pistil compatibility phenotypes of heterozygote plants are
204	reported in Table S3.

205 RNA extraction and reverse transcription

206 On each plant, we collected flower buds at four developmental stages: 1) five highly 207 immature inflorescence extremities (more than 2.5 days before opening, buds below 208 0.5mm, stages 1-10 in A. thaliana according to Smyth et al., 1990); 2) ten immature buds 209 (2.5 days before opening, between 0.5 and 1mm, approximately stage 11); 3) ten mature 210 buds (one day before opening, longer than 1mm, approximately stage 12); and 4) ten open 211 flowers (approximately stages 13-15). These stages were characterized by establishing the 212 size distribution within each stage and measuring the time to flower opening based on ten 213 buds. Samples collected were flash-frozen in liquid nitrogen, then stored at - 80°C before

214	RNA extraction. Tissues were finely ground with a FastPrep-24 5G Benchtop
215	Homogenizer (MP Biomedicals, Model #6004-500) equipped with Coolprep 24 x 2mL
216	adapter (6002-528) and FastPrep Lysis Beads & Matrix tube D. Total RNAs were
217	extracted with the Arcturus "Picopure RNA isolation" kit from Life Science (PN:
218	KIT0204) according to the manufacturer's protocol, including a step of incubation with
219	DNAse to remove gDNA contamination. We normalized samples by using 1 mg of total
220	RNA to perform reverse-transcription (RT) using the RevertAid Fermentas enzyme
221	following the manufacturer's instructions.
222	Primer design
223	A major challenge to study expression of multiple S-alleles is the very high levels of
224	nucleotide sequence divergence among them, precluding the possibility of designing qPCR
225	primers that would amplify all alleles of the allelic series (both for SRK and SCR). Hence,
226	we rather designed qPCR primers specifically targeted towards each of the SCR and SRK
227	alleles, and for each heterozygote genotype we independently measured expression of both
228	alleles of each gene. Primers were designed based on genomic sequences from BAC clones
229	(Goubet et al. 2012; Durand et al. 2014; Novikova et al. 2017), with a length of ~20
230	nucleotides, a GC content around 50% and a target amplicon size around 150nt (Figure
231	S1). For SCR, we focused on a set of 9 S-alleles. Whenever possible, we placed primers on
232	either side of the SCR intron to identify and discard amplification from residual gDNA.
233	However, because the coding sequence of the SCR gene is short, the number of possible
234	primers was limited and this was not always possible. In two cases (SCR01 and SCR20),
235	both primers were thus located within the same exon. For SRK alleles, the primers were
236	also designed on either side of the first intron to avoid genomic contamination (Figure S2).

237	Because no differences in transcript levels were previously observed between dominant
238	and recessive SRK alleles (Suzuki et al. 1999; Kusaba et al. 2002), and given the effort
239	required to optimize new qPCR primers, we decided to place more effort on SCR and
240	focused on a more limited number of <i>SRK</i> alleles ($n=5$). To obtain relative expression
241	levels across samples, we used <i>actin</i> 8 (At1g49240) as a housekeeping gene for
242	standardization after we verified that the A. <i>thaliana</i> and A. <i>halleri</i> sequences are identical
243	at the primer positions (An et al. 1996). Primer sequences are reported in Table S4.
244	Quantitative real-time PCR
245	On each cDNA sample, at least three qPCR reactions (referred to below as "technical"
246	replicates) were performed for actin 8 and for each of the S-alleles contained in the
247	genotype (one S-allele for homozygotes, two S-alleles for heterozygotes). The runs were
248	made on a LightCycler480 (Roche) with iTaq Universal SYBR Green Supermix (Bio-rad,
249	ref 172-5121). Amplified cDNA was quantified by the number of cycles at which the
250	fluorescence signal was greater than a defined threshold during the logarithmic phase of
251	amplification using the LightCycler 480 software release 1.5.0 SP3. The relative transcript
252	levels are shown after normalisation with actin amplification through the comparative $2^{-\Delta Ch}$
253	method (Livak & Schmittgen, 2001). The Ct_{SCR} and Ct_{SRK} values of each technical
254	replicate were normalized relative to the average Ct_{actin} measure across the three replicates.
255	Validation of qPCR primers at the dilution limits
256	Given the very large nucleotide divergence between alleles of either SCR or SRK, cross-
257	amplification is unlikely. However, to formally exclude that possibility, we first performed
258	cross-amplification experiments by using each pair of SCR primers on a set of cDNA

samples that did not contain that target SCR allele but instead contained two other SCR

260	alleles in various heterozygous genotypic combinations ($n=7$ on average). In order to					
261	evaluate our ability to measure expression of SCR alleles in biological situations where					
262	they are expected to be transcriptionally silenced, we then used a series of limit dilutions to					
263	explore the loss of linearity of the relationship between Ct and the dilution factor (six to					
264	eight replicates per dilution level). Then we examined the shape of the melting curves to					
265	determine whether our measures at this limit dilution reflected proper PCR amplification or					
266	the formation of primer dimers. Finally, we used water in place of cDNA to evaluate the					
267	formation of primer dimers in complete absence of the target template DNA.					
268	Expression dynamics and the effect of dominance					
269	We used generalized linear mixed models (lme4 package in <i>R</i> ; Bates <i>et al.</i> , 2014) to					
270	decompose Ct values normalized by the actin 8 control (as the dependent variable) into the					
271	effects of five explanatory variables. Two of them were treated as fixed effects:					
272	developmental stage (4 categories) and relative dominance of the allele studied in the					
273	genotype (3 categories: recessive, dominant, homozygous). Because expression of the					
274	different SCR (and SRK) alleles was quantified by different primer pairs with inevitably					
275	different amplification efficiencies, Ct values cannot be directly compared across alleles					
276	and accordingly we included the identity of SCR or SRK alleles as random effects.					
277	Biological and clone replicates were also treated as random effects, with clones nested					
278	within biological replicates (Table S5). We visually examined normality of the residuals of					
279	the model under different distributions of $2^{-\Delta Ct}$, including Gaussian, Gamma and Gaussian					
280	with logarithmic transformations. We tested whether the different S-alleles have different					
281	expression profiles across developmental stages, as suggested by Kusaba et al. (2002) for					
282	SCR in A. lyrata, by using ANOVA to compare nested models in which a random effect					

for the interaction between the "allele measured" and "stage" effects was either absent

284 (model 1) or introduced (model 2, Table S5b) in addition to the fixed effect of stage. The

- existence of this interaction was tested for *SCR* and *SRK* separately.
- 286 Target features and silencing effect.

287 Expression of SCR in heterozygote genotypes in A. halleri is controlled by a small RNA-288 based regulatory machinery (Durand *et al.* 2014). We then sought to determine how SCR 289 transcript levels were affected by specific features of the small RNA-target interactions 290 between S-alleles. We retrieved sRNA sequencing data from individuals carrying eight of 291 the nine S-alleles considered (S01, S03, S04, S10, S12, S13 and Ah20 from Durand et al. 292 (2014) and S02 from Novikova et al. (2017)). No sRNA sequencing data were available 293 for the last S-allele (S29). We used these sRNA sequencing data to determine the complete 294 set of sRNA molecules uniquely produced by the annotated sRNA precursors of each of 295 these eight S-alleles. To do that, we mapped the sRNA reads to the sRNA precursor 296 sequences carried by the respective S-alleles after excluding those that mapped to other 297 locations in the closely related A. lyrata genome (Durand et al. 2014). For each sRNA 298 produced by a given S-allele, we then predicted putative target sites on the SCR gene of all 299 other S-alleles including 2kb of genomic nucleotide sequence both upstream and 300 downstream of SCR using a dedicated alignment algorithm and scoring matrix, as 301 described in Durand et al. (2014). Briefly, alignment quality was assessed by a scoring 302 system based on the addition of positive or negative values for matching nucleotides (+1), 303 mismatches and gaps (-1), taking into account the non-canonical G:U interaction (-0.5).

- 304 For each pair of alleles considered, only the sRNA/target combination with the highest
- 305 score was selected for further analysis (Table S6). The analysis was performed regardless

306 of the dominance relationship (i.e. we predicted putative target sites of sRNAs produced by 307 dominant S-alleles onto recessive S-alleles, and reciprocally from recessive S-alleles onto 308 dominant S-alleles). Because the mechanisms by which silencing is achieved remain 309 unclear at this stage, we did not filter these sRNA further in terms of length or identity of 310 the 5' nucleotide, in line with Durand et al. (2014). In the cases where the target with the 311 highest score was due to a sRNA with non-canonical size (anything but 21 or 24nt), we 312 also reported the best target score among the set of 21 and 24nt sRNA molecules produced 313 by the same S-allele (Table S6). We used Akaike Information Criteria (AIC) to compare 314 how well different base-pairing scores for target site identification predicted the level of 315 SCR expression (and hence the silencing phenomenon), varying the threshold from 14 to 316 22. Lower values of AIC are associated with a best fit of the model. We then added a new 317 fixed effect in our basal model to test whether targets at different positions along the SCR 318 gene (5 categories: 5' portion, exons, intron, overlapping the exon-intron boundary or 3' 319 portion of the gene) are associated with different strengths of silencing. For this analysis, 320 we included only targets above the threshold identified (score ≥ 18). 321 Effective silencing of recessive SCR alleles in Brassica rapa is dependent upon 322 combinations of individual sequence mismatches between the Smi & Smi2 small RNAs and 323 their target sites in the class II alleles (Yasuda *et al.*, 2016), but interaction in this study 324 relied on raw counts of nucleotide mismatches and were thus not directly comparable to 325 our results. To determine whether the base-pair requirements for silencing are similar, we 326 thus reanalysed these interactions using our scoring system to compare the small RNA-327 target alignment scores between Brassica and Arabidopsis (Tarutani et al., 2010, Yasuda et 328 al., 2016).

329	Finally, we used the phylogeny in Durand et al. (2014) to classify sRNA/target interactions
330	into "recent" (mir867 and mirS4) and "ancient" (mirS1, mirS2 mirS3, mirS5, mir1887 and
331	mir4239). Based on this classification, we used a linear regression to compare the
332	alignment score for recent and ancient sRNAs and tested the hypothesis that interactions
333	with base-pairing scores above the threshold at which silencing was complete correspond
334	to recently emerged interactions that have not yet accumulated mismatches.

335

336

Results

337 Validation of the qPCR protocol and the allele-specific primers 338 The specificity test confirmed the absence of cross-amplification between alleles, as the Ct 339 measures for water control and cross amplification were comparably high (around Ct=34) 340 and both were higher than the positive controls (median *Ct*=22, Figure S3). Overall, serial 341 dilutions of the template cDNA confirmed linearity of the Ct measure within the range of 342 values observed for a given allele across the different conditions examined (Figure S4a). 343 Because we study a silencing phenomenon, we then explored how signal was lost at the 344 dilution limits. As expected, linearity started to be lost at very low cDNA concentrations 345 (in particular for alleles SCR01, SCR02, SCR04, SCR13 and SCR20, Figure S4a), and 346 examination of melting curves under these conditions indicated the formation of primer 347 dimers rather than the expected transcripts. Hence, we note that comparing levels of 348 expression for a given allele between different recessive contexts (e.g. when silenced by 349 different sRNAs) should be challenging, especially for the above-mentioned alleles.

Linearity was good for most *SRK* alleles (Figure S4b) except for *SRK12* (data not shown),
so this allele was excluded from further analyses.

352	SCR and SRK expression dynamics across flower development stages
353	In total, we performed 344 RNA extractions and RT-PCR from the 37 different S-locus
354	genotypes sampled at four developmental stages. For SCR, we measured 1,838 Ct_{SCR}/Ct_{actin}
355	expression ratios (i.e. an average of 26.9 expression measures per S-allele in each diploid
356	genotype, Table S1). For SRK, we measured 480 Ct_{SRK}/Ct_{actin} ratios (<i>i.e.</i> an average of 11.1
357	expression measures per S-allele in each diploid genotype, Table S2). Distribution of the
358	residuals of the generalized mixed linear model was closest to normality after log-
359	transformation of the ratios (Figure S6). As expected, measured expression levels were
360	more highly repeatable across clones than across biological replicates for a given S-locus
361	genotype (deviance estimates of 0.40, 1.08, respectively, Table S5a). The deviance
362	associated with the allele's expression dynamic was higher (deviance = 4.56), although we
363	note that the technical error was also important (deviance = 6.08 , Table S5a). We first
364	examined the expression dynamics of the different SCR alleles. Because recessive SCR
365	alleles were consistently silenced (see below), we isolated the effect of developmental
366	stages by focusing only on genotypes in which each focal allele was known to be dominant
367	at the phenotypic level (Figure 1a). Overall, we observed a strong pattern of variation
368	among stages (F-value: 10.76, p-value: 5.7e-5, Table S5c) with high expression of SCR in
369	buds at early developmental stages (<0.5 to 1mm), and low expression in late buds right
370	before opening and in open flowers. This pattern is consistent with degeneration in these
371	stages of the anther tapetum, the cellular layer where SCR is expected to be expressed. The
372	expression dynamics of SRK was sharply different from that of SCR, with monotonously

373	increasing expression in the course of flower development, with lowest expression in
374	immature buds (<0.5mm) and highest expression in open flowers (Figure 1b, F-value:
375	4.411, <i>p</i> -value: 0.007, Table S5h). We found evidence that the expression dynamics varied
376	across S-alleles, not only for SCR (Chi ² : 308.19, <i>p</i> -value < 2.2e-16, Table S5b) in line with
377	Kusaba et al., (2002), but also for SRK (Chi ² : 6.9103, p-value 0.00857, Table S5g).

378

379 Transcriptional control

Based on these results, we averaged $2^{-\Delta Ct}$ values across <0.5mm to 1mm stages to compare 380 381 expression of a given focal SCR allele between genotypic contexts where it was either 382 dominant or recessive relative to the other allele present in the diploid genotype. Of the 54 383 pairwise interactions for which the dominance phenotype had been firmly established by 384 controlled crosses and the qPCR assay had been performed for both SCR alleles (Table 385 S3), as many as 51 (94.4%) are associated with strong asymmetries in transcript levels, 386 with high expression of the dominant SCR allele and low expression of the recessive SCR 387 allele (Figure 2). Hence, our expression data were largely consistent with the hypothesis of 388 transcriptional control of the dominance hierarchy in pollen genotypic combinations. SCR 389 transcripts of the most recessive allele (S1) were only detected in an S1S1 homozygote 390 genotype, but not in any other genotypic combination. Climbing up the dominance 391 hierarchy from most recessive to most dominant, expression of SCR was detected in an increasing number of heterozygous combinations, in strong agreement with phenotypic 392 393 dominance (Figure 2). At the top of the dominance hierarchy, the two most dominant 394 alleles, SCR13 and SCR20, were expressed in all heterozygous contexts, including when 395 they formed a heterozygote combination with one another (S13S20), also as expected

396	given the codominance observed between them at the phenotypic level (Durand et al.,
397	2014). This general rule had a few exceptions however (indicated by arrows on Figure 2).
398	Specifically, we observed low expression for both SCR01 and SCR12 when in
399	heterozygote combination (S01S12 genotypes) and for both SCR10 and SCR12 in
400	heterozygote combination (S10S12 genotype), which is not consistent with the
401	documented phenotypic dominance of these alleles in pollen ($S12 > S01$ and $S12 > S10$; see
402	Table S3). We also detected expression of both SCR02 and SCR29 when placed in
403	heterozygote combination, which might explain the unusual phenotypic data indicating
404	robust rejection of pollen from this heterozygote genotype on the [S02] tester line, but only
405	partial compatibility on the [S29] tester line (Table S3). Hence, the dominance interaction
406	between these two alleles may be partial, both at the transcriptional and phenotypic levels.
407	Interestingly, these two alleles belong to class III, which in A. lyrata tend to show
408	inconsistent (or leaky) SI responses (Kusaba et al. 2001).
409	Overall, in spite of these three exceptions, we observed a striking contrast in transcript
410	levels for a given allele according to its relative phenotypic dominance status in the
411	genotype (<i>F</i> -value = 19.538; <i>p</i> -value < 2.2e-16, Table S5c), suggesting complete silencing
412	of recessive alleles. Specifically, we observed an average 145-fold decrease in transcript
413	abundance in genotypes where a given focal allele was phenotypically recessive as
414	compared to genotypes in which the same focal allele was dominant. We note that the
415	silencing was so strong that the Ct values associated with recessive SCR transcripts were
416	comparable with those of the negative controls (Figure S3) and close to the detection limits
417	of our method as determined by the break of linearity of the dilution experiment, such that
418	the magnitude of the calculated fold-change value is probably under-estimated (Figure S1).

419	In strong contrast.	we found r	no significant	effect of	dominance in	pistils on SRK
11/	in buong continuot,	ne rouna i	io significant		aominance m	pionio on orni

- 420 expression (F-value: 6.8884 p-value: 0.068244; Figure 3, Table S5h), confirming the
- 421 absence of transcriptional control of dominance for *SRK*.
- 422 Target features and silencing effect

423 Levels of *SCR* expression of any given focal allele varied sharply with the alignment score

424 of the "best" target available for the repertoire of canonical sRNAs produced by the other

425 allele present in the genotype (Figure 4a). Specifically, we observed on average high levels

426 of *SCR* transcripts when the score of their best predicted target was low, but consistently

427 low levels of *SCR* transcripts when the score of the best target was high (Figure 4a, Table

428 S5d). Strikingly, the transition between high expression and low expression was abrupt

429 (around an alignment score of 18), suggesting a sharp threshold effect rather than a

430 quantitative model for transcriptional silencing.

431 In two cases, the presence of a target with a high score within the *SCR* gene of the

432 dominant allele was associated with high relative SCR expression (in agreement with the

433 dominant phenotype established by controlled crosses), confirming the absence of

434 silencing (target of Ah04mir4239 on *SCR20*, score=20; and target of Ah10mir4239 on

435 *SCR20*, score =21; Figure 5a) and suggesting that these interactions are not functional.

436 Examining in detail these two exceptions did not reveal mismatches at the 10-11th

437 nucleotide position, suggesting that mismatches at other positions have rendered these

438 sRNA-target interactions inactive (Figure 5a). We note that the target of Ah10mir4239

- 439 with the highest score is predicted for a sRNA with non-canonical size (25nt), but this
- 440 precursor also produces a canonical 24nt isomir with a score above the threshold (score =
- 441 20, Table S6). These two sRNAs (Ah04mir4239 and Ah10mir4239) have a 5' nucleotide

different from the expected "A" for 24nt sRNAs, possibly suggesting that loading into an
improper AGO protein may have rendered these predicted interactions inactive. Another
exception concerns the observed low score (15.5) for the best match between a sRNA from
the dominant allele Ah04mirS4 and its best putative target at the recessive *SCR03* (Figure
5b). Whether *SCR*04 silences *SCR03* through this unusual target or through another elusive
mechanism remains to be discovered.

448 In spite of the generally very low expression of all recessive alleles, we found marginal

449 evidence that the strength of silencing experienced by a given *SCR* allele varies across

450 genotypic combinations for a given allele (F-value=2.221, *p*-value = 0.0756, Table S5i).

451 However, there was no evidence that the position of the target site on the measured allele

452 (promoter; intron; intron-exon boundary; upstream vs. downstream) could explain this

453 variation (F-value=1.7061, *p*-value = 0.1928, TableS5e). We also found no effect of the

454 inferred age of the miRNA on the mean alignment score (mean= 20.41 and 20.22 for

455 recent or ancient miRNAs, respectively; F-value: 0.0362; *p*-value = 0.8504, Table S5j).

456 Finally, we compared the alignment scores observed here in Arabidopsis with those in

457 Brassica for *Smi & Smi2* on their *SCR* target sequences. A clear threshold was also

458 observed, but in Brassica the alignment score threshold distinguishing dominant from

459 recessive interactions was 16.5 instead of 18 (Table S6), suggesting distinct base-pairing

460 requirements for effective silencing in these two systems.

462

Discussion

463 Determining the base-pairing requirement for sRNA silencing in plants has remained 464 challenging because the "rules" used for target prediction have typically been deduced 465 from observations that conflate distinct microRNA genes and their distinct mRNA targets 466 over different genes. Moreover, detailed evaluations of the functional consequences of 467 mismatches have relied on heterologous reporter systems (typically GFP in transient 468 tobacco assays), hence limiting the scope of the phenotypic consequences that can be 469 studied. Here, we build upon the inter-allelic regulatory system controlling transcriptional 470 activity of alleles of the SI system in Arabidopsis revealed in Durand et al. (2014), where 471 multiple sRNAs regulate target sites on alleles of a single gene (SCR), and in which we are 472 able to make a direct link between the sRNA-target interactions, the level of SCR transcript 473 and the encoded phenotype (dominance/recessivity interaction). The first step was to 474 clarify several aspects of the expression pattern of the genes controlling SI in A. halleri, 475 which was necessary to confirm that the dominance interactions in Arabidopsis involve 476 transcriptional regulation.

477 Expression profile

478 Earlier accounts had suggested that alleles of the allelic series may differ from one another

- 479 in their expression profile (Kusaba et al., 2002). In line with Kakizaki et al., (2003),
- 480 Suzuki et al., (1999); Schopfer et al., (1999); Takayama et al., (2000) and Shiba et al.,
- 481 (2002), we found maximal expression of *SCR* in early buds but low or no expression at the
- 482 open flower stage. This expression pattern is consistent with in situ hybridization
- 483 experiments showing that SCR transcripts are localized in the tapetum, a specialized layer
- 484 of cells involved in pollen grains coating (Iwano *et al.*, 2003) which undergoes apoptosis

485	and is quickly degraded as the development of pollen grains inside the anther progresses
486	(Murphy & Ross, 1998; Takayama et al., 2000). We confirmed that differences exist in the
487	temporal dynamics of expression among alleles, as suggested by Kusaba et al. (2002) in A.
488	lyrata, possibly as the result of strong sequence divergence of the promotor sequences of
489	the different SCR alleles. Finally, we confirmed that SCR and SRK have sharply distinct
490	expression dynamics throughout flower development. Indeed, transcript levels of SRK
491	increased steadily along development and were very low in early buds, consistent with the
492	observation that SI can be experimentally overcome to obtain selfed progenies by "bud-
493	pollination" (Llaurens et al. 2009).

494 Generality of the transcriptional control of dominance in Arabidopsis

495 Based on this clarified transcriptional dynamics, we confirmed the generality of the 496 transcriptional control of dominance for SCR, with as much as 96.3% of the documented 497 dominance interactions associated with mono-allelic expression of the dominant SCR allele 498 and complete silencing of the recessive SCR allele in heterozygote genotypes. Even in the 499 single heterozygote genotype where in our previous study (Durand et al., 2014) no sRNA 500 produced by the phenotypically dominant allele was predicted to target the sequence of the 501 phenotypically recessive SCR allele (e.g. S04>S03), transcripts from the recessive SCR03 502 allele were undetected. This suggests either that some functional sRNAs or targets have 503 remained undetected by previous sequencing and/or by our *in silico* prediction procedures, or that mechanisms other than sRNAs may cause transcriptional silencing for some S-allele 504 505 combinations. Regardless of the underlying cause, the generality of the transcriptional 506 control of dominance suggests that the simple comparison of transcript levels between the 507 two alleles in a heterozygote genotype could be used as a first approximation to determine

their relative dominance levels. In contrast, we confirmed the absence of transcriptional
control for *SRK*, for which both alleles were consistently expressed at similar levels in all
heterozygote genotypes examined, irrespective of the (pistil) dominance phenotype. For *SRK*, other dominance mechanisms must therefore be acting, which are yet to be

512 discovered (*e.g.* Naithani *et al.*, 2007).

513 Variation in the strength of silencing

514 An important feature of the silencing phenomenon is that the decrease of transcript levels

515 for recessive *SCR* alleles was very strong in heterozygous genotypes, bringing down

516 transcript levels below the limits of detection in most cases. This is in line with the

517 intensity of transcriptional silencing by heterochromatic siRNAs (typically very strong for

518 transposable element sequences, see Marí-Ordóñez et al., 2013), while post-transcriptional

519 gene silencing by microRNAs can be more quantitative (Liu *et al.*, 2014). As a result of

520 this strong decrease of transcript levels, the strength of silencing appeared independent

521 from the position of the sRNA target along the SCR gene (promoter vs. intron), although

522 we note that our power to distinguish among levels of transcripts of recessive alleles,

523 which were all extremely low, is itself fairly low. It remains to be discovered whether the

524 different positions of the sRNA targets (Durand *et al.*, 2014) do indeed imply different

525 transcriptional silencing mechanisms.

526 A simple threshold model for sRNA-based silencing

527 Based on the many allelic combinations where we could compare the agnostic prediction

528 of putative target sites with the level of transcriptional silencing, we find that a simple

529 threshold model for base-pairing between sRNAs and their target sites captures most of the

530 variation in SCR expression in heterozygotes. This result provides a direct experimental

531 validation of the *ad-hoc* criteria used in Durand *et al.*, (2014). However, our results also 532 indicate that this quantitative threshold is not entirely sufficient to capture the complexity 533 of targeting interactions. Indeed, in three of the 54 cases tested this simple threshold model 534 would inappropriately predict targeting of a dominant SCR allele by a sRNAs from a more 535 recessive allele, yet the dominant SCR allele was expressed at normal levels with no sign 536 of silencing in these heterozygote genotypes (Figure 5a). The targeting interaction may be 537 abolished either by defects in the sRNA itself (e.g. for Ah04mir4239 the 5' nucleotide is a 538 G, while the majority of functional 24 nt small RNA molecules end with a 5'A, which may 539 interfere with loading in the appropriate AGO protein). Alternatively, the targeting 540 interactions may be abolished by the position of the mismatches (at position 14 and 22 of 541 the Ah10mir4239 and at position 13 and 21 of the Ah04mir4239, both on SCR20). 542 Similarly, a single mismatch at position 10 in the Smi interaction in Brassica (Tarutani et 543 al., 2010) and in other microRNA-targets interactions (Franco-Zorrilla et al., 2007) was 544 shown to result in loss of function of the interaction (Table S6). Interestingly, quantitative 545 differences may exist between Arabidopsis and Brassica, as the experimentally validated 546 targets in Brassica (Tarutani et al., 2010; Yasuda et al., 2016) correspond to base-pairing 547 threshold below the one that we find in Arabidopsis (*i.e.* a target score of 16.5 seems 548 sufficient for silencing in Brassica vs. 18 in Arabidopsis). For Brassica, both class I and class II alleles have *Smi*, but a mismatch at the 10th position was proposed to explain why 549 550 the class II Smi is not functional. Here, we show that this mismatch drives the alignment 551 score below the 16.5 threshold and could be sufficient to explain the loss of function, 552 regardless of its position. Overall, although these small RNAs achieve their function in a 553 way that may be sharply different from classical microRNAs (DNA methylation vs.

mRNA cleavage), our results suggest that the sRNA-target complementarity rules for silencing in both cases are qualitatively consistent (Liu *et al.*, 2014). Better understanding the molecular pathway by which these sRNAs epigenetically silence their target gene (*SCR*) will now be key to determine whether this threshold model can be generalized to more classical siRNAs found across the genome, as evidence is still missing for such classes of sRNAs.

560 Implications for the evolution of the dominance hierarchy

561 The existence of a threshold model has important implications for how the dominance 562 hierarchy can evolve. In fact, our model suggests that a single SNP can be sufficient to turn 563 a codominance interaction into a dominance interaction (and vice-versa), making this a 564 relatively trivial molecular event. This is actually what Yasuda et al., (2016) observed in B. 565 rapa, where the combination of single SNPs at the sRNA Smi2 and its SCR target 566 sequences resulted in a linear dominance hierarchy among the four class II S-alleles found 567 in that species. Strikingly, in some cases, we observed base pairing at sRNA-target 568 interactions with very high alignment scores (up to 22), *i.e.* above the threshold at which 569 transcriptional silencing was already complete (score =18). Under our simple threshold 570 model, such interactions are not expected since complete silencing is already achieved at 571 the threshold, and no further fitness gain is therefore to be expected by acquiring a more 572 perfect target. A first possibility is that these interactions reflect the recent emergence of 573 these silencing interactions. In fact, one of the models for the emergence of new 574 microRNAs in plant genomes involves a partial duplication of the target gene, hence 575 entailing perfect complementarity at the time of origin that becomes degraded over time by 576 the accumulation of mutations (Allen et al., 2004). Under this scenario, the higher-than577 expected levels of sRNA-target complementarity could reflect the recent origin of these 578 sRNAs but we found no evidence of a difference in alignment score for young vs. old 579 sRNA precursors. A second possibility is that selection for developmental robustness is 580 acting to prevent the phenotypic switch from mono- to bi-allelic expression of SCR 581 (especially during stress events, Boukhibar & Barkoulas, 2016) that could be devastating 582 for the plant reproductive fitness (Llaurens et al. 2009). Indeed, we observed strong 583 variation in overall SCR expression when the sRNA target score of the companion allele is 584 below the threshold in the benign greenhouse conditions under which we grew our plants, 585 and it is possible that under stress conditions the epigenetic machinery may be less 586 efficient, hence requiring stronger base-pairing to achieve proper silencing. Finally, a third 587 possibility is that sRNA-target complementarity above the threshold reflects the pleiotropic 588 constraint of having a given sRNA from a dominant allele control silencing of the 589 complete set of target sequences from the multiple recessive alleles segregating, and 590 reciprocally of having a given SCR target in a recessive allele maintaining molecular match 591 with a given sRNA distributed among a variety of dominant alleles. Comparing the 592 complementarity score of sRNA/target interactions among sRNAs or targets that contribute 593 to high versus low numbers of dominance/recessive interactions will now require a more 594 complete depiction of the sRNA-target regulatory network among the larger set of S-alleles 595 segregating in natural populations.

597	Acknowledgments
598	We thank Sylvain Billiard and Isabelle de Cauwer for statistical advice and discussions,
599	Romuald Rouger and Anne Duputié for help with producing figures and Alexis Sarazin
600	and three anonymous reviewers for comments on the manuscript. This work was funded by
601	the European Research Council (NOVEL project, grant #648321). N.B. was supported by a
602	doctoral grant from the president of Université de Lille-Sciences et Technologies and the
603	French ministry of research. The authors also thank the Région Hauts-de-France, and the
604	Ministère de l'Enseignement Supérieur et de la Recherche (CPER Climibio), and the
605	European Fund for Regional Economic Development for their financial support
606	
607	Author Contribution
608	
	NB, SS, SB, ACH performed the molecular biology experiments. CP and ES obtained and
609	NB, SS, SB, ACH performed the molecular biology experiments. CP and ES obtained and took care of the plants. SS, IFL and XV provided advice on the experimental strategy and
609 610	NB, SS, SB, ACH performed the molecular biology experiments. CP and ES obtained and took care of the plants. SS, IFL and XV provided advice on the experimental strategy and interpretations. NB performed the statistical analyses. VC supervised the work. NB and
609 610 611	NB, SS, SB, ACH performed the molecular biology experiments. CP and ES obtained and took care of the plants. SS, IFL and XV provided advice on the experimental strategy and interpretations. NB performed the statistical analyses. VC supervised the work. NB and VC wrote the manuscript.
609610611612	NB, SS, SB, ACH performed the molecular biology experiments. CP and ES obtained and took care of the plants. SS, IFL and XV provided advice on the experimental strategy and interpretations. NB performed the statistical analyses. VC supervised the work. NB and VC wrote the manuscript.
 609 610 611 612 613 	NB, SS, SB, ACH performed the molecular biology experiments. CP and ES obtained and took care of the plants. SS, IFL and XV provided advice on the experimental strategy and interpretations. NB performed the statistical analyses. VC supervised the work. NB and VC wrote the manuscript. References cited
 609 610 611 612 613 614 	NB, SS, SB, ACH performed the molecular biology experiments. CP and ES obtained and took care of the plants. SS, IFL and XV provided advice on the experimental strategy and interpretations. NB performed the statistical analyses. VC supervised the work. NB and VC wrote the manuscript. References cited Aalto, A. P. , Pasquinelli, A. E. 2012. Small non-coding RNAs mount a silent revolution

- 616 Allen, E., Xie, Z., Gustafson, A. M., Sung, G. H., Spatafora, J. W., et al. 2004. Evolution
- 617 of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis*
- 618 *thaliana. Nature Genetics* 36: 1282–1290.
- An, Y. Q., McDowell, J. M., Huang, S., McKinney, E. C., Chambliss, S., et al. 1996.
- 620 Strong, constitutive expression of the Arabidopsis ACT2/ACT8 actin subclass in
- 621 vegetative tissues. *The Plant journal: for cell and molecular biology* 10: 107–121.
- 622 Axtell, M.J., Meyers, B.C., 2018. Revisiting criteria for plant miRNA annotation in the era
- 623 of big data. *The Plant Cell*: tpc.00851.2017.
- 624 Bates, D., Mächler, M., Bolker, B., Walker, S., 2014. Fitting Linear Mixed-Effects Models
- 625 using lme4. *Journal of Statistical Software* 67: 1–48.
- 626 Boukhibar, L. M., Barkoulas, M., 2016. The developmental genetics of biological
- 627 robustness. Annals of Botany 117: 699–707.
- 628 Castric, V., Bechsgaard, J., Schierup, M. H., Vekemans, X., 2008. Repeated adaptive
- 629 introgression at a gene under multiallelic balancing selection. *PLoS Genetics* 4.
- 630 Castric, V., Vekemans, X., 2004. Plant self-incompatibility in natural populations: A
- 631 critical assessment of recent theoretical and empirical advances. *Molecular Ecology* 13:
- 632 2873–2889.
- 633 Cuerda-Gil, D., Slotkin, R. K., 2016. Non-canonical RNA-directed DNA methylation.
- 634 *Nature Plants* 2: 16163.
- 635 Ding, J., Zhou, S., Guan, J., 2012. Finding MicroRNA Targets in Plants: Current Status
- and Perspectives. *Genomics, Proteomics and Bioinformatics* 10: 264–275.

- 637 Durand, E., Méheust, R., Soucaze, M., Goubet, P. M., Gallina, S., et al. 2014. Dominance
- 638 hierarchy arising from the evolution of a complex small RNA regulatory network.
- 639 *Science* 346: 1200–1205.
- 640 Fei, Q., Xia, R., Meyers, B. C., 2013. Phased, secondary, small interfering RNAs in
- 641 posttranscriptional regulatory networks. *The Plant Cell* 25: 2400–2415.
- 642 Finnegan, E. J., Liang, D., Wang, M., 2011. Self-incompatibility : Smi silences through a
- 643 novel sRNA pathway. *Trends in Plant Science* 16: 238–241.
- 644 Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., et al. 2007. Target
- 645 mimicry provides a new mechanism for regulation of microRNA activity. *Nature*
- 646 *Genetics* 39: 1033–1037.
- 647 Genete, M., Castric, V., Vekemans, X., 2020, Genotyping and *de novo* discovery of allelic
- 648 variants at the Brassicaceae self-incompatibility locus from short read sequencing data,
- 649 *Molecular Biology and Evolution*, msz258, https://doi.org/10.1093/molbev/msz258
- 650 Goubet, P. M., Bergès, H., Bellec, A., Prat, E., Helmstetter, N., et al. 2012. Contrasted
- 651 pattern of molecular evolution in dominant and recessive self-incompatibility
- haplotypes in Arabidopsis. *PLoS genetics* 8.
- 653 Hatakeyama, K., Takasaki, T., Suzuki, G., Nishio, T., Watanabe, M., et al. 2001. The S
- 654 Receptor Kinase gene determines dominance relationships in stigma expression of self-
- 655 incompatibility in Brassica. *Plant Journal* 26: 69–76.
- Iwano, M., Shiba, H., Funato, M., Shimosato, H., Takayama, S., et al. 2003.
- 657 Immunohistochemical studies on translocation of pollen S-haplotype determinant in

658	self-incompati	bility of <i>Bras</i>	sica rapa. P	Plant and Cell	Physiology 44	4: 428–436.

- Jones-Rhoades, M. W., Bartel, D. P., Bartel, B., 2006. MicroRNAs and their regulatory
- roles in plants. *Annual Review of Plant Biology* 57: 19–53.
- 661 Kakizaki, T., Takada, Y., Ito, A., Suzuki, G., Shiba, H., et al. 2003. Linear dominance
- relationship among four class-II S haplotypes in pollen is determined by the expression
- of SP11 in Brassica self-incompatibility. *Plant & cell physiology* 44: 70–75.
- 664 Kusaba, M., Dwyer, K., Hendershot, J., Vrebalov, J., Nasrallah, J.B., et al. 2001. Self-
- 665 incompatibility in the genus Arabidopsis: characterization of the S locus in the
- outcrossing A. lyrata and its autogamous relative A. thaliana. *The Plant cell* 13: 627–
- 667 643.
- Kusaba, M., Tung, C-W., Nasrallah, M. E., Nasrallah, J. B., 2002. Monoallelic expression
- and dominance interactions in anthers of self-incompatible *Arabidopsis lyrata*. *Plant*
- 670 *physiology* 128: 17–20.
- 671 Leducq, J-B., Gosset, C. C., Gries, R., Calin, K., Schmitt, É., et al. 2014. Self-
- 672 incompatibility in Brassicaceae: identification and characterization of SRK -like

673 sequences linked to the S-Locus in the tribe Biscutelleae. *Genes/Genetics* 4:

- 674983–992.
- Li, J., Reichel, M., Li, Y., Millar, A.A., 2014. The functional scope of plant microRNA-
- 676 mediated silencing. *Trends in Plant Science* 19: 750–756.
- 677 Liu, Q., Wang, F., Axtell, M. J., 2014. Analysis of complementarity requirements for plant
- 678 MicroRNA targeting using a Nicotiana benthamiana quantitative transient assay. *The*

- 679 *Plant Cell* 26: 741–753.
- Livak, K. J., Schmittgen, T. D., 2001. Analysis of relative gene expression data using realtime quantitative PCR and. *Methods* 25: 402–408.
- 682 Llaurens, V., Billiard, S., Leducq, J-B., Castric, V., Klein, E. K., et al. 2008. Does
- 683 frequency-dependent selection with complex dominance interactions accurately predict
- allelic frequencies at the self-incompatibility locus in *Arabidopsis halleri? Evolution* 62:
 2545–2557.
- 686 Ma, R., Han, Z., Hu, Z., Lin, G., Gong, X., et al. 2016. Structural basis for specific self-
- 687 incompatibility response in Brassica. *Nature Publishing Group* 26: 1320–1329.
- Mable, B. K., Schierup, M. H., Charlesworth, D., 2003. Estimating the number, frequency,
- and dominance of S-alleles in a natural population of *Arabidopsis lyrata* (Brassicaceae)

690 with sporophytic control of self-incompatibility. *Heredity* 90: 422–31.

- Mallory, A. C., Reinhart, B. J., Jones-Rhoades, M. W., Tang, G., Zamore, P. D., et al.
- 692 2004. MicroRNA control of PHABULOSA in leaf development: importance of pairing
- to the microRNA 5' region. *The EMBO Journal* 23: 3356–3364.
- 694 Marí-Ordóñez, A., Marchais, A., Etcheverry, M., Martin, A., Colot, V., et al. 2013.
- Reconstructing *de novo* silencing of an active plant retrotransposon. *Nature genetics* 45:
 1029–1039.
- 697 Matzke, M., Kanno, T., Daxinger, L., Huettel, B., Matzke, A. J., 2009. RNA-mediated
- 698 chromatin-based silencing in plants. *Current Opinion in Cell Biology* 21: 367–376.
- Murphy, D. J., Ross, J. H. 1998. Biosynthesis, targeting and processing of oleosin-like

- proteins, which are major pollen coat components in *Brassica napus*. *Plant J* 13: 1–16.
- 701 Naithani, S., Chookajorn, T., Ripoll, D. R., Nasrallah, J. B., 2007. Structural modules for
- 702 receptor dimerization in the S-locus receptor kinase extracellular domain. *Proceedings*
- 703 *of the National Academy of Sciences* 104: 12211–6.
- 704 De Nettancourt, D., 2001. Incompatibility and Incongruity in Wild and Cultivated Plants.
- 705 (BY Springer-Verlag., Ed.).
- 706 Novikova. P. Y., Hohmann, N., Nizhynska, V., Tsuchimatsu, T., Ali, J., et al. 2016.
- 707 Sequencing of the genus Arabidopsis identifies a complex history of nonbifurcating
- speciation and abundant trans-specific polymorphism. *Nature Genetics* 48: 1077–1082.
- 709 Palazzo, A. F., Lee, E. S., 2015. Non-coding RNA: What is functional and what is junk?
- 710 *Frontiers in Genetics* 5: 1–11.
- 711 Parizotto, E. A., Parizotto, E. A., Dunoyer, P., Dunoyer, P., Rahm, N., et al. 2004. In vivo
- 712 investigation of the transcription, processing, endonucleolytic activity, and functional
- relevance of the spatial distribution of a plant miRNA. *Genes & Development* 18(18):
- 714 2237–2242.
- 715 Remans, T., Smeets, K., Opdenakker, K., Mathijsen, D., Vangronsveld, J., et al. 2008.
- 716 Normalisation of real-time RT-PCR gene expression measurements in Arabidopsis
- thaliana exposed to increased metal concentrations. *Planta* 227: 1343–1349.
- 718 Schopfer, C. R., Nasrallah, M. E., Nasrallah, J-B., 1999. The male determinant of self-
- 719 incompatibility in Brassica. *Science* 286: 1697 LP-1700.
- 720 Schwab, R., Palatnik, J. F., Riester, M., Schommer, C., Schmid, M., et al. 2005. Specific

701	offects of micro DNAs on the	plant transprintom	Davalonmental	Call 9, 517 527
121	effects of microkinas on the	plant transcriptom	e. Developmental	$Cell \delta: 517-527.$

- 722 Shiba, H., Iwano, M., Entani, T., Ishimoto, K., Shimosato, H., et al. 2002. The dominance
- of alleles controlling self-incompatibility in Brassica pollen is regulated at the RNA
- 724 level. *The Plant cell* 14: 491–504.
- 725 Shiba, H., Kakizaki, T., Iwano, M., Tarutani, Y., Watanabe, M., et al. 2006. Dominance
- relationships between self-incompatibility alleles controlled by DNA methylation.
- 727 *Nature Genetics* 38: 297–9.
- 728 Smyth, D. R., Bowman, J. L., Meyerowitz, E. M., 1990. Early flower development in
- 729 Arabidopsis. *The Plant Cell* 2: 755–767.
- 730 Suzuki, G., Kai, N., Hirose, T., Fukui, K., Nishio, T., et al. 1999. Genomic organization of
- the S locus : identification and characterization of genes in SLG / SRK region of S9

haplotype of Brassica campestris (syn . rapa). *Genetics* 153(1): 391–400.

- 733 Takayama, S., Shiba, H., Iwano, M., Shimosato, H., Che, F. S., et al. 2000. The pollen
- determinant of self-incompatibility in *Brassica campestris*. *Proceedings of the National*

Academy of Sciences of the United States of America 97: 1920–1925.

- 736 Tarutani, Y., Shiba, H., Iwano, M., Kakizaki, T., Suzuki, G., et al. 2010. Trans-acting
- small RNA determines dominance relationships in Brassica self-incompatibility. *Nature*466: 983–986.
- 739 Vaucheret, H., Béclin, C., Elmayan, T., Feuerbach, F., Godon, C., et al. 1998. Transgene-

induced gene silencing in plants. *The Plant journal* 16: 651–659.

741 Vazquez, F., Legrand, S., Windels, D., 2010. The biosynthetic pathways and biological

742	scopes of	plant small RNAs.	Trends in Plant	Science 15: 337–345.

- 743 Wang, F., Polydore, S., Axtell, M. J., 2015. More than meets the eye? Factors that affect
- target selection by plant miRNAs and heterochromatic siRNAs. *Current Opinion in*
- 745 *Plant Biology* 27: 118–124.
- 746 Wickham, H., 2009. ggplot2: Elegant Graphics for Data Analysis. New-York: Springer-
- 747 Verlag.
- 748 Yasuda, S., Wada, Y., Kakizaki, T., Tarutani, Y., Miura-uno, E., et al. 2016. A complex
- dominance hierarchy is controlled by polymorphism of small RNAs and their targets.
- 750 *Nature Plants* 16206: 1–6.
- 751 Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., Smith, G. M., 2009. Mixed Effects
- 752 *Models and Extensions in Ecology with R.* New York: Springer-Verlag.
- 753
- 754

Figure legends

755 Figure 1: Expression dynamics of **a.** SCR and **b.** SRK during flower development, from early buds (<0.5mm) to open flowers. For SCR, only genotypes in which a given allele was 756 757 either dominant or co-dominant were included (recessive SCR alleles were strongly 758 silenced at all stages and were therefore not informative here). For each allele, $2^{-\Delta Ct}$ values 759 were normalized relative to the developmental stage with the highest expression. For each stage, the thick horizontal line represents the median, the box represents the 1st and 3rd 760 761 quartiles. The upper whisker extends from the hinge to the largest value no further than 1.5 762 * Inter Quartile Range from the hinge (or distance between the first and third quartiles).

The lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of thehinge and the black dots represents outlier values.

765	Figure 2: Expression of individual SCR alleles in different genotypic contexts. Pollen
766	dominance status of the S-allele whose expression is measured relative to the other allele in
767	the genotype as determined by controlled crosses are represented by different letters (D :
768	dominant; C: codominant; R: recessive; U: unknown; H: Homozygote, Table S3). In a few
769	instances, relative dominance status of the two alleles had not been resolved
770	phenotypically and were inferred from the phylogeny (marked by asterisks). Thick
771	horizontal bars represent the median of $2^{-\Delta Ct}$ values, 1^{st} and 3^{rd} quartile are indicated by the
772	upper and lower limits of the boxes. The upper whisker extends from the hinge to the
773	largest value no further than 1.5 * Inter Quartile Range from the hinge (or distance
774	between the first and third quartiles). The lower whisker extends from the hinge to the
775	smallest value at most 1.5 * IQR of the hinge and the black dots represents outlier values.
776	We normalized values relative to the highest median across heterozygous combinations
777	within each panel. Alleles are ordered from left to right and from top to bottom according
778	to their position along the dominance hierarchy, with SCR01 the most recessive and
779	SCR13 and SCR20 the most dominant alleles. Under a model of transcriptional control of
780	dominance, high expression is expected when a given allele is either dominant or co-
781	dominant and low expression when it is recessive. Exceptions to this model are marked by
782	black vertical arrows and discussed in the text. "Na" marks homozygote or heterozygote
783	genotypes that were not available.
784	Figure 3: Expression of individual SRK alleles in different genotypic contexts. Putative

785 pistil dominance status of the S-allele whose expression is measured relative to the other

786 allele in the genotype is represented by different letters (**D**: dominant; **R**: recessive; **U**: 787 unknown; **H**: Homozygote). Note that the pistil dominance hierarchy of the S-allele have 788 been less precisely determined than the pollen hierarchy, and so many of the pairwise 789 dominance interactions were indirectly inferred from the phylogenetic relationships (and 790 marked by an asterisk) rather than directly measured phenotypically. Thick horizontal bars represent the median of $2^{-\Delta Ct}$ values, 1^{st} and 3^{rd} quartile are indicated by the upper and 791 792 lower limits of the boxes. The upper whisker extends from the hinge to the largest value no 793 further than 1.5 * Inter Quartile Range from the hinge (or distance between the first and 794 third quartiles). The lower whisker extends from the hinge to the smallest value at most 1.5 795 * IQR of the hinge and the black dots represents outlier values. We normalized the values 796 for each allele relative to the higher median across heterozygous combination. We 797 normalized values relative to the highest median across heterozygous combinations within 798 each panel. Alleles are ordered from left to right and from top to bottom according to their 799 position in the pistil dominance hierarchy, with SRK01 the most recessive and SRK04 the 800 most dominant allele in our sample, based on the phenotypic determination in Llaurens et 801 al. (2008).

Figure 4: Base-pairing requirements for the transcriptional control of *SCR* alleles by sRNAs suggest a threshold model. **a.** Relative expression of *SCR* alleles as a function of the alignment score of the "best" interaction between the focal allele (including 2kb of sequence upstream and downstream of *SCR*) and the population of sRNAs produced by sRNA precursors of the other allele in the genotype. For each allele, expression was normalized relative to the genotype in which the $2^{-\Delta Ct}$ value was highest. Dots are coloured according to the dominance status of the focal *SCR* allele in each genotypic context (black:

809	dominant; white: recessive; grey: undetermined). The black line corresponds to a local
810	regression obtained by a smooth function (loess function, span=0.5) in the ggplot2 package
811	(Wickham, 2009) and the grey area covers the 95% confidence interval. Vertical arrows
812	point to observations that do not fit the threshold model of transcriptional control and are
813	represented individually on Figure 5. b. Barplots of the Akaike Information Criteria (AIC)
814	quantifying the fit of the generalized linear model for different target alignment scores
815	used to define functional targets. Lower AIC values indicate a better fit.
816	Figure 5: Predicted sRNA/target interactions that do not fit with the documented
817	dominance phenotype or the measured expression. For each alignment, the sequence on top
818	is the sRNA and the bottom sequence is the best predicted target site on the SCR gene
819	sequence (including 2kb of sequence upstream and downstream of SCR). a. sRNA targets
820	with a score above 18, while the S-allele producing the sRNA is phenotypically recessive
821	over the S-allele containing the SCR sequence. b. sRNA target with a score below 18,
822	while the S-allele producing the sRNA (S04) is phenotypically dominant over the S-allele
823	containing the SCR sequence and transcript levels of the SCR03 allele is accordingly very
824	low. This is the best target we could identify on SCR03 for sRNAs produced by S04.

825

826 Supplementary figures

Figure S1. Position of the *SCR* qPCR primers. Each primer is represented by a black
arrow, according to its relative position to the gene. Positions indicated correspond to base
pairs from the start codon on the cDNA.

Figure S2. Position of the *SRK* qPCR primers. Each primer is represented by a black

arrow, according to its relative position to the gene. Positions indicated correspond to base

- pairs from the start codon on the cDNA.
- 833 **Figure S3.** Validation of the *SCR* qPCR primers. "Positive control": amplification assay
- 834 with primers for *SCR* alleles that are expressed in the cDNA used. "Cross Amplification":

amplification assay with primers for SCR alleles that are different from the ones in the

836 cDNA used. "Water" : amplification assay with primers but water instead of cDNA.

837 "Recessive": amplification with primers for the phenotypically recessive SCR allele in the

838 cDNA (mean *Ct* values of the biological and technical replicates). "Dominant":

amplification with primers for the phenotypically dominant SCR allele in the cDNA (mean

840 *Ct* values of the biological and technical replicates). Thick horizontal bars represent the

841 median of $2^{-\Delta Ct}$ values, 1^{st} and 3^{rd} quartile are indicated by the upper and lower limits of

the. The upper whisker extends from the hinge to the largest value no further than 1.5 *

843 Inter Quartile Range from the hinge (or distance between the first and third quartiles). The

844 lower whisker extends from the hinge to the smallest value at most 1.5 * IQR of the hinge

and the black dots represents outlier values.

846 **Figure S4:** qPCR amplification (non-transformed *Ct* values) in serial dilutions for each

847 SCR (a) and SRK (b) allele. Solid lines are the linear regressions over all Ct values. Dashed

848 lines are linear regressions excluding the highest dilution level.

849 **Figure S5.** Expression of individual *SCR* alleles in different genotypic contexts,

850 representing each biological and clone replicate separately. Symbols on top of the boxes

- 851 indicate measures from identical clone replicates. See legend of Figure 2 for a full
- 852 description.

853 Figure S6. Generalized linear mixed model used to test the effect of developmental stage

- and dominance status on the expression of *SCR* alleles (*Ct* values). The distribution shows
- that the residues of the full model are approximately normally distributed when taking
- allele identity, developmental stage and dominance status into account and using a
- 857 logarithmic transformation of the Ct_{SCR} / Ct_{actin} ratios.
- 858 **Table S1**. *SCR* samples analysed for each S-locus genotype, showing the number of
- biological and clone replicates over the four developmental stages sampled. "Allele 1"
- refers to the first allele noted in the genotype (for example in the S1S2 genotype, "allele 1"
- 861 is S1 and "allele 2" is S2).
- **Table S2**: *SRK* samples analysed for each S-locus genotype, showing the number of

biological and clone replicates over the four developmental stages sampled. The alleles are

named accordingly to the Table S1.

865 Table S3: Dominance relationships between alleles from the different genotypes included866 in this study as determined by controlled crosses.

Table S4: qPCR primer sequences for each *SCR* and *SRK* alleles studied.

868 **Table S5:** Detailed results from the generalized linear mixed models. **a**. Decomposition of

the sources of variance across allele identity and the hierarchical levels biological, clones

and technical replicates for SCR. b. Test of the variation of expression dynamic across SCR

- alleles. c. Test of the dominance and stage effects on SCR transcript levels, showing a
- 872 significant interaction. **d**. Comparison of the fit of the model under different base-pairing
- 873 score thresholds. e. Test of the effect of the position of the target on the strength of
- silencing. **f**. Decomposition of the source of variance across the technical replicates and the

875	allele identity for SRK. g. Test of the variation of expression dynamic across SRK alleles.
876	h. Test of the effect of stage and dominance on <i>SRK</i> transcript levels. i . Test of the effect
877	of the identity of the companion allele on SCR transcript levels. j : Test of the effect of age
878	on alignment score above the threshold of 18.
879	Table S6: sRNA and SCR target identified as the best match for every pair of S-alleles.
880	Ct_{SCR}/Ct_{actin} ratios are given for the target S-allele in the interaction and is calculated as the
881	mean value across the two earliest developmental stages (see Figure 1). The positions of
882	the targets are given relative to the beginning of the nearest exon of SCR for targets
883	upstream from the gene or in the intron, and relative to the stop codon for downstream
884	targets. R: Recessive; D: dominant as phenotypically determined by controlled crosses; H:
885	homozygote; R* or D*: dominance as indirectly inferred from the phylogeny of S-alleles.
886	The size of the canonical (21 or 24nt-long) isomiR with the highest targeting score is given
887	in parentheses. The targeting score of the best canonical isomiR is given in parentheses).

Figure 1











895 Figure4



897 Figure5

а			
			20 15 10 5 1
	SO4 Mir4239	3'	CCUCGUACACCUUUAUUGCCUUUG 5'
			X X
	S20 target	5'	GGAACATGTGGCAATAACGGAAAC 3'
	Score = 20		
			25 20 15 10 5 1
	S10 Mir4239	3'	CCUCGUACACCUUUAUUGCCUUUGU 5'
			X X
	S20 target	5'	GGAACATGTGGCAATAACGGAAACA 3'
	Score = 21		
b	00010 21		
			20 15 10 5 1
	S04 Mirs4	γ '	GUAUGAUUCUUGUUAGAUUCA 5
	001 111101	0	~~~
	CO2 target	51	
	SUS Largel	5	ACTACTAAGAATAATCTAAGA 3
	score = 15, 5		

899



Positions indicated correspond to base pairs from the start codon on the cDNA

Figure S2





Figure S4



а





Figure S6

```
lmer(log(s1$Ct_SCR.actine) ~
(1|allele_measured:stage)+stage*dom_phenotype+
(1|replicatBiol_genotype/replicat_Techclone), na.action=na.omit)
```

