



HAL
open science

NATURAL EPIALLELIC VARIATION IS ASSOCIATED WITH QUANTITATIVE RESISTANCE TO THE PATHOGEN PLASMODIOPHORA BRASSICAE

Benjamin Liegard, Antoine Gravot, Leandro Quadrana, Yoann Aigu, Juliette Bénéjam, Christine Lariagon, Jocelyne Lemoine, Vincent Colot, Maria M. Manzanares-Dauleux, Mélanie Jubault

► To cite this version:

Benjamin Liegard, Antoine Gravot, Leandro Quadrana, Yoann Aigu, Juliette Bénéjam, et al.. NATURAL EPIALLELIC VARIATION IS ASSOCIATED WITH QUANTITATIVE RESISTANCE TO THE PATHOGEN PLASMODIOPHORA BRASSICAE. 2019. hal-02398784

HAL Id: hal-02398784

<https://hal.science/hal-02398784v1>

Preprint submitted on 8 Dec 2019

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 **NATURAL EPIALLELIC VARIATION IS ASSOCIATED WITH QUANTITATIVE**
2 **RESISTANCE TO THE PATHOGEN *PLASMODIOPHORA BRASSICAE***

3 Benjamin Liégard¹, Antoine Gravot¹, Leandro Quadrana², Yoann Aigu¹, Juliette Bénéjam¹, Christine
4 Lariagon¹, Jocelyne Lemoine¹, Vincent Colot², Maria J. Manzanares-Dauleux¹ and Mélanie Jubault¹

5 These authors contributed equally: Benjamin Liégard, Antoine Gravot

6 ¹AGROCAMPUS OUEST – INRA – Université de Rennes UMR 1349 IGEPP, F-35000 Rennes,
7 France

8 ²Institut de Biologie de l’Ecole Normale Supérieure (IBENS), Ecole Normale Supérieure,
9 Centre National de la Recherche Scientifique (CNRS), Institut National de la Santé et de la
10 Recherche Médicale (INSERM), F-75005 Paris, France

11 **Running title:** Epialleles associated with quantitative resistance

12 **Corresponding author:** Mélanie Jubault

13 Mailing address: IGEPP. AGROCAMPUS OUEST – INRA – Université de Rennes. 35653 Le Rheu.
14 France

15 Phone number: + (33) 2 23 48 57 36

16 Email address: melanie.jubault@agrocampus-ouest.fr

17 **Abstract**

18 Clubroot caused by the protist *Plasmodiophora brassicae* is a major disease affecting cultivated
19 *Brassicaceae*. Here, we uncover the existence of a natural epigenetic variation that is associated with
20 partial resistance to clubroot in *Arabidopsis*, by using QTL fine mapping followed by extensive DNA
21 sequence and methylation analyses. We show that at QTL *Pb-At5.2*, DNA methylation variation is
22 extensive across accessions and strictly correlates with expression variation of the two neighboring
23 genes *At5g47260* and *At5g47280*, which encode NLR-immune receptors. Moreover, these natural
24 variants are stably inherited and are not consistently associated with any nucleotide variation. These
25 findings suggest a direct role for epigenetic variation in quantitative resistance of plants to pathogen
26 attacks.

27 Intraspecific diversity in plant immune interactions is associated with the high level of sequence
28 variations at hundreds of NLR, one of the largest and most rapidly evolving plant gene families (Stahl
29 et al., 1999; Meyers et al., 2003; Clark et al., 2004; Innes et al., 2008; Yue et al., 2012; Shao et al., 2016).
30 Each NLR protein can be involved in the direct or indirect recognition of a small range of effector
31 proteins secreted by specific strains of plant pathogens, potentially triggering the induction of strong
32 plant defense responses that can rapidly stops pathogen invasion (Maekawa et al., 2011; Jones et al.,
33 2016). Thus, with only a few exceptions (including non-NLR driven resistances (Thomas, 1998; Xiao
34 et al., 2001; Larkan et al., 2013) and NLR-driven broad-spectrum resistances (Ernst et al., 2002; Qu et
35 al., 2006)), the catalogue of NLR encoding *R*-genes expressed in a given plant genotype shapes the range
36 of isolate-specific full resistances (incompatible interactions).

37 In contrast to *R*-gene driven resistance, quantitative resistance is polygenic, i.e. it involves allelic
38 variation at several Quantitative Trait Loci (QTL), which collectively contribute to post-invasive partial
39 resistance in compatible plant-pathogen interactions. An emerging list of cloned resistance QTL
40 supports the premise that Quantitative Resistance Genes (QRG) are functionally more diverse than *R*
41 genes (Pilet-Nayel et al., 2017; Nelson et al., 2017). This short list however still includes genes encoding
42 NLR (Hayashi et al., 2010; Fukuoka et al., 2014; Xu et al., 2014; Debieu et al., 2016) and other receptors
43 (Diener and Ausubel, 2005; Hurni et al., 2015) or co-receptors (Huard-Chauveau et al., 2013). Thus,
44 variation in NLR genes (or other non-self-recognition loci) appears to contribute to the variations in
45 basal resistance levels in compatible interactions.

46 The triggering of effective resistance requires that cellular levels of NLR proteins reach minimum
47 thresholds. However, elevated expression levels of NLR can also lead to autoimmunity drawbacks,
48 including spontaneous HR and retarded plant growth (Li et al., 2015; Lai and Eulgem, 2018). Their
49 abundance is thus tightly controlled by multiple mechanisms, at the transcriptional, post-transcriptional
50 (Zhang and Gassmann, 2007) (*i.e.* alternative splicing) and post-translational levels (*i.e.* Ubiquitin-
51 dependent proteolytic regulation). NLR regulation also involves a wealth of epigenetic-related cellular
52 processes, including redundant networks of small RNA (Shivaprasad et al., 2012; Fei et al., 2013; Deng

53 et al., 2018), histone modifications (Palma et al., 2010; Xia et al., 2013; Zou et al., 2014), histone-mark
54 dependent-alternative splicing (Tsuchiya and Eulgem, 2013), regulation of chromatin structure and
55 DNA methylation (Li et al., 2010). There is increasing evidence that epigenetic processes can play roles
56 in the transitory imprinting of some plant biotic stress responses, at least for a few generations (Molinier
57 et al., 2006; Slaughter et al., 2011; Luna et al., 2012). It is however not yet clear to which extent the
58 stable transgenerational inheritance of epigenetically regulated gene expression contributes to the
59 natural intraspecific diversity of plant-pathogen interactions.

60 The few available examples of transgenerational epigenetically controlled traits are mostly found in
61 plant species, where the association between natural or induced differentially methylated regions (DMR)
62 and phenotypic traits were shown stably or (most often) metastably inherited across the generations
63 (Quadrana and Colot, 2016; Furci et al., 2019; Liégard et al., 2019). Such regions, designated as
64 epialleles, can affect agronomically relevant traits: compatibility, accumulation of vitamin E and fruit
65 ripening in tomato, disease resistance, sex determination in melon, and fruit productivity in oil palm.

66 In plants, DNA methylation can occur at cytosines in the three sequence contexts, CG, CHG and CHH
67 (Henderson and Jacobsen, 2007) (where H could be A, C or T) and its impact varies depending on the
68 targeted genomic features (i.e. transposable elements, gene promoters or gene bodies). DNA methylation
69 patterns result from the dynamic combination of *de novo* methylation, maintenance methylation and
70 demethylation. The *de novo* DNA methylation is catalyzed by the canonical and non-canonical RNA-
71 directed DNA methylation (RdDM) pathways, which are both guided by siRNA (Zhang et al., 2018).
72 Maintenance of DNA methylation mainly relies on RNA-independent pathways and require the activity
73 of DDM1, MET1 and VIM proteins at CG sites, and of DDM1, KYP, CMT2/3 and histone mark
74 HK9me2 at CHG and CHH sites (Law and Jacobsen, 2010; Matzke and Mosher, 2014)). Previous
75 studies highlighted that natural DMRs were overrepresented over genes belonging to the *NLR* disease
76 resistance gene family (Kawakatsu et al., 2016). However, it remains unclear if natural epigenetic
77 variation at *NLR* genes can shape plant pathogen interactions.

78 Here, we report the identification of naturally occurring stable epigenetic variation underlying a QTL
79 involved in partial resistance to clubroot in *Arabidopsis*. Clubroot is a root gall disease caused by the
80 telluric biotrophic pathogen *Plasmodiophora brassicae* (Rhizaria), affecting all *Brassicaceae* crops. The
81 infection process involves a primary infection in root hairs for only a few days. Then secondary
82 plasmodia develop in root cortical cells, causing hyperplasia and hypertrophy that ultimately impair
83 plant water and nutrient uptake. The reference accession Col-0 and Bur-0 are fully susceptible and
84 partially resistant to *P. brassicae* isolate eH, respectively (Alix et al., 2007; Jubault et al., 2008b)
85 (**Supplementary Figure S1**). Four main QTL determine this difference, which act additively⁴³. Here
86 we identify by fine mapping of the largest effect resistance QTL *Pb-At5.2* a strong association between
87 partial resistance and the expression level of the two *NLR* genes *At5g47260* and *At5g47260*, linked to
88 the DNA methylation status of the small region including those two genes and a neighboring
89 transposable element (TE) sequence. Furthermore, we show that epiallelic variation at this locus is
90 frequent among natural *Arabidopsis* accessions and that the highly methylated state is negatively
91 correlated with the expression of the two *NLR*-genes as well as with quantitative resistance to *P.*
92 *brassicae*. In contrast, there is no correlation between the trait and any specific nucleotide variant across
93 the entire QTL interval. We further show that the RNA independent pathway involving DDM1, MET1,
94 VIM and CMT2/3 maintains the hypermethylated epiallele in Col-0. Overall, our findings demonstrate
95 that the quantitative resistance to a major root disease affecting *Brassicaceae* is associated, in
96 *Arabidopsis*, with the stable inheritance of a natural epigenetic variation involved in the control of the
97 constitutive expression of a pair of *NLR*-genes.

98 **Results**

99 ***Fine mapping of the *Pb-At5.2* locus responsible for clubroot resistance.***

100 Using a population of F7 recombinant inbred lines (RILs) between the partially resistant accession Bur-
101 0 and the susceptible Col-0, we previously mapped a QTL (*Pb-At5.2*), located on chromosome 5
102 between 67.5 and 71.8 cM, explaining a significant fraction ($R^2=20\%$) of the resistance (**Fig. 1a**) (Alix
103 et al., 2007; Jubault et al., 2008b). This interval contained 157 annotated sequences between At5g46260

104 and At5g47690. The effect and confidence interval of this QTL was also previously confirmed in
105 Heterogeneous Inbred Family (HIF) lines 10499 and 13499 (Lemarié et al., 2015), both derived from
106 the RIL 499 which harbored residual heterozygosity in the *Pb-At5.2* region (**Fig. 1b-c, Supplementary**
107 **Text S1**). The initial aim of the present work was to fine map *Pb-At5.2*, starting with reciprocal crosses
108 between HIF lines 10499 and 13499. Clubroot symptoms in individuals of the F1 progeny were equally
109 severe to those in the susceptible parental line HIF 13499, suggesting that the resistance allele Bur-0
110 was recessive (**Supplementary Figure S2**). The boundaries of the *Pb-At5.2* resistance locus was further
111 refined through several rounds of genotyping and clubroot phenotyping (F3 to F5 generations
112 downstream 10499/13499 crosses), details in **Fig. 1d-f, Supplementary Figure S3, Supplementary**
113 **Text S1, Supplementary Data 1**). This enabled us to narrow down the confidence interval to 26 kb
114 between the markers CLG4 (19,182,401 bp, in the promoter region of At5g47240), and the marker K64
115 (19,208,823 bp, in At5g47330). This region contained eight annotated open reading frames (ORFs),
116 including the three NLR-encoding genes At5g47250/At5g47260/At5g47280, six annotated TE
117 sequences and one lncRNA gene (**Fig. 1f**). The two F5 homozygous progeny lines 1381-2 and 2313-15,
118 harboring the closest recombination events from both sides of the 26 kb interval, (see **Fig. 1e**) also
119 showed partial resistance to a series of additional *P. brassicae* isolates (from pathotypes 1, 4 and 7
120 following the classification of Some et al. (1996). This highlighted the broad spectrum of the resistance
121 conferred by the Bur-0 allele of *Pb-At5.2* (**Supplementary Figure S4**).

122 ***Possible causal role for a stably inherited expression/methylation polymorphism affecting two NLR***
123 ***genes***. RNA-seq analysis was carried out on the Bur-0 and Col-0 accessions and the recombinant HIF
124 lines 10499 and 13499. Pathogen-induced gene patterns markedly differed in genotypes harboring
125 alleles *Pb-At5.2_{BUR}* or *Pb-At5.2_{COL}* (**Supplementary Figure S5, Supplementary Data 2**). Those
126 regulations were consistent with our previously published studies *i.e.* a role for camalexin biosynthesis
127 and SA-mediated responses in *Pb-At5.2_{BUR}*-mediated resistance; a role for JA-driven induction of
128 *ARGAH2* in *Pb-At5.2_{COL}*-mediated basal resistance (details in **Supplementary Text 2**). We then focused
129 on the eight ORFs in *Pb-At5.2*. In Col-0, sequenced reads were only found for four of the genes:
130 At5g47240/At5g47250/At5g47310/At5g47320. At5g47310 and At5g47320 only had one SNP and one

131 indel in their promoter regions in Bur-0, respectively. The possible causal role of these variations was
132 discarded, as both genes were expressed similarly from the Bur-0 and Col-0 alleles (**Fig. 2a**). Most of
133 the other SNPs in the 26kb region were associated with the NUDX8 encoding gene At5g47240 and to a
134 smaller extent with the adjacent NLR-gene At5g47250. However, clubroot symptoms in the
135 homozygous T-DNA mutant lines SALK_092325C (T-DNA in the At5g47240 gene) and
136 WiscDsLoxHs110_09B (T-DNA in the At5g47250 gene) were the same as in Col-0 (**Supplementary**
137 **Figure S6**). At5g47260 and At5g47280, both encoding proteins belonging to the family of non-TIR-
138 NLR immune receptors, displayed one single non-synonymous SNP, and no SNP, respectively.
139 However, RNA-seq analysis indicated that these two genes were constitutively expressed in the roots of
140 Bur-0, 10499 and 1381-2 (with the Bur-0 allele, **Fig. 2a** and **Supplementary Figure S4**), but their
141 expression was almost undetectable in Col-0, 13499 and 2313-15 (with the Col-0 allele).

142 To understand why these two NLR genes At5g47260 and At5g47280 were differentially expressed in
143 Bur-0 and Col-0, even in the absence of any sequence variation within the putative promoter regions of
144 these genes, we analyzed the DNA methylation level at the region in these two accessions using public
145 methylome data (Kawakatsu et al., 2016). The genomic interval between 19,188,411 and 19,196,559,
146 which includes, in Col-0 and Bur-0, the two genes At5g47260 and At5g47280 and the transposon
147 At5TE69050 in-between, were hypermethylated and hypomethylated in Col-0 and Bur-0, respectively
148 (**Fig. 2b**). This contrasting methylation state was experimentally confirmed using DNA extracted from
149 infected or control roots of Col-0 and Bur-0 plants (**Fig. 2c**) and CHOP qPCR. These DNA methylation
150 differences were also found between the progeny HIF lines 10499 and 13499 and in the pair of HIF-
151 derived homozygous near-isogenic lines 1381-2/2313-15 (**Fig. 2c**), thus indicating that they are stably
152 inherited independent of any DNA sequence polymorphism outside of the locus. Moreover, the ‘Col-
153 like’ hypermethylation of At5g47260 and At5g47280 was systematically associated with a low
154 expression of the two NLR genes and a lower level of partial resistance to *P. brassicae* infection. To
155 further investigate the inheritance of this epiallelic variation and its penetrance on gene expression and
156 clubroot resistance, we then investigated two series of 100 individual plants, corresponding to the
157 progenies derived from selfing the heterozygous 2509 and 1381 lines (harbouring heterozygosity at the

158 locus). The evaluation of plant disease for each individual plant in the two progenies indicated a 3:1
159 mendelian segregation of the partial resistance phenotype. Clubroot symptoms in individuals with only
160 one Bur-0 resistance allele were the same as in individuals with the two susceptible Col-0 allele
161 (**Fig. 3a**). In clubroot-inoculated roots of each of those individual plants from the 2509 progeny, the
162 methylation state of the *Pb-At5.2* region was monitored by CHOP qPCR on At5g47260. In addition, the
163 SNP allele status at *Pb-At5.2* was investigated for each individual plant (details of markers are given in
164 **Supplementary Data 1**). Heterozygous Bur/Col individuals displayed intermediate parental
165 methylation and expression values (**Fig. 3b-d**), thus providing a molecular explanation for the
166 recessivity of the Bur-0 resistance allele. Altogether, these results suggested a link between partial
167 resistance to *P. brassicae* and a stably inherited epiallelic variation at *Pb-At5.2*, which controls the
168 expression of two NLR genes.

169 ***The Bur-like hypo-methylated epiallele is well represented among Arabidopsis accessions and***
170 ***contributes to reduce clubroot symptoms.*** To assess the relative contribution of DNA sequence and
171 DNA methylation changes at *Pb-At5.2* to the variable resistance to clubroot we investigated the natural
172 allelic and epiallelic diversity across *Arabidopsis* accessions. We took advantage of recently
173 published Illumina short genome sequence reads obtained from 1135 *Arabidopsis* accessions (1001
174 Genomes Consortium, 2016) to document the species-wide molecular diversity of the *Pb-At5.2* genomic
175 region. Based on quantitative horizontal and vertical coverage of short-reads aligned to the Col-0
176 reference genome sequences we identified two discrete groups of accessions. One of such groups,
177 containing 401 accessions, was characterized by high vertical and horizontal coverage (>0.75) and
178 included the reference accession Col-0 as well as the partially resistance Bur-0 (**Figure 4a**; detailed list
179 of genotypes in **Supplementary Data 3, sheet 1**). Conversely, the remaining 734 accessions present
180 diverse structural rearrangements, principally long deletions that translates in poor horizontal and
181 vertical coverage compared to the reference Col-0 genome. Closer examination of coverage plots for
182 the 401 accessions with Col-0/Bur-0-like revealed a uniform haplotype structure, which was present at
183 high frequency at the species level (Minor Allele Frequency MAF ~0.37). Nonetheless, the haplotype
184 frequency varied depending on geographic groups, ranging from 52.7 % in Spain to 17.7 % in Asia

185 **(Supplementary Figure S7)**. We then analyzed DNA methylation levels in 287 accessions belonging
186 to the 401 accessions containing the Col-0/Bur-0-like *Pb-At5.2* and for which public bisulfite data is
187 available (Kawakatsu et al., 2016). Based on this data, we could distinguish a group of 228 accessions
188 showing hypomethylation of *Pb-At5.2*, including Bur-0, and another group of 59 accessions, which
189 includes Col-0, that displays hypermethylation (**Fig. 4b**). The prevalence of accessions showing the Col-
190 like (epi)haplotype varied considerably depending on geographic origin, ranging from 1.8 % in Spain
191 to 16.8 % in Central Europe (**Supplementary Data 3, and Supplementary Figure S7**). Consistent with
192 a causal role for DNA methylation in the transcriptional regulation At5g47260 and At5g47280,
193 reanalysis of publicly available RNA-seq data revealed a pronounced negative correlation between
194 methylation level and At5g47260 and At5g47280 gene expression (**Fig. 4c**). These results were further
195 validated in infected roots of 20 natural accessions (**Fig. 4d**).

196 Both Col-like and Bur-like epialleles were significantly represented among natural accessions, thus
197 offering interesting genetic material to determine the actual contribution of DNA sequence and DNA
198 methylation in the control of clubroot partial resistance. A selection of 126 accessions, including 42
199 accessions with the Col-like epiallele and 85 accessions with the Bur-like epiallele, were assessed for
200 their resistance to *P. brassicae* isolate eH. While no DNA sequence polymorphism within *Pb-At5.2*
201 shows association with clubroot resistance (**Supplementary Data 4**), the low DNA methylation state of
202 the At5g47260/At5g47280 locus was significantly associated with enhanced resistance levels (**Fig. 5**).
203 Altogether, the results corroborate and extend the conclusions obtained by fine mapping of *Pb-At5.2*
204 and provide strong evidence that natural epiallelic variation contribute to the quantitative differences
205 observed in clubroot resistance among Arabidopsis accessions.

206 ***Pb-At5.2* epivariation is independent of cis- genetic variations.** At the locus *Pb-At5.2*, the transposon
207 AT5TE69050 was present in both parental genotypes, with no sequence variation that might have been
208 the primary cause of the variation of DNA methylation on the two adjacent genes. Analysis of 34 out of
209 the 287 accessions with the Col-0/Bur-0-like haplotype, did not reveal the presence/absence of TE
210 insertion variant within the 26kb *Pb-At5.2* region (Quadrana et al., 2016; Stuart et al.), with the exception

211 of a private helitron insertion in the accession NFA-10. Moreover, 18 and 16 of these accessions
212 displayed hypermethylated and hypomethylated epialleles, respectively, indicating that variation in
213 DNA methylation is not associated with TE presence/absence variants. In addition, the cis-nucleotide
214 polymorphism located within the coding sequence of At5g47260 and detected in Bur-0 was absent in at
215 least five others accessions sharing the hypomethylated epiallele (**Supplementary Figure S8**),
216 indicating that the hypomethylated state of *Pb-At5.2* is not correlated with any specific DNA sequence
217 polymorphism at the locus.

218 ***The hypermethylated epigenetic variant is maintained by the RNA-independent pathway.*** Analysis of
219 sRNAs identified in Col-0 (Stroud et al., 2014) revealed that the At547260/At5g47280 region is targeted
220 mostly by 24-nt sRNA, which prompted us to generate sRNA profiles from non-inoculated roots of Col-
221 0 and Bur-0 as well as to 7 and 14 dpi with *P. brassicae* isolate eH. Consistent with the pattern of DNA
222 methylation observed before, we found high levels of sRNAs only in Col-0 (**Fig. 6a**). To explore further
223 the mechanisms involved in the maintenance of methylation at this locus, we made use of publicly
224 available methylomes of Col-0 mutant plants defective in one or several DNA methylation pathways
225 (Stroud et al., 2013). Despite the high levels of sRNAs detected over the At547260/At5g47280 region,
226 mutations affecting the RdDM pathway did not influence its DNA methylation level (**Supplementary**
227 **Figure S9**). Conversely, DNA methylation was largely lost in mutants defective in sRNA-independent
228 maintenance of DNA methylation, i.e. *ddm1*, *cmt2/3*, *met1*, and *svh456*. (**Fig. 6b**). These results raise
229 the question of the role of the sRNA targeted to the Col-like hypermethylated region, whereas the
230 methylation maintenance solely depends on the RNA-independent pathway.

231 **Discussion**

232 To date, only a very small number of resistance QTL have been characterized at the molecular level
233 (Nelson et al., 2017). Detection and fine mapping of resistance QTL is typically challenging not only
234 because of the difficulties associated with measuring small variations in partial resistance in a large
235 number of individual progeny, but also because resistance QTL can be sensitive to environmental
236 changes (Laperche et al., 2017; Aoun et al., 2017; Aigu et al., 2018). However, technical issues may

237 have only been part of the problem. Recent developments in the field of epigenetics suggest that the
238 detection of some inherited resistance factors could be missed due to classic genetic approaches being
239 based only on DNA sequence variation. In the present work, genome-wide association study failed at
240 identifying any nucleotide variation in the 26 kb interval of *Pb-At5.2* associated with clubroot response.
241 By contrast, clubroot resistance was clearly related to the epigenetic variation at two NLR-coding genes
242 in this interval. This work thus uncovers for the first time an epigenetically driven expression
243 polymorphism substantially contributing to the natural diversity of plant immune response.

244 Many examples of epialleles are metastable, *i.e.* they can be reversed by stochastic or non-identified
245 factors (Weigel and Colot, 2012). Stability over multiple generations is a primary concern from both
246 evolutionary and breeding perspectives. The epiallele described here seems to be extremely stable, as it
247 was robustly detected in all our previous QTL investigations in *Arabidopsis*. This included two
248 independent segregating progenies derived from Bur-0 and Col-0 (Jubault et al., 2008a) and additional
249 studies with HIF lines 10499/13499 (Lemarié et al., 2015; Gravot et al., 2016). The high level of
250 methylation and absence of *At5g47260* and *At5g47280* expression in Col-0 was found in a series of
251 publicly available data obtained in different laboratories, with a diversity of plant tissues and conditions
252 (Winter et al., 2007; Stroud et al., 2013; Klepikova et al., 2016; 1001 Genomes Consortium, 2016;
253 Kawakatsu et al., 2016). It was also confirmed by our own data generated on inoculated roots, non-
254 inoculated roots and leaf samples. Finally, this methylation pattern was also robustly found in multiple
255 replicates of individual plants. Thus, *Pb-At5.2* can be unreservedly classified as a stable epialleles.

256 There has only been one recent report of a plant disease resistance caused by an inherited methylation
257 variant affecting the expression of a resistance-related gene (Nishimura et al., 2017). In that study, a
258 stable expression polymorphism (between Ler-0 and Ag-0 accessions) on the TIR-only encoding gene
259 *RBA1* (*At1g47370*) affected ETI-responses to the *Pseudomonas syringae* effector *hopBA1*. This
260 expression polymorphism was linked to the nearby presence/absence of a TE sequence in the promoter
261 region of the gene and to *MET1*-dependent DNA methylation variation. However, because DNA
262 methylation was reversed when the TE sequence was segregated away (**Figure S2** in Nishimura et al.

263 (2017), DNA methylation variation is not epigenetic as it is an obligate consequence of sequence
264 variation (i.e. presence/absence of the TE sequence). In the present study, we showed that DNA
265 methylation variation in the region between At5g47260 and At5g47280, including the TE sequence
266 *At5TE69050*, is not linked to any nucleotide/structural variation at the locus or elsewhere in the genome.
267 Thus, *Pb-At5.2_{COL}* and *Pb-At5.2_{BUR}* can be considered as ‘pure epialleles’, as defined by Richards (2006).
268 From available genomic and epigenomic data from the ‘1001 genomes project’, it can be extrapolated
269 that the Bur-like clubroot resistance epiallele is present in about half of the accessions from the ‘Relict’,
270 ‘Spain’, and ‘Italy/Balkans/Caucasus’ groups and 39 % of the accessions from the ‘North Sweden’
271 group (**Supplementary Data 3**). In contrast, the Bur-like epiallele is likely (taking into account missing
272 methylation data) present at about 10 % in the ‘Germany’ group. On the other hand, the clubroot
273 susceptibility of the Col-like epiallele was absent from the accessions in the ‘Relict’ and ‘North Sweden’
274 groups, but reached at least 16.8 % in the ‘Central Europe’ group. This sharp geographical structuration
275 suggests that both epialleles confer fitness gains depending on environmental contexts. However, it does
276 not appear to be obviously related to the distribution of clubroot incidence in *Brassica* cultures (usually
277 low in the warm southern European regions). Keeping in mind that NLR can detect unrelated effectors
278 from distinct microbial species (Narusaka et al., 2009), and echoing previous work (Karasov et al.,
279 2014), we hypothesize that the maintenance of this epivariation in natural populations may reflect
280 additional roles that *Pb-At5.2* plays against other plant pathogens (besides the control of clubroot
281 infection).

282 At5g47260 and At5g47280 belong to a small heterogeneous cluster of three non-TIR-NLR located on
283 chromosome 5, not far from the largest NLR-hot spot in the *Arabidopsis* genome (Meyers et al., 2003).
284 None of the three genes have been previously shown to be involved in plant pathogen interactions. There
285 are a few examples of tandem NLR genes coding for pairs of proteins that function as heterodimers
286 (Cesari et al., 2014; Williams et al., 2014; Saucet et al., 2015). Similarly, the proteins encoded by these
287 two epigenetically joint-regulated genes may function together in the control of cell defense responses
288 during clubroot infection. Although the underlying molecular mechanisms are not known, the canonical

289 example of the TIR-NLR heterodimer *RRS1/RPS4* corresponds to a recessive resistance locus, similar
290 to *Pb-At5.2*. Among the 287 genotypes with Col/Bur-like genomic structure and available methylation
291 data, both were genes were always methylated and it was thus not possible to use these genetic resources
292 to gain information on the impact of silencing only one of the two genes. Additional molecular work
293 (including CRISPR-Cas9-based nucleotide deletions or demethylation, in the Bur-0 accession, and the
294 analysis of protein-protein interactions) is thus necessary to assess the individual role and possible
295 partnership of these two genes.

296 Using public data (Zilberman et al., 2007), we found that 32 and 77 genes from the NLR family are
297 methylated and unmethylated in Col-0, respectively (**Supplementary Data 5**). This suggests that plant
298 genomes contain series of functional resistance genes, whose possible roles in biotic interactions are
299 locked by epigenetic processes. This hypothesis is also supported by our recent study, where we
300 demonstrated that the *ddm1*-triggered hypomethylation at different genomic loci resulted in unlocking
301 genetic factors ultimately exerting a significant control over the development of clubroot symptoms
302 (Liégard et al., 2019). It would now be interesting to carry out a careful genome wide analysis of
303 methylation profiles of all NLR-genes among *Arabidopsis* accessions, which would take into account
304 the structural diversity of all these individual genes (supported by additional targeted resequencing of
305 NLR-loci). The intraspecific diversity of the methylation patterns of NLR and RLK/RLP genes in plants,
306 their heritability, and their consequences on plant biotic interactions, may also deserve further attention
307 in future studies.

308 **Methods**

309 ***Plant materials and growth conditions.*** Heterogeneous Inbred Family (HIF) lines 10499 and 13499 and
310 their parental accessions Col-0 (186AV) and Bur-0 (172AV) were obtained from the Institute Jean Pierre
311 Bourgin (INRA Versailles, France). *Arabidopsis thaliana* accessions were all purchased from the
312 Nottingham Stock Center. The panel of 126 accessions was selected according to their methylation level
313 at the region of interest (Kawakatsu et al., 2016). All accessions and in-house generated recombinant
314 lines used in this study are listed in **Supplementary Data 1** and **Supplementary Data 3**. Seed

315 germination was synchronized by placing seeds on wet blotting paper in Petri dishes for two days at
316 4 °C. Seeds were sown individually in pots (four cm diameter) containing a sterilized mix composed of
317 2/3 compost and 1/3 vermiculite. Growth chamber controlled conditions of 16 h light ($110\mu\text{mol m}^{-2}\text{s}^{-1}$)
318 at 20 °C and 8 h dark at 18 °C were used to grow plants. The 126 *Arabidopsis* accessions and HIF were
319 challenged by *P. brassicae* in two biological replicates in a completely randomised block designs (with
320 two blocks per replicate, each block consisting in 6 plants by genotype). The *Arabidopsis* accessions
321 Col-0, Bur-0 and the HIF 10499, 13499, 1381-2 and 2313-15 used in RNA-seq and smRNA-seq
322 approaches were assessed when infected with *P. brassicae* or in uninfected condition in three
323 randomised blocks. Almost all clubroot tests were performed with the isolate eH of *P. brassicae*
324 described by Fählng (2003) which belongs to the most virulent pathotype P1. The resistance spectrum
325 of *Pb-At5.2* was also assessed with a series of additional isolates Pb137-522, Ms6, K92-16 and P1⁺. For
326 every isolate used in this study, the pathotype was validated in every experiment using the differential
327 host set according to Some et al. (1996), also including two genotypes of *B. oleracea* ssp. *acephala* C10
328 and CB151. One ml of resting spore suspension (10^7 spores.ml⁻¹) prepared according to Manzanares-
329 Dauleux et al (2000) was used for pathogen inoculation 10 days (d) after germination (stage 1.04 (Boyes
330 et al., 2001)). This inoculum was applied onto the crown of each seedling.

331 **Phenotyping.** HIF and *Arabidopsis* accessions were phenotyped three weeks after inoculation (21 days
332 post inoculation (dpi) for their susceptibility to *P. brassicae*. Plants were thoroughly rinsed with water
333 and photographed. Infected roots were removed and frozen in liquid nitrogen. Clubroot symptoms were
334 evaluated through image analysis using the GA/LA index calculated according to Gravot et al. (2011).

335 **Fine mapping of the loci responsible for clubroot resistance.** Fine mapping of *Pb-At5.2* was performed
336 starting from crosses between HIF lines 10499 and 13499, followed by successive rounds of genotyping
337 and clubroot phenotyping in subsequent plant generations (full details are given in **Supplementary Text**
338 **S1**).

339 **RNA isolation, mRNA Sequencing and differentially gene expression determination.** Total RNA was
340 extracted from frozen and lyophilized roots (collected 14 days after inoculation) using the mirVana™

341 miRNA Isolation Kit (Invitrogen) according to the manufacturer's instructions. After extraction, the
342 RNA were quantified using a NanoDrop ND-1000 technologies and their quality was controlled using
343 the RNA 6000 assay kit total RNA (Agilent). Samples with a RIN greater or equal to seven were used
344 for sequencing. cDNA sequencing library construction and the sequencing were performed by the NGS
345 platform at the Marie Curie Institute of Paris. Each library was paired-end sequenced on an Illumina
346 HiSeq 2500 technology. Reads were aligned to the TAIR10 genome annotation and assembly of Col-0
347 *Arabidopsis thaliana* genome concatenated with the *P. brassicae* genome using STAR software (Dobin
348 et al., 2013) (Version 2.5.3.a). Alignment conditions were selected according to the *Arabidopsis*
349 genome. A maximum of five multiple read alignments were accepted and for each alignment no more
350 than three mismatches were allowed. The resulting BAM files were used to determine read counts using
351 the function Counts of Feature Counts software (Version 1.4.6) and the TAIR10 gff of *Arabidopsis*
352 concatenated with the gff of *P. brassicae*. Differentially expressed genes were determined using the
353 package EdgeR (Robinson et al., 2010) in R software (Team, 2013) (version 3.3.0). Raw counts obtained
354 as described previously were used as entry data in EdgeR. Once CPM (Counts Per Million) was
355 determined, only genes with at least one CPM in three samples were retained. Expression signals were
356 normalized using the TMM method (Trimmed Mean of M-values) of the CalcNormFactors function of
357 Edge R. Finally, the differentially expressed genes were determined using the decideTests function of
358 Edge R using one minimum fold change between -1.5 and 1.5.

359 ***GWAS analyses***

360 A conventional GWA study on GA/LA data from 126 accessions was performed with easyGWAS
361 (Grimm et al., 2017) (<https://easygwas.ethz.ch/>). Association analysis was performed with EMMAX
362 (Kang et al., 2010) using 1 806 554 SNP with a MAF >0.05, after correction for population structure by
363 including the three first principal components in the additive model.

364 ***Small RNA isolation, sequencing, clusterization and differential presence determination.*** Small RNA
365 was extracted from frozen and lyophilized roots (collected 14 days after inoculation) using the
366 mirVana™ miRNA Isolation Kit (Invitrogen) according to the manufacturer's instructions. After

367 extraction, the small RNA was quantified using a NanoDrop ND-1000 and quality controlled using the
368 Small RNA assay kit (Agilent). Samples with a RIN greater or equal to seven were used for sequencing.
369 Construction of cDNA sequencing libraries and the sequencing were performed by the NGS platform
370 of Marie Curie Institute of Paris. For each sample, single-end (50 bp) sequencing was carried out using
371 Illumina HiSeq 2500 technology. Reads were aligned to the TAIR10 genome annotation and assembly
372 of Col-0 *Arabidopsis thaliana* genome concatenated with *P. brassicae* genome using STAR software
373 (Dobin et al., 2013) (Version 2.5.3.a), counts and clustered using Shortstack software (Axtell, 2013).
374 The presence of differentially expressed sRNA was determined using the EdgeR (Robinson et al., 2010)
375 package in R software (Team, 2013) (version 3.3.0). Raw counts obtained as described previously were
376 used as entry data in EdgeR. Once CPM (Counts Per Million) were determined, only genes with at least
377 one CPM in three samples were retained. Expression signals were normalized using the TMM method
378 (Trimmed Mean of M-values) of the CalcNormFactors function of Edge R. Finally, the differentially
379 expressed sRNA were determined using the decideTests function of Edge R using one minimum fold
380 change between -1.5 and 1.5.

381 ***RNA isolation and RT-qPCR analysis.*** Total RNA was extracted from the lyophilised root of accessions
382 and HIF 21 days post infection using the trizol extraction protocol. Samples with residual traces of DNA
383 were treated with DNase (Promega ref M6 10A). Before reverse transcription of RNA to cDNA by the
384 SuperScript II (Invitrogen), the RNA quality was verified by agarose gel electrophoresis gel. RT-qPCR
385 was carried out in a LightCycler® 480 thermocycler (Roche) on cDNA obtained as described above.
386 Gene expression was normalized using as reference two *Arabidopsis* genes defined as stable during
387 infection using RNA-seq data (At1g54610, At5g38470) following Pfaffl's method (Pfaffl, 2001). Primer
388 sets were designed for each gene and are listed in **Supplementary Table S1**.

389 ***CHOP PCR and qPCR assays.*** Gene methylation profiles were investigated using the enzyme McrBC
390 (M0272L, BioLabs®)(Zhang et al., 2014). Forty ng of DNA was incubated with 0.5 µL BSA
391 (20 mg/mL), 0.5 µL GTP (20 mM), 5 µL NEB (10X) and 0.2 µL McrBC (10000 U/mL). For CHOP
392 PCR and qPCR 2 ng of digested and undigested DNA was used. For the CHOP PCR the temperature

393 conditions were adjusted according to primer design and 35 amplification cycles were used. To
394 determine the methylation state of the targeted region, each sample was digested or not (control) with
395 McrBC before amplification. For the CHOP qPCR the temperature conditions were adjusted according
396 to primer design and 30 amplification cycles of were used. Methylation levels of the target region were
397 calculated as the percentage of molecules lost through McrBC digestion according to Silveira et al
398 (2013) with the formula: $(1 - (2^{-(Ct \text{ digested sample} - Ct \text{ undigested sample})})) * 100$. The percentage of
399 DNA methylation for At5g13440 and At5g47400, was calculated in all CHOP qPCR as controls. Primer
400 sets designed for each gene are listed in **Supplementary Table S1**.

401 **Published data.** The DNaseq data, RNAseq data, variant sequences and bisulfite data of the natural
402 accessions studied were obtained from previous studies (1001 Genomes Consortium, 2016; Kawakatsu
403 et al., 2016) archived at the NCBI SRA number SRP056687 and the NCBI GEO reference: GSE43857,
404 GSE80744. The bisulfite data and small RNA data of *Arabidopsis* mutants studied were obtained from
405 previous study (Stroud et al., 2014) archived at the NCBI GEO reference: GSE39901.

406 **Statistical analysis.** Data were statistically analyzed using the R program (Team, 2013).

407 **Data availability.** The data supporting the findings of this study are available within the paper and its
408 supplementary information files. All unique materials used are readily available from the authors.

409 References

410 **1001 Genomes Consortium** (2016). 1,135 Genomes Reveal the Global Pattern of Polymorphism in
411 *Arabidopsis thaliana*. *Cell* **166**:481–491.

412 **Aigu, Y., Laperche, A., Mendes, J., Lariagon, C., Guichard, S., Gravot, A., and Manzaneres-Dauleux,**
413 **M. J.** (2018). Nitrogen supply exerts a major/minor switch between two QTLs controlling
414 *Plasmodiophora brassicae* spore content in rapeseed. *Plant Pathol.* **67**:1574–1581.

415 **Alix, K., Lariagon, C., Delourme, R., and Manzaneres-Dauleux, M. J.** (2007). Exploiting natural
416 genetic diversity and mutant resources of *Arabidopsis thaliana* to study the *A. thaliana*–
417 *Plasmodiophora brassicae* interaction. *Plant Breed.* **126**:218–221.

418 **Aoun, N., Tauleigne, L., Lonjon, F., Deslandes, L., Vailleau, F., Roux, F., and Berthomé, R.** (2017).
419 Quantitative Disease Resistance under Elevated Temperature: Genetic Basis of New
420 Resistance Mechanisms to *Ralstonia solanacearum*. *Front. Plant Sci.* **8**.

- 421 **Axtell, M. J.** (2013). ShortStack: comprehensive annotation and quantification of small RNA genes.
422 *Rna* Advance Access published 2013.
- 423 **Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., and Görlach, J.**
424 (2001). Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput
425 functional genomics in plants. *Plant Cell* **13**:1499–1510.
- 426 **Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., and Dodds, P. N.** (2014). A novel conserved
427 mechanism for plant NLR protein pairs: the “integrated decoy” hypothesis. *Front. Plant Sci.*
428 **5**:606.
- 429 **Clark, L. B., Viswanathan, P., Quigley, G., Chiang, Y.-C., McMahon, J. S., Yao, G., Chen, J., Nelsbach,**
430 **A., and Denis, C. L.** (2004). Systematic mutagenesis of the leucine-rich repeat (LRR) domain
431 of CCR4 reveals specific sites for binding to CAF1 and a separate critical role for the LRR in
432 CCR4 deadenylase activity. *J. Biol. Chem.* Advance Access published 2004.
- 433 **Debieu, M., Huard-Chauveau, C., Genissel, A., Roux, F., and Roby, D.** (2016). Quantitative disease
434 resistance to the bacterial pathogen *Xanthomonas campestris* involves an Arabidopsis
435 immune receptor pair and a gene of unknown function. *Mol. Plant Pathol.* **17**:510–520.
- 436 **Deng, Y., Liu, M., Li, X., and Li, F.** (2018). microRNA-mediated R gene regulation: molecular scabbards
437 for double-edged swords. *Sci. China Life Sci.* Advance Access published 2018.
- 438 **Diener, A. C., and Ausubel, F. M.** (2005). RESISTANCE TO FUSARIUM OXYSPORUM 1, a Dominant
439 Arabidopsis Disease-Resistance Gene, Is Not Race Specific. *Genetics* **171**:305.
- 440 **Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and**
441 **Gingeras, T. R.** (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21.
- 442 **Ernst, K., Kumar, A., Kriseleit, D., Kloos, D.-U., Phillips, M. S., and Ganai, M. W.** (2002). The broad-
443 spectrum potato cyst nematode resistance gene (Hero) from tomato is the only member of a
444 large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region.
445 *Plant J.* **31**:127–136.
- 446 **Fähling, M., Graf, H., and Siemens, J.** (2003). Pathotype Separation of *Plasmodiophora brassicae* by
447 the Host Plant. *J. Phytopathol.* **151**:425–430.
- 448 **Fei, Q., Xia, R., and Meyers, B. C.** (2013). Phased, Secondary, Small Interfering RNAs in
449 Posttranscriptional Regulatory Networks. *Plant Cell* **25**:2400.
- 450 **Fukuoka, S., Yamamoto, S.-I., Mizobuchi, R., Yamanouchi, U., Ono, K., Kitazawa, N., Yasuda, N.,**
451 **Fujita, Y., Nguyen, T. T. T., and Koizumi, S.** (2014). Multiple functional polymorphisms in a
452 single disease resistance gene in rice enhance durable resistance to blast. *Sci. Rep.* **4**:4550.
- 453 **Furci, L., Jain, R., Stassen, J., Berkowitz, O., Whelan, J., Roquis, D., Baillet, V., Colot, V., Johannes, F.,**
454 **and Ton, J.** (2019). Identification and characterisation of hypomethylated DNA loci
455 controlling quantitative resistance in Arabidopsis. *Elife* **8**:e40655.
- 456 **Gravot, A., Grillet, L., Wagner, G., Jubault, M., Lariagon, C., Baron, C., Deleu, C., Delourme, R.,**
457 **Bouchereau, A., and Manzaneres-Dauleux, M. J.** (2011). Genetic and physiological analysis
458 of the relationship between partial resistance to clubroot and tolerance to trehalose in
459 Arabidopsis thaliana. *New Phytol.* **191**:1083–1094.

- 460 **Gravot, A., Richard, G., Lime, T., Lemarié, S., Jubault, M., Lariagon, C., Lemoine, J., Vicente, J.,**
461 **Robert-Seilaniantz, A., Holdsworth, M. J., et al. (2016).** Hypoxia response in Arabidopsis
462 roots infected by *Plasmodiophora brassicae* supports the development of clubroot. *BMC*
463 *Plant Biol.* **16**:251.
- 464 **Grimm, D. G., Roqueiro, D., Salomé, P. A., Kleeberger, S., Greshake, B., Zhu, W., Liu, C., Lippert, C.,**
465 **Stegle, O., Schölkopf, B., et al. (2017).** easyGWAS: A Cloud-Based Platform for Comparing the
466 Results of Genome-Wide Association Studies. *Plant Cell* **29**:5–19.
- 467 **Hayashi, N., Inoue, H., Kato, T., Funao, T., Shirota, M., Shimizu, T., Kanamori, H., Yamane, H.,**
468 **Hayano-Saito, Y., and Matsumoto, T. (2010).** Durable panicle blast-resistance gene Pb1
469 encodes an atypical CC-NBS-LRR protein and was generated by acquiring a promoter through
470 local genome duplication. *Plant J.* **64**:498–510.
- 471 **Henderson, I. R., and Jacobsen, S. E. (2007).** Epigenetic inheritance in plants. *Nature* **447**:418–424.
- 472 **Huard-Chauveau, C., Percepied, L., Debieu, M., Rivas, S., Kroj, T., Kars, I., Bergelson, J., Roux, F.,**
473 **and Roby, D. (2013).** An atypical kinase under balancing selection confers broad-spectrum
474 disease resistance in Arabidopsis. *PLoS Genet.* **9**:e1003766.
- 475 **Hurni, S., Scheuermann, D., Krattinger, S. G., Kessel, B., Wicker, T., Herren, G., Fitze, M. N., Breen,**
476 **J., Presterl, T., and Ouzunova, M. (2015).** The maize disease resistance gene Htn1 against
477 northern corn leaf blight encodes a wall-associated receptor-like kinase. *Proc. Natl. Acad. Sci.*
478 Advance Access published 2015.
- 479 **Innes, R. W., Ameline-Torregrosa, C., Ashfield, T., Cannon, E., Cannon, S. B., Chacko, B., Chen, N.**
480 **W., Couloux, A., Dalwani, A., and Denny, R. (2008).** Differential accumulation of
481 retroelements and diversification of NB-LRR disease resistance genes in duplicated regions
482 following polyploidy in the ancestor of soybean. *Plant Physiol.* Advance Access published
483 2008.
- 484 **Jones, J. D. G., Vance, R. E., and Dangl, J. L. (2016).** Intracellular innate immune surveillance devices
485 in plants and animals. *Science* **354**.
- 486 **Jubault, M., Hamon, C., Gravot, A., Lariagon, C., Delourme, R., Bouchereau, A., and Manzanares-**
487 **Dauleux, M. J. (2008a).** Differential Regulation of Root Arginine Catabolism and Polyamine
488 Metabolism in Clubroot-Susceptible and Partially Resistant Arabidopsis Genotypes. *Plant*
489 *Physiol.* **146**:2008–2019.
- 490 **Jubault, M., Lariagon, C., Simon, M., Delourme, R., and Manzanares-Dauleux, M. J. (2008b).**
491 Identification of quantitative trait loci controlling partial clubroot resistance in new mapping
492 populations of Arabidopsis thaliana. *Theor. Appl. Genet.* **117**:191–202.
- 493 **Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-Y., Freimer, N. B., Sabatti, C., and Eskin,**
494 **E. (2010).** Variance component model to account for sample structure in genome-wide
495 association studies. *Nat. Genet.* **42**:348–354.
- 496 **Karasov, T. L., Kniskern, J. M., Gao, L., DeYoung, B. J., Ding, J., Dubiella, U., Lastra, R. O., Nallu, S.,**
497 **Roux, F., and Innes, R. W. (2014).** The long-term maintenance of a resistance polymorphism
498 through diffuse interactions. *Nature* **512**:436.

- 499 **Kawakatsu, T., Huang, S.-S. C., Jupe, F., Sasaki, E., Schmitz, R. J., Urich, M. A., Castanon, R., Nery, J.**
500 **R., Barragan, C., He, Y., et al.** (2016). Epigenomic Diversity in a Global Collection of
501 *Arabidopsis thaliana* Accessions. *Cell* **166**:492–505.
- 502 **Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D., and Penin, A. A.** (2016). A high
503 resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq
504 profiling. *Plant J. Cell Mol. Biol.* **88**:1058–1070.
- 505 **Lai, Y., and Eulgem, T.** (2018). Transcript-level expression control of plant NLR genes. *Mol. Plant*
506 *Pathol.* **19**:1267–1281.
- 507 **Laperche, A., Aigu, Y., Jubault, M., Ollier, M., Guichard, S., Glory, P., Strelkov, S., Gravot, A., and**
508 **Manzanares-Dauleux, M.** (2017). Clubroot resistance QTL are modulated by nitrogen input in
509 *Brassica napus*. *Theor. Appl. Genet.* **130**:669–684.
- 510 **Larkan, N. J., Lydiate, D. J., Parkin, I. A. P., Nelson, M. N., Epp, D. J., Cowling, W. A., Rimmer, S. R.,**
511 **and Borhan, M. H.** (2013). The *Brassica napus* blackleg resistance gene LepR3 encodes a
512 receptor-like protein triggered by the *Leptosphaeria maculans* effector AVRML1. *New Phytol.*
513 **197**:595–605.
- 514 **Law, J. A., and Jacobsen, S. E.** (2010). Establishing, maintaining and modifying DNA methylation
515 patterns in plants and animals. *Nat. Rev. Genet.* **11**:204–220.
- 516 **Lemarié, S., Robert-Seilaniantz, A., Lariagon, C., Lemoine, J., Marnet, N., Jubault, M., Manzanares-**
517 **Dauleux, M. J., and Gravot, A.** (2015). Both the Jasmonic Acid and the Salicylic Acid Pathways
518 Contribute to Resistance to the Biotrophic Clubroot Agent *Plasmodiophora brassicae* in
519 *Arabidopsis*. *Plant Cell Physiol.* **56**:2158–2168.
- 520 **Li, Y., Tessaro, M. J., Li, X., and Zhang, Y.** (2010). Regulation of the Expression of Plant Resistance
521 Gene SNC1 by a Protein with a Conserved BAT2 Domain. *Plant Physiol.* **153**:1425.
- 522 **Li, X., Kapos, P., and Zhang, Y.** (2015). NLRs in plants. *Curr. Opin. Immunol.* **32**:114–121.
- 523 **Liégard, B., Baillet, V., Etcheverry, M., Joseph, E., Lariagon, C., Lemoine, J., Evrard, A., Colot, V.,**
524 **Gravot, A., Manzanares-Dauleux, M. J., et al.** (2019). Quantitative resistance to clubroot
525 infection mediated by transgenerational epigenetic variation in *Arabidopsis*. *New Phytol.*
526 **222**:468–479.
- 527 **Luna, E., Bruce, T. J. A., Roberts, M. R., Flors, V., and Ton, J.** (2012). Next-Generation Systemic
528 Acquired Resistance. *Plant Physiol.* **158**:844.
- 529 **Maekawa, T., Kufer, T. A., and Schulze-Lefert, P.** (2011). NLR functions in plant and animal immune
530 systems: so far and yet so close. *Nat. Immunol.* **12**:817–826.
- 531 **Manzanares-Dauleux, M. J., Divaret, I., Baron, F., and Thomas, G.** (2000). Evaluation of French
532 *Brassica oleracea* landraces for resistance to *Plasmodiophora brassicae*. *Euphytica* **113**:211–
533 218.
- 534 **Matzke, M. A., and Mosher, R. A.** (2014). RNA-directed DNA methylation: an epigenetic pathway of
535 increasing complexity. *Nat. Rev. Genet.* Advance Access published 2014.

- 536 **Meyers, B. C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R. W.** (2003). Genome-Wide
537 Analysis of NBS-LRR–Encoding Genes in Arabidopsis. *Plant Cell* **15**:809.
- 538 **Molinier, J., Ries, G., Zipfel, C., and Hohn, B.** (2006). Transgeneration memory of stress in plants.
539 *Nature* **442**:1046.
- 540 **Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., and Narusaka, Y.**
541 (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial
542 pathogens. *Plant J.* **60**:218–226.
- 543 **Nelson, R., Wiesner-Hanks, T., Wissner, R., and Balint-Kurti, P.** (2017). Navigating complexity to breed
544 disease-resistant crops. *Nat. Rev. Genet.* **19**:21.
- 545 **Nishimura, M. T., Anderson, R. G., Cherkis, K. A., Law, T. F., Liu, Q. L., Machius, M., Nimchuk, Z. L.,**
546 **Yang, L., Chung, E.-H., and El Kasmi, F.** (2017). TIR-only protein RBA1 recognizes a pathogen
547 effector to regulate cell death in Arabidopsis. *Proc. Natl. Acad. Sci.* **114**:E2053–E2062.
- 548 **Palma, K., Thorgrimsen, S., Malinovsky, F. G., Fiil, B. K., Nielsen, H. B., Brodersen, P., Hofius, D.,**
549 **Petersen, M., and Mundy, J.** (2010). Autoimmunity in Arabidopsis *acd11* is mediated by
550 epigenetic regulation of an immune receptor. *PLoS Pathog.* **6**:e1001137.
- 551 **Pfaffl, M. W.** (2001). A new mathematical model for relative quantification in real-time RT-PCR.
552 *Nucleic Acids Res.* **29**:e45.
- 553 **Pilet-Nayel, M.-L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.-C., Fournet, S., Durel, C.-E., and**
554 **Delourme, R.** (2017). Quantitative resistance to plant pathogens in pyramiding strategies for
555 durable crop protection. *Front. Plant Sci.* **8**:1838.
- 556 **Qu, S., Liu, G., Zhou, B., Bellizzi, M., Zeng, L., Dai, L., Han, B., and Wang, G.-L.** (2006). The Broad-
557 Spectrum Blast Resistance Gene Pi9 Encodes a Nucleotide-Binding Site–Leucine-Rich Repeat
558 Protein and Is a Member of a Multigene Family in Rice. *Genetics* **172**:1901.
- 559 **Quadrana, L., and Colot, V.** (2016). Plant transgenerational epigenetics. *Annu. Rev. Genet.* **50**:467–
560 491.
- 561 **Quadrana, L., Bortolini Silveira, A., Mayhew, G. F., LeBlanc, C., Martienssen, R. A., Jeddloh, J. A.,**
562 **and Colot, V.** (2016). The Arabidopsis thaliana mobilome and its impact at the species level.
563 *eLife* **5**:e15716.
- 564 **Richards, E. J.** (2006). Inherited epigenetic variation--revisiting soft inheritance. *Nat. Rev. Genet.*
565 **7**:395–401.
- 566 **Robinson, M. D., McCarthy, D. J., and Smyth, G. K.** (2010). edgeR: a Bioconductor package for
567 differential expression analysis of digital gene expression data. *Bioinformatics* **26**:139–140.
- 568 **Saucet, S. B., Ma, Y., Sarris, P. F., Furzer, O. J., Sohn, K. H., and Jones, J. D.** (2015). Two linked pairs
569 of Arabidopsis TNL resistance genes independently confer recognition of bacterial effector
570 AvrRps4. *Nat. Commun.* **6**:6338.
- 571 **Shao, Z.-Q., Xue, J.-Y., Wu, P., Zhang, Y.-M., Wu, Y., Hang, Y.-Y., Wang, B., and Chen, J.-Q.** (2016).
572 Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat (NBS-LRR)

- 573 genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant*
574 *Physiol.* Advance Access published 2016.
- 575 **Shivaprasad, P. V., Chen, H.-M., Patel, K., Bond, D. M., Santos, B. A., and Baulcombe, D. C.** (2012). A
576 microRNA superfamily regulates nucleotide binding site–leucine-rich repeats and other
577 mRNAs. *Plant Cell* Advance Access published 2012.
- 578 **Silveira, A. B., Trontin, C., Cortijo, S., Barau, J., Del Bem, L. E. V., Loudet, O., Colot, V., and Vincentz,**
579 **M.** (2013). Extensive Natural Epigenetic Variation at a De Novo Originated Gene. *PLOS Genet.*
580 **9:e1003437.**
- 581 **Slaughter, A., Daniel, X., Flors, V., Luna, E., Hohn, B., and Mauch-Mani, B.** (2011). Descendants of
582 primed Arabidopsis plants exhibit resistance to biotic stress. *Plant Physiol.* Advance Access
583 published 2011.
- 584 **Some, A., Manzanares, M. J., Laurens, F., Baron, F., Thomas, G., and Rouxel, F.** (1996). Variation for
585 virulence on *Brassica napus* L. amongst Plasmodiophora brassicae collections from France
586 and derived single-spore isolates. *Plant Pathol.* **45**:432–439.
- 587 **Stahl, E. A., Dwyer, G., Mauricio, R., Kreitman, M., and Bergelson, J.** (1999). Dynamics of disease
588 resistance polymorphism at the Rpm1 locus of Arabidopsis. *Nature* **400**:667.
- 589 **Stroud, H., Greenberg, M. V. C., Feng, S., Bernatavichute, Y. V., and Jacobsen, S. E.** (2013).
590 Comprehensive Analysis of Silencing Mutants Reveals Complex Regulation of the Arabidopsis
591 Methylome. *Cell* **152**:352–364.
- 592 **Stroud, H., Do, T., Du, J., Zhong, X., Feng, S., Johnson, L., Patel, D. J., and Jacobsen, S. E.** (2014).
593 Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nat. Struct.*
594 *Mol. Biol.* **21**:64–72.
- 595 **Stuart, T., Eichten, S. R., Cahn, J., Karpievitch, Y. V., Borevitz, J. O., and Lister, R.** Population scale
596 mapping of transposable element diversity reveals links to gene regulation and epigenomic
597 variation. *eLife* **5**.
- 598 **Team, R. C.** (2013). R: A language and environment for statistical computing Advance Access
599 published 2013.
- 600 **Thomas, C. M.** (1998). Genetic and molecular analysis of tomato Cf genes for resistance to
601 *Cladosporium fulvum*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**:1413–1424.
- 602 **Tsuchiya, T., and Eulgem, T.** (2013). An alternative polyadenylation mechanism coopted to the
603 Arabidopsis RPP7 gene through intronic retrotransposon domestication. *Proc. Natl. Acad. Sci.*
604 **110**:E3535.
- 605 **Weigel, D., and Colot, V.** (2012). Epialleles in plant evolution. *Genome Biol.* **13**.
- 606 **Williams, S. J., Sohn, K. H., Wan, L., Bernoux, M., Sarris, P. F., Segonzac, C., Ve, T., Ma, Y., Saucet, S.**
607 **B., and Ericsson, D. J.** (2014). Structural basis for assembly and function of a heterodimeric
608 plant immune receptor. *Science* **344**:299–303.

- 609 **Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J.** (2007). An “Electronic
610 Fluorescent Pictograph” browser for exploring and analyzing large-scale biological data sets.
611 *PloS One* **2**:e718.
- 612 **Xia, S., Cheng, Y. T., Huang, S., Win, J., Soards, A., Jinn, T.-L., Jones, J. D. G., Kamoun, S., Chen, S.,**
613 **Zhang, Y., et al.** (2013). Regulation of Transcription of Nucleotide-Binding Leucine-Rich
614 Repeat-Encoding Genes SNC1 and RPP4 via H3K4 Trimethylation. *Plant Physiol.* **162**:1694.
- 615 **Xiao, S., Ellwood, S., Calis, O., Patrick, E., Li, T., Coleman, M., and Turner, J. G.** (2001). Broad-
616 spectrum mildew resistance in *Arabidopsis thaliana* mediated by RPW8. *Science* **291**:118–
617 120.
- 618 **Xu, X., Hayashi, N., Wang, C.-T., Fukuoka, S., Kawasaki, S., Takatsuji, H., and Jiang, C.-J.** (2014). Rice
619 blast resistance gene Pikahei-1 (t), a member of a resistance gene cluster on chromosome 4,
620 encodes a nucleotide-binding site and leucine-rich repeat protein. *Mol. Breed.* **34**:691–700.
- 621 **Yue, J.-X., Meyers, B. C., Chen, J.-Q., Tian, D., and Yang, S.** (2012). Tracing the origin and
622 evolutionary history of plant nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes.
623 *New Phytol.* **193**:1049–1063.
- 624 **Zhang, X.-C., and Gassmann, W.** (2007). Alternative splicing and mRNA levels of the disease
625 resistance gene RPS4 are induced during defense responses. *Plant Physiol.* **145**:1577–1587.
- 626 **Zhang, H., Tang, K., Wang, B., Duan, C.-G., Lang, Z., and Zhu, J.-K.** (2014). Protocol: a beginner’s
627 guide to the analysis of RNA-directed DNA methylation in plants. *Plant Methods* **10**:18.
- 628 **Zhang, H., Lang, Z., and Zhu, J.-K.** (2018). Dynamics and function of DNA methylation in plants. *Nat.*
629 *Rev. Mol. Cell Biol.* **19**:489–506.
- 630 **Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T., and Henikoff, S.** (2007). Genome-wide analysis
631 of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation
632 and transcription. *Nat. Genet.* **39**:61–69.
- 633 **Zou, B., Yang, D.-L., Shi, Z., Dong, H., and Hua, J.** (2014). Monoubiquitination of Histone 2B at the
634 disease resistance gene locus regulates its expression and impacts immune responses in
635 *Arabidopsis*. *Plant Physiol.* Advance Access published 2014.

636 **Acknowledgments**

637 We acknowledge the Versailles Arabidopsis Stock Center for providing HIF lines, the biological
638 resource center BrACySol for furnishing Brassica accessions, the Gentyane Platform for their
639 contribution in genotyping the fine-mapping population, the Institute Marie Curie for sequencing mRNA
640 and sRNA. Cyril Falentin is acknowledged for his help in the design of Kaspar markers. IGEPP
641 colleagues are acknowledged for their technical support for clubroot phenotyping and sampling.

642 **Author contributions**

643 BL, AG, MMD and MJ designed and conducted the experiments. CL and AG carried out the fine
644 mapping. CL, JL, JB, YA, BL, AG, MJ performed the phenotyping and sampling. BL, AG, JB, JL and
645 MJ carried out epigenetic and gene expression studies. BL, YA, LQ and VC conducted bioinformatics
646 analyses. AG, BL, LQ, VC, MMD and MJ participated in drafting and revisions of the manuscript.

647 **Competing Interests statement**

648 The authors declare no competing interests.

649 **Figure legends**

650 **Fig. 1 | Fine mapping of *Pb-At5.2*.** **a** Genetic map and residual heterozygosity in the Recombinant
651 Inbred Line (RIL) 499 derived from Bur-0 and Col-0 and position of clubroot resistance QTL (from
652 Jubault et al. (2008b)). Black: Bur-0 allele, White: Col-0 allele, Hatched: Heterozygous (Col-0/Bur-0).
653 **b** Allele configuration at *Pb-At5.2* in the two derived HIF lines 10499 and 13499. **c** Photos showing *Pb-*
654 *At5.2* conferred partial resistance to the eH isolate conferred by *Pb-At5.2* in the RIL499 genetic
655 background. Observations were made at 21 days post-inoculation. The white arrow indicates the
656 presence of a limited number of galls in inoculated 10499. **d** First round of fine mapping: allelic structure
657 in the F1 lines derived from reciprocal crosses between 10499 and 13499. 554 F3 lines with
658 recombination in the confidence interval were screened using 10 SNP markers between AG_14959 and
659 AG_20993. High density genotyping (94 SNP from K1 to K94), in a series of 88 recombinant F3 lines,
660 and clubroot phenotyping in their bulked segregating F4 progenies, led to a new interval between
661 markers K58 (At5g47230) and K65 (At5g47360). **e** Second round of fine mapping: Recombination
662 positions in homozygous individuals obtained from selected recombinant lines. For each line, the
663 GA/LA index (disease symptoms) is indicated on the right panel. Center lines show the medians; box
664 limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the
665 interquartile range from the 25th and 75th percentiles, outliers are represented by dots; data points are
666 plotted as open circles. Number of individual plants analysed for each genotype is indicated (n). The
667 notches are defined as $\pm 1.58 \cdot \text{IQR} / \sqrt{n}$ and represent the 95% confidence interval for each median.

668 Genetic markers are indicated for each recombination position. Markers between CLG5 and 2017-C
669 were used in every line, but only shown for 2313-15. **f** New 26 kb interval of *Pb-At5.2* between markers
670 CL4 (excluded) and K64 (excluded), containing eight annotated ORF, six transposons and one lncRNA.
671 Yellow and red diamonds indicate SNP and nucleotide deletions, respectively.

672 **Fig. 2 | Identification of two candidate NLR-encoding genes with contrasted expression and**
673 **methylation profiles at QTL *Pb-At5.2*.** **a** Sequence variations and expression levels of genes in the *Pb-*
674 *At5.2* region. Yellow and red diamonds indicate SNP and INDEL variations, respectively. Gene
675 expression values are from RNAseq analyses conducted in inoculated and control conditions at 14 dpi
676 (\log_2 normalized cpm), with parental lines Col-0/Bur-0 and HIF lines 13499/10499 (the last two were
677 derived from the RIL 499, and homozygous *Pb-At5.2*_{Col/Col} and *Pb-At5.2*_{Bur/Bur}, respectively)
678 (**Supplementary Data 2**). FDR adjusted *p-values* are shown if less than 0.05. **b** Methylation profiles in
679 the At5g47260 and At5g47280 region, in Col-0 and Bur-0 accessions, inferred from bisulfite data
680 previously reported in Kawakatsu et al. (2016) Average methylation level was calculated within non-
681 overlapping 100-bp windows starting 1 kb before the TSS site of At5g47260 and stopping 1 kb after the
682 TSE site of At5g47280. In red: methylation in the CG context. In green: methylation in the CHG context.
683 In blue: methylation in the CHH context. **c** Methylation profiles obtained by CHOP-qPCR on
684 At5g47260 in inoculated and non-inoculated roots of Bur-0/Col-0, 10499/13499, and in the homozygous
685 recombinant lines 2313-15 (*Pb-At5.2*_{Col/Col}) and 1381-2 (*Pb-At5.2*_{Bur/Bur}). Those two last genotypes
686 harbor the narrowest recombination events from either side of *Pb-At5.2* (between markers CLG4 and
687 K64, details in Fig. 1). Center lines show the medians; box limits indicate the 25th and 75th percentiles
688 as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th
689 percentiles, outliers are represented by dots; data points are plotted as open circles. *n* = 4 bulks of 6
690 plants and *p-values* are shown (sided t-test).

691 **Fig. 3 | Intermediate methylation and transcription levels of candidate genes in heterozygous**
692 **plants are associated with full clubroot susceptibility.** Eighty-three individual plants of the
693 segregating progeny from the recombinant line 2509 (heterozygous in Chr.5 regions between genetic

694 markers K58 and K93) were sampled at 21 dpi. Leaves from each individual plant were used for
695 genotyping (PCR marker CL_N8), which defined n pools of >3 plants of each zygosity profile: Bur/Bur
696 (n=5), Col/Bur (n=13) and Col/Col (n=6) (black, grey and white boxes, respectively). Each plant pool
697 was evaluated for **a** clubroot resistance (GA/LA), **b** % methylation at the locus, and **c-d** candidate gene
698 expression (At5g47260 and At5g47280). Gene expression was quantified through RT-qPCR, data were
699 normalized over mean-Cp from the pools Bur/Bur, following Pfaffl's method with two reference genes
700 (Pfaffl, 2001). Center lines show the medians; box limits indicate the 25th and 75th percentiles as
701 determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th
702 percentiles, outliers are represented by dots; data points are plotted as open circles.

703 **Fig. 4 | Natural epigenetic variation at *Pb-At5.2* affects the expression of At5g47260/Atg47280**
704 **among Arabidopsis accessions. a** Screening for 1001 genome Arabidopsis accessions displaying a
705 Col/Bur-like genomic structure at *Pb-At5.2* (chr5: 19185600 - 19200600). X axis: Horizontal coverage
706 region covered by at least one read. Y axis: Vertical coverage in read percentage. The 401 accessions
707 framed in the northeast intercardinal region delimited by dotted lines harbour a vertical read coverage
708 >0.8 and an horizontal DNA-seq >0.5 (DNA-seq data from⁴⁶). **b** Clustering of a series of accessions
709 harbouring Col/Bur-like genomic structure at *Pb-At5.2*, by their level of methylation on At5g47260 and
710 At5g47280. Bisulfite data were obtained from the 1001 genome data project (**Supplementary Data 3**
711 **sheet 2**). Average methylation level was calculated 1 kb before the TSS site of At5g47260 and stopping
712 1 kb after the TSE site of At5g47280 for each context. **c** Spearman correlation between the methylation
713 and gene expression of At5g47260 and At5g47280 in a subset of 253 *Arabidopsis* accessions for which
714 expression data was available (RNAseq data from⁴¹). Correlation between gene expression and
715 methylation level is given for all three DNA-methylation contexts in the interval from 1 kb before the
716 TSS site of At5g47260, to 1 kb after the TSE site of At5g47280. **d** Confirmation of methylation profiles
717 at At5g47260 in inoculated roots from 20 ecotypes. Methylation level was obtained using CHOP-qPCR.
718 Black and white bars indicate genotypes with Bur-like and Col-like methylation patterns, respectively.
719 Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R

720 software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers
721 are represented by dots; data points are plotted as open circles. n = 4 bulks of 6 plants.

722 **Fig. 5 | Clubroot symptom variation among Arabidopsis accessions is linked to epivariation at *Pb-***
723 ***At5.2*.** Effect of *Pb-At5.2* epiallele variation on clubroot susceptibility, evaluated in a series of 126
724 Arabidopsis accessions harbouring similar Bur/Col-like genomic structure at the locus. In total 42
725 accessions have Col-0 like epiallele and 84 have Bur-0 like epiallele. For each accession, the mean
726 GA/LA was obtained by modelling raw data of two biological replicates with two blocks (6 individual
727 plants in each block). Center lines show the medians; box limits indicate the 25th and 75th percentiles
728 as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th
729 percentiles, outliers are represented by dots; data points are plotted as open circles. The notches are
730 defined as $\pm 1.58 \cdot \text{IQR} / \sqrt{n}$ and represent the 95% confidence interval for each median. *p*-value
731 (Wilcoxon test) is indicated.

732 **Fig. 6. Epigenetic variation at *Pb-At5.2* correlates with the abundance of locus-targeted sRNA, but**
733 **is maintained by the RNA-independent methylation pathway. a** Mapping of sRNA-seq reads. Reads
734 were obtained from roots of Col-0 and Bur-0 accessions at two time points at 14 and 21 days after
735 sowing. For each condition, 3 bulks (numbered from Rep 1 to Rep 3) of 6 plants were used **b** Methylation
736 state at the *Pb-At5.2* locus in knock-out lines (Col-0 genomic background) defective for the RdDM or
737 non-RdDM pathways (Stroud et al., 2014). In red the methylation in CG context. In green the
738 methylation in CHG context; in blue the methylation in CHH context.

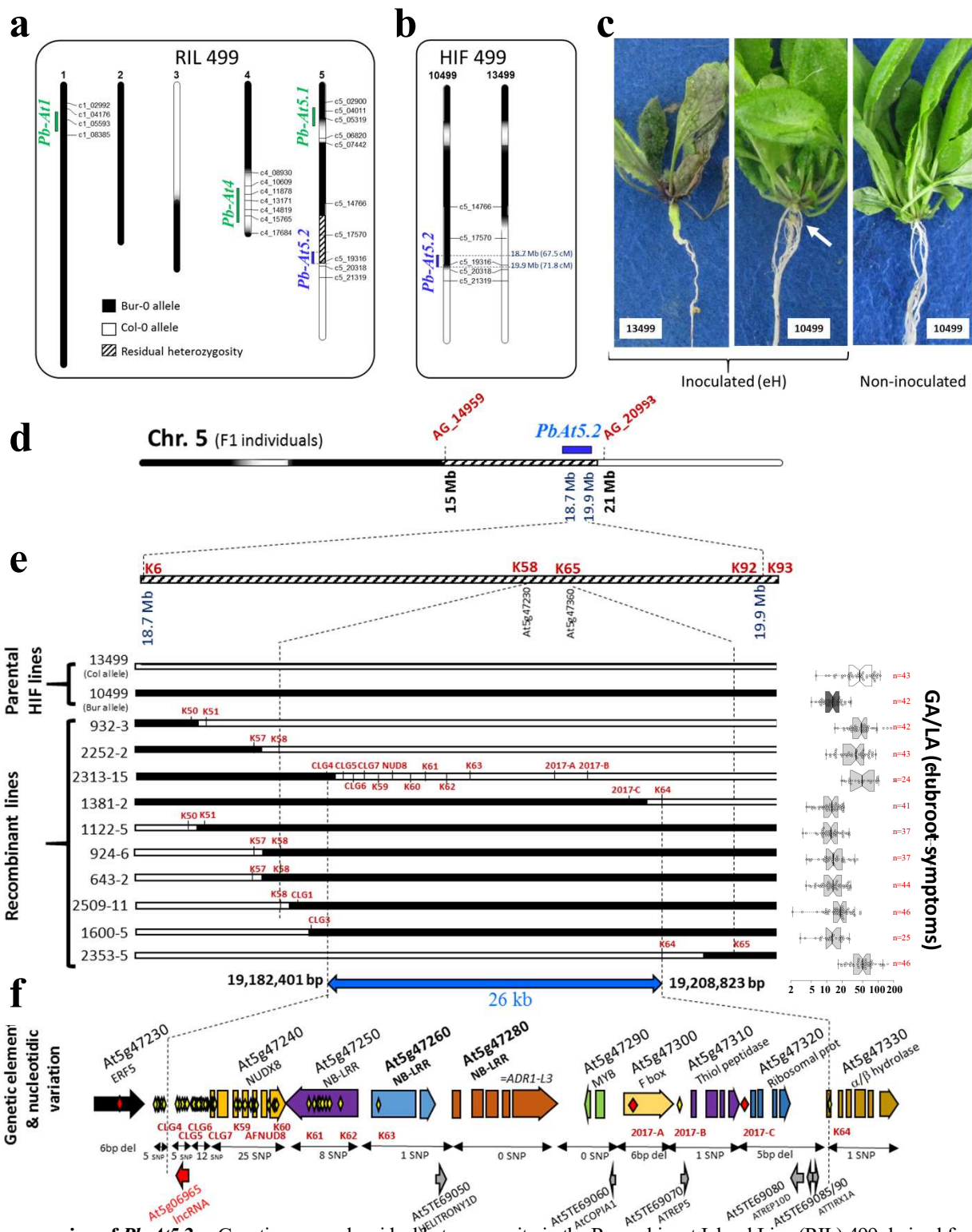
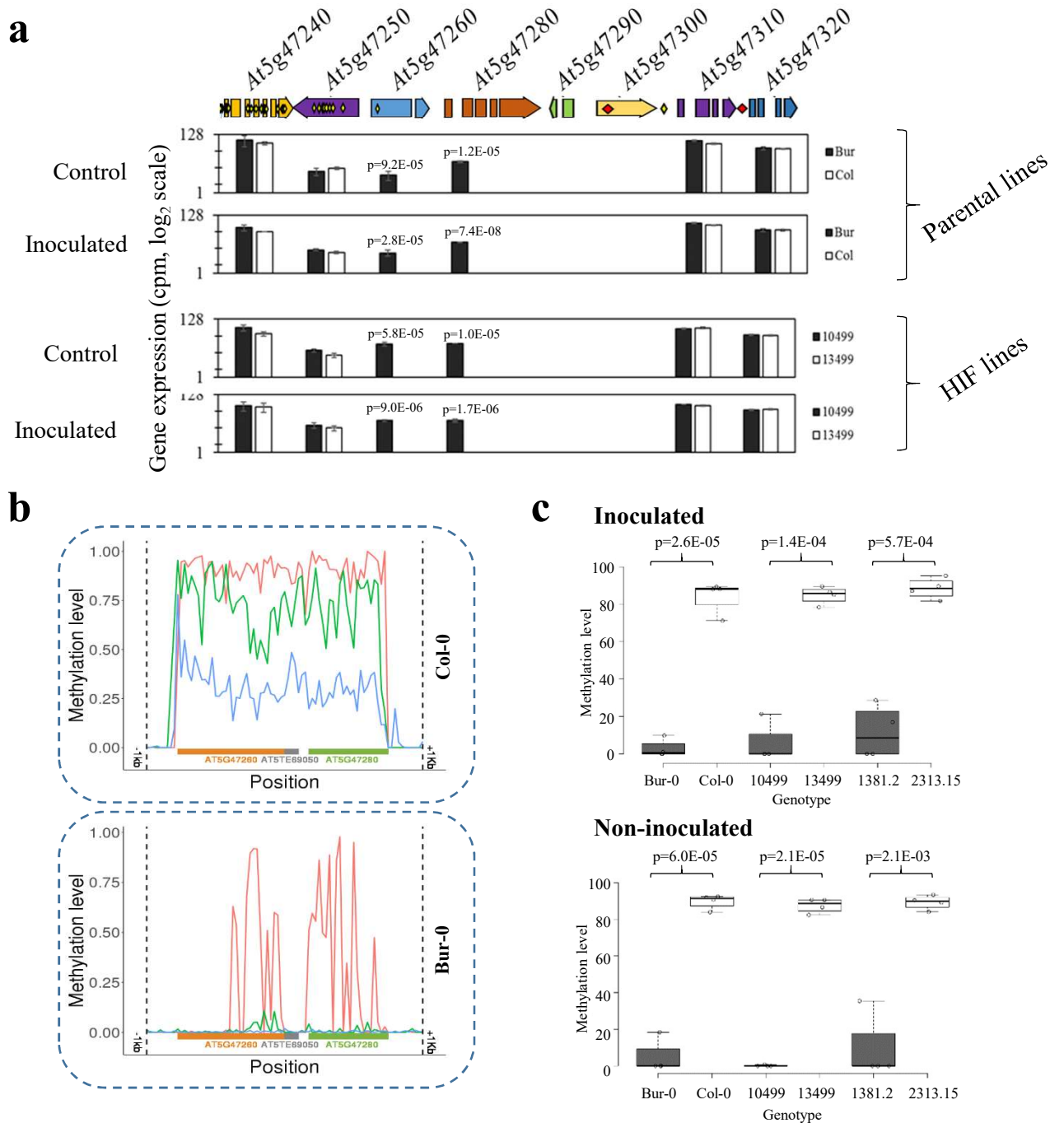


Fig. 1 | Fine mapping of *Pb-At5.2*. **a** Genetic map and residual heterozygosity in the Recombinant Inbred Line (RIL) 499 derived from Bur-0 and Col-0 (from⁴²) and position of clubroot resistance QTL (from⁴³). Black: Bur-0 allele, White: Col-0 allele, Hatched: Heterozygous (Col-0/Bur-0). **b** Allele configuration at *Pb-At5.2* in the two derived HIF lines 10499 and 13499. **c** Photos showing *Pb-At5.2* conferred partial resistance to the eH isolate conferred by *Pb-At5.2* in the RIL499 genetic background. Observations were made at 21 days post-inoculation. The white arrow indicates the presence of a limited number of galls in inoculated 10499. **d** First round of fine mapping: allelic structure in the F1 lines derived from reciprocal crosses between 10499 and 13499. 554 F3 lines with recombination in the confidence interval were screened using 10 SNP markers between AG_14959 and AG_20993. High density genotyping (94 SNP from K1 to K94), in a series of 88 recombinant F3 lines, and clubroot phenotyping in their bulked segregating F4 progenies, led to a new interval between markers K58 (At5g47230) and K65 (At5g47360). **e** Second round of fine mapping: Recombination positions in homozygous individuals obtained from selected recombinant lines. For each line, the GA/LA index (disease symptoms) is indicated on the right panel. Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots; data points are plotted as open circles. Number of individual plants analysed for each genotype is indicated (n). The notches are defined as $\pm 1.58 \cdot \text{IQR} / \sqrt{n}$ and represent the 95% confidence interval for each median. Genetic markers are indicated for each recombination position. Markers between CLG5 and 2017-C were used in every line, but only shown for 2313-15. **f** New 26 kb interval of *Pb-At5.2* between markers CL4 (excluded) and K64 (excluded), containing eight annotated ORF, six transposons and one lncRNA. Yellow and red diamonds indicate SNP and nucleotide deletions, respectively.



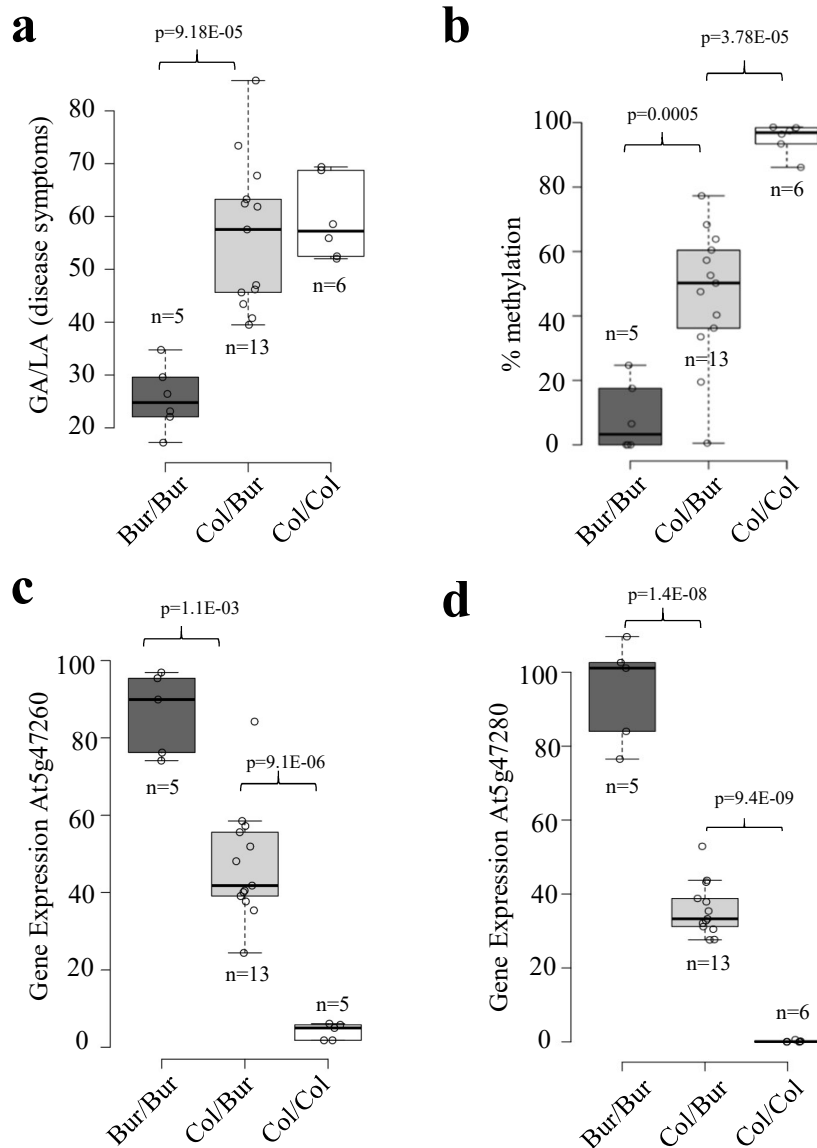


Fig. 3 | Intermediate methylation and transcription levels of candidate genes in heterozygous plants are associated with full clubroot susceptibility. Eighty-three individual plants of the segregating progeny from the recombinant line 2509 (heterozygous in Chr.5 regions between genetic markers K58 and K93) were sampled at 21 dpi. Leaves from each individual plant were used for genotyping (PCR marker CL_N8), which defined n pools of >3 plants of each zygosity profile: Bur/Bur ($n=5$), Col/Bur ($n=13$) and Col/Col ($n=6$) (black, grey and white boxes, respectively). Each plant pool was evaluated for **a** clubroot resistance (GA/LA), **b** % methylation at the locus, and **c-d** candidate gene expression (At5g47260 and At5g47280). Gene expression was quantified through RT-qPCR, data were normalized over mean-Cp from the pools Bur/Bur, following Pfaffl's method with two reference genes⁷⁹. Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots; data points are plotted as open circles.

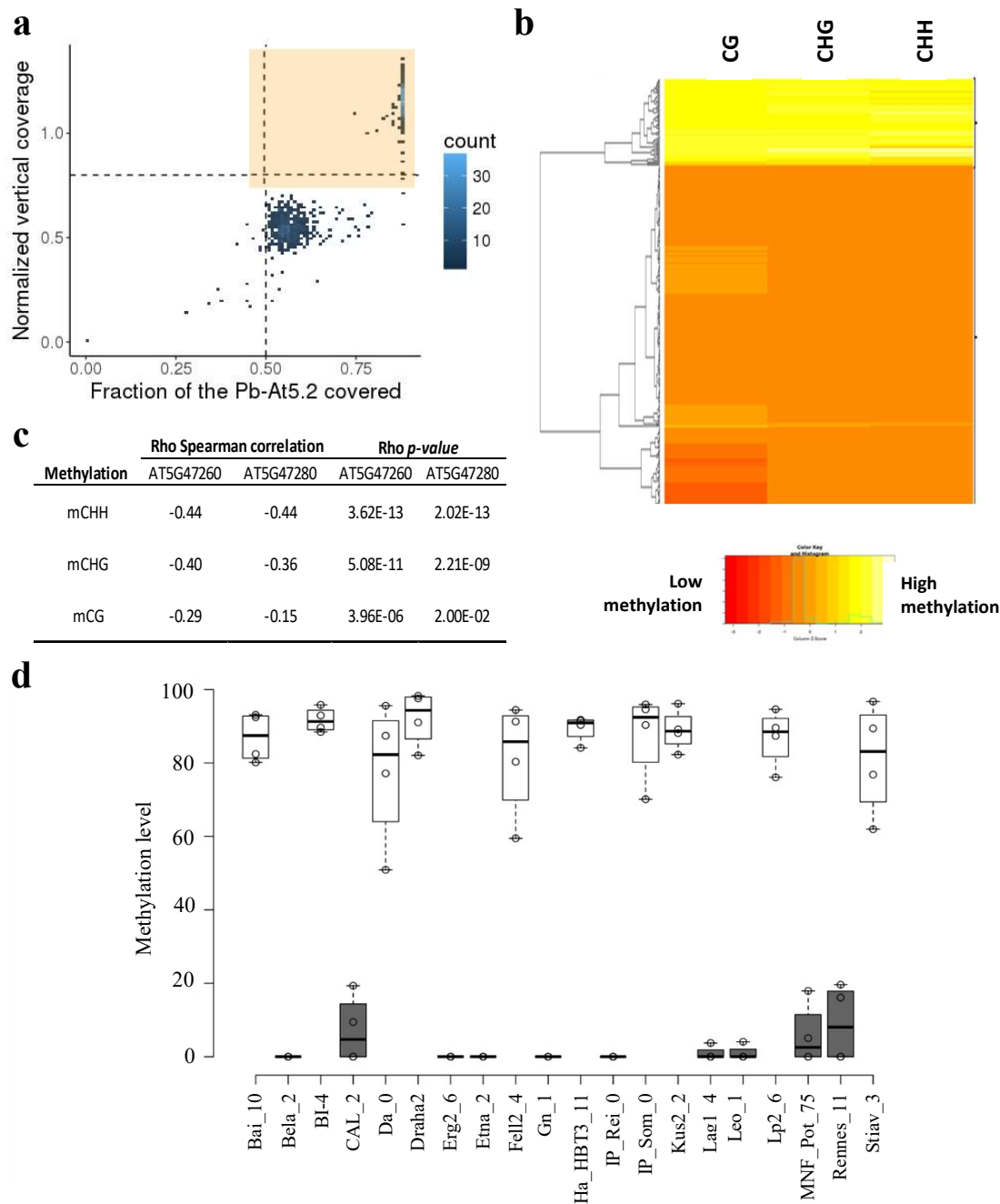


Fig. 4 | Natural epigenetic variation at *Pb-At5.2* affects the expression of *At5g47260/Atg47280* among *Arabidopsis* accessions. **a Screening for 1001 genome *Arabidopsis* accessions displaying a Col/Bur-like genomic structure at *Pb-At5.2* (chr5: 19185600 - 19200600). X axis: Horizontal coverage region covered by at least one read. Y axis: Vertical coverage in read percentage. The 401 accessions framed in the northeast intercardinal region delimited by dotted lines harbour a vertical read coverage >0.8 and an horizontal DNA-seq >0.5 (DNA-seq data from⁴⁶). **b** Clustering of a series of accessions harbouring Col/Bur-like genomic structure at *Pb-At5.2*, by their level of methylation on *At5g47260* and *At5g47280*. Bisulfite data were obtained from the 1001 genome data project (**Supplementary Data 3 sheet 2**). Average methylation level was calculated 1 kb before the TSS site of *At5g47260* and stopping 1 kb after the TSE site of *At5g47280* for each context. **c** Spearman correlation between the methylation and gene expression of *At5g47260* and *At5g47280* in a subset of 253 *Arabidopsis* accessions for which expression data was available (RNAseq data from⁴¹). Correlation between gene expression and methylation level is given for all three DNA-methylation contexts in the interval from 1 kb before the TSS site of *At5g47260*, to 1 kb after the TSE site of *At5g47280*. **d** Confirmation of methylation profiles at *At5g47260* in inoculated roots from 20 ecotypes. Methylation level was obtained using CHOP-qPCR. Black and white bars indicate genotypes with Bur-like and Col-like methylation patterns, respectively. Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots; data points are plotted as open circles. $n = 4$ bulks of 6 plants.**

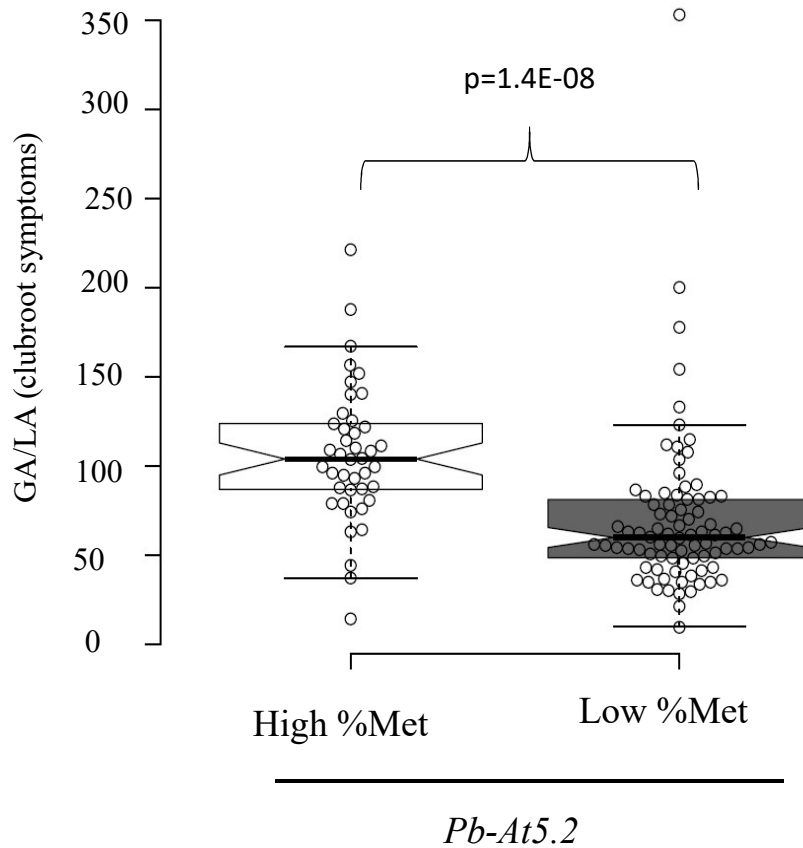


Fig. 5 | Clubroot symptom variation among Arabidopsis accessions is linked to epivariation at *Pb-At5.2*. Effect of *Pb-At5.2* epiallele variation on clubroot susceptibility, evaluated in a series of 126 Arabidopsis accessions harbouring similar Bur/Col-like genomic structure at the locus. In total 42 accessions have Col-0 like epiallele and 84 have Bur-0 like epiallele. For each accession, the mean GA/LA was obtained by modelling raw data of two biological replicates with two blocks (6 individual plants in each block). Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots; data points are plotted as open circles. The notches are defined as $\pm 1.58 \cdot \text{IQR} / \sqrt{n}$ and represent the 95% confidence interval for each median. *p*-value (Wilcoxon test) is indicated.

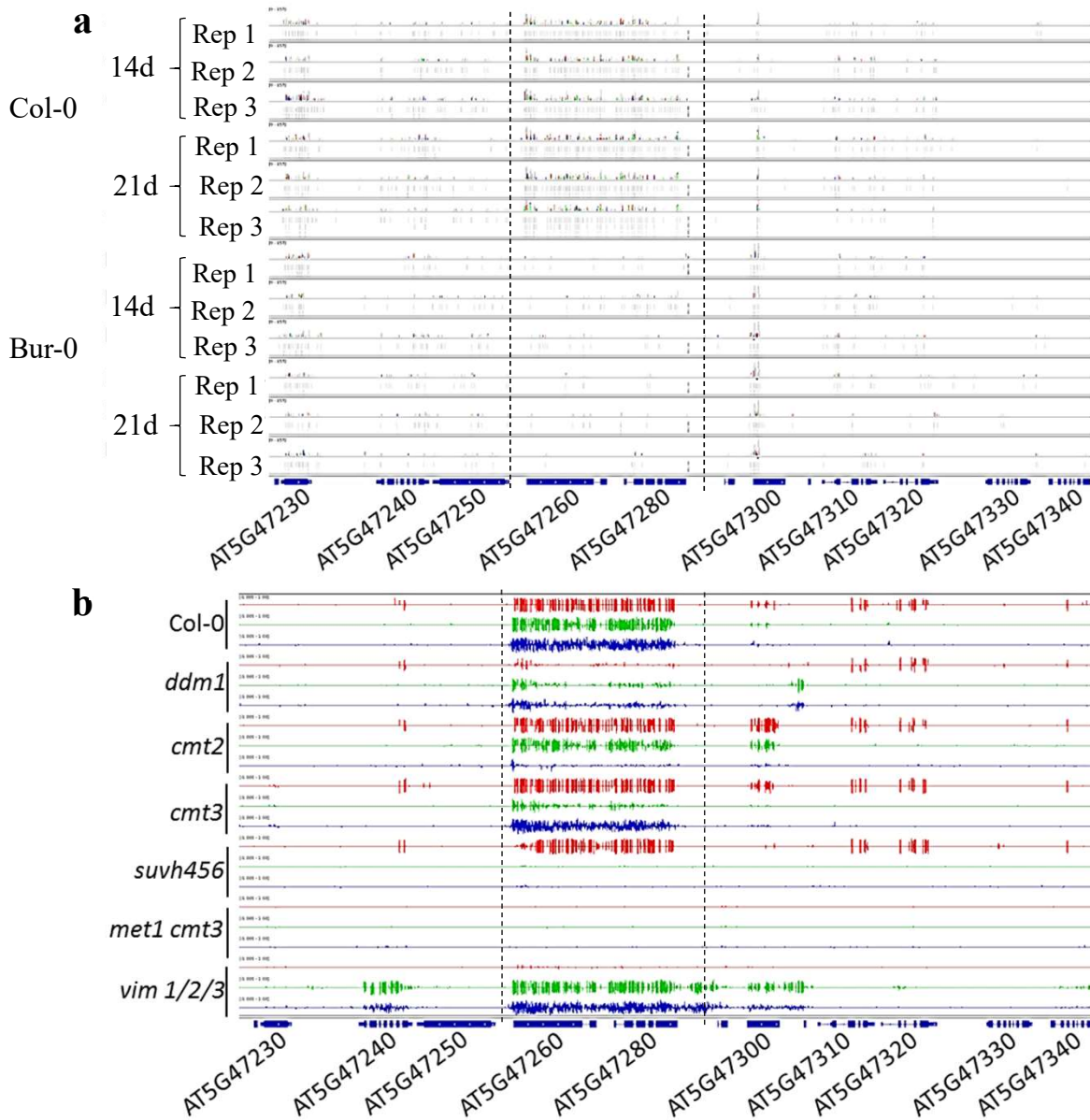


Fig. 6. Epigenetic variation at *Pb-At5.2* correlates with the abundance of locus-targeted sRNA, but is maintained by the RNA-independent methylation pathway. a Mapping of sRNA-seq reads. Reads were obtained from roots of Col-0 and Bur-0 accessions at two time points at 14 and 21 days after sowing. For each condition, 3 bulks (numbered from Rep 1 to Rep 3) of 6 plants were used **b** Methylation state at the *Pb-At5.2* locus in knock-out lines (Col-0 genomic background) defective for the RdDM or non-RdDM pathways⁴⁹. In red the methylation in CG context. In green the methylation in CHG context; in blue the methylation in CHH context.