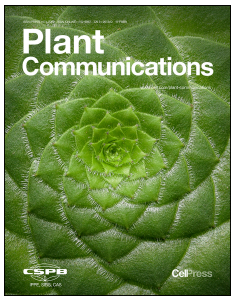
Two adjacent NLR genes conferring quantitative resistance to clubroot disease in Arabidopsis are regulated by a stably inherited epiallelic variation

Antoine Gravot, Benjamin Liégard, Leandro Quadrana, Florian Veillet, Yoann Aigu, Tristan Bargain, Juliette Bénéjam, Christine Lariagon, Jocelyne Lemoine, Vincent Colot, et al.

Journal Pre-proof



Two adjacent *NLR* genes conferring quantitative resistance to clubroot disease in *Arabidopsis* are regulated by a stably inherited epiallelic variation

Antoine Gravot, Benjamin Liégard, Leandro Quadrana, Florian Veillet, Yoann Aigu, Tristan Bargain, Juliette Bénéjam, Christine Lariagon, Jocelyne Lemoine, Vincent Colot, Maria J. Manzanares-Dauleux, Mélanie Jubault

PII: S2590-3462(24)00044-0

DOI: https://doi.org/10.1016/j.xplc.2024.100824

Reference: XPLC 100824

To appear in: *PLANT COMMUNICATIONS*

Received Date: 5 September 2023

Revised Date: 21 December 2023

Accepted Date: 19 January 2024

Please cite this article as: Gravot, A., Liégard, B., Quadrana, L., Veillet, F., Aigu, Y., Bargain, T., Bénéjam, J., Lariagon, C., Lemoine, J., Colot, V., Manzanares-Dauleux, M.J., Jubault, M., Two adjacent *NLR* genes conferring quantitative resistance to clubroot disease in *Arabidopsis* are regulated by a stably inherited epiallelic variation, *PLANT COMMUNICATIONS* (2024), doi: https://doi.org/10.1016/j.xplc.2024.100824.

Two adjacent *NLR* genes conferring quantitative resistance to clubroot disease in *Arabidopsis* are regulated by a stably inherited epiallelic variation

Antoine Gravot[1], Benjamin Liégard[1], Leandro Quadrana[2], Florian Veillet[3*], Yoann Aigu[1], Tristan Bargain[1], Juliette Bénéjam[1], Christine Lariagon[1], Jocelyne Lemoine[1], Vincent Colot[2], Maria J. Manzanares-Dauleux[1] and Mélanie Jubault[1*]

These authors contributed equally: Antoine Gravot, Benjamin Liégard

[1] IGEPP Institut Agro, INRAE, Université de Rennes, 35650 Le Rheu, France

[2] Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole Normale Supérieure, Centre National de la Recherche Scientifique (CNRS), Institut National de la Santé et de la Recherche Médicale (INSERM), F-75005 Paris, France

[3]IGEPP INRAE, Institut Agro, Université de Rennes, 29260 Ploudaniel, France

Present address: Aix-Marseille University, CEA, CNRS, Institute of Biosciences and Biotechnologies, BIAM Cadarache, 13108 Saint-Paul-lez-Durance, France

Running title: Two adjacent NLR genes are controlled by an epimutation

Corresponding author: Mélanie Jubault

Mailing address: IGEPP Institut Agro, INRAE, Université de Rennes, 35650 Le Rheu, France

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

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

We report a natural, stable, transgenerationally inherited epimutation in *Arabidopsis* that affects two adjacent genes encoding immune receptors, AT5G47260 and AT5G47280, which control broad-spectrum partial resistance to the root pathogen *Plasmodiophora brassicae*. This methylation variation is widespread across Arabidopsis ecotypes and correlates negatively with the expression of both genes.

Abstract

Clubroot caused by the protist *Plasmodiophora brassicae* is a major disease affecting cultivated *Brassicaceae*. Using a combination of QTL fine mapping, CRISPR/Cas9 validation and extensive

analyses of DNA sequence and methylation patterns, we uncovered that the two adjacent neighboring NLR (nucleotide-binding and leucine-rich repeat) genes *AT5G47260* and *AT5G47280* cooperate in controlling broad-spectrum quantitative partial resistance to the root pathogen *P. brassicae* in Arabidopsis and that they are epigenetically regulated. The variation in DNA methylation is not associated with any nucleotide variation nor any TE presence/absence variants and is stably inherited. Variations in DNA methylation at the *Pb-At5.2* QTL are widespread across Arabidopsis accessions and correlate negatively with variations in the expression of the two genes. Our study highlights that natural, stable and transgenerationally inherited epigenetic variations can play an important role in shaping the resistance to plant pathogens by modulating the expression of immune receptors.

Keywords: methylation, clubroot, *Plasmodiophora brassicae*, *AT5G47260*, *AT5G47280* (ADR1-L3), *ddm1*

Intraspecific diversity in plant immune interactions is associated with a high level of sequence variations at hundreds of NLR (nucleotide-binding and leucine-rich repeat), one of the largest and most rapidly evolving plant gene families (Meyers et al., 2003; Yue et al., 2012; Shao et al., 2016). Based on the nature of their N-terminal domain, NLR have been classified into 4 subclasses: toll/interleukin-1 receptor-type (TIR-NLR or TNL), coil-coiled type (CC-NLR or CNL), RPW8-type CC NLR (CCRPW8 NLR or RNL), and G10-type CC-NLR (CCG10 NLR) (Contreras et al., 2023). Many NLR proteins are involved in the recognition of a small range of effector proteins secreted by specific strains of plant pathogens, potentially triggering the induction of strong plant defense responses that can rapidly stop pathogen invasion (Maekawa et al., 2011; Jones et al., 2016). The catalogue of NLR genes expressed in a given plant genotype thus globally shapes the range of isolate-specific total resistances (incompatible interactions). This general rule has however few exceptions, including the existence of non-NLR driven resistances (Thomas, 1998; Xiao et al., 2001; Larkan et al., 2013) and broad-spectrum NLR-driven resistances (Ernst et al., 2002; Qu et al., 2006).

Effectors may be recognized in different ways: 1/ NLR can monitor the impact of pathogen effectors on their cellular targets 2/ pathogen effectors can be recognized by their direct interaction with one canonical NLR-domain, or alternatively 3/ effectors can be recognized by one non-canonical NLR-domain, called Integrated Decoy (ID) that mimics a protein domain of the effector target (Kourelis and van der Hoorn, 2018). Then, effector-activated CNL assemble into pentameric oligomers called resistosomes, driving a rapid intracellular Ca^{2+} inward flux that triggers downstream defence cellular responses (Förderer et al., 2022). Besides, activated TNL drive similar Ca^{2+}-mediated defence responses by an indirect pathway: assembled into tetrameric oligomers, their TIR domain mediates the biosynthesis of small signaling molecules leading to the downstream assembling of pentameric CCRPW8 NLR-based resistosomes, which mediate Ca^{2+}-mediated defences (Essuman et al., 2022). CCRPW8 NLR thus play a central role in the integration of hub-connected TNL-based non-self-recognition processes and have been therefore called ‘helper NLR’ (Wu et al., 2017).

In contrast to *R*-gene driven resistance, quantitative resistance is polygenic, i.e. it involves allelic variation at several Quantitative Trait Loci (QTL), which collectively contribute to post-invasive partial resistance in compatible plant-pathogen interactions. The nature of the few resistance QTL cloned so far supports the premise that Quantitative Resistance Genes (QRG) are functionally more diverse than *R* genes (Pilet-Nayel et al., 2017; Nelson et al., 2017; Delplace et al., 2022). Among these QRG, however, there are still genes encoding NLR (Hayashi et al., 2010; Fukuoka et al., 2014; Xu et al., 2014; Debieu et al., 2016) and other receptors (Diener and Ausubel, 2005; Hurni et al., 2015) or co-receptors (Huard-Chauveau et al., 2013). Thus, variation in NLR genes (or other non-self-recognition loci) also appears to contribute to variations in basal resistance levels during compatible interactions.

To trigger effective resistance, cellular levels of NLR proteins must reach minimum thresholds. However, high levels of NLR can also lead to autoimmunity drawbacks, including spontaneous HR and retarded plant growth (Li et al., 2015; Lai and Eulgem, 2018). Their abundance is thus tightly controlled by multiple mechanisms, at the transcriptional, post-transcriptional (*i.e.* alternative splicing) and post-translational levels (*i.e.* Ubiquitin-dependent proteolytic regulation) (Zhang and Gassmann, 2007; Li et al., 2015; Lai and Eulgem, 2018). NLR regulation also involves a multitude of epigenetic-related cellular processes, including redundant networks of small RNA (Shivaprasad et al., 2012; Fei et al., 2013; Deng et al., 2018; Huang et al., 2019), histone modifications (Palma et al., 2010; Xia et al., 2013; Zou et al., 2014; Ramirez-Prado et al., 2018), histone-mark dependent-alternative splicing (Tsuchiya and Eulgem, 2013), regulation of chromatin structure and DNA methylation (Li et al., 2010; Deleris et al., 2016). There is increasing evidence that epigenetic processes can play a role in the transitory imprinting of some plant biotic stress responses, at least for a few generations (Molinier et al., 2006; Slaughter et al., 2011; Luna et al., 2012; López Sánchez et al., 2016; López Sánchez et al., 2021; Morán-Diez et al., 2021). It is however not yet clear to which extent the stable transgenerational inheritance of epigenetically regulated gene expression contributes to the natural intraspecific diversity of plant-pathogen interactions.

The few available examples of transgenerational epigenetically controlled traits are mostly found in plant species, where the association between natural or induced differentially methylated regions (DMR) and phenotypic traits were shown stably or (most often) metastably inherited across the generations (Quadrana and Colot, 2016; Furci et al., 2019; Liégard et al., 2019). Such regions, designated as epialleles, can have an impact on relevant agronomic traits: compatibility, accumulation of vitamin E and fruit ripening in tomato, starch metabolism, disease resistance, sex determination in melon and fruit productivity in oil palm (Manning et al., 2006; Martin et al., 2009; Durand et al., 2012; Silveira et al., 2013; Quadrana et al., 2014; Ong-Abdullah et al., 2015; He et al., 2018; Bhat et al., 2020).

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

Here, we report the identification of two adjacent *NLR* genes controlled by a naturally occurring stable epigenetic variation underlying a QTL involved in partial resistance to clubroot in *Arabidopsis*. Clubroot is a root gall disease caused by the telluric biotrophic pathogen *Plasmodiophora brassicae* (Rhizaria), affecting all *Brassicaceae* crops such as oilseed rape, kales and turnips. The infection process involves a primary infection in root hairs for only a few days. Then secondary plasmodia develop in root cortical

cells, causing hyperplasia and hypertrophy that ultimately impair plant water and nutrient uptake. The reference accession Col-0 and Bur-0 are fully susceptible and partially resistant to *P. brassicae* isolate eH, respectively (Alix et al., 2007; Jubault et al., 2008b) (**Supplementary Figure S1**). Four main QTL determine this difference, which act additively. Here, combining fine mapping of the QTL *Pb-At5.2*, which had the strongest effect on resistance, and CRISPR/Cas9 validation, we identified two adjacent *NLR* genes *AT5G47260* and *AT5G47280* both involved in the control of clubroot partial resistance. The expression level of the two genes varies between the susceptible and resistant parents and is linked to the DNA methylation status of the small region including these two genes and a neighboring transposable element (TE) sequence. The methylation status of the two resistance genes is stable over generations and is not associated with any structural variation on the transposon in-between. Epiallelic variation at this locus is frequent among natural *Arabidopsis* accessions and the low methylated state was found correlated with the expression of the two *NLR*-genes as well as with increased quantitative resistance to *P. brassicae* among 126 accessions. We further showed that the RNA independent pathway involving DDM1, MET1, VIM and CMT2/3 maintains the hypermethylated epiallele in the Col-0 clubroot-susceptible accession. Overall, our findings demonstrate that the quantitative resistance to a major root disease affecting *Brassicaceae* is associated, in *Arabidopsis*, with the stable inheritance of a natural epigenetic variation involved in the control of the constitutive expression of a pair of NLR-genes.

Results

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

In previous work, we mapped, using a population of F7 recombinant inbred lines (RIL) between the partially resistant accession Bur-0 and the susceptible Col-0, a QTL (*Pb-At5.2)*, located on chromosome 5 between 67.5 and 71.8 cM, explaining a significant proportion (R^2=20 %) of the resistance (**Fig. 1A**) (Alix et al., 2007; Jubault et al., 2008b). In TAIR10, this interval (between *AT5G46260* and *AT5G47690*) contained 158 annotated sequences including protein coding genes, TE genes, pretRNA and small nuclear RNA. The effect and confidence interval of this QTL was also previously confirmed in Heterogeneous Inbred Family (HIF) lines 10499 and 13499 (Lemarié et al., 2015). Both lines derived

from the RIL 499, which harbored the homozygous Bur-0 (resistance) allele on the QTL *Pb-At1* and *Pb-At5.1*, the homozygous Col-0 (susceptibility) allele on the QTL *Pb-At4* and finally a residual heterozygosity in the *Pb-At5.2* region. The lines 10499 and 13499 inherited, at the QTL *Pb-At5.2*, the homozygous Bur-0 (resistance) allele and Col-0 (susceptibility), respectively (**Fig. 1B-C, Supplementary Text S1)**.

The initial aim of the present work was to fine map *Pb-At5.2*, starting with reciprocal crosses between HIF lines 10499 and 13499. Clubroot symptoms in individuals of the F1 progeny were as severe as in the susceptible parental line HIF 13499, suggesting that the resistance allele Bur-0 was recessive (**Supplementary Figure S2**). The boundaries of the *Pb-At5.2* resistance locus was further refined through several rounds of genotyping and clubroot phenotyping (generations F3 to F5 downstream of crosses 10499/13499) (details are given in **Fig. 1D-F**, **Supplementary Figure S3, Supplementary Text S1, Supplementary Data 1, Supplementary Text S3).** This enabled us to narrow down the confidence interval to 26 kb between the markers CLG4 (19,182,401 bp, in the promoter region of *AT5G47240*), and the marker K64 (19,208,823 bp, in *AT5G47330*), genetic markers being defined using the available *de-novo* genome assembly of Bur-0 (Schneeberger et al., 2011). The comparison of genetic sequences between Bur-0 and Col-0 in this 26 kb region revealed the absence of any structural variation and a low frequency of SNP (**Fig 1F**, details in **Supplementary Text S3**). This region contained eight annotated open reading frames (ORFs), including the three NLR-encoding genes *AT5G47250*/*AT5G47260*/*AT5G47280*, six annotated TE sequences and one lncRNA gene (**Fig. 1F**). The two F5 homozygous progeny lines 1381-2 and 2313-15, harboring the closest recombination events from both sides of the 26 kb interval, (see **Fig. 1E**) also showed partial resistance to a series of additional *P. brassicae* isolates (pathotypes 1, 4 and 7 following the classification of Some et al. (1996) and P1(+), which is representative of the new virulent strains that are emerging in Europe following clubroot resistance breaking of oilseed-rape varieties derived from the cultivar ‘Mendel’ (Zamani-Noor et al., 2022). This highlighted the broad-spectrum of resistance conferred by the Bur-0 allele of *Pb-At5.2* (**Supplementary Figure S4**).

RNA-seq analysis revealed a constitutive expression polymorphism of two NLR genes in the 26kb QTL confidence interval

RNA-seq analysis was carried out on Bur-0 and Col-0 accessions and on recombinant HIF lines 10499 and 13499. Pathogen-induced gene expression patterns markedly differed between genotypes harboring alleles *Pb-At5.2*$_{BUR}$ or *Pb-At5.2*$_{COL}$ (**Supplementary Figure S5, Supplementary Data 2**). Those regulations were consistent with our previously published studies *i.e.* a role for camalexin biosynthesis and SA-mediated responses in *Pb-At5.2*$_{BUR}$-mediated resistance and a role for JA-driven induction of *ARGAH2* in *Pb-At5.2*$_{COL}$-mediated basal resistance (details in **Supplementary Text S2**). We then focused on the eight ORFs in *Pb-At5.2*. Genes *AT5G47290* and *AT5G47300* showed no expression and genes *AT5G47240*, *AT5G47250*, *AT5G47310* and *AT5G47320* showed similar expression in all four accessions (**Fig. 2A**). In the 26kb interval, only two genes *AT5G47260* and *AT5G47280*, both encoding proteins belonging to the family of non-TIR-NLR genes, are differentially expressed between resistant and susceptible accessions: these two genes are constitutively expressed in Bur-0 and 10499 roots (i.e. harboring Bur-0 allele) but their expression was undetectable in Col-0 and 13499 (i.e. with Col-0 allele) (**Fig. 2A**). *AT5G47280* encodes ADR1-L3, an NBS-LRR protein related to the small family of ADR1-type RNL (although the encoded protein lacks the N-ter RPW8 domain). *AT5G47260*, encodes a CC-NBS-LRR-X protein, displaying an integrated decoy-like C-ter extension domain (**Fig. 2B**) homologous with members of the IAN family of AIG1(=*AvrRpt2-Induced Gene1*, *AT1G33960*)-related proteins (Martin et al., 2023). These two adjacent genes are separated by an helitron *AT5TE69050* (**Fig. 1f** and **Fig. 3a**). *AT5G47280* displayed no SNP and *AT5G47260* displayed only one single non-synonymous SNP (**Supplementary Figures S7 and S8**, **Supplementary Text S3**). There was also no sequence variation on the helitron *AT5TE69050* located in between the two genes.

Validation by CRIPSR/Cas9 of the role of the two NLR genes AT5G47260 and AT5G47280 in clubroot resistance

Based on the contrasted expression of those two genes between Bur-0 and Col-0, their functional significance in clubroot resistance was then addressed by generating knockout lines via CRISPR

(clustered regularly interspaced short palindromic repeats)/Cas9 (CRISPR-associated protein 9) technology targeting the region coding for the NB-ARC domain of both genes with two single-guide RNA (sgRNA) in both resistant Bur-0 and HIF 10499 (i.e. with Bur-0 allele) accessions. The CRISPR/Cas9 generated mutations in the *AT5G47260* and *AT5G47280* genes gave rise to premature stop codons in most mutant lines, and to the substitution of a stretch of 14 aa in the ARC1 domain for the line 160-2 (**Fig. 2B, Supplementary Figure S6 to S8** and **Supplementary Text S4**). For the following experiments, we kept 3 mutants from separate transformation events for each gene and each background (Bur-0 or HIF 10499). For each gene, one mutant without T-DNA insertion was obtained in the background Bur-0. The *AT5G47260* and *AT5G47280* CRISPR-KO mutants were then evaluated for clubroot resistance in a complete randomised design. For both genes, clubroot symptoms in all lines edited in the Bur-0 genetic background were significantly higher than in the wild type resistant Bur-0 accession and reached those found in the susceptible accession Col-0 (**Fig. 2B**), demonstrating the involvement of both genes *AT5G47260* and *AT5G47280* in clubroot resistance. Similar results were found with the CRISPR-edited lines in the 10499 HIF genetic background (**Supplementary Figure S9**).

Expression polymorphism of both NLR genes is associated with stably inherited methylation variation

To understand why these two NLR genes *AT5G47260* and *AT5G47280* were differentially expressed in Bur-0 and Col-0, even in the absence of any sequence variation within the putative promoter regions of these genes, we analyzed the DNA methylation level at the region in these two accessions using public methylome data (Kawakatsu et al., 2016). The genomic interval between 19,188,411 and 19,196,559, which includes, in Col-0 and Bur-0, the two genes *AT5G47260* and *AT5G47280* and the transposon *AT5TE69050* in-between, was hypermethylated and hypomethylated in Col-0 and Bur-0, respectively (**Fig. 3A**). This contrasting methylation state was experimentally confirmed using DNA extracted from infected and non-infected roots of Col-0 and Bur-0 plants (**Fig. 3B**) and CHOP qPCR. These DNA methylation differences were also found between the progeny HIF lines 10499 and 13499 and in the pair of HIF-derived homozygous near-isogenic lines 1381-2/2313-15 (**Fig. 3B**), thus indicating that they are stably inherited independently of any DNA sequence polymorphism outside the locus. Moreover,

the 'Col-like' hypermethylation of *AT5G47260* and *AT5G47280* was systematically associated with a low expression of the two NLR genes and a lower level of partial resistance to *P. brassicae* infection. To further investigate the inheritance of this epiallelic variation and its penetrance on gene expression and clubroot resistance, we then investigated two series of 100 individual plants, corresponding to the progenies derived from selfing the heterozygous 2509 and 1381 lines (harbouring heterozygosity at the locus). The evaluation of plant disease for each individual plant in the two progenies indicated a 3:1 mendelian segregation of the partial resistance phenotype. Clubroot symptoms in individuals with only one Bur-0 resistance allele were the same as in individuals with the two susceptible Col-0 allele (**Fig. 4A**). In clubroot-inoculated roots of each of those individual plants from the 2509 progeny, the methylation state of the *Pb-At5.2* region was monitored by CHOP qPCR on *AT5G47260*. In addition, the SNP allele status at *Pb-At5.2* was investigated for each individual plant (details of markers are given in **Supplementary Data 1**). Heterozygous Bur/Col individuals displayed intermediate parental methylation and expression values (**Fig. 4B-D**), thus providing a molecular explanation for the recessivity of the Bur-0 resistance allele. Altogether, these results suggested a link between partial resistance to *P. brassicae* and a stably inherited epiallelic variation at *Pb-At5.2*, which controls the expression of two NLR genes.

The Bur-like hypo-methylated epiallele is well represented among Arabidopsis accessions and contributes to reduce clubroot symptoms

To assess the relative contribution of DNA sequence and DNA methylation changes at *Pb-At5.2* to clubroot resistance, we investigated the natural allelic and epiallelic diversity across *Arabidopsis* accessions. We took advantage of recently published Illumina short genome sequence reads obtained from 1135 Arabidopsis accessions (1001 Genomes Consortium, 2016) to document the species-wide molecular diversity of the *Pb-At5.2* genomic region. Based on quantitative horizontal and vertical coverage of short-reads aligned to the Col-0 reference genome sequences, we identified two discrete groups of accessions. One of such groups, containing 401 accessions, was characterized by high vertical and horizontal coverage (>0.75) and included the reference accession Col-0 as well as the partially

clubroot resistant Bur-0 (**Figure 5A**; detailed list of genotypes in **Supplementary Data 3, sheet 1**). Conversely, the remaining 734 accessions present diverse structural rearrangements, principally long deletions that translates in poor horizontal and vertical coverage compared to the reference Col-0 genome. Closer examination of coverage plots for the 401 accessions with Col-0/Bur-0-like revealed a uniform haplotype structure, which was present at high frequency at the species level (Minor Allele Frequency MAF ~0.37). Nonetheless, the haplotype frequency varied depending on geographic groups, ranging from 52.7 % in Spain to 17.7 % in Asia (**Supplementary Figure S10**). We then analyzed DNA methylation levels in 287 accessions belonging to the 401 accessions containing the Col-0/Bur-0-like *Pb-At5.2* and for which public bisulfite data is available (Kawakatsu et al., 2016). Based on this data, we could distinguish a group of 228 accessions showing hypomethylation of *Pb-At5.2*, including Bur-0, and another group of 59 accessions, which includes Col-0, that displays hypermethylation (**Fig. 5B**). The prevalence of accessions showing the Col-like (epi)haplotype varied considerably depending on geographic origin, ranging from 1.8 % in Spain to 16.8 % in Central Europe (**Supplementary Data 3 and Supplementary Figure S10**). Consistent with a causal role for DNA methylation in the transcriptional regulation *AT5G47260* and *AT5G47280*, reanalysis of publicly available RNA-seq data revealed a pronounced negative correlation between methylation level and *AT5G47260* and *AT5G47280* gene expression (**Fig. 5C**). These results were further validated in infected roots of 20 natural accessions (**Fig. 5D**).

Both Col-like and Bur-like epialleles were significantly represented among the natural accessions, thus offering interesting genetic material to determine the actual contribution of DNA sequence and DNA methylation in the control of clubroot partial resistance. One hundred and twenty-six accessions, selected for their methylation levels at the *AT5G47260-AT5G47280* region in data from Kawakatsu et al. (2016), including 42 accessions with the Col-like epiallele and 85 accessions with the Bur-like epiallele, were assessed for their resistance to *P. brassicae* isolate eH. While no DNA sequence polymorphism within *Pb-At5.2* shows association with clubroot resistance (**Supplementary Data 4**), the low DNA methylation state of the *AT5G47260*/*AT5G47280* locus was significantly

associated with enhanced resistance levels (**Fig. 6**). Altogether, the results corroborate and extend the conclusions obtained by fine mapping of *Pb-At5.2* and provide strong evidence that natural epiallelic variation contribute to the quantitative differences observed in clubroot resistance among Arabidopsis accessions.

Pb-At5.2 epivariation is independent of cis- genetic variations

At the locus *Pb-At5.2*, the transposon *AT5TE69050* was present in both parental genotypes, with no sequence variation that might have been the primary cause of the variation of DNA methylation on the two adjacent genes. Analysis of 34 out of the 287 accessions with the Col-0/Bur-0-like haplotype, did not reveal the presence/absence of TE insertion variant within the 26kb *Pb-At5.2* region (Quadrana et al., 2016; Stuart et al.), with the exception of an helitron insertion in the accession NFA-10. Moreover, 18 and 16 of these accessions displayed hypermethylated and hypomethylated epialleles, respectively, indicating that variation in DNA methylation is not associated with TE presence/absence variants. In addition, the cis-nucleotide polymorphism located within the coding sequence of *AT5G47260* and detected in Bur-0 was absent in at least five other accessions sharing the hypomethylated epiallele (**Supplementary Figure S11**), indicating that the hypomethylated state of *Pb-At5.2* is not correlated with any specific DNA sequence polymorphism at the locus.

The hypermethylated epigenetic variant is maintained by the RNA-independent pathway

Analysis of sRNAs identified in Col-0 (Stroud et al., 2013) revealed that the *AT5G47260*/*AT5G47280* region is targeted mostly by 24-nt sRNA, which prompted us to generate sRNA profiles from non-inoculated roots of Col-0 and Bur-0 and from roots inoculated with *P. brassicae* isolate eH 7 and 14 days after inoculation. Consistent with the pattern of DNA methylation observed before, we found high levels of sRNAs only in Col-0 (**Fig. 7A**). To explore further the mechanisms involved in the maintenance of methylation at this locus, we made use of publicly available methylomes of Col-0 mutant plants defective in one or several DNA methylation pathways (Stroud et al., 2013). Despite the high levels of sRNAs detected over the *AT5G47260*/*AT5G47280* region, mutations affecting the RdDM pathway did

not influence its DNA methylation level (**Supplementary Figure S12**). Conversely, DNA methylation was largely lost in mutants defective in sRNA-independent maintenance of DNA methylation, i.e. *ddm1*, *cmt2/3*, *met1*, and *suvh456*. (**Fig. 7B)**. These results raise the question of the role of the sRNA targeted to the Col-like hypermethylated region, whereas the methylation maintenance solely depends on the RNA-independent pathway.

Discussion

To date, only a very small number of resistance QTL have been characterized at the molecular level (Delplace et al., 2022). Detection and fine mapping of resistance QTL is typically challenging not only because of the difficulties associated with measuring small variations in partial resistance in large numbers of individual progeny, but also because resistance QTL can be sensitive to environmental changes (Laperche et al., 2017; Aoun et al., 2017; Aigu et al., 2018). However, technical issues may have only been part of the problem. Recent developments in the field of epigenetics suggest that the detection of some inherited resistance factors could be missed by classic genetic approaches based solely on DNA sequence variation. In the present work, genome-wide association study failed at identifying any nucleotide variation in the 26 kb interval of *Pb-At5.2* associated with clubroot response. By contrast, clubroot resistance was clearly related to the epigenetic variation at two NLR-coding genes in this interval. This work thus uncovers for the first time an epigenetically driven expression polymorphism that makes a substantial contribution to the natural diversity of the plant immune response.

Many examples of epialleles are metastable, *i.e.* they can be reversed by stochastic or non-identified factors (Weigel and Colot, 2012). Stability over multiple generations is a primary concern from both evolutionary and breeding perspectives. The epiallele described here seems to be extremely stable, as it was robustly detected in all our previous QTL investigations in *Arabidopsis*. This included two independent segregating progenies derived from Bur-0 and Col-0 (Jubault et al., 2008a) and additional studies with HIF lines 10499/13499 (Lemarié et al., 2015; Gravot et al., 2016). The high level of methylation and absence of *AT5G47260* and *AT5G47280* expression in Col-0 was found in a series of publicly available data obtained in different laboratories, with a diversity of plant tissues and conditions

(Winter et al., 2007; Stroud et al., 2013; Klepikova et al., 2016; 1001 Genomes Consortium, 2016; Kawakatsu et al., 2016). It was also confirmed by our own data generated on inoculated roots, non-inoculated roots and leaf samples. Finally, this methylation pattern was also robustly found in multiple replicates of individual plants. Thus, *Pb-At5.2* can therefore be classified as a stable epiallele without reservation.

There has only been one report of a plant disease resistance caused by an inherited methylation variant affecting the expression of a resistance-related gene (Nishimura et al., 2017). In this study, a stable expression polymorphism (between Ler-0 and Ag-0 accessions) on the TIR-only encoding gene *RBA1* (*AT1G47370*) affected ETI-responses to the *Pseudomonas syringae* effector *hopBA1*. This expression polymorphism was linked to the nearby presence/absence of a TE sequence in the promoter region of the gene and to *MET1*-dependent DNA methylation variation. However, because DNA methylation was reversed when the TE sequence was segregated away (Figure S2 in Nishimura et al. (2017), DNA methylation variation is not "epigenetic" as it is an obligate consequence of sequence variation (i.e. presence/absence of the TE sequence). In the present study, we showed that DNA methylation variation in the region between *AT5G47260* and *AT5G47280*, including the TE sequence *AT5TE69050*, is not linked to any nucleotide/structural variation at the locus or elsewhere in the genome. Thus, *Pb-At5.2*$_{COL}$ and *Pb-At5.2*$_{BUR}$ can be considered as 'pure epialleles', as defined by Richards (2006).

From available genomic and epigenomic data from the '1001 genomes project', it can be extrapolated that the Bur-like clubroot resistance epiallele is present in about half of the accessions from the 'Relict', 'Spain', and 'Italy/Balkans/Caucasus' groups and 39 % of the accessions from the 'North Sweden' group (**Supplementary Data 3**). In contrast, the Bur-like epiallele is likely (taking into account missing methylation data) present at about 10 % in the 'Germany' group. On the other hand, the clubroot susceptibility of the Col-like epiallele was absent from the accessions in the 'Relict' and 'North Sweden' groups but reached at least 16.8 % in the 'Central Europe' group. This sharp geographical structuration suggests that both epialleles confer fitness gains depending on environmental contexts. However, it does not appear to be obviously related to the distribution of clubroot incidence in *Brassica* cultures (usually

low in the warm southern European regions). Keeping in mind that NLR can detect unrelated effectors from distinct microbial species (Narusaka et al., 2009) and echoing previous work (Karasov et al., 2014), we hypothesize that the maintenance of this epivariation in natural populations may reflect additional roles that *Pb-At5.2* plays against other plant pathogens (besides the control of clubroot infection).

AT5G47280 has been annotated as *ADR1-Like 3*, based on its phylogenetical relationship with the small family of the helper CC_{RPW8}-NLR including *ADR1*, *ADR1-L1* and *ADR1-L2* (Saile et al., 2020; Saile et al., 2021). However, the absence of *RPW8* in *ADR1-L3* and the absence of expression of this gene in Col-0 has questioned its actual role in plant immunity. *AT5G47260* and *AT5G47280* belong to a small heterogeneous cluster of three non-TIR-NLR, including also *AT5G47250*. This small cluster is located on chromosome 5, not far from the largest NLR-hot spot in the *Arabidopsis* genome (Meyers et al., 2003). None of these three genes have been previously shown to be involved in plant pathogen interactions. We showed in this work that the expression of both *AT5G47260 and AT5G47280* is necessary to the partial resistance to *P. brassicae.* There are a few examples of tandem NLR genes coding for pairs of proteins that function as heterodimers (Cesari et al., 2014; Williams et al., 2014; Saucet et al., 2015). Similarly, the proteins encoded by these two epigenetically joint-regulated genes may function together in the control of cell defense responses during clubroot infection. Although the underlying molecular mechanisms are not known, the canonical example of the TIR-NLR heterodimer *RRS1*/*RPS4* corresponds to a recessive resistance locus, similar to *Pb-At5.2*.

In Arabidopsis, DNA methylation is widely distributed in both the promoters and the bodies of most NB-LRR-encoding genes (Kong et al., 2018; Kong et al., 2020), predominantly in the CG sequence context. This suggests that plant genomes contain series of functional resistance genes, whose possible roles in biotic interactions are locked by epigenetic processes. This hypothesis is also supported by our previous study, where we demonstrated that the *ddm1*-triggered hypomethylation at different genomic loci resulted in unlocking genetic factors ultimately exerting a significant control over the development of clubroot symptoms (Liégard et al., 2019). It would now be interesting to carry out a careful genome wide analysis of methylation profiles of all NLR-genes among *Arabidopsis* accessions, which would

take into account the structural diversity of all these individual genes (supported by additional targeted resequencing of NLR-loci). The intraspecific diversity of the methylation patterns of NLR and RLK/RLP genes in plants, their heritability, and their consequences on plant biotic interactions, may also deserve further attention in future studies.

Methods

Plant materials and growth conditions. Heterogeneous Inbred Family (HIF) lines 10499 and 13499 and their parental accessions Col-0 (186AV) and Bur-0 (172AV) were obtained from the Institute Jean Pierre Bourgin (INRAE Versailles, France). *Arabidopsis thaliana* accessions were all purchased from the Nottingham Stock Center. The panel of 126 accessions was selected according to their methylation level at the region of interest (Kawakatsu et al., 2016). All accessions and in-house generated recombinant lines used in this study are listed in **Supplementary Data 1** and **Supplementary Data 3**. Seed germination was synchronized by placing seeds on wet blotting paper in Petri dishes for two days at 4 °C. Seeds were sown individually in pots (four cm diameter) containing a sterilized mix composed of 2/3 compost and 1/3 vermiculite. Growth chamber control conditions of 16 h light (110µmol $m^{-2}s^{-1}$) at 20 °C and 8 h dark at 18 °C were used to grow plants. The 126 *Arabidopsis* accessions and HIF were challenged by *P. brassicae* in two biological replicates in a completely randomised block designs (with two blocks per replicate, each block consisting in 6 plants per genotype). The *Arabidopsis* accessions Col-0, Bur-0 and the HIF 10499, 13499, 1381-2 and 2313-15 used in RNA-seq and smRNA-seq approaches were assessed when infected with *P. brassicae* or in uninfected condition in three randomised blocks. The CRISPR/Cas9 edited lines and corresponding wild-type lines were challenged by *P. brassicae* in 8 replicates in a completely randomized block designs (each replicate consisting in 10-12 plants per genotype). Almost all clubroot tests were performed with the isolate eH of *P. brassicae* described by Fähling (2003) which belongs to the most virulent pathotype P1. The resistance spectrum of *Pb-At5.2* was also assessed with a series of additional isolates Pb137-522, Ms6, K92-16 and P1(+). For every isolate used in this study, the pathotype was validated in every experiment using the differential host set according to Some et al. (1996), also including two genotypes of *B. oleracea* ssp. *acephala* C10 and CB151. One ml of resting spore suspension (10^7 spores.ml^{-1}) prepared according to

399 Manzanares-Dauleux et al. (2000) was used for pathogen inoculation 10 days (d) after germination
400 (stage 1.04 (Boyes et al., 2001)). This inoculum was applied onto the crown of each seedling.

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

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

409 ***RNA isolation, mRNA Sequencing and differentially gene expression determination.*** Total RNA was
410 extracted from frozen and lyophilized roots (collected 14 days after inoculation) using the mirVana™
411 miRNA Isolation Kit (Invitrogen) according to the manufacturer's instructions. After extraction, the
412 RNA were quantified using a NanoDrop ND-1000 technologies and their quality was controlled using
413 the RNA 6000 assay kit total RNA (Agilent). Samples with a RIN greater or equal to seven were used
414 for sequencing. cDNA sequencing library construction and the sequencing were performed by the NGS
415 platform at the Marie Curie Institute of Paris. Each library was paired-end sequenced on an Illumina
416 HiSeq 2500 technology. Reads were aligned to the TAIR10 genome annotation and assembly of Col-0
417 *Arabidopsis thaliana* genome concatenated with the *P. brassicae* genome using STAR software (Dobin
418 et al., 2013) (Version 2.5.3.a). Alignment conditions were selected according to the *Arabidopsis*
419 genome. A maximum of five multiple read alignments were accepted and for each alignment no more
420 than three mismatches were allowed. The resulting BAM files were used to determine read counts using
421 the function Counts of Feature Counts software (Version 1.4.6) and the TAIR10 gff of *Arabidopsis*
422 concatenated with the gff of *P. brassicae*. Differentially expressed genes were determined using the
423 package EdgeR (Robinson et al., 2010) in R software (Team, 2013) (version 3.3.0). Raw counts obtained
424 as described previously were used as entry data in EdgeR. Once CPM (Counts Per Million) was

425 determined, only genes with at least one CPM in three samples were retained. Expression signals were
426 normalized using the TMM method (Trimmed Mean of M-values) of the CalcNormFactors function of
427 Edge R. Finally, the differentially expressed genes were determined using the decideTests function of
428 Edge R using one minimum fold change between -1.5 and 1.5.

429 ***CRISPR-Cas9 constructs and plant transformation.*** Two guide sequences were designed using the
430 CRISPOR software (Concordet and Haeussler, 2018) for each targeted gene (i.e. *AT5G47260* and
431 *AT5G47280*), taking care of selecting sequences with a very high specificity score (**Supplementary**
432 **Data 5**). Guide sequences were ordered as oligonucleotides (IDT) and cloned downstream of the
433 *Arabidopsis* U6-26 promoter and upstream of an enhanced single guide RNA scaffold, as previously
434 reported (Chauvin et al., 2021), producing individual guide modules. Assembly of guide modules for
435 single gene was performed using PCR amplification with specific primers followed by classical
436 restriction/ligation cloning (**Supplementary Data 5 sheet 1**). Guide assemblies were then cloned
437 through a Gateway LR reaction (Thermo Fisher Scientific) into the pDe-Cas9 backbone (Fauser et al.,
438 2014) harboring a *nptII* resistance cassette (Chauvin et al., 2021), resulting in two binary plasmids for
439 CRISPR-mediated targeting of *AT5G47260* (pDe-Cas9_T79-80) and *AT5G47280* (pDe-Cas9_T81-82),
440 respectively (**Supplementary Data 5 sheet 1**). All the constructs were checked by Sanger sequencing.
441 The resulting plasmids were then transferred into *A. tumefaciens* C58/GV3101*::pMP90* strain by heat
442 shock, then used for the transformation of Arabidopsis plants *via* the floral dip method (Clough and
443 Bent, 1998). Transgenic plants were screened on solid plates with half-strength MS medium containing
444 50 mg l−1 kanamycin (Yeasen, cat. no. 60206ES10). A first screening was performed on the T1
445 generation, using PCR and sequencing, to identify plants harbouring mutations in *AT5G47260* or
446 *AT5G47280* (primer pairs listed in **Supplementary Data 5 sheet 2)**. A second round of screening
447 allowed the identification of T2 plants homozygous for the mutations and free from CRISPR/Cas9
448 cassette T-DNA. T3 seedlings were used for the evaluation of the resistance to *P. brassicae*.

449 ***GWAS analyses.*** A conventional GWA study on GA/LA data from 126 accessions was performed with
450 easyGWAS (Grimm et al., 2017) (https://easygwas.ethz.ch/). Association analysis was performed with

EMMAX (Kang et al., 2010) using 1 806 554 SNP with a MAF >0.05, after correction for population structure by including the three first principal components in the additive model.

Small RNA isolation, sequencing, clusterization and differential presence determination. Small RNA was extracted from frozen and lyophilized roots (collected 14 days after inoculation) using the mirVana™ miRNA Isolation Kit (Invitrogen) according to the manufacturer's instructions. After extraction, the small RNA were quantified using a NanoDrop ND-1000 and quality controlled using the Small RNA assay kit (Agilent). Samples with a RIN greater or equal to seven were used for sequencing. Construction of cDNA sequencing libraries and the sequencing were performed by the NGS platform of Marie Curie Institute of Paris. For each sample, single-end (50 bp) sequencing was carried out using Illumina HiSeq 2500 technology. Reads were aligned to the TAIR10 genome annotation and assembly of Col-0 *Arabidopsis thaliana* genome concatenated with *P. brassicae* genome using STAR software (Dobin et al., 2013) (Version 2.5.3.a), counts and clustered using Shortstack software (Axtell, 2013). The presence of differentially expressed sRNA was determined using the EdgeR (Robinson et al., 2010) package in R software (Team, 2013) (version 3.3.0). Raw counts obtained as described previously were used as entry data in EdgeR. Once CPM (Counts Per Million) were determined, only genes with at least one CPM in three samples were retained. Expression signals were normalized using the TMM method (Trimmed Mean of M-values) of the CalcNormFactors function of Edge R. Finally, the differentially expressed sRNA were determined using the decideTests function of Edge R using one minimum fold change between -1.5 and 1.5.

RNA isolation and RT-qPCR analysis. Total RNA was extracted from the lyophilised root of accessions and HIF 21 days post infection using the trizol extraction protocol. Samples with residual traces of DNA were treated with DNAse (Promega ref M6 10A). Before reverse transcription of RNA to cDNA by the SuperScript II (Invitrogen), the RNA quality was verified by agarose gel electrophoresis. RT-qPCR was carried out in a LightCycler® 480 thermocycler (Roche) on cDNA obtained as described above. Gene expression was normalized using as reference two *Arabidopsis* genes defined as stable during infection

using RNA-seq data (*AT1G54610*, *AT5G38470*) following Pfaffl's method (Pfaffl, 2001). Primer sets were designed for each gene and are listed in **Supplementary Table S1**.

CHOP PCR and qPCR assays. Gene methylation profiles were investigated using the enzyme McrBC (M0272L, BioLabs®)(Zhang et al., 2014). Forty ng of DNA was incubated with 0.5 µL BSA (20 mg/mL), 0.5 µL GTP (20 mM), 5 µL NEB (10X) and 0.2 µL McrBC (10000 U/mL). For CHOP PCR and qPCR, 2 ng of digested and undigested DNA was used. For the CHOP PCR the temperature conditions were adjusted according to primer design and 35 amplification cycles were used. To determine the methylation state of the targeted region, each sample was digested or not (control) with McrBC before amplification. For the CHOP qPCR, the temperature conditions were adjusted according to primer design and 30 amplification cycles were used. Methylation levels of the target region were calculated as the percentage of molecules lost through McrBC digestion according to Silveira et al. (2013) with the formula: (1-(2-(Ct digested sample - Ct undigested sample)))*100. The percentage of DNA methylation for *AT5G13440* and *AT5G47400*, was calculated in all CHOP qPCR as controls. *AT5G13440* and *AT5G47400* genes were selected from 1001 genome data as hypomethylated and hypermethylated respectively in most Arabidopsis accessions and their expression did not vary during clubroot infection. Primer sets designed for each gene are listed in **Supplementary Table S1**.

Published data. The DNAseq data, RNAseq data, variant sequences and bisulfite data of the natural accessions studied were obtained from previous studies (1001 Genomes Consortium, 2016; Kawakatsu et al., 2016) archived at the NCBI SRA number SRP056687 and the NCBI GEO reference: GSE43857, GSE80744. The bisulfite data and small RNA data of *Arabidopsis* mutants studied were obtained from previous study (Stroud et al., 2013) archived at the NCBI GEO reference: GSE39901.

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

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

Acknowledgments

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

Author contributions

BL, AG, MMD and MJ designed and conducted the experiments. CL and AG carried out the fine mapping. CL, JL, JB, TB, BL, AG, MJ performed the phenotyping and sampling. BL, AG, JB, JL and MJ carried out epigenetic and gene expression studies. FV designed and performed the CRISPR-Cas9 constructs and CL and TB carried out the transformation, selection and handling of edited plants. BL, YA, LQ and VC conducted bioinformatics analyses. AG, BL, LQ, FV, VC, MMD and MJ participated in drafting and revisions of the manuscript.

Competing Interests statement

The authors declare no competing interests.

References

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

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

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

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

Axtell, M. J. (2013). ShortStack: comprehensive annotation and quantification of small RNA genes. *Rna* Advance Access published 2013.

Bhat, R. S., Rockey, J., Shirasawa, K., Tilak, I. S., Brijesh Patil, M. P., and Reddy Lachagari, V. B. (2020). DNA methylation and expression analyses reveal epialleles for the foliar disease resistance genes in peanut (Arachis hypogaea L.). *BMC Res. Notes* **13**:20.

Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., and Görlach, J. (2001). Growth stage–based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. *Plant Cell* **13**:1499–1510.

Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T., and Dodds, P. N. (2014). A novel conserved mechanism for plant NLR protein pairs: the "integrated decoy" hypothesis. *Front. Plant Sci.* **5**:606.

Chauvin, L., Sevestre, F., Lukan, T., Nogué, F., Gallois, J.-L., Chauvin, J.-E., and Veillet, F. (2021). Gene Editing in Potato Using CRISPR-Cas9 Technology. *Methods Mol. Biol. Clifton NJ* **2354**:331–351.

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium -mediated transformation of Arabidopsis thaliana. *Plant J.* **16**:735–743.

Concordet, J.-P., and Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res.* **46**:W242–W245.

Contreras, M. P., Pai, H., Tumtas, Y., Duggan, C., Yuen, E. L. H., Cruces, A. V., Kourelis, J., Ahn, H.-K., Lee, K.-T., Wu, C.-H., et al. (2023). Sensor NLR immune proteins activate oligomerization of their NRC helpers in response to plant pathogens. *EMBO J.* **42**:e111519.

Cuerda-Gil, D., and Slotkin, R. K. (2016). Non-canonical RNA-directed DNA methylation. *Nat. Plants* **2**:1–8.

Debieu, M., Huard-Chauveau, C., Genissel, A., Roux, F., and Roby, D. (2016). Quantitative disease resistance to the bacterial pathogen X anthomonas campestris involves an A rabidopsis immune receptor pair and a gene of unknown function. *Mol. Plant Pathol.* **17**:510–520.

Deleris, A., Halter, T., and Navarro, L. (2016). DNA Methylation and Demethylation in Plant Immunity. *Annu. Rev. Phytopathol.* **54**:579–603.

Delplace, F., Huard-Chauveau, C., Berthomé, R., and Roby, D. (2022). Network organization of the plant immune system: from pathogen perception to robust defense induction. *Plant J.* **109**:447–470.

Deng, Y., Liu, M., Li, X., and Li, F. (2018). microRNA-mediated R gene regulation: molecular scabbards for double-edged swords. *Sci. China Life Sci.* Advance Access published 2018.

Diener, A. C., and Ausubel, F. M. (2005). RESISTANCE TO FUSARIUM OXYSPORUM 1, a Dominant Arabidopsis Disease-Resistance Gene, Is Not Race Specific. *Genetics* **171**:305.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21.

Durand, S., Bouché, N., Perez Strand, E., Loudet, O., and Camilleri, C. (2012). Rapid Establishment of Genetic Incompatibility through Natural Epigenetic Variation. *Curr. Biol.* **22**:326–331.

Ernst, K., Kumar, A., Kriseleit, D., Kloos, D.-U., Phillips, M. S., and Ganal, M. W. (2002). The broad-spectrum potato cyst nematode resistance gene (Hero) from tomato is the only member of a large gene family of NBS-LRR genes with an unusual amino acid repeat in the LRR region. *Plant J.* **31**:127–136.

Essuman, K., Milbrandt, J., Dangl, J. L., and Nishimura, M. T. (2022). Shared TIR enzymatic functions regulate cell death and immunity across the tree of life. *Science* **377**:eabo0001.

Fähling, M., Graf, H., and Siemens, J. (2003). Pathotype Separation of Plasmodiophora brassicae by the Host Plant. *J. Phytopathol.* **151**:425–430.

Fauser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. *Plant J. Cell Mol. Biol.* **79**:348–359.

Fei, Q., Xia, R., and Meyers, B. C. (2013). Phased, Secondary, Small Interfering RNAs in Posttranscriptional Regulatory Networks. *Plant Cell* **25**:2400.

Förderer, A., Yu, D., Li, E., and Chai, J. (2022). Resistosomes at the interface of pathogens and plants. *Curr. Opin. Plant Biol.* **67**:102212.

Fukuoka, S., Yamamoto, S.-I., Mizobuchi, R., Yamanouchi, U., Ono, K., Kitazawa, N., Yasuda, N., Fujita, Y., Nguyen, T. T. T., and Koizumi, S. (2014). Multiple functional polymorphisms in a single disease resistance gene in rice enhance durable resistance to blast. *Sci. Rep.* **4**:4550.

Furci, L., Jain, R., Stassen, J., Berkowitz, O., Whelan, J., Roquis, D., Baillet, V., Colot, V., Johannes, F., and Ton, J. (2019). Identification and characterisation of hypomethylated DNA loci controlling quantitative resistance in Arabidopsis. *Elife* **8**:e40655.

Gravot, A., Grillet, L., Wagner, G., Jubault, M., Lariagon, C., Baron, C., Deleu, C., Delourme, R., Bouchereau, A., and Manzanares-Dauleux, M. J. (2011). Genetic and physiological analysis of the relationship between partial resistance to clubroot and tolerance to trehalose in Arabidopsis thaliana. *New Phytol.* **191**:1083–1094.

Gravot, A., Richard, G., Lime, T., Lemarié, S., Jubault, M., Lariagon, C., Lemoine, J., Vicente, J., Robert-Seilaniantz, A., Holdsworth, M. J., et al. (2016). Hypoxia response in Arabidopsis roots infected by Plasmodiophora brassicae supports the development of clubroot. *BMC Plant Biol.* **16**:251.

Grimm, D. G., Roqueiro, D., Salomé, P. A., Kleeberger, S., Greshake, B., Zhu, W., Liu, C., Lippert, C., Stegle, O., Schölkopf, B., et al. (2017). easyGWAS: A Cloud-Based Platform for Comparing the Results of Genome-Wide Association Studies. *Plant Cell* **29**:5–19.

Hayashi, N., Inoue, H., Kato, T., Funao, T., Shirota, M., Shimizu, T., Kanamori, H., Yamane, H., Hayano-Saito, Y., and Matsumoto, T. (2010). Durable panicle blast-resistance gene Pb1 encodes an atypical CC-NBS-LRR protein and was generated by acquiring a promoter through local genome duplication. *Plant J.* **64**:498–510.

He, L., Wu, W., Zinta, G., Yang, L., Wang, D., Liu, R., Zhang, H., Zheng, Z., Huang, H., Zhang, Q., et al. (2018). A naturally occurring epiallele associates with leaf senescence and local climate adaptation in Arabidopsis accessions. *Nat. Commun.* **9**:460.

Henderson, I. R., and Jacobsen, S. E. (2007). Epigenetic inheritance in plants. *Nature* **447**:418–424.

Huang, C.-Y., Wang, H., Hu, P., Hamby, R., and Jin, H. (2019). Small RNAs – Big Players in Plant-Microbe Interactions. *Cell Host Microbe* **26**:173–182.

Huard-Chauveau, C., Perchepied, L., Debieu, M., Rivas, S., Kroj, T., Kars, I., Bergelson, J., Roux, F., and Roby, D. (2013). An atypical kinase under balancing selection confers broad-spectrum disease resistance in Arabidopsis. *PLoS Genet.* **9**:e1003766.

Hurni, S., Scheuermann, D., Krattinger, S. G., Kessel, B., Wicker, T., Herren, G., Fitze, M. N., Breen, J., Presterl, T., and Ouzunova, M. (2015). The maize disease resistance gene Htn1 against northern corn leaf blight encodes a wall-associated receptor-like kinase. *Proc. Natl. Acad. Sci.* Advance Access published 2015.

Jones, J. D. G., Vance, R. E., and Dangl, J. L. (2016). Intracellular innate immune surveillance devices in plants and animals. *Science* **354**.

Jubault, M., Hamon, C., Gravot, A., Lariagon, C., Delourme, R., Bouchereau, A., and Manzanares-Dauleux, M. J. (2008a). Differential Regulation of Root Arginine Catabolism and Polyamine Metabolism in Clubroot-Susceptible and Partially Resistant Arabidopsis Genotypes. *Plant Physiol.* **146**:2008–2019.

Jubault, M., Lariagon, C., Simon, M., Delourme, R., and Manzanares-Dauleux, M. J. (2008b). Identification of quantitative trait loci controlling partial clubroot resistance in new mapping populations of Arabidopsis thaliana. *Theor. Appl. Genet.* **117**:191–202.

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-Y., Freimer, N. B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**:348–354.

Karasov, T. L., Kniskern, J. M., Gao, L., DeYoung, B. J., Ding, J., Dubiella, U., Lastra, R. O., Nallu, S., Roux, F., and Innes, R. W. (2014). The long-term maintenance of a resistance polymorphism through diffuse interactions. *Nature* **512**:436.

Kawakatsu, T., Huang, S.-S. C., Jupe, F., Sasaki, E., Schmitz, R. J., Urich, M. A., Castanon, R., Nery, J. R., Barragan, C., He, Y., et al. (2016). Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions. *Cell* **166**:492–505.

Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D., and Penin, A. A. (2016). A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *Plant J. Cell Mol. Biol.* **88**:1058–1070.

Kong, W., Li, B., Wang, Q., Wang, B., Duan, X., Ding, L., Lu, Y., Liu, L.-W., and La, H. (2018). Analysis of the DNA methylation patterns and transcriptional regulation of the NB-LRR-encoding gene family in Arabidopsis thaliana. *Plant Mol. Biol.* **96**:563–575.

Kong, W., Xia, X., Wang, Q., Liu, L.-W., Zhang, S., Ding, L., Liu, A., and La, H. (2020). Impact of DNA Demethylases on the DNA Methylation and Transcription of Arabidopsis NLR Genes. *Front. Genet.* **11**:460.

Kourelis, J., and van der Hoorn, R. A. L. (2018). Defended to the Nines: 25 Years of Resistance Gene Cloning Identifies Nine Mechanisms for R Protein Function. *Plant Cell* **30**:285–299.

644 **Lai, Y., and Eulgem, T.** (2018). Transcript-level expression control of plant NLR genes. *Mol. Plant*
645 *Pathol.* **19**:1267–1281.

646 **Laperche, A., Aigu, Y., Jubault, M., Ollier, M., Guichard, S., Glory, P., Strelkov, SE., Gravot, A., and**
647 **Manzanares-Dauleux, MJ.** (2017). Clubroot resistance QTL are modulated by nitrogen input
648 in Brassica napus. *Theor. Appl. Genet.* **130**:669–684.

649 **Larkan, N. J., Lydiate, D. J., Parkin, I. A. P., Nelson, M. N., Epp, D. J., Cowling, W. A., Rimmer, S. R.,**
650 **and Borhan, M. H.** (2013). The Brassica napus blackleg resistance gene LepR3 encodes a
651 receptor-like protein triggered by the Leptosphaeria maculans effector AVRLM1. *New Phytol.*
652 **197**:595–605.

653 **Law, J. A., and Jacobsen, S. E.** (2010). Establishing, maintaining and modifying DNA methylation
654 patterns in plants and animals. *Nat. Rev. Genet.* **11**:204–220.

655 **Lemarié, S., Robert-Seilaniantz, A., Lariagon, C., Lemoine, J., Marnet, N., Jubault, M., Manzanares-**
656 **Dauleux, M. J., and Gravot, A.** (2015). Both the Jasmonic Acid and the Salicylic Acid Pathways
657 Contribute to Resistance to the Biotrophic Clubroot Agent Plasmodiophora brassicae in
658 Arabidopsis. *Plant Cell Physiol.* **56**:2158–2168.

659 **Li, Y., Tessaro, M. J., Li, X., and Zhang, Y.** (2010). Regulation of the Expression of Plant Resistance
660 Gene SNC1 by a Protein with a Conserved BAT2 Domain. *Plant Physiol.* **153**:1425.

661 **Li, X., Kapos, P., and Zhang, Y.** (2015). NLRs in plants. *Curr. Opin. Immunol.* **32**:114–121.

662 **Liégard, B., Baillet, V., Etcheverry, M., Joseph, E., Lariagon, C., Lemoine, J., Evrard, A., Colot, V.,**
663 **Gravot, A., Manzanares-Dauleux, M. J., et al.** (2019). Quantitative resistance to clubroot
664 infection mediated by transgenerational epigenetic variation in Arabidopsis. *New Phytol.*
665 **222**:468–479.

666 **López Sánchez, A., Stassen, J. H. M., Furci, L., Smith, L. M., and Ton, J.** (2016). The role of DNA
667 (de)methylation in immune responsiveness of Arabidopsis. *Plant J. Cell Mol. Biol.* **88**:361–
668 374.

669 **López Sánchez, A., Pascual-Pardo, D., Furci, L., Roberts, M. R., and Ton, J.** (2021). Costs and Benefits
670 of Transgenerational Induced Resistance in Arabidopsis. *Front. Plant Sci.* **12**.

671 **Luna, E., Bruce, T. J. A., Roberts, M. R., Flors, V., and Ton, J.** (2012). Next-Generation Systemic
672 Acquired Resistance. *Plant Physiol.* **158**:844.

673 **Maekawa, T., Kufer, T. A., and Schulze-Lefert, P.** (2011). NLR functions in plant and animal immune
674 systems: so far and yet so close. *Nat. Immunol.* **12**:817–826.

675 **Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A. J., King, G. J., Giovannoni, J. J., and**
676 **Seymour, G. B.** (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-
677 box transcription factor inhibits tomato fruit ripening. *Nat. Genet.* **38**:948.

678 **Manzanares-Dauleux, M. J., Divaret, I., Baron, F., and Thomas, G.** (2000). Evaluation of French
679 Brassica oleracea landraces for resistance to Plasmodiophora brassicae. *Euphytica* **113**:211–
680 218.

Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., Pitrat, M., Dogimont, C., and Bendahmane, A. (2009). A transposon-induced epigenetic change leads to sex determination in melon. *Nature* **461**:1135.

Martin, E. C., Ion, C. F., Ifrimescu, F., Spiridon, L., Bakker, J., Goverse, A., and Petrescu, A.-J. (2023). NLRscape: an atlas of plant NLR proteins. *Nucleic Acids Res.* **51**:D1470–D1482.

Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* Advance Access published 2014.

Meyers, B. C., Kozik, A., Griego, A., Kuang, H., and Michelmore, R. W. (2003). Genome-Wide Analysis of NBS-LRR–Encoding Genes in Arabidopsis. *Plant Cell* **15**:809.

Molinier, J., Ries, G., Zipfel, C., and Hohn, B. (2006). Transgeneration memory of stress in plants. *Nature* **442**:1046.

Morán-Diez, M. E., Martínez de Alba, Á. E., Rubio, M. B., Hermosa, R., and Monte, E. (2021). Trichoderma and the Plant Heritable Priming Responses. *J. Fungi* **7**:318.

Narusaka, M., Shirasu, K., Noutoshi, Y., Kubo, Y., Shiraishi, T., Iwabuchi, M., and Narusaka, Y. (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *Plant J.* **60**:218–226.

Nelson, R., Wiesner-Hanks, T., Wisser, R., and Balint-Kurti, P. (2017). Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* **19**:21.

Nishimura, M. T., Anderson, R. G., Cherkis, K. A., Law, T. F., Liu, Q. L., Machius, M., Nimchuk, Z. L., Yang, L., Chung, E.-H., and El Kasmi, F. (2017). TIR-only protein RBA1 recognizes a pathogen effector to regulate cell death in Arabidopsis. *Proc. Natl. Acad. Sci.* **114**:E2053–E2062.

Ong-Abdullah, M., Ordway, J. M., Jiang, N., Ooi, S.-E., Kok, S.-Y., Sarpan, N., Azimi, N., Hashim, A. T., Ishak, Z., Rosli, S. K., et al. (2015). Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm. *Nature* **525**:533.

Palma, K., Thorgrimsen, S., Malinovsky, F. G., Fiil, B. K., Nielsen, H. B., Brodersen, P., Hofius, D., Petersen, M., and Mundy, J. (2010). Autoimmunity in Arabidopsis acd11 is mediated by epigenetic regulation of an immune receptor. *PLoS Pathog.* **6**:e1001137.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**:e45.

Pilet-Nayel, M.-L., Moury, B., Caffier, V., Montarry, J., Kerlan, M.-C., Fournet, S., Durel, C.-E., and Delourme, R. (2017). Quantitative resistance to plant pathogens in pyramiding strategies for durable crop protection. *Front. Plant Sci.* **8**:1838.

Qu, S., Liu, G., Zhou, B., Bellizzi, M., Zeng, L., Dai, L., Han, B., and Wang, G.-L. (2006). The Broad-Spectrum Blast Resistance Gene Pi9 Encodes a Nucleotide-Binding Site–Leucine-Rich Repeat Protein and Is a Member of a Multigene Family in Rice. *Genetics* **172**:1901.

Quadrana, L., and Colot, V. (2016). Plant transgenerational epigenetics. *Annu. Rev. Genet.* **50**:467–491.

Quadrana, L., Almeida, J., Asís, R., Duffy, T., Dominguez, P. G., Bermúdez, L., Conti, G., Corrêa da Silva, J. V., Peralta, I. E., Colot, V., et al. (2014). Natural occurring epialleles determine vitamin E accumulation in tomato fruits. *Nat. Commun.* **5**:4027.

Quadrana, L., Bortolini Silveira, A., Mayhew, G. F., LeBlanc, C., Martienssen, R. A., Jeddeloh, J. A., and Colot, V. (2016). The Arabidopsis thaliana mobilome and its impact at the species level. *eLife* **5**:e15716.

Ramirez-Prado, J. S., Piquerez, S. J. M., Bendahmane, A., Hirt, H., Raynaud, C., and Benhamed, M. (2018). Modify the Histone to Win the Battle: Chromatin Dynamics in Plant–Pathogen Interactions. *Front. Plant Sci.* **9**.

Richards, E. J. (2006). Inherited epigenetic variation--revisiting soft inheritance. *Nat. Rev. Genet.* **7**:395–401.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**:139–140.

Saile, S. C., Jacob, P., Castel, B., Jubic, L. M., Salas-Gonzáles, I., Bäcker, M., Jones, J. D. G., Dangl, J. L., and Kasmi, F. E. (2020). Two unequally redundant "helper" immune receptor families mediate Arabidopsis thaliana intracellular "sensor" immune receptor functions. *PLOS Biol.* **18**:e3000783.

Saile, S. C., Ackermann, F. M., Sunil, S., Keicher, J., Bayless, A., Bonardi, V., Wan, L., Doumane, M., Stöbbe, E., Jaillais, Y., et al. (2021). Arabidopsis ADR1 helper NLR immune receptors localize and function at the plasma membrane in a phospholipid dependent manner. *New Phytol.* **232**:2440–2456.

Saucet, S. B., Ma, Y., Sarris, P. F., Furzer, O. J., Sohn, K. H., and Jones, J. D. (2015). Two linked pairs of Arabidopsis TNL resistance genes independently confer recognition of bacterial effector AvrRps4. *Nat. Commun.* **6**:6338.

Schneeberger, K., Ossowski, S., Ott, F., Klein, J. D., Wang, X., Lanz, C., Smith, L. M., Cao, J., Fitz, J., Warthmann, N., et al. (2011). Reference-guided assembly of four diverse Arabidopsis thaliana genomes. *Proc. Natl. Acad. Sci.* **108**:10249–10254.

Shao, Z.-Q., Xue, J.-Y., Wu, P., Zhang, Y.-M., Wu, Y., Hang, Y.-Y., Wang, B., and Chen, J.-Q. (2016). Large-scale analyses of angiosperm nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes reveal three anciently diverged classes with distinct evolutionary patterns. *Plant Physiol.* Advance Access published 2016.

Shivaprasad, P. V., Chen, H.-M., Patel, K., Bond, D. M., Santos, B. A., and Baulcombe, D. C. (2012). A microRNA superfamily regulates nucleotide binding site–leucine-rich repeats and other mRNAs. *Plant Cell* Advance Access published 2012.

Silveira, A. B., Trontin, C., Cortijo, S., Barau, J., Del Bem, L. E. V., Loudet, O., Colot, V., and Vincentz, M. (2013). Extensive Natural Epigenetic Variation at a De Novo Originated Gene. *PLOS Genet.* **9**:e1003437.

Slaughter, A., Daniel, X., Flors, V., Luna, E., Hohn, B., and Mauch-Mani, B. (2011). Descendants of primed Arabidopsis plants exhibit resistance to biotic stress. *Plant Physiol.* Advance Access published 2011.

Some, A., Manzanares, M. J., Laurens, F., Baron, F., Thomas, G., and Rouxel, F. (1996). Variation for virulence on Brassica napus L. amongst Plasmodiophora brassicae collections from France and derived single-spore isolates. *Plant Pathol.* **45**:432–439.

Stroud, H., Greenberg, M. V. C., Feng, S., Bernatavichute, Y. V., and Jacobsen, S. E. (2013). Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**:352–364.

Stuart, T., Eichten, S. R., Cahn, J., Karpievitch, Y. V., Borevitz, J. O., and Lister, R. Population scale mapping of transposable element diversity reveals links to gene regulation and epigenomic variation. *eLife* **5**.

Team, R. C. (2013). R: A language and environment for statistical computing Advance Access published 2013.

Thomas, C. M. (1998). Genetic and molecular analysis of tomato Cf genes for resistance to Cladosporium fulvum. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **353**:1413–1424.

Tsuchiya, T., and Eulgem, T. (2013). An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication. *Proc. Natl. Acad. Sci.* **110**:E3535.

Weigel, D., and Colot, V. (2012). Epialleles in plant evolution. *Genome Biol.* **13**.

Williams, S. J., Sohn, K. H., Wan, L., Bernoux, M., Sarris, P. F., Segonzac, C., Ve, T., Ma, Y., Saucet, S. B., and Ericsson, D. J. (2014). Structural basis for assembly and function of a heterodimeric plant immune receptor. *Science* **344**:299–303.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PloS One* **2**:e718.

Wu, C.-H., Abd-El-Haliem, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., and Kamoun, S. (2017). NLR network mediates immunity to diverse plant pathogens. *Proc. Natl. Acad. Sci.* **114**:8113–8118.

Xia, S., Cheng, Y. T., Huang, S., Win, J., Soards, A., Jinn, T.-L., Jones, J. D. G., Kamoun, S., Chen, S., Zhang, Y., et al. (2013). Regulation of Transcription of Nucleotide-Binding Leucine-Rich Repeat-Encoding Genes SNC1 and RPP4 via H3K4 Trimethylation. *Plant Physiol.* **162**:1694.

Xiao, S., Ellwood, S., Calis, O., Patrick, E., Li, T., Coleman, M., and Turner, J. G. (2001). Broad-spectrum mildew resistance in Arabidopsis thaliana mediated by RPW8. *Science* **291**:118–120.

Xu, X., Hayashi, N., Wang, C.-T., Fukuoka, S., Kawasaki, S., Takatsuji, H., and Jiang, C.-J. (2014). Rice blast resistance gene Pikahei-1 (t), a member of a resistance gene cluster on chromosome 4, encodes a nucleotide-binding site and leucine-rich repeat protein. *Mol. Breed.* **34**:691–700.

Yue, J.-X., Meyers, B. C., Chen, J.-Q., Tian, D., and Yang, S. (2012). Tracing the origin and evolutionary history of plant nucleotide-binding site–leucine-rich repeat (NBS-LRR) genes. *New Phytol.* **193**:1049–1063.

Zamani-Noor, N., Wallenhammar, A.-C., Kaczmarek, J., Patar, U. R., Zouhar, M., Manasova, M., and Jędryczka, M. (2022). Pathotype Characterization of Plasmodiophora brassicae, the Cause of Clubroot in Central Europe and Sweden (2016–2020). *Pathogens* **11**:1440.

Zhang, X.-C., and Gassmann, W. (2007). Alternative splicing and mRNA levels of the disease resistance gene RPS4 are induced during defense responses. *Plant Physiol.* **145**:1577–1587.

Zhang, H., Tang, K., Wang, B., Duan, C.-G., Lang, Z., and Zhu, J.-K. (2014). Protocol: a beginner's guide to the analysis of RNA-directed DNA methylation in plants. *Plant Methods* **10**:18.

Zhang, H., Lang, Z., and Zhu, J.-K. (2018). Dynamics and function of DNA methylation in plants. *Nat. Rev. Mol. Cell Biol.* **19**:489–506.

Zou, B., Yang, D.-L., Shi, Z., Dong, H., and Hua, J. (2014). Monoubiquitination of Histone 2B at the disease resistance gene locus regulates its expression and impacts immune responses in Arabidopsis. *Plant Physiol.* Advance Access published 2014.

809 **Figure legends**

810 **Fig. 1 | Fine mapping of *Pb-At5.2*. A** Genetic map and residual heterozygosity in the Recombinant Inbred Line
811 (RIL) 499 derived from Bur-0 and Col-0 (Alix et al., 2007) and genetic and physical positions of clubroot
812 resistance QTL (Jubault et al., 2008b). Black: Bur-0 allele, White: Col-0 allele, Hatched: Heterozygous (Col-
813 0/Bur-0). **B** Allele configuration at *Pb-At5.2* in the two derived HIF lines 10499 and 13499. **C** Photos showing
814 *Pb-At5.2* conferred partial resistance to the eH isolate in the HIF 499 genetic background. HIF 10499 and 10499
815 harbored at *Pb-At5.2* Bur-0 and Col-0 alleles, respectively. Observations were made at 21 days post-inoculation.
816 Observations were made at 21 days post-inoculation. The white arrow indicates the presence of a limited number
817 of galls in inoculated 10499. **D** First round of fine mapping: allelic structure in the F1 lines derived from reciprocal
818 crosses between 10499 and 13499. 554 F3 lines with recombination in the confidence interval were screened using
819 10 SNP markers between AG_14959 and AG_20993. High density genotyping (94 SNP from K1 to K94), in a
820 series of 88 recombinant F3 lines and clubroot phenotyping in their bulked segregating F4 progenies led to a new
821 interval between markers K58 (*AT5G47230*) and K65 (*AT5G47360*). **E** Second round of fine mapping:
822 Recombination positions in homozygous individuals obtained from selected recombinant lines. For each line, the
823 GA/LA index (disease symptoms, log scale) is indicated on the right panel. Center lines show the medians; box
824 limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the
825 interquartile range from the 25th and 75th percentiles, outliers are represented by dots; data points are plotted as
826 open circles. Number of individual plants analysed for each genotype is indicated (n). The notches are defined
827 as +/-1.58*IQR/sqrt(n) and represent the 95% confidence interval for each median. GA/LA values
828 statistically different from 13499 GA/LA value are indicated by stars (from T-Test, ***p-value <0.001). Genetic
829 markers are indicated for each recombination position. Markers between CL5 and 2017-C were used in every line
830 but are only shown for 2313-15. **F** New 26 kb interval of *Pb-At5.2* between markers CL4 (excluded) and K64
831 (excluded), containing eight annotated ORF, six transposons and one lncRNA. Yellow and red diamonds indicate
832 SNP and nucleotide deletions, respectively.

833 **Fig. 2 | Identification and CRISPR/Cas9 validation of two NLR-encoding genes controlling**
834 **clubroot resistance at QTL *Pb-At5.2*. A** Sequence variations and expression levels of genes in the *Pb-*
835 *At5.2* region. Gene expression values were from RNAseq analyses conducted in inoculated and control
836 conditions at 14 dpi (log_2 normalized cpm), with parental lines Col-0/Bur-0 and HIF lines 13499/10499
837 (the last two were derived from the RIL 499 and are homozygous *Pb-At5.2*$_{Col/Col}$ and *Pb-At5.2*$_{Bur/Bur}$,
838 respectively) (**Supplementary Data 2**). FDR adjusted *p-values* are shown if less than 0.05. Yellow and
839 red diamonds indicate SNP and INDEL variations, respectively. **B** Effect of *AT5G47260* or *AT5G47280*
840 knock-out on GA/LA index (disease symptoms) in Bur-0 background. Cas9-mediated mutations were obtained
841 in the Bur-0 genetic background. For each targeted gene, three independent lines harbouring independent
842 homozygous mutations were used. Lines 117-1 and 21-20 no longer have the CRISPR-Cas9 cassette. For each
843 line, the mean clubroot symptoms score (GA/LA, log scale) was obtained by modelling raw data of eight
844 biological replicates (with 10 to 12 individual plants per replicate). Center lines show the medians; box
845 limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the
846 interquartile range from the 25th and 75th percentiles, data points are plotted as open circles. Edited line GA/LA
847 values statistically different from Bur-0 GA/LA value are indicated by stars (from Dunnett 's test) with the
848 following code * p-value <0.05; ** p-value <0.01; ***p-value <0.001.

849 **Fig. 3 | Methylation of the region surrounding the two NLR-encoding genes controlling clubroot**
850 **resistance at QTL *Pb-At5.2* A** Methylation profiles in the *AT5G47260* and *AT5G47280* region, in Col-
851 0 and Bur-0 accessions, inferred from bisulfite data previously reported in Kawakatsu et al. (2016).
852 Average methylation level was calculated within non-overlapping 100-bp windows starting 1 kb before
853 the TSS site of *AT5G47260* and stopping 1 kb after the TSE site of *AT5G47280*. In red: methylation in
854 the CG context. In green: methylation in the CHG context. In blue: methylation in the CHH context. **B**
855 Methylation profiles obtained by CHOP-qPCR on *AT5G47260* in inoculated and non-inoculated roots
856 of Bur-0/Col-0, 10499/13499, and in the homozygous recombinant lines 2313-15 (*Pb-At5.2*$_{Col/Col}$) and
857 1381-2 (*Pb-At5.2*$_{Bur/Bur}$). Those two last genotypes harbor the narrowest recombination events from
858 either side of *Pb-At5.2* (between markers CLG4 and K64, details in Fig. 1). Center lines show the

859 medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend
860 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots;
861 data points are plotted as open circles. n = 4 bulks of 6 plants and *p-values* are shown (sided t-test).

862 **Fig. 4 | Intermediate methylation and transcription levels of candidate genes in heterozygous**
863 **plants are associated with full clubroot susceptibility.** Eighty-three individual plants of the
864 segregating progeny from the recombinant line 2509 (heterozygous in Chr.5 regions between genetic
865 markers K58 and K93) were sampled at 21 dpi. Leaves from each individual plant were used for
866 genotyping (PCR marker CL_N8), which defined n pools of >3 plants of each zygosity profile: Bur/Bur
867 (n=5), Col/Bur (n=13) and Col/Col (n=6) (black, grey and white boxes, respectively). Each plant pool
868 was evaluated for **A** clubroot resistance (GA/LA), **B** % methylation at the locus, and **C-D** candidate
869 gene expression (*AT5G47260* and *AT5G47280*). Gene expression was quantified through RT-qPCR,
870 data were normalized over mean-Cp from the pools Bur/Bur, following Pfaffl's method with two
871 reference genes[78]. Center lines show the medians; box limits indicate the 25th and 75th percentiles as
872 determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th
873 percentiles, outliers are represented by dots; data points are plotted as open circles.

874 **Fig. 5 | Natural epigenetic variation at *Pb-At5.2* affects the expression of *AT5G47260*/*AT5G47280***
875 **among Arabidopsis accessions. A** Screening for 1001 genome Arabidopsis accessions displaying a Col/Bur-
876 like genomic structure at *Pb-At5.2* (chr5: 19185600 - 19200600). X axis: Horizontal coverage region covered
877 by at least one read. Y axis: Vertical coverage in read percentage. The 401 accessions framed in the northeast
878 intercardinal region delimited by dotted lines harbour a vertical read coverage >0.8 and a horizontal DNA-
879 seq >0.5 (DNA-seq data from[46]). **B** Clustering of a series of accessions harbouring Col/Bur-like genomic
880 structure at *Pb-At5.2*, by their level of methylation on *AT5G47260* and *AT5G47280*. Bisulfite data were
881 obtained from the 1001 genome data project **(Supplementary Data 3 sheet 2)**. Average methylation level
882 was calculated 1 kb before the TSS site of *AT5G47260* and stopping 1 kb after the TSE site of *AT5G47280*
883 for each context. **C** Spearman correlation between the methylation and gene expression of *AT5G47260* and
884 *AT5G47280* in a subset of 253 *Arabidopsis* accessions for which expression data was available (RNAseq

data from Kawakatsu et al. (2016)). 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. 6 | 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. Each open circle represents one accession. In total 42 accessions have Col-like epiallele (high percentage of methylation, High % Met) and 84 have Bur-like epiallele (low percentage of methylation, Low % Met). 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 +/-1.58*IQR/sqrt(n) and represent the 95% confidence interval for each median. *p*-value (Wilcoxon test) is indicated.

Fig. 7| 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 pathways (Stroud et al., 2013). In red the methylation in CG context. In green the methylation in CHG context; in blue the methylation in CHH context.

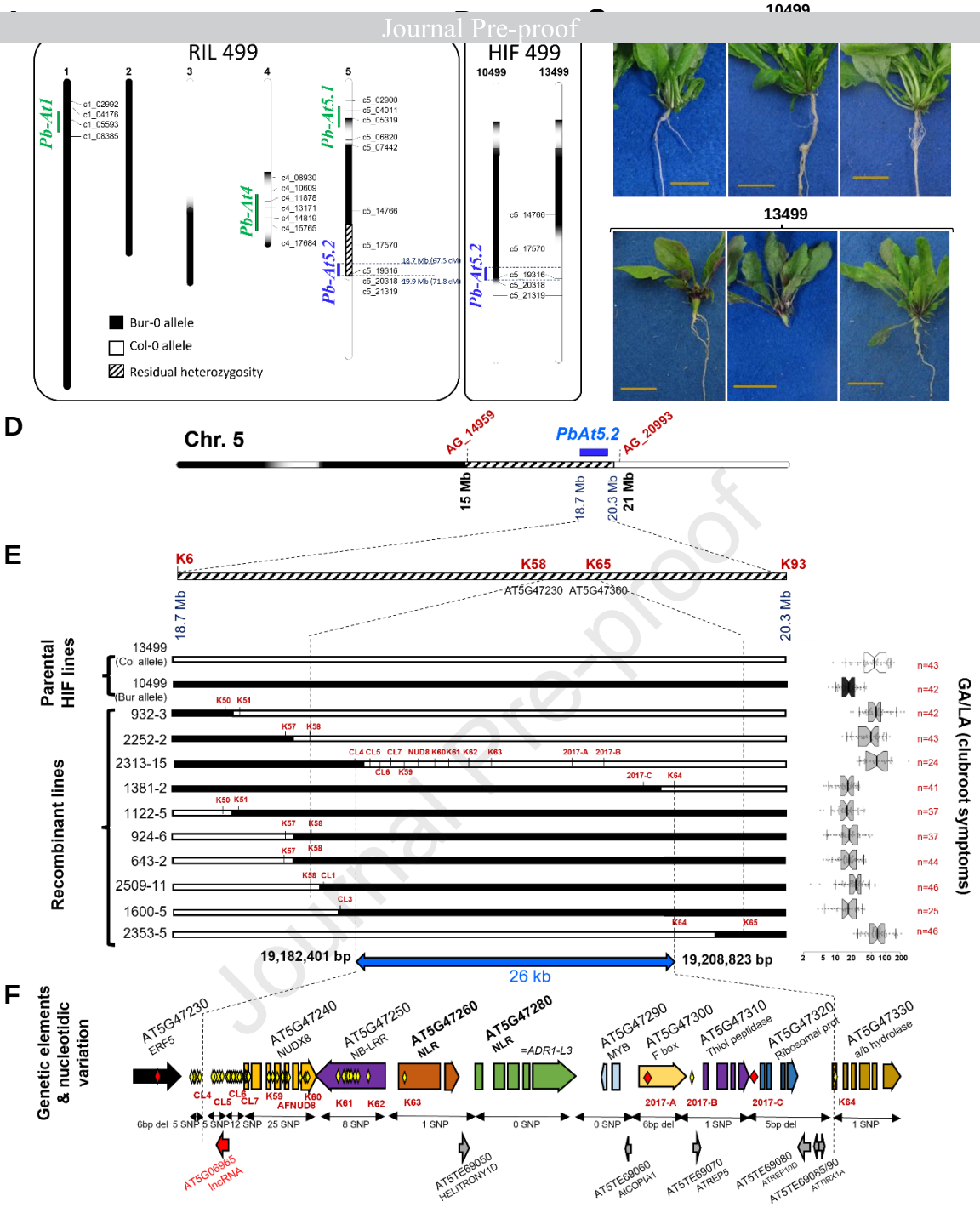


Fig. 1 | Fine mapping of *Pb-At5.2*.

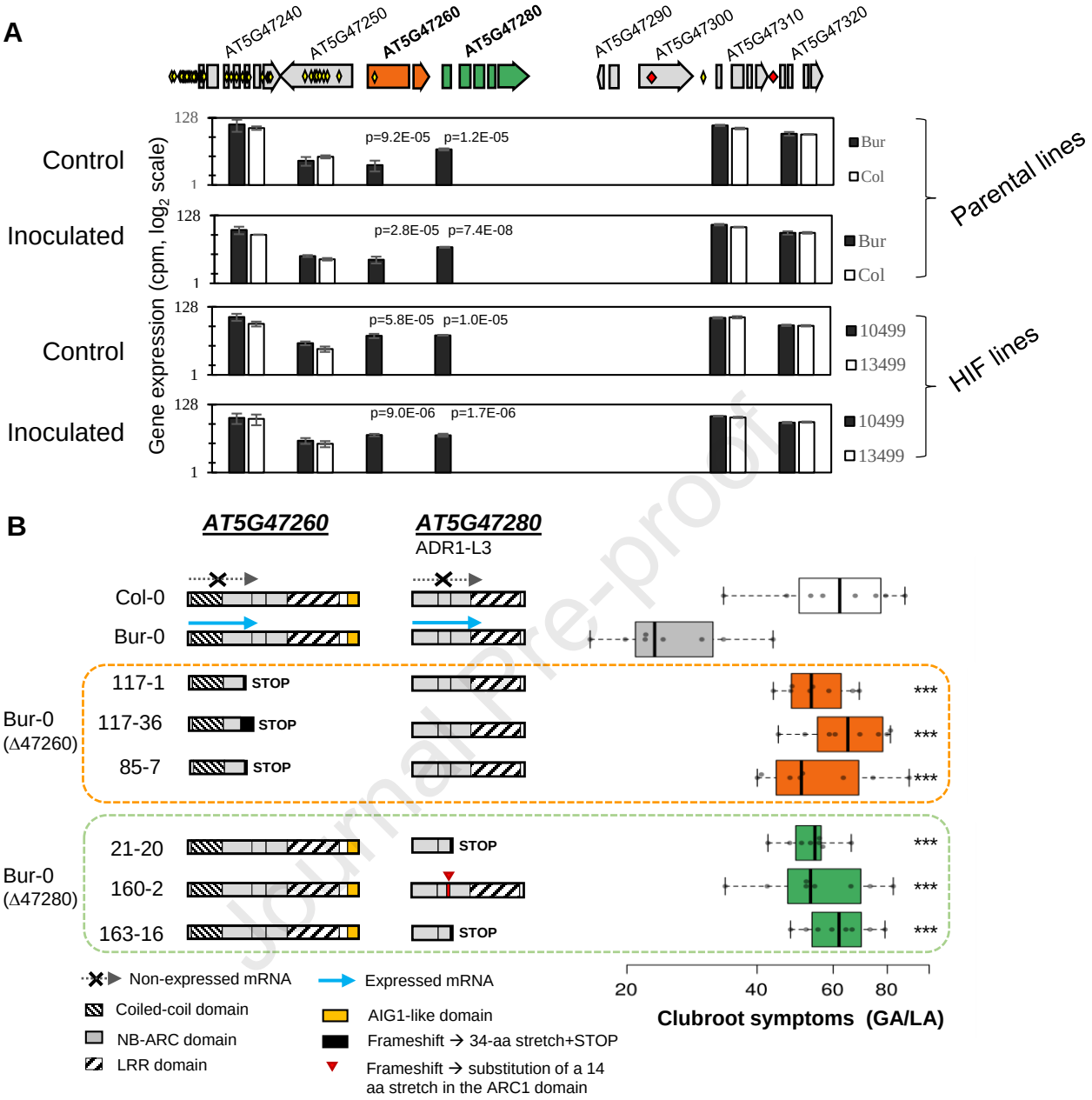


Fig. 2 | Identification and CRISPR/Cas9 validation of two NLR-encoding genes controlling clubroot resistance at QTL *Pb-At5.2*.

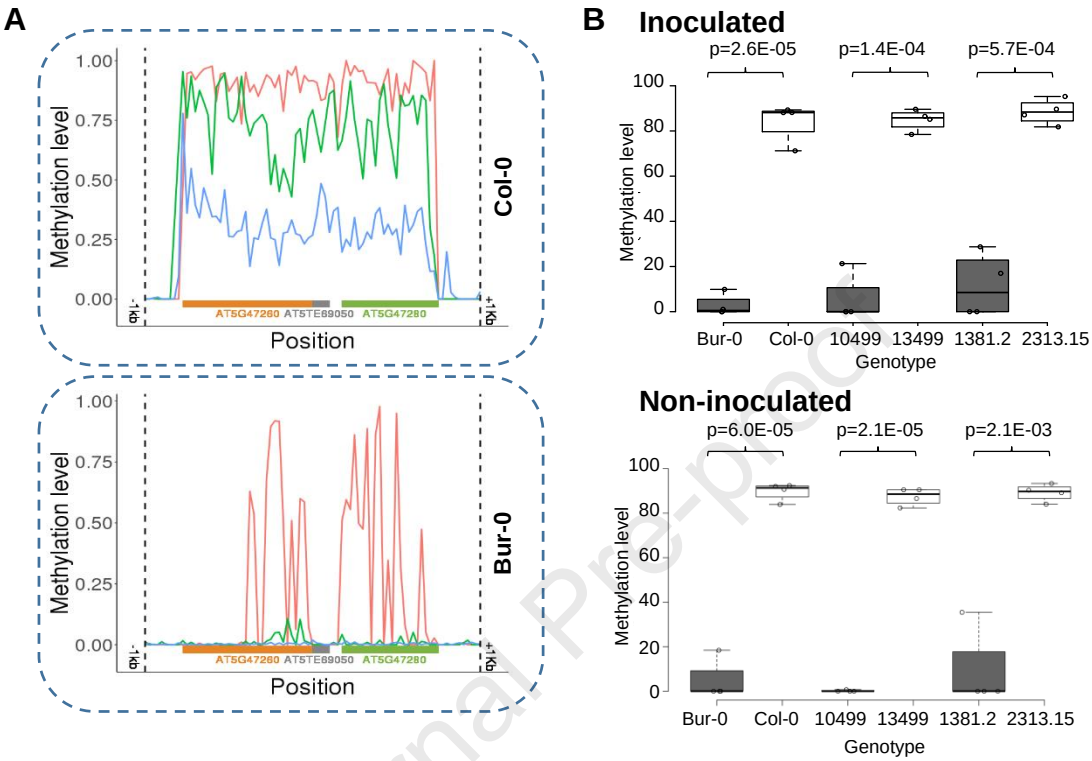


Fig. 3 | Methylation of the region surrounding the two NLR-encoding genes controlling clubroot resistance at QTL *Pb-At5.2*

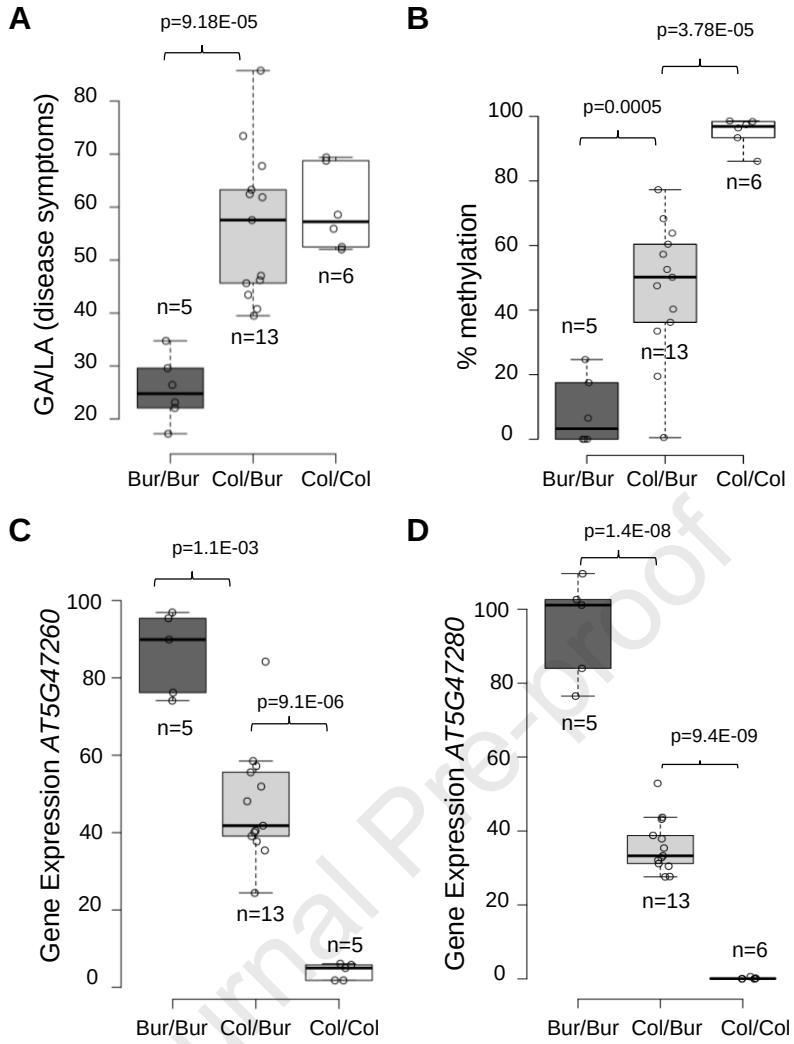
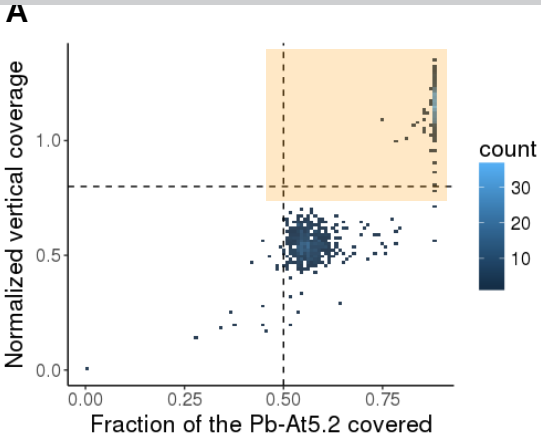
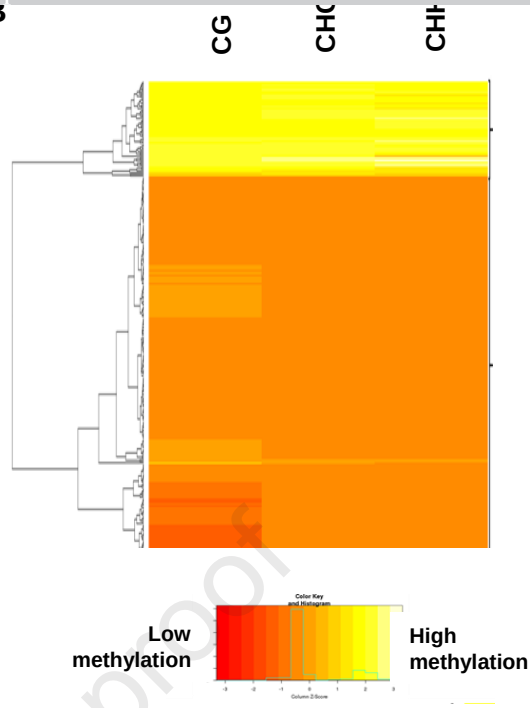


Fig. 4 | Intermediate methylation and transcription levels of candidate genes in heterozygous plants are associated with full clubroot susceptibility.





C

	Rho Spearman correlation		Rho *p-value*	
Methylation	AT5G47260	AT5G47280	AT5G47260	AT5G47280
mCHH	-0.44	-0.44	3.62E-13	2.02E-13
mCHG	-0.40	-0.36	5.08E-11	2.21E-09
mCG	-0.29	-0.15	3.96E-06	2.00E-02

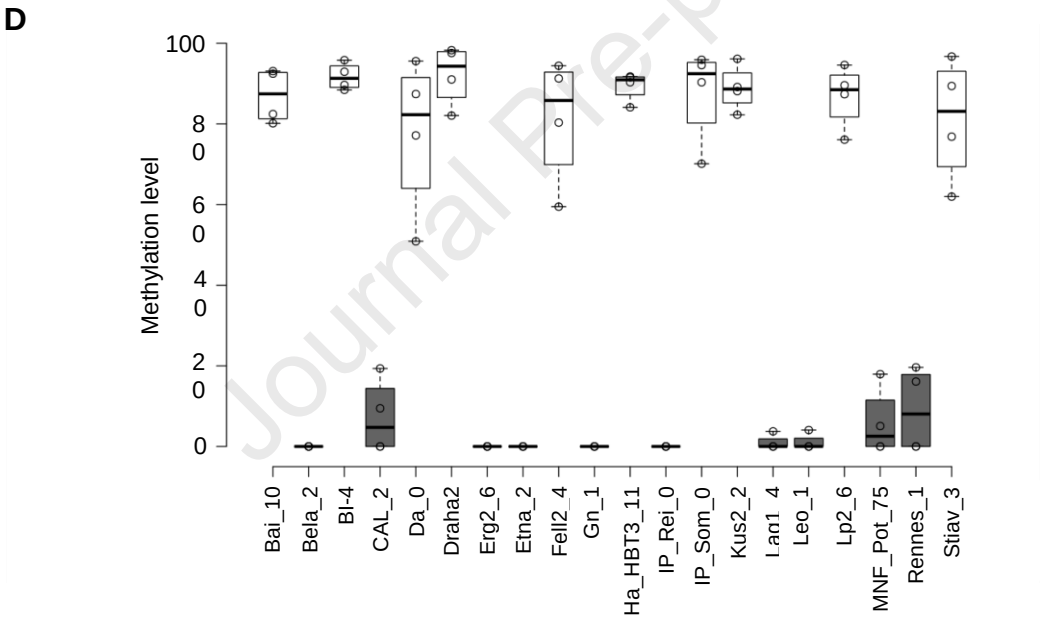


Fig. 5 | Natural epigenetic variation at *Pb-At5.2* affects the expression of *AT5G47260/AT5G47280* among Arabidopsis accessions.

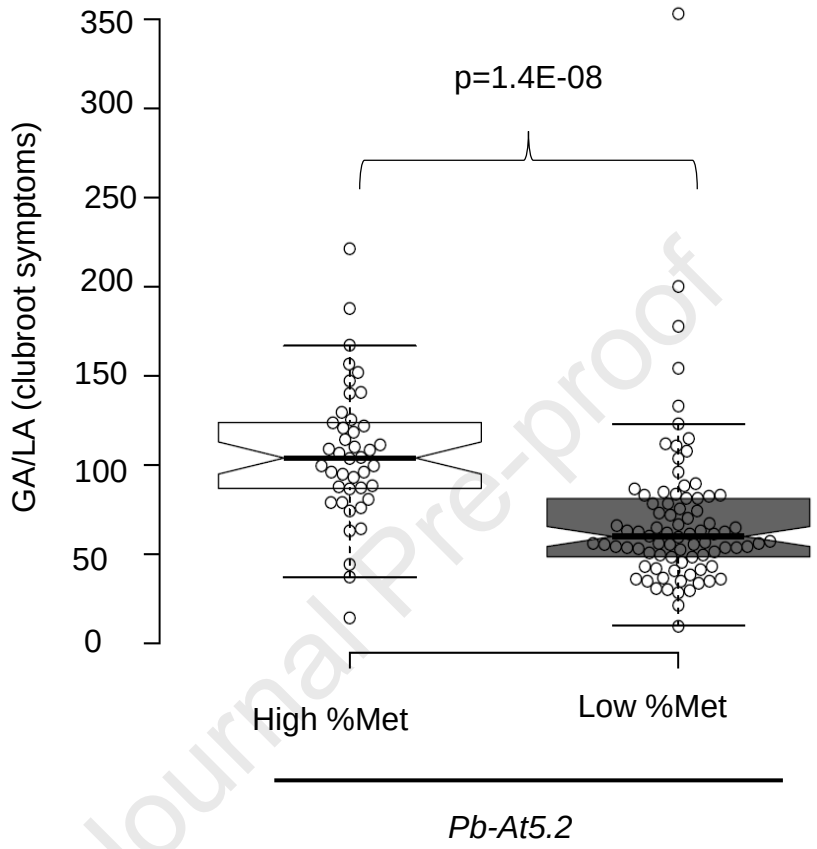


Fig. 6 | Clubroot symptom variation among Arabidopsis accessions is linked to epivariation at *Pb-At5.2*.

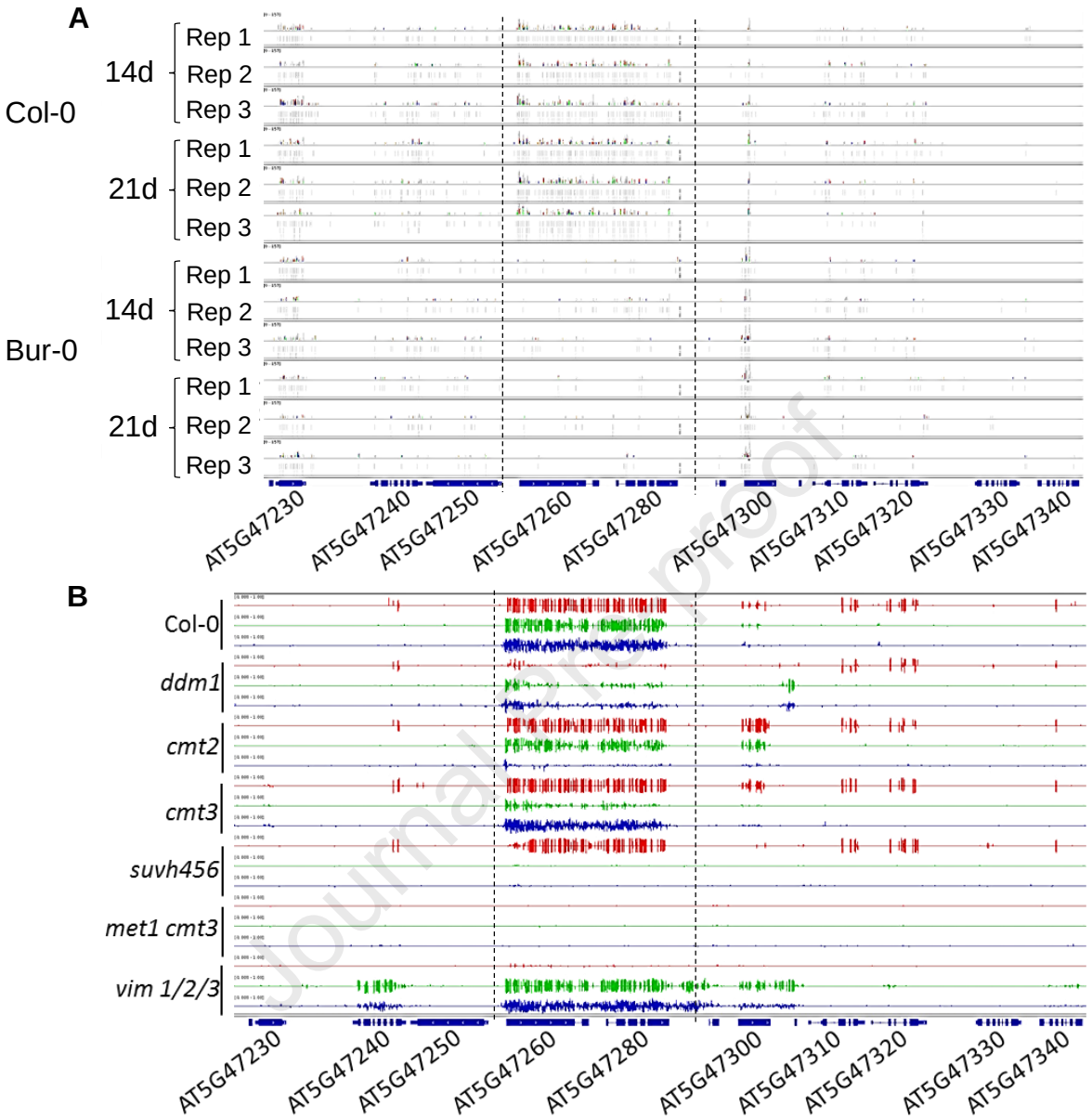


Fig. 7 | Epigenetic variation at *Pb-At5.2* correlates with the abundance of locus-targeted sRNA but is maintained by the RNA-independent methylation pathway.