Aphid feeding induces the relaxation of epigenetic control and the associated regulation of the defense response in Arabidopsis

Maria Luz Annacondia, Dimitrije Markovic, Juan Luis Reig-Valiente, Vassilis Scaltsoyiannes, Corné Pieterse, Velemir Ninkovic, R. Keith Slotkin, German Martinez

New Phytologist

Research

Aphid feeding induces the relaxation of epigenetic control and the associated regulation of the defense response in *Arabidopsis*

Maria Luz Annacondia[1]*, Dimitrije Markovic[2,3]*, Juan Luis Reig-Valiente[1], Vassilis Scaltsoyiannes[1,4], Corné M. J. Pieterse[5], Velemir Ninkovic[6], R. Keith Slotkin[7,8] and German Martinez[1]

[1]Department of Plant Biology, Uppsala BioCenter, Swedish University of Agricultural Sciences and Linnean Center for Plant Biology, Uppsala 75007, Sweden; [2]Department of Crop Production Ecology, Swedish University of Agricultural Sciences, Uppsala 75007, Sweden; [3]Faculty of Agriculture, University of Banja Luka, Banja Luka 78000, Bosnia and Herzegovina; [4]Institut de Biologie de Moléculaire des Plantes, UPR 2357 du CNRS, Strasbourg University, Strasbourg 67000, France; [5]Department of Biology, Science4Life, Utrecht University, Utrecht 3584 CS, the Netherlands; [6]Department of Ecology, Swedish University of Agricultural Sciences, Uppsala 75007, Sweden; [7]Donald Danforth Plant Science Center, St Louis, MO 63132, USA; [8]Division of Biological Sciences, University of Missouri-Columbia, Columbia, MO 65021, USA

Author for correspondence:
German Martinez
Email: german.martinez.arias@slu.se

Received: *11 November 2020*
Accepted: *14 January 2021*

New Phytologist (2021) **230:** 1185–1200
doi: 10.1111/nph.17226

Key words: aphids, *Arabidopsis*, defense response, epigenetics, transcriptional regulation, transposable elements.

Summary

- Environmentally induced changes in the epigenome help individuals to quickly adapt to fluctuations in the conditions of their habitats.
- We explored those changes in *Arabidopsis thaliana* plants subjected to multiple biotic and abiotic stresses, and identified transposable element (TE) activation in plants infested with the green peach aphid, *Myzus persicae*. We performed a genome-wide analysis mRNA expression, small RNA accumulation and DNA methylation
- Our results demonstrate that aphid feeding induces loss of methylation of hundreds of loci, mainly TEs. This loss of methylation has the potential to regulate gene expression and we found evidence that it is involved in the control of plant immunity genes. Accordingly, mutant plants deficient in DNA and H3K9 methylation (*kyp*) showed increased resistance to *M. persicae* infestation.
- Collectively, our results show that changes in DNA methylation play a significant role in the regulation of the plant transcriptional response and induction of defense response against aphid feeding.

Introduction

While adaptation to long-term environmental changes involves genetic variation, fluctuating stresses are normally coped with through the modulation of the transcription machinery (Lamke & Baurle, 2017). Several mechanisms govern the transcriptional response during stress, including transcription factors (TFs) and epigenetic regulation (Gutzat & Mittelsten, 2012). In eukaryotic organisms, epigenetic modifications of chromatin and DNA are the core of genome stability regulation through the control of transposable element (TE) expression and transposition (Law & Jacobsen, 2010). Epigenetic modifications consist of covalent and reversible marks that are deposited on both the DNA and the histones. DNA methylation constitutes a vital and widespread mark in plant genomes, where it can happened in three different sequence combinations: the symmetric contexts CG and CHG, and the asymmetric CHH (where H can be A, C or T) (Law & Jacobsen, 2010). This mark is established by the action of small RNAs (sRNAs) through a pathway named RNA-directed DNA methylation (RdDM) and can be actively removed from any context by the action of DNA glycosylases (Matzke & Mosher, 2014; Zhang *et al.*, 2018). The modifications that occur in the tails of histones can be active or repressive marks. For example, H3K4 mono-, di- and tri-methylation (H3K4me1, H3K4me2 and H3K4me3) are associated with highly transcribed genes (Zhang *et al.*, 2009), H3K27 tri-methylation (H3K27me3) is mainly found in silenced genes (Zhang *et al.*, 2007) and H3K9 di-methylation (H3K9me2) is rarely seen in genes while being predominantly present in TEs, where it correlates with the presence of DNA methylation, leading to transcriptional silencing and the formation of heterochromatin (Zhou *et al.*, 2010).

Transposable elements are a source of new mutations and genetic/genomic variation and of new regulatory regions for genes (Kidwell & Lisch, 1997; Lisch, 2009). Several agricultural traits like orange, maize and apple color or pepper pungency are regulated by TEs inserted in new locations, creating new expression patterns for the gene(s) in the vicinity of the insertion (Dooner *et al.*, 1991; Butelli *et al.*, 2012; Tanaka *et al.*, 2019; Zhang *et al.*, 2019). These TE domestication events are especially important for plant interaction with their environment (Annacondia *et al.*, 2018). Different abiotic and biotic stresses (including drought, salinity, heat, cold, UV radiation, chemical agents

*These authors contributed equally to this work.

and viral, viroid, bacterial and fungal infections) show examples of TE domestication events that influence gene expression and/or induce changes in the epigenetic regulation of repeats (Annacondia *et al.*, 2018; Mozgova *et al.*, 2019). Defense genes are interesting examples of the interaction between epigenetic regulation and gene regulation and evolution, as most nucleotide-binding site and leucine-rich repeat domain protein (NBS-LRR) genes accumulate in heterochromatic clusters populated by TEs (Meyers *et al.*, 2003). As an example of the role of epigenetic regulation in their transcriptional control, several defense genes, such as *RECOGNITION OF PERONOSPORA PARASITICA 7* (*RPP7*), *RPP4* and *RESISTANCE METHYLATED GENE 1* (*RMG1*), are transcriptionally regulated by domesticated TEs (Tsuchiya & Eulgem, 2013; Yu *et al.*, 2013; Zervudacki *et al.*, 2018). Additionally, mutants of different DNA methylation, RdDM and small RNA pathways regulate immunity to bacterial and fungal infection (Agorio & Vera, 2007; Lopez *et al.*, 2011; Dowen *et al.*, 2012; Yu *et al.*, 2013). Intriguingly, some biotic stresses can induce tolerance towards the pathogen in the subsequent generation (Boyko *et al.*, 2007; De Vos & Jander, 2009; Boyko *et al.*, 2010; Kathiria *et al.*, 2010; Luna *et al.*, 2012; Slaughter *et al.*, 2012), a phenomenon that could be explained by changes in the methylation status of the DNA or chromatin rather than by spontaneous mutagenesis and reversion (Boyko & Kovalchuk, 2011; Luna & Ton, 2012; Annacondia & Martinez, 2019).

The relationship between pathogens and host plants involves an interaction between both genomes and leads to events of coevolution. An example of this interaction takes place between plants and insects. Both groups interact in different ways and have influenced each other during evolution (e.g. the appearance in land plants of entomophily (Darwin, 1899) or carnivory (Renner & Specht, 2013) or the artificial selection of insects that evolve resistance to plants with defense genes (Bown *et al.*, 1997)). Plant–insect interactions are classified as mutualistic, antagonistic or commensalistic. Although they are basic for the ecological equilibrium, some of them can be a threat to the agricultural ecosystems and, hence, to food production. Herbivory insects represent *c.* 50% of the total insect species (Schoonhoven *et al.*, 2005) and are considered a threat to plant productivity. They are among the stresses that induce parental transmission of acquired resistance to the next generation, pointing to a potential role of epigenetic regulation of plant defense (Rasmann *et al.*, 2012). Nevertheless, how this epigenetic response is established during insect infestation is poorly characterized.

Here, we report that epigenetic control is an important part of the *Arabidopsis thaliana* defense response against the infestation by the green peach aphid *Myzus persicae*. Our analysis of DNA methylation, mRNA and sRNA changes induced in plants exposed to aphid feeding shows that the response of the plant is characterized by a transcriptional reprogramming and methylation changes in TEs. These TEs are normally associated with repressive/heterochromatic marks and are dependent on the RdDM pathway for their silencing. Along with this, we find that upon infestation, certain differentially methylated regions (DMRs) are associated with infestation-responsive genes and TF binding sites. Finally, we find that mutant plants deficient in epigenetic silencing show increased resistance to *M. persicae* infestation. Together, our data uncover a novel role for plant epigenetic control in the induction of the transcriptional response to aphid feeding.

Materials and Methods

Plant and insect material

Arabidopsis thaliana (Columbia wild-type Col-0, *ddm1-2, ddc, nrpd1a-4* and *kyp-6*) were sown into potting soil (P-Jord, Hasselfors Garden, Örebro, Sweden). At the four-leaf stage, seedlings were selected by uniformity and carefully replanted into plastic pots (9 × 9 × 7 cm) with one plant per pot at temperature 20–22°C and 70% relative humidity. Plants were grown under a 16 h : 8 h, light : dark photoperiod. The light was provided by FQ, 80 W, Hoconstant lumix (Osram, Munich, Germany) with a light intensity of 220 μmol photons m^{-2} s^{-1}. Green peach aphid *M. persicae* (Sulzer) was reared in cultures on potted rapeseed plants *Brassica napus* L. under the same climate conditions as the test plants but in different climate chambers.

Aphid settling test

An aphid no-choice settling test (Ninkovic *et al.*, 2002) was used to investigate aphid behavioral response to different *Arabidopsis* mutants. One randomly chosen leaf was placed inside a transparent polystyrene tube (diameter 1.5 cm, length 5 cm). The lower end of the tube was plugged with a plastic sponge through which the leaf entered via a slit. Ten wingless second- to fourth-instar larvae of *M. persicae* were placed inside the tube. The upper end of the tube was sealed with nylon net. A leaf of each treatment plant placed inside the tube represented a replicate. The number of aphids that settled on the leaf was recorded after 2 h, which is sufficient time for aphids to settle and reach the phloem (Prado & Tjallingii, 1997).

Tissue collection for sRNA, RNA and bisulfite sequencing

Five-week-old plants were infested with 40 wingless second- to fourth-instar larvae of *M. persicae* and covered with a net cage. After 72 h, all aphids were carefully removed using a brush and all the leaves from the *Arabidopsis* rosette (between eight and 10 leaves) were sampled into Falcon tubes and placed in liquid nitrogen for nucleic acid extraction. Four plants were pooled on each bioreplicate. Frozen plant tissue was stored at −70°C before being used for RNA and DNA extraction. The same tissue was used for sRNA, mRNA and genome-wide bisulfite sequencing.

DNA and RNA extraction

Total RNA was extracted using TRIzol reagent (Life Technologies) following the manufacturer's instructions. mRNA for RNA sequencing was obtained by purification with the NEB mRNA isolation kit (New England Biolabs, Ipswich, MA, USA). RNA for sRNA library preparation was enriched with the mirVana

miRNA Isolation Kit (Life Technologies). Genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen).

Small RNA, RNA sequencing and analysis

Small RNA libraries were produced using the TruSeq Small RNA Sample Preparation Kit (Illumina Inc., San Diego, CA, USA). Each library was barcoded and sequenced in one lane of an Illumina HiSeq 2000. RNA libraries were produced using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs). Each library was barcoded and sequenced in one lane of an Illumina HiSeq 2500. The resulting sequences were de-multiplexed, adapter trimmed, and filtered on length and quality. Three bioreplicates were sequenced for sRNA analysis. sRNAs were matched to the *Arabidopsis* genome. Library size was normalized by calculating reads per million of 18–28 nt genome-matched sRNAs. sRNA alignments were performed using BOWTIE (Langmead *et al.*, 2009) with the parameters –t –v2, which allow two mismatches to the alignments. For gene expression analysis, two bioreplicates from each treatment were sequenced. RNA-sequencing paired reads were aligned to the Arabidopsis TAIR10 genome using BOWTIE2 (Langmead & Salzberg, 2012) with default parameters. HTSEQ-COUNTS (Anders *et al.*, 2014) was used to count reads per gene with the parameters --mode union --stranded no --minequal 10 and --nonunique none. For TE expression analysis, RNA sequencing paired reads were aligned to the Arabidopsis TAIR10 genome using STAR (Dobin *et al.*, 2013), allowing mapping to at most 100 'best' matching loci with the following parameters, --outMultimapperOrder Random --outSAMmultNmax -1 --outFilterMultimapNmax 100, used previously for TE analysis (Warman *et al.*, 2020). HTSEQ-COUNTS was used to count reads per TE with the parameters --mode union --stranded no --minequal 0 and --nonunique all. Count tables obtained were used in DESEQ2 (Love *et al.*, 2014) to infer significant expression with fit type set to parametric. Volcano plots were created using GGPLOT2 (Wickham, 2009). All these tools were used through the Galaxy platform (Afgan *et al.*, 2018).

RT–qPCR

For quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis, total RNA was DNAseI-treated and reverse-transcribed using the RevertAid RT Reverse Transcription Kit (Thermo Fisher Scientific, Waltham, MA, USA). qPCR was performed with 5× HOT FIREPol EvaGreen qPCR Mix Plus (Solis Biodyne, Tartu, Estonia) using three technical replicates from three biological replicates each. qPCR was performed on a CFX Connect Real-Time Detection System and the results analyzed on the CFX MANAGER software package (BioRad, Hercules, CA, USA). The relative expression values for all experiments were calculated based on the expression of the control housekeeping gene AT4G05320 (UBIQUITIN 10). Relative expression was calculated using the 'delta-delta method' formula $2^{-[\Delta CP\ sample - \Delta CP\ control]}$, where 2 represents perfect PCR efficiency. Statistical significance was calculated using unpaired *t*-tests. RT-qPCR primers are shown in Supporting Information Table S1.

Bisulfite library preparation and sequencing analysis

Bisulfite libraries were produced from genomic DNA at Novogene (Beijing, China) and sequenced as paired-end 150 bp fragments in an Illumina Novaseq 6000. Two bioreplicates from each treatment were sequenced. Raw reads were trimmed using TRIMGALORE 0.6.1 for removal of adapters and 10 bases from 5′ ends. Clean reads were mapped to the reference *Arabidopsis* genome TAIR 10 using BISMARK (Krueger & Andrews, 2011), allowing one mismatch per 25 nt seed. Forward and reverse reads were mapped independently. Alignments at the same position were removed using deduplicate_bismark script, including alignments of reads 1 and 2 together. Conversion rates of cytosines were obtained using bismark_methylation_extractor; the first seven bases from the 5′ end and 13 from the 3′ end of each read were ignored. The mean conversion rate based on the cytosine methylation levels in the chloroplast genome for the four samples was 99.76%, and the estimated false-positive methylation rates were 0.24% (Fig. S4e; see later). Tile values for genomic DNA methylation were obtained using the Circos: Interval to Tiles pipeline in the Galaxy platform (Afgan *et al.*, 2018). Circular plots were obtained using J-CIRCOS (An *et al.*, 2015).

DMR identification

The DMR analysis was carried on with the R package DMRCALLER (Catoni *et al.*, 2018); biological replicates from control and infected samples were pooled and compared between treatments. In order to compare both pools, the genome was divided in equal bins of 50 bp size. The DMRs were then computed by performing Fisher's exact test between the number of methylated reads and the total number of reads in both conditions for each bin. The obtained *P*-values were then adjusted for multiple testing using Benjamini and Hochberg's (Benjamini & Hochberg, 1995) method to control the false discovery. Bins with fewer than three cytosines in the specified context or < 0.25 difference in methylation proportion between the two conditions or an average number of reads lower than 8 were discarded. Finally, bins that were at < 300 bp were joined.

Microarray analysis

Microarray analysis was performed for the datasets indicated in Table S2 and retrieved from the NCBI Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The expression values were normalized by quantifying the ratio of the mean value for each treatment to the mean value of its respective control. A heat map for the analysis of microarray data was produced using HEATMAPPER (Babicki *et al.*, 2016).

Transcription factor binding site prediction

Transcription factor binding site prediction was performed using the plant transcription factor database (http://planttfdb.cbi.pku.edu.cn/). The prediction tool was used against the nucleotide sequences of the CHH DMRs indicated.

Gene ontology (GO) term analysis

Gene ontology term analysis was carried out using the *GO annotation search, functional categorization and download* tool from the TAIR website (www.arabidopsis.org). In the different analysis, the whole genome categorization was compared to the categorization for the specific group of genes selected for the analysis. Bubble graphs were produced in Microsoft Excel. A biomaps graph was obtained using VirtualPlant 1.3 (http://virtualplant.bio.nyu.edu/cgi-bin/vpweb/).

Results

Analysis of TE activation under different stresses identifies *M. persicae* as a potential inducer of epigenetic changes

To identify stresses that alter the epigenetic regulation in *A. thaliana*, we performed an analysis of TE expression from ATH1 microarray datasets, which have been widely used by the community. The ATH1 microarray contains 1155 TE probes used to track changes in transcript abundance influenced by epigenetic reprogramming (Slotkin *et al.*, 2009). We investigated TE expression under different stresses, including abiotic (heavy metal presence, exposure to heat, cold, spaceflight or UV light among others) and biotic (viral, oomycete, bacterial and insect infection/infestation) (Fig. 1a; Table S2). We found that, in general, these stresses can induce a modest reactivation of TEs, although this response is dependent on the specific stress (Fig. 1a,b). Biotic stress seemed to activate TE expression more consistently than the abiotic stresses analyzed here (Fig. 1a,b). This analysis identified that among the stresses inducing TE reactivation, *M. persicae* infestation after 72 h induced the highest TE transcription. *Myzus persicae* is a major agricultural pest to a large variety of plants that include stone fruits, potato and horticultural crops (Louis & Shah, 2013). A high number of TEs (533 TEs, 46.1% of all the TEs represented in the ATH1 microarray; Fig. 1b) showed evidence of transcriptional activation when plants were under attack from *M. persicae* as compared with control plants. This reactivation included > 40% of all the DNA transposons and retrotransposons represented in the ATH1 microarray, is overrepresented by *Gypsy* and *Copia* retrotransposons and TIR DNA transposons, and is significantly enriched in upregulated members of the *MuDR* nonTIR transposon class ($P = 0.039$, Fisher's exact test, TEs upregulated more than two-fold; Fig. 1c). Analysis of the reactivation indicated that TE activation takes place at 48 h and increases by 72 h post-infestation (pi) (Figs 1b, S1; average fold-change values for retrotransposons at 72 hpi are 3.64- and 4.4-fold for DNA transposons). Other cases of large-scale TE activation are seen when DNA methylation, histone modification and/or heterochromatin formation are lost (Lippman *et al.*, 2003; Lippman *et al.*, 2004; Zilberman *et al.*, 2007; Panda *et al.*, 2016). Together, these results indicate that *M. persicae* infestation results in TE activation, potentially as a result of a large-scale change in the epigenome.

Transcriptional response to aphid feeding in *Arabidopsis* is characterized by transcription factor activity

The extent of TE reactivation observed in our previous analysis could be biased by the presence of TE probes on the ATH1 microarray. To monitor the transcriptional changes under aphid infestation, we repeated the experiment described in De Vos *et al.* (2005) (*Arabidopsis plants* infested with *M. persicae* for 72 h, for details see the Materials and Methods section) and prepared and sequenced high-throughput mRNA libraries (Table S3). First, we focused on understanding the genic transcriptional changes taking place in our libraries. Principal component analysis of gene expression in mRNA libraries generated from control and infested tissue demonstrated that biological replicates clustered together (Fig. S2a). Our differential expression analysis identified 267 genes that were significantly differentially expressed (adjusted $P < 0.05$), with almost all of these being upregulated (265 genes; Fig. 2a; Table S4). Differentially expressed genes contained a significant overrepresentation of mobile mRNAs (24.34% of differentially expressed genes; two-tailed $P < 0.0001$ calculated with a χ^2 test with Yates correction) (Thieme *et al.*, 2015) (Fig. S2c). As expected, the analysis of the GO categories for significantly upregulated genes indicated that these genes were associated with the response to stress or environmental stimuli (Figs 2b,c, S2B).

We further analyzed the molecular functions of these stress-responsive genes by checking the GO term enrichment according to molecular function (Fig. 2d). This revealed an overrepresentation of DNA-binding/transcription factor categories (GO terms 'nucleic acid binding', 'DNA-binding transcription factor activity' and 'DNA binding' were significantly enriched with $P < 0.000\,01$, calculated with Fisher's exact test; Fig. 2d), indicating that these transcriptional regulators are an important part of the response to aphid feeding. Several well-studied TFs showed a strong upregulation (higher than 1.5 $\log_2$fold-change) including members of the WRKY and ERF families (Fig. 2e), which have previously been associated with the response to aphid feeding (Gao *et al.*, 2010). Furthermore, we identified the overexpression of a single component of the epigenetic regulatory pathways that was overexpressed under aphid attack, *HIKESHI-LIKE PROTEIN1* (*HLP1,* significantly overexpressed), a promoter-binding protein that promotes chromatin acetylation (Sharma *et al.*, 2019) (Fig. 2f). In summary, the transcriptional response against aphids showed an overrepresentation of TF activity.

Aphid infestation induces transcriptional activation of TEs

Our previous analysis of ATH1 public datasets indicated a potential reactivation of TEs during aphid infestation. However, the TE probes on the ATH1 array do not represent the genomic distribution of TEs, and favor *Helitron* elements that resemble genes. Accordingly, we explored TE transcriptional and post-transcriptional regulation by performing RNA and sRNA sequencing, which target (respectively) mRNAs and sRNAs derived from Pol II and Pol IV activity (Fig. 3). Principal component analysis of TE expression in mRNA libraries generated from control and infested tissue demonstrated that biological replicates

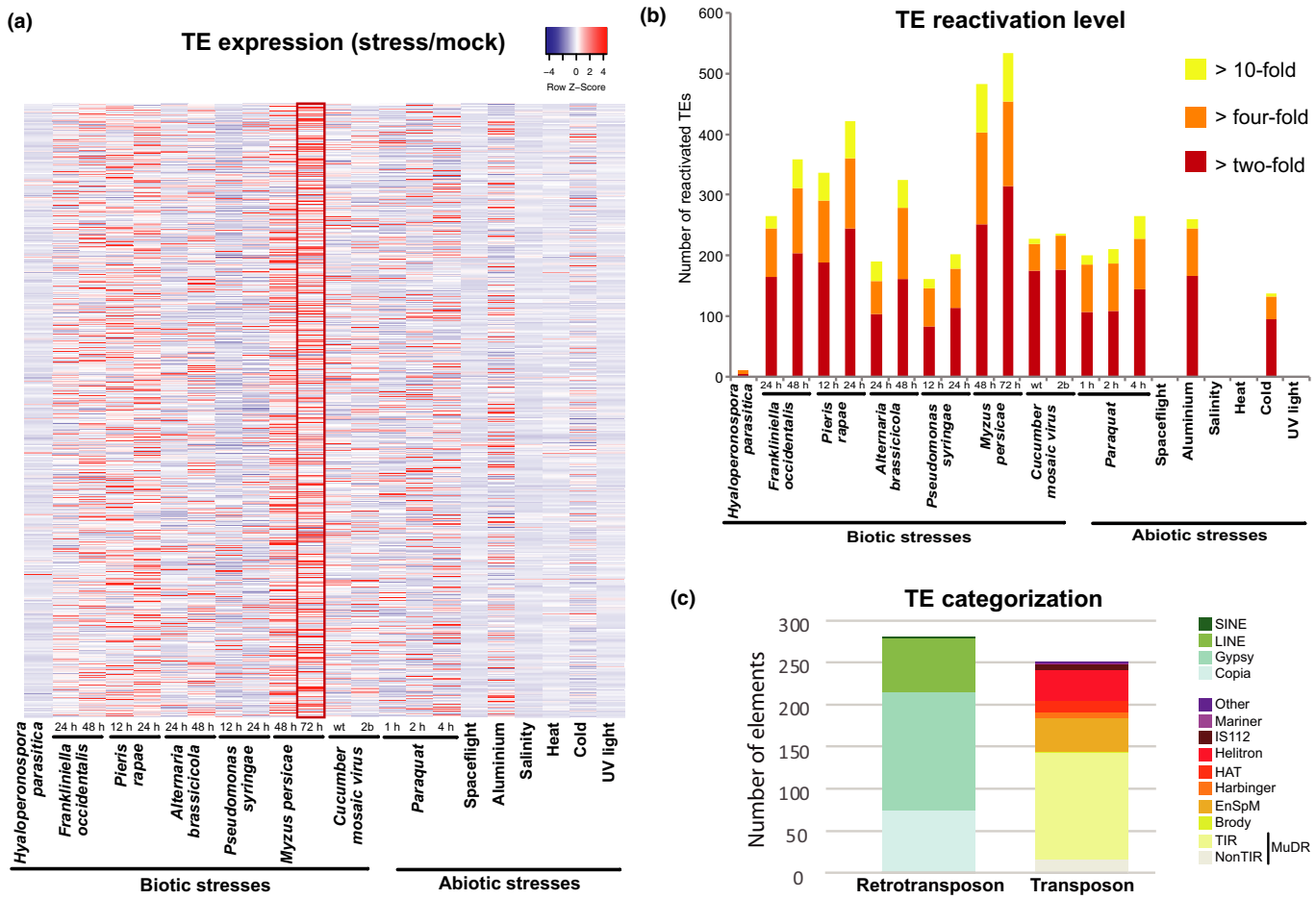


Fig. 1 Aphid infestation induces transposable element (TE) reactivation. (a) Analysis of *Arabidopsis* TE expression in the ATH1 microarray under several stresses. Heat map of the expression values of the indicated treatment relative to its respective control. In experiments with several bioreplicates, the mean values between bioreplicates were used. (b) Number of TEs reactivated in the analyzed stresses grouped by fold categories. (c) Percentage of reactivated TEs belonging to different categories in the ATH1 microarray.

clustered together (Fig. S3a). Analysis of RNA sequencing indicated that 71 TEs suffered a transcriptional reactivation upon aphid infestation (Fig. 3a; Table S5). Reactivated TEs included several members of the ATDNA12T3 family, which cluster in the centromeric regions of chromosomes 3, 4 and 5 and other TEs known to be reactivated under other stresses like ATCOPIA78/ONSEN (Fig. 3a). The DNA TE superfamily (to which ATDNA12T3 TEs belong) was indeed significantly over-represented in the reactivated TEs population (Fisher's exact test, $P < 0.000\,01$; Fig. 3b).

Next, the analysis of our sRNA sequencing revealed dramatic differences taking place almost exclusively at 24 nt TE-derived sRNAs (Figs 3c,d, S3b–i). Principal component analysis of sRNA libraries generated from control and infested tissue demonstrated that biological replicates clustered together (Fig. S3b). Loss of 24 nt sRNAs was significant at both total sRNAs and TE-derived sRNA populations (*P*-value calculated trough an unpaired *t*-test; Fig. 3c,d). This loss of 24 nt sRNAs was slightly more pronounced in long transposons of almost all TE families (Figs 3e, S3c). Long retrotransposons are located in centromeric and pericentromeric regions, which are the genomic habitats of *Gypsy* and *Copia/LINE* elements, respectively (Underwood *et al.*, 2017). The subgroup of RNA sequencing-identified reactivated TEs also experienced changes at the sRNA level with significant increased levels of 21 nt sRNAs and significant loss of 24 nt sRNAs (Fig. 3f). Indeed, 21 nt sRNAs showed a significant increase of their accumulation levels between control and *M. persicae* sRNA libraries (*P*-value calculated trough a paired *t*-test; Fig. 3g), which are dependent on Pol II and, subsequently, their overaccumulation is a common signature of TE transcriptional reactivation in *Arabidopsis* epigenetic mutants like *ddm1* or *met1* (McCue *et al.*, 2012). In summary, our RNA- and sRNA-sequencing data indicated that during aphid infestation plants reduced the activity of the RdDM pathway, leading to the transcriptional reactivation of centromeric TEs.

Differential methylation of the *Arabidopsis* genome upon aphid infestation

The transcriptional changes observed and the loss of TE-derived 24 nt sRNAs lead us to analyze the levels of DNA methylation. Genomic DNA was isolated, treated with sodium bisulfite and

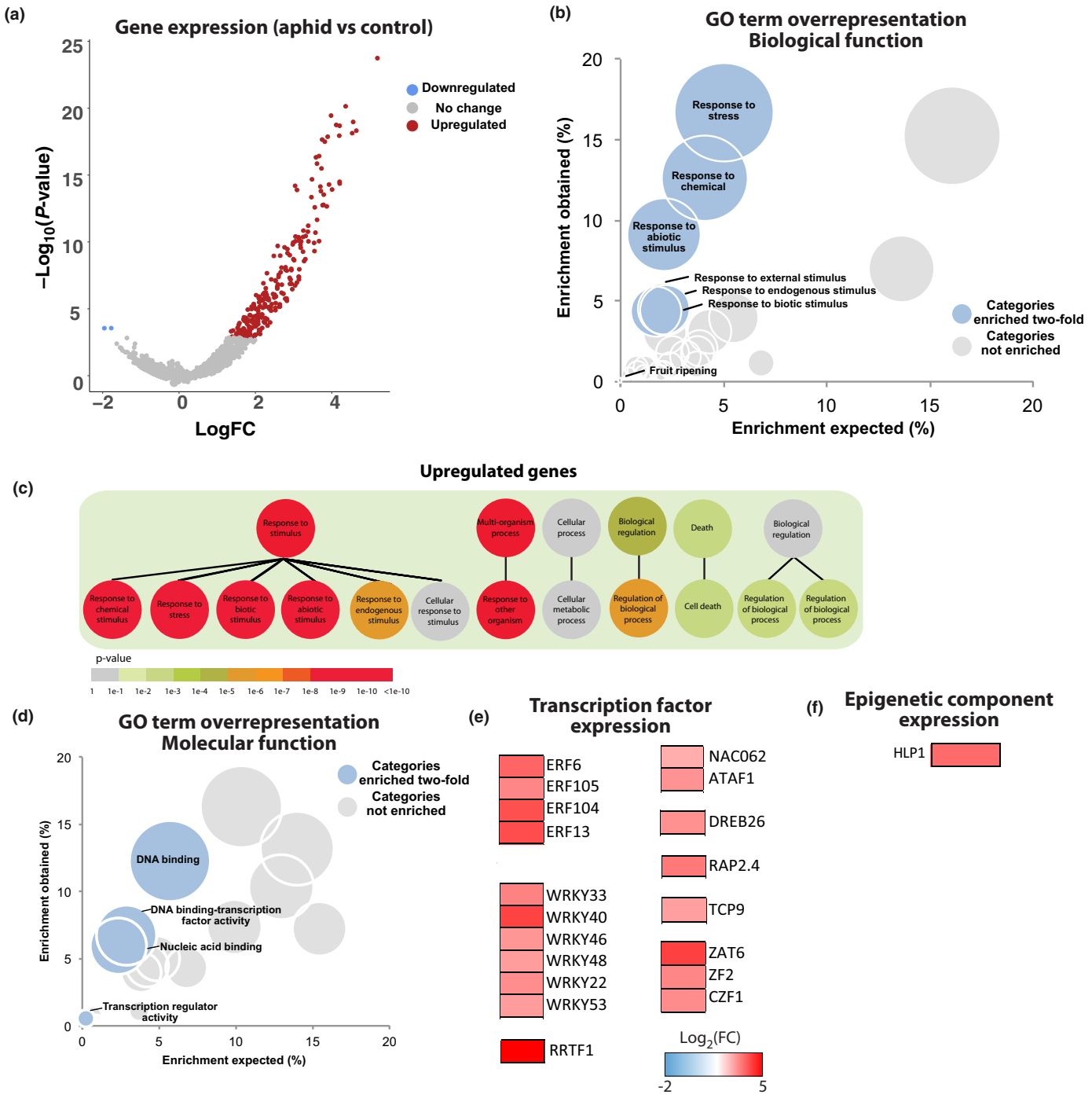


Fig. 2 Aphid feeding-induced changes in gene expression. (a) Volcano plot depicting *Arabidopsis* gene expression in the comparison between aphid-infested and control samples. Red dots indicate genes with significant upregulation. (b) Bubble graph depicting the gene ontology (GO) term overrepresentation test for upregulated genes grouped by biological function. Bubbles in blue show GO categories upregulated two-fold or more. (c) Biomaps of upregulated genes. The colors indicate the statistical significance of the overrepresentation as indicated in the legend. (d) Bubble graph depicting the GO term overrepresentation test for upregulated genes grouped by molecular function. Bubbles in blue show GO categories upregulated two-fold or more. (e) Examples of different transcription factors showing significant upregulation during aphid infestation. (f) A single epigenetic component is upregulated upon aphid infestation. FC, fold-change.

sequenced at 26.4 times average coverage (Table S3; mean conversion rate based on the cytosine methylation levels in the chloroplast genome for the four samples was 99.76%; Fig. S4e). The data were plotted as a heat map on all five chromosomes comparing the control and aphid-infested samples (Fig. 4a). These data revealed a strong enrichment of DNA methylation in the pericentromeric heterochromatin, as expected from somatic tissues. A global analysis of the methylation level at genes and

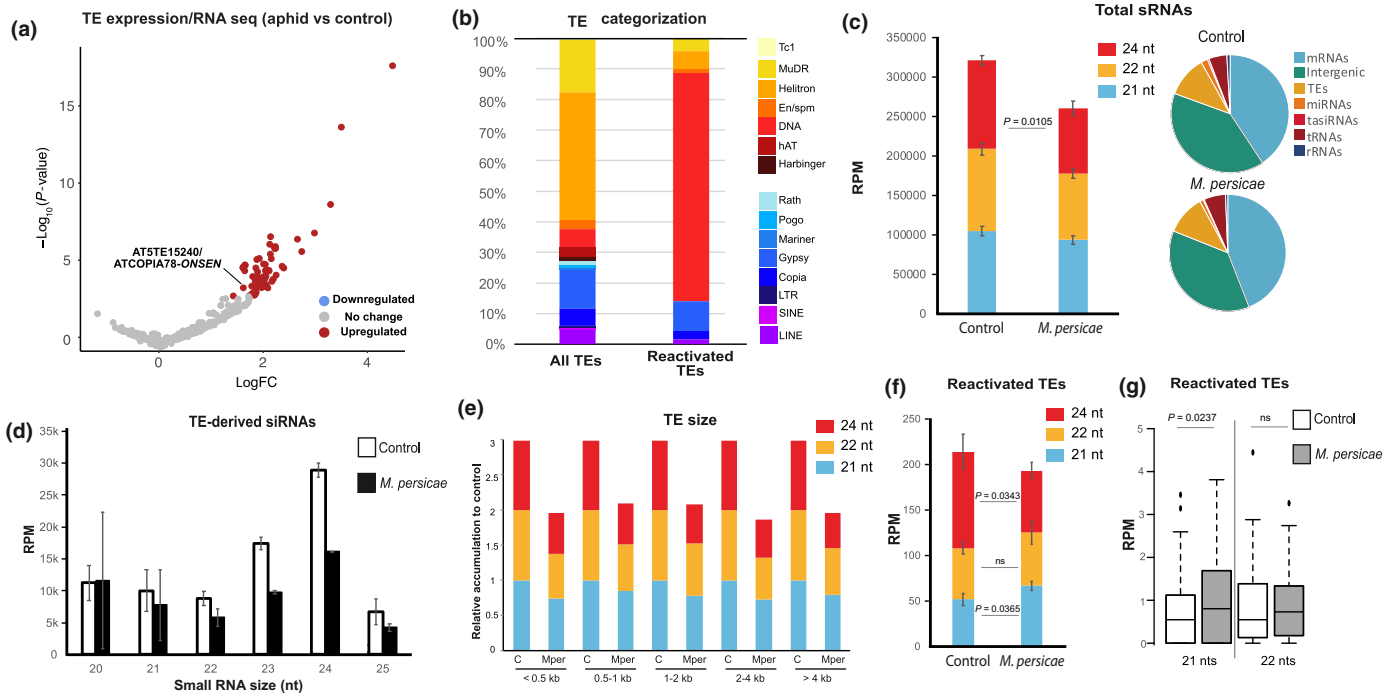


Fig. 3 Changes induced in transposable element (TE) expression by aphid feeding. (a) Volcano plot depicting *Arabidopsis* TE mRNA-seq expression in the comparison between aphid-infested and control RNA samples. Red dots indicate genes with significant upregulation. (b) Categorization of reactivated TEs (right column) compared with categorization of all TEs in the *Arabidopsis* genome (left column). (c) Accumulation of 21, 22 and 24 nt small RNAs (sRNAs) in control and aphid-infested samples from total sRNAs mapping to the *Arabidopsis* chromosomes and normalized to reads per million (RPM). Error bars indicate standard deviation (SD) of three bioreplicates. Pie charts indicate the categorization of total sRNAs from 18 to 28 nt for the categories indicated. The *P*-value was calculated using an unpaired *t*-test. (d) TE-derived sRNA profiles of control and stressed samples normalized to RPM. Error bars indicate SD of three bioreplicates. (e) Relative accumulation of 21, 22 and 24 nt sRNAs in control (C) and aphid-infested samples (Mper) for TEs of different sizes. Values shown are relative to control, where accumulation values for each sRNA category were set to 1. (f) Accumulation of 21, 22 and 24 nt sRNAs in control (C) and aphid-infested samples (Mper) for reactivated TEs. Values are shown in RPM. Error bars indicate the SD of three bioreplicates. The *P*-value was calculated using an unpaired *t*-test. (g) Box plot of 21 and 22 nt sRNA accumulation values per TE member for reactivated TEs in control and *Myzus persicae* samples. Whiskers extend to 5th and 95th percentiles. *P*-values were calculated using a paired *t*-test. FC, fold-change.

TEs for each methylation context revealed that, overall, no dramatic differences existed between the control and aphid-infested samples in any of the profiles for each methylation context (Fig. 4b). This is expected, as aphids cause very subtle wounding as a result of their feeding strategy.

To identify specific regions in the genome harboring differential methylation upon aphid feeding, we determined DMRs (Catoni *et al.*, 2018). This analysis revealed the presence of 2125 statistically significant DMRs for all the DNA methylation contexts and associated both with hypo- and hypermethylation (false discovery rate < 0.05; Figs S4a, 4e; Table S6). The CHG context had the greatest amount of DMRs (1123) followed by CG (691) and CHH (311). Furthermore, while CG DMRs were both present at genes and TEs, most of the CHG and CHH DMRs were associated with TEs (Fig. 4c). TEs located at DMRs were mostly the same TEs that lose 24 nt sRNAs (Fig. 4d). DMRs in the CG context have low CHG and CHH methylation values and the changes observed in these contexts during aphid feeding were not significant (Fig. 4e), pointing to their association with gene body methylation (Fig. 4c). On the other hand, DMRs in the CHG and CHH contexts are highly dynamic and experienced significant changes in different methylation contexts (especially in the CHG and CHH contexts) in the regions that experienced hypo- and hypermethylation (Fig. 4e). Owing to the tight association between CHG and CHH methylation with H3K9me2 (Du *et al.*, 2015), this might indicate that a strong reorganization of heterochromatin takes place in these regions upon aphid feeding.

The relative low number of DMRs identified and the lack of overall changes in the global profiles of DNA methylation indicated that methylation changes could take place only in specific regions. To test if DMRs were associated with particular histone marks, we retrieved public datasets of different histone modification coverage in *Arabidopsis* somatic tissues (Luo *et al.*, 2013) and checked the enrichment of those histone marks in our identified DMRs. Hypomethylated DMRs in the CHH context showed enrichment in the permissive mark H3K18ac ($P = 0.0174$, calculated using an unpaired *t*-test) while simultaneously showing low amounts of the repressive marks H3K27me3 and H3K9me2 (although these changes were not statistically significant) when compared with hypermethylated DMRs (Figs 4f, S4b–d). This indicated that removal of CHH methylation during aphid infestation only took place at regions of the genome that had a high level of permissive histone marks and a low level of repressive

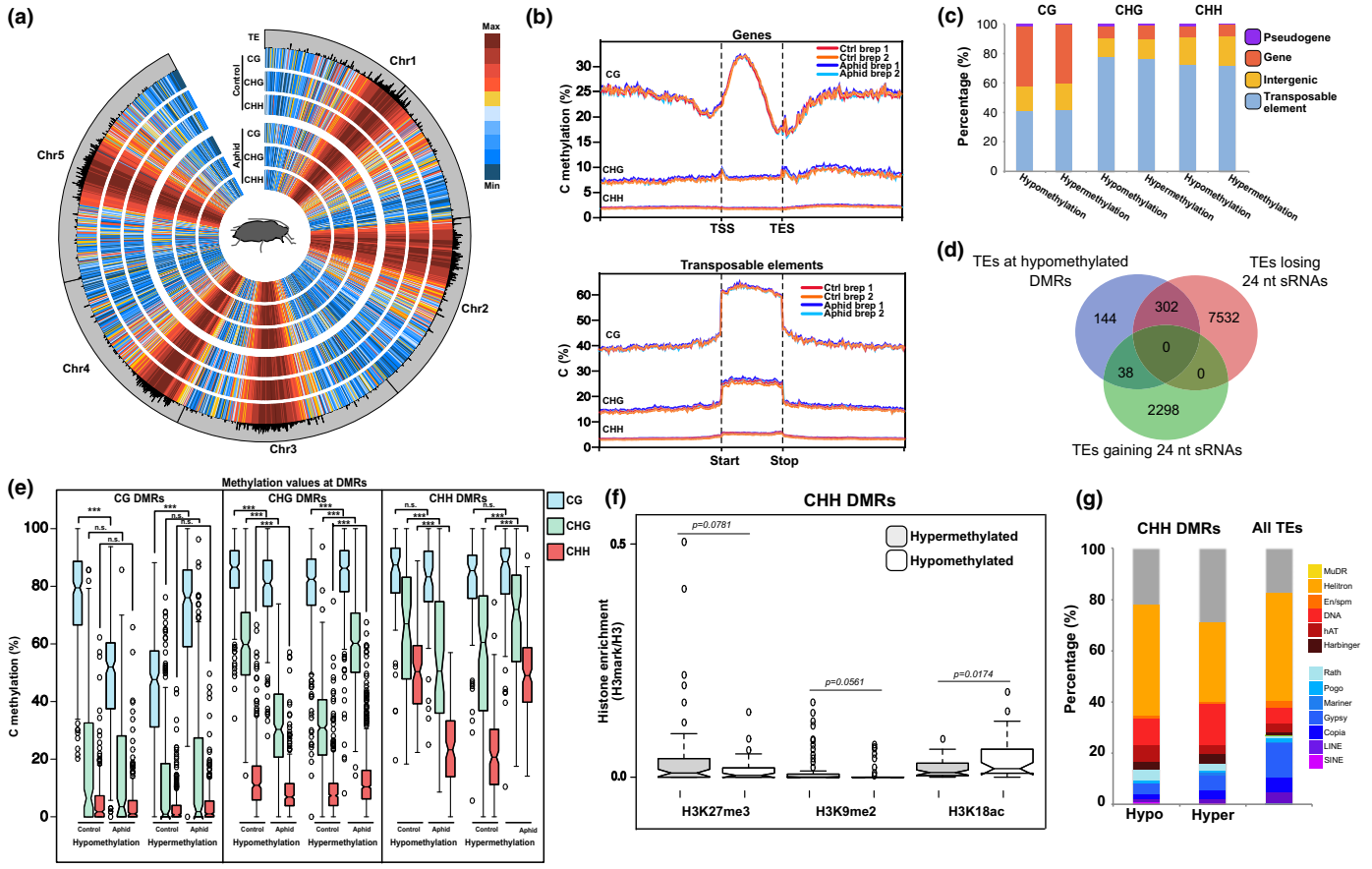


Fig. 4 DNA methylation changes induced by aphid feeding. (a) *Arabidopsis* genome-wide methylation levels for each of the cytosine (C) methylation contexts (CG, CHG and CHH) in control and aphid-infested samples. (b) DNA methylation coverage for genes and transposable elements (TEs) for each C methylation context. (c) Hypermethylation and hypomethylation differentially methylated regions (DMRs) identified for each C methylation context. (d) DMR colocalization with different genomic entities. (e) C methylation values at hypermethylation and hypomethylation DMRs for each methylation context. Asterisks indicate the different levels of significance between the comparisons (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$). The *P*-value was calculated using an unpaired *t*-test. (f) H3K27me3, H3K9me2 and H3K18ac enrichment relative to H3 for hypermethylated and hypomethylated DMRs. Whiskers extend to 5th and 95th percentiles. *P*-values were calculated using an unpaired *t*-test. (g) Categorization of TEs colocalizing with CHH hypermethylation and hypomethylation DMRs in comparison to all the TEs in the TAIR10 *Arabidopsis* genome.

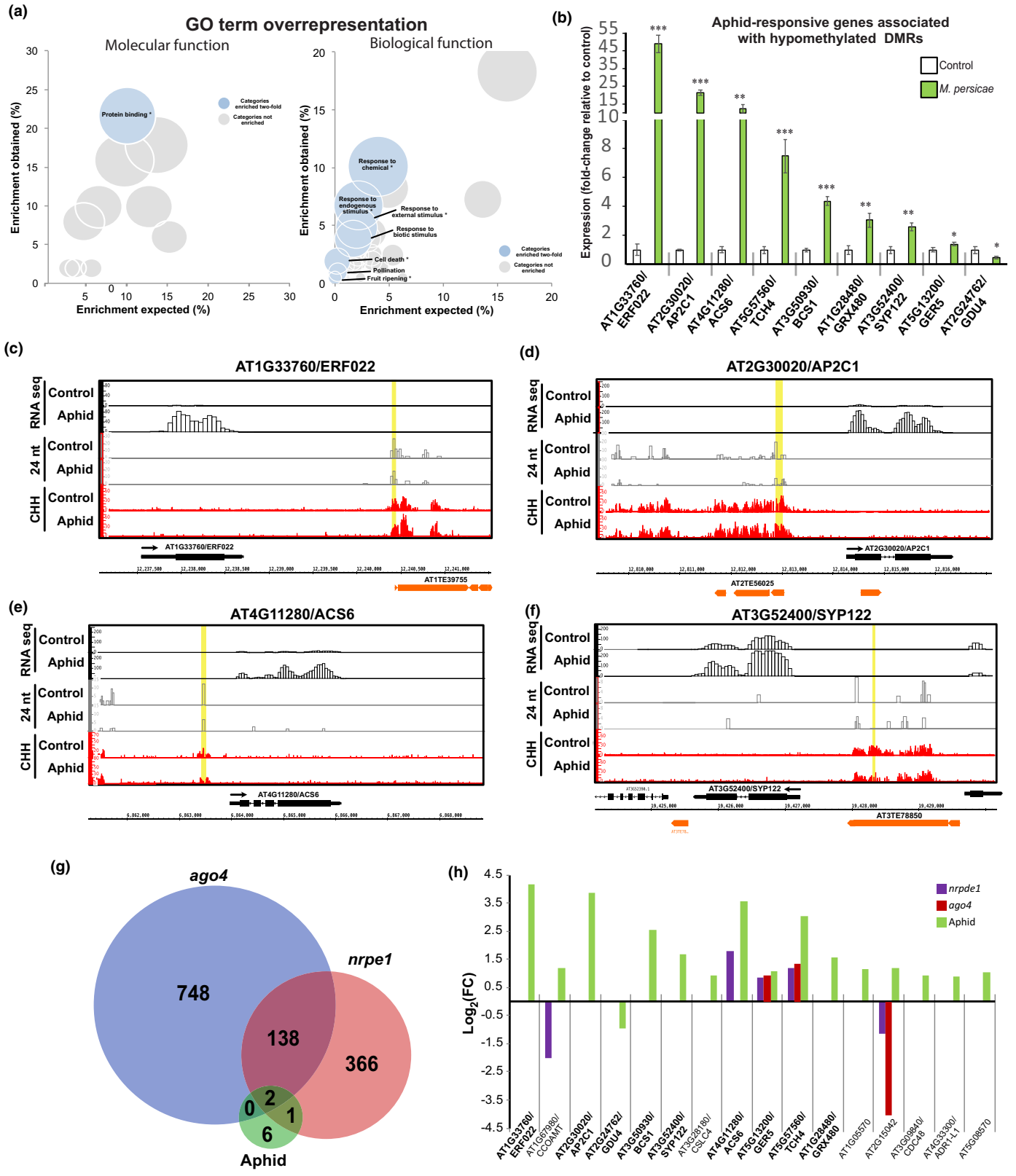
histone marks. Furthermore, hypomethylated CHH DMRs showed an enrichment in *Rath* elements and significant depletion of *Gypsy* elements (two-tailed $P < 0.05$ calculated by a Fisher exact test compared with presence of those elements against the whole genome; Fig. 4g). Therefore, upon aphid feeding, very localized methylation changes take place, mainly associated with epigenetic labile TE regions.

Stress-induced changes in methylation are associated with expression changes in defense-associated genes

Changes in TE methylation can influence the expression of neighboring genes (Wang *et al.*, 2013). To test if the identified DMRs could influence gene expression during aphid feeding, we obtained the list of neighbor genes within a 4 kb window (2 kb

Fig. 5 Transcriptional changes associated with differentially methylated regions (DMRs). (a) Bubble graph depicting the gene ontology (GO) term overrepresentation test for upregulated *Arabidopsis* genes grouped by molecular (left panel) or biological function (right panel). Bubbles in blue show GO categories enriched two-fold or more. Statistically significant categories ($P < 0.05$ determined by Fisher's exact test) are indicated with an asterisk. (b) Quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis of mRNA expression in control and aphid samples for RNA sequencing-identified aphid-responsive genes. Error bars represent the SD values for the three bioreplicates analyzed. The *P*-value is the result of a standard *t*-test with two tails and unequal variance. Asterisks indicate the different levels of significance between the comparisons (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$). (c–f) Examples of upregulated genes associated with DMRs and confirmed by RT-qPCR. The localization of statistically significant DMRs is highlighted in yellow. (g) Venn diagram depicting overlap between differentially and significantly expressed genes in *polv (nrpe1)*, *ago4* and DMR-associated genes in aphid-infested samples corroborated by RT-qPCR. (h) Expression of DMR-associated differentially expressed genes in *polv (nrpe1)*, *ago4* and aphid-infested samples. Only values with significant difference in expression ($P < 0.05$) in RNA-sequencing data are shown. Genes highlighted in bold were confirmed as significantly overexpressed in aphid samples through RT-qPCR. FC, fold-change.

upstream and downstream) for each DMR. This strategy identified 1010 genes associated with hypermethylated DMRs and 661 genes associated with hypomethylated DMRs (Table S7). As hypomethylation is expected to affect gene expression we focused our analysis on this category. Genes located in the proximities of hypomethylated DMRs were associated with oxygen binding,



translation regulator activity, nuclease and motor activity, and fruit ripening and cell death when associated by biological function (> 1.5-fold upregulation; Fig. S6a,b; see later). When the GO categories were restrained to genes that showed a significant change of expression ($P < 0.05$, 16 genes), we obtained an enrichment of genes with protein-binding activity functions when grouped by molecular function (fruit ripening, cell death, pollination) (> two-fold upregulation and $P = 0.0075$ calculated with Fisher's exact test), and response to endogenous, chemical, external and biotic stimulus when grouped by biological function (> two-fold upregulation categories with significant enrichment; $P < 0.05$ values are indicated in the figure with an asterisk; Fig. 5a). We further confirmed the significant change in expression of nine of those 16 genes by RT-qPCR (Fig. 5b).

We identified several significantly overexpressed genes ($P < 0.05$) located in the proximity of CHH hypomethylated DMRs that were related to plant defense (Figs 5c–f, S5; Table S8). These genes included *AP2C1*, a PP2C-type phosphatase that modulates innate immunity (Schweighofer *et al.*, 2007); *ACS6*, a 1-aminocyclopropane-1-carboxylic acid synthase that is a rate-limiting enzyme that catalyzes the committing step of ethylene biosynthesis (Joo *et al.*, 2008); *SYP122*, a Qa-SNARE protein that drives vesicle fusion and is important for cell growth and expansion and pathogen defense (Waghmare *et al.*, 2018); *GER5*, a stress-responsive glucosyltransferase, rab-like GTPase activator and myotubularin domain protein involved in ABA-mediated stress responses (Baron *et al.*, 2014); and the ethylene response factor *ERF022*, which belongs to the IIIa subgroup of the ERF subfamily which is associated with the response to stress (Nakano *et al.*, 2006).

Next, to explore the potential epigenetic regulation of these genes, we analyzed their expression in epigenetic mutants (not exposed to aphid feeding). We used RNA-sequencing public datasets from Pol V and AGO4 mutants (Zhu *et al.*, 2013; Rowley *et al.*, 2017). Pol V and AGO4 are components of the RdDM pathway that produces sRNAs to target genomic regions and introduces DNA methylation (Matzke & Mosher, 2014). Pol V produces long noncoding transcripts that guide Pol IV-derived 24 nt sRNAs loaded into AGO4 to chromatin (Wierzbicki *et al.*, 2009). Mutations in AGO4 or PolV impair RdDM-dependent methylation, especially in the CHH context, and 82% of loci regulated by Pol V or Pol IV are also regulated by AGO4/AGO6 (Duan *et al.*, 2015). Differentially expressed genes associated with DMRs and confirmed by RT-qPCR were significantly enriched in the portion of genes regulated by the RdDM pathway components AGO4 and/or Pol V (33.3% overlap, two-tailed $P < 0.0001$ calculated by a χ^2 test with Yates correction; Fig. 5g, h). Although some genes showed a similar expression pattern between the RdDM mutants and the aphid-infested samples (e.g. *GER5*, *ACS6*; Fig. 5h) others showed opposing patterns of expression between the aphid-infested samples and the RdDM mutants (notably *CCOAMT* and *AT2G15042*). This different expression pattern led us to question whether the expression of these genes could be regulated by TFs that were not overexpressed in the RdDM mutants. The analysis of TF-binding motifs present in the DMRs of differentially expressed genes showed that several TF-binding motifs were highly enriched, including B3 binding domain-containing TFs such as B3/ARF, AP2/B3 and B3 (20.65-, 11.8- and 8.6-fold enrichment, respectively; Fig. S6c,e). Several TFs of the B3 subfamily belonging to the ERF/AP2 TF family were differentially expressed in the aphid-infested samples, while they did not show this pattern of expression in RdDM mutants (Table S4; Fig. S6d). This indicated that differential expression of TFs probably leads to the observed differences in the expression pattern between aphid-infested samples and RdDM mutants. Overall, our data indicate that DNA methylation changes are associated with gene expression changes, probably in combination with TF-induced expression.

Epigenetic mutants show enhanced defense against aphids

Finally, we tested whether different *Arabidopsis* mutants defective in epigenetic regulation were resistant to aphid infestation. For this, we analyzed aphid no-choice settling where 10 aphids were transferred to a random caged leaf (Fig. 6a). We performed this test in different mutants, including the histone remodeler *DDM1*, the triple mutant defective in maintenance of nonCG methylation *ddc* (*drm1 drm2 cmt3*), the main subunit of the principal factor of the RdDM pathway RNA Pol IV (*nrpd1*), the main subunit of the other principal factor of the RdDM pathway RNA Pol V (*nrpe1*), the main ARGONAUTE protein introducing methylation in the DNA through the canonical RdDM pathway ARGONAUTE 4 (*ago4*), and the H3K9me2 methyltransferase *KYP* (Fig. 6b). All these mutants are known to affect DNA methylation/histone modifications genome-wide and a preliminary analysis of CHH methylation changes in our identified DMRs indicated that, indeed, all of them affect CHH methylation levels in CHH DMRs and in DMRs associated with differentially expressed genes (Fig. S7b,c). Our aphid-settling analysis indicated that, from these components, mutations in *nrpd1* (the largest subunit of Pol IV) and *kyp* show a reduced number of aphids settled, although only *kyp* had a significant decrease (Fig. 6b). Interestingly, we observed the natural overexpression of aphid-resistance genes in *kyp* (Fig. 6c). Furthermore, this resistance is not connected to the leaf phenotype for each of the mutants analyzed here (Fig. S7a).

This indicated that, first, heterochromatin maintenance (regulated by DDM1) and maintenance of nonCG methylation (*ddc*) were not fundamental to elicit a defense response against aphid feeding. Second, our result indicated that the roles of KYP in the regulation of H3K9me2 and CHG methylation (Jackson *et al.*, 2002) and/or its uncharacterized role in the maintenance of CHH methylation (Stroud *et al.*, 2013) were an important part of the defense response against aphid infestation. This result correlates with our observed increase in transcription of centromeric TEs and reduction of sRNAs in centromeric and pericentromeric regions (rich in H3K9me2), and the observed changes in CHH and CHG methylation (tightly associated with H3K9me2). KYP has been previously associated with the regulation of the defense against geminiviruses (Raja *et al.*, 2008; Castillo-Gonzalez *et al.*, 2015; Sun *et al.*, 2015) and the maintenance of β-aminobutyric

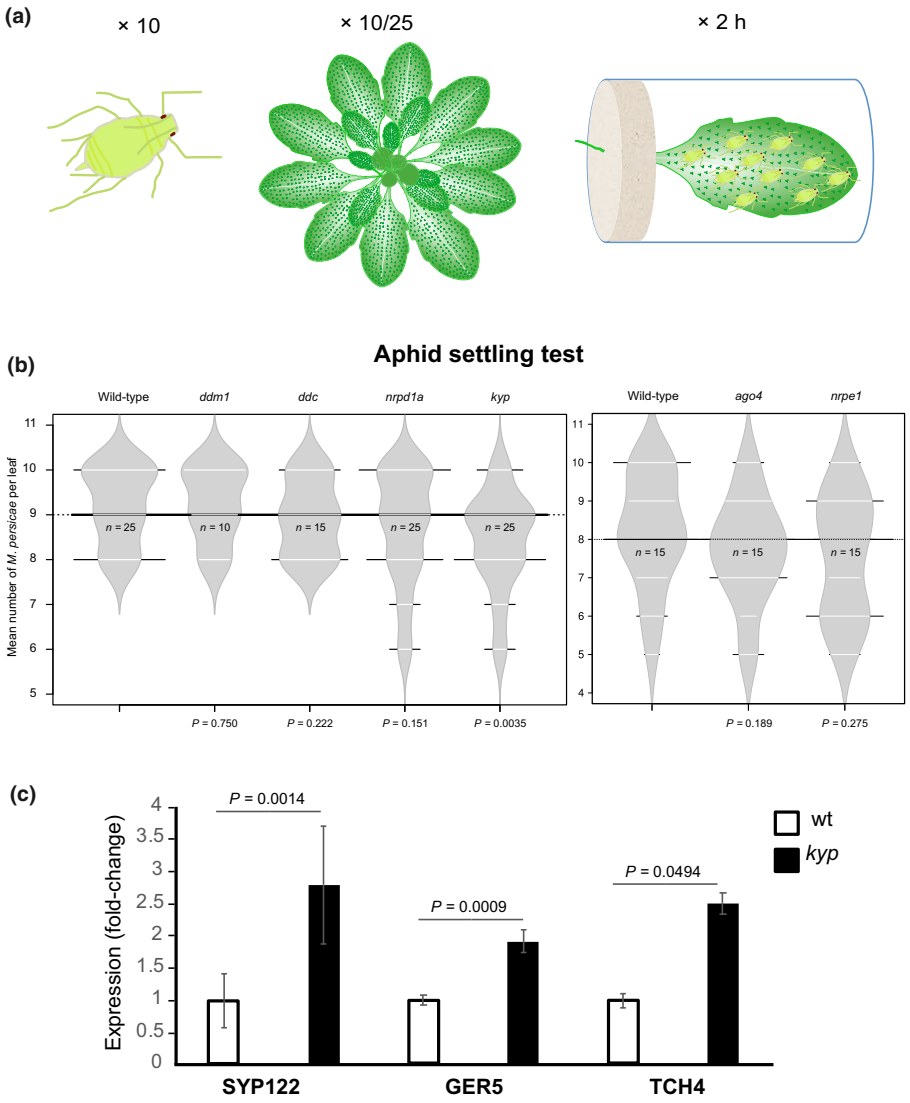


Fig. 6 Epigenetic mutants are resistant to aphid settlement. (a) Depiction of the aphid settlement experiment carried out in our analysis. In brief, 10 aphids were moved to a single caged leaf (attached to the plant) from 10/25 individual *Arabidopsis* plants. (b) Aphid settlement test in different epigenetic mutants. The *P*-values shown were calculated using an unpaired *t*-test. *n*, number of individuals analyzed. (c) Quantitative reverse transcription polymerase chain reaction (RT-qPCR) analysis of mRNA expression in wild-type and *kyp* for RNA sequencing-identified aphid-responsive genes. Error bars represent the SD values for the three bioreplicates analyzed. The *P*-value is the result of a standard *t*-test with two tails and unequal variance.

acid (BABA)-induced priming of the salicylic acid (SA)-dependent defense response (Luna *et al.*, 2014). In summary, our proof-of-concept analysis indicates that mutants in different layers of epigenetic regulation do indeed show enhanced resistance against aphid settlement.

Discussion

Organisms monitor environmental conditions and adapt their development according to them. Plants have developed elegant mechanisms of gene regulation adapted to their sessile nature. One such mechanism is epigenetic regulation, which could maintain modified transcriptional states through cell division and be reversible once the trigger condition disappears. Although it has been widely proposed that epigenetic regulation is an important part of the stress response, we lack a comprehensive knowledge of the genomic loci that are susceptible to those epigenetic changes and their variability between stresses. Here, we demonstrated that aphid feeding induced changes in the epigenetic regulation of the plant genome and that these changes correlated with the transcriptional response. Our data suggest that these epigenetic changes were taking place mainly in TEs. We hypothesize that these changes could be important for recruiting TFs that in turn affect the expression of a specific set of defense genes. This will explain why, despite having a relatively high number of DMRs (Fig. 4), only a very small subset enriched in specific TF-binding motifs was associated with transcriptional changes (Figs 5, S5, S6). An alternative hypothesis to this is that DNA methylation

changes took place downstream of TF binding, a situation that has been described in human dendritic cells (Pacis *et al.*, 2019). Nevertheless, the presence in our analysis of a high number of DMRs without effects at the transcriptional level argues against this hypothesis. As our analysis focused only on a single data point, it is also plausible that the low correlation between the transcriptional and DNA methylation changes could be a result of a temporal separation between both events, as previously described under phosphate starvation in rice (Secco *et al.*, 2015).

Despite their subtle wounding strategy, aphid feeding activates hormonal signals that trigger the reprogramming of the plant transcriptome (Moran *et al.*, 2002; De Vos *et al.*, 2005; Couldridge *et al.*, 2007; Kusnierczyk *et al.*, 2007; Gao *et al.*, 2010). In our study, the transcriptional changes identified by RNA sequencing showed enrichment in genes associated with TF-related activities (Fig. 2). These TFs include AR2/ERF and WRKY TFs, which have previously been associated with the transcriptional response against aphid infestation (Foyer *et al.*, 2015; Kloth *et al.*, 2016). Our analysis of the transcriptional and post-transcriptional regulation of TEs during aphid infestation indicated that TEs were reactivated during aphid feeding, although to a lower extent than expected from our initial study of similar experiments analyzed with the ATH1 microarray data (Fig. 3). One of the reasons for this divergence in the number of reactivated TEs between both analyses could be a result of the nature of the RNA used in the two experiments, that is, total RNA in De Vos *et al.* (2005) against purified mRNA in our study. This aphid-reactivated TE group included the reactivation of the Copia retrotransposon ONSEN, which is known to activate and transpose in *Arabidopsis* plants exposed to different stresses (Cavrak *et al.*, 2014; Matsunaga *et al.*, 2015; Ito *et al.*, 2016). We further detected that TEs experienced a decrease in the activity of the RdDM pathway translated in a loss of 24 nt sRNAs, which led to their transcriptional reactivation. Several of these TEs have a centromeric localization, which correlated with the transposon families losing the majority of 24 nt sRNAs. Furthermore, reactivated TEs accumulated higher levels of 21 and 22 nt sRNAs, which is a signature of TE transcriptional reactivation in epigenetic mutants like *met1* or *ddm1* (McCue *et al.*, 2012).

The changes of TE activity at the transcriptional level prompted us to profile the genome-wide methylation changes under aphid infestation (Fig. 4). Our genome-wide analysis of DNA methylation changes induced by aphid feeding showed that methylation changes happened primarily at genes (in the CG context) and TEs (in the CHG and CHH contexts). CHH hypomethylated DMRs took place only at epigenetically labile regions characterized by low levels of the repressive histone marks H3K27me3 and H3K9me2 and high levels of the transcriptionally permissive mark H3K18ac. As expected, CHH hypomethylated DMRs were predominantly depleted of *Gypsy* TEs, which are long centromeric elements with relative low influence on gene expression (Lermontova *et al.*, 2015). An analysis of the presence of genes in a 4 kb window for CHH hypomethylated DMRs showed the potential transcriptional changes associated with these DMRs. Between differentially expressed genes associated with DRMs, we found several genes related to the defense response at different levels, such as *AP2C1* (Schweighofer *et al.*, 2007), *ACS6* (Joo *et al.*, 2008), *SYP122* (Waghmare *et al.*, 2018), *GER5* (Baron *et al.*, 2014) and the ethylene response factor *ERF022* (Nakano *et al.*, 2006) (Figs 5, S5). A percentage (33.3%) of the differentially expressed genes associated with DMRs were also differentially expressed in *nrpe1* and/or *ago4* mutants, indicating an influence of the RdDM pathway in the regulation of this response (exemplified by *SYP122* in the data showed in Fig. 5f). Together with this observation, we found that DMRs associated with differentially expressed genes showed an enrichment in binding motifs for certain families of TFs, including the AP2-ERF/B3, which has seven members significantly upregulated upon aphid infestation (Fig. S6d). These TFs showed a modest upregulation in the *nrpe1* mutant and none in an *ago4* mutant, which could be one of the reasons why the transcriptional response differed between aphid-infested samples and RdDM mutants. While aphid feeding induced the expression of several TFs, RdDM mutants lack the presence of aphid-induced TFs that would stimulate the defense transcriptional response. As a proof-of-concept, we tested whether *Arabidopsis* mutants defective in DNA and histone methylation had a differential susceptibility to aphid infestation (Fig. 6). Our analysis indicated that mutations in *Pol IV* and *KYP* show increased resistance to aphid settling, confirming the importance of epigenetic regulation in the response against aphids. In *Arabidopsis,* defense genes are located in pericentromeric regions which are densely populated by TEs (Meyers *et al.*, 2003). KYP and Pol IV have a known role in the repression of TEs, so we speculate that their lack of function can also facilitate the transcription of genes located in the proximities of TEs. Our data also indicate that *kyp* has a natural reactivation of some of the aphid-responsive genes. In *kyp* and *nrpd1* mutants, the enhanced activation of defense genes (via transcription or binding of TFs) will explain the increased defense against aphid feeding. Indeed, most of the differentially expressed genes with a proximal CHH DMR identified in our analysis had a TE in the proximities of their regulatory regions (Fig. 5). We hypothesize that *kyp* might show increased resistance compared with other mutants as a result of its reduced methylation level genome-wide, but also as a result of its reduced level of the repressive histone mark H3K9me2. Chromatin marks are known to modulate transcription through influence over TF binding sites (Wu *et al.*, 2019). In *kyp* the low levels of H3K9me2 might allow for a more favorable environment for aphid-responsive TF binding.

It is tempting to speculate that together with the downregulation of the epigenetic silencing at DMRs, the observed overexpression of mobile mRNAs and decrease of 24nt sRNAs would trigger transcriptional or post-transcriptional changes on gene expression at distal tissues, other than leaves, including the precursors of the reproductive structures. Some herbivore insects, like *Pieris rapae*, are known to trigger a defense phenotype in the next generation (Rasmann *et al.*, 2012). TE silencing is reinforced in the shoot apical meristem (SAM) by the RdDM pathway, which leads to the correct transmission of the right epigenetic states for TEs during vegetative growth (Baubec *et al.*, 2014). A potential lack of mobile 24 nt (Molnar *et al.*, 2010) or

21 nt (Dunoyer *et al.*, 2010) TE-derived small interfering RNAs in the SAM or the reproductive structures could lead to epigenetic states that could be inherited. Further analysis of the effect of localized stresses on distal tissues and their offspring could shed light onto the existence of such an elegant overlapping of pathways potentially regulating transgenerational inheritance.

In summary, the evidence presented in our work indicates that changes in epigenetic control were associated with the defense response against aphid infestation in *A. thaliana*. Intriguingly, this response is more complex than previously thought and may involve the interplay between epigenetic and transcriptional regulation. Our work exemplifies the importance of epigenetic regulation in the stress response and the epigenetic plasticity of plant genomes subjected to stress.

Acknowledgements

Research in the Martinez group is supported by SLU, the Carl Tryggers Foundation and the Swedish Research Council (VR 2016-05410). Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory. The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation.

Author contributions

RKS, VN and GM designed the experiment. CMP contributed material. MLA, DM, VS and GM performed the experiments. JLR-V and GM carried out bioinformatic processing of the data. GM did the data analysis. GM wrote the manuscript. All authors interpreted the data and thoroughly checked the manuscript. MLA and DM contributed equally to this work.

ORCID

Maria Luz Annacondia https://orcid.org/0000-0001-7998-8362
Dimitrije Markovic https://orcid.org/0000-0003-3030-4155
German Martinez https://orcid.org/0000-0002-5215-0866
Velemir Ninkovic https://orcid.org/0000-0001-9276-7169
Corné M. J. Pieterse https://orcid.org/0000-0002-5473-4646
Juan Luis Reig-Valiente https://orcid.org/0000-0001-9894-0196
Vassilis Scaltsoyiannes https://orcid.org/0000-0003-4155-9540
R. Keith Slotkin https://orcid.org/0000-0001-9582-3533

Data availability

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession no. GSE144181. Reviewers can access the data using the reviewer token: qvqdmcwefbolhkf.

References

Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Gruning BA *et al.* 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research* **46**: W537–W544.

Agorio A, Vera P. 2007. ARGONAUTE4 is required for resistance to *Pseudomonas syringae* in *Arabidopsis*. *Plant Cell* **19**: 3778–3790.

An J, Lai J, Sajjanhar A, Batra J, Wang C, Nelson CC. 2015. J-Circos: an interactive Circos plotter. *Bioinformatics* **31**: 1463–1465.

Anders S, Pyl PT, Huber W. 2014. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.

Annacondia ML, Mageroy MH, Martinez G. 2018. Stress response regulation by epigenetic mechanisms: changing of the guards. *Physiologia Plantarum* **162**: 239–250.

Annacondia ML, Martinez G. 2019. Chapter 12 – Plant models of transgenerational epigenetic inheritance. In: Tollefsbol TO, ed. *Transgenerational epigenetics, 2nd edn.* Cambridge, UK: Academic Press, 263–282.

Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS. 2016. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Research* **44**: W147–W153.

Baron KN, Schroeder DF, Stasolla C. 2014. GEm-Related 5 (GER5), an ABA and stress-responsive GRAM domain protein regulating seed development and inflorescence architecture. *Plant Science* **223**: 153–166.

Baubec T, Finke A, Mittelsten Scheid O, Pecinka A. 2014. Meristem-specific expression of epigenetic regulators safeguards transposon silencing in *Arabidopsis*. *EMBO Reports* **15**: 446–452.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate – a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B–Statistical Methodology* **57**: 289–300.

Bown DP, Wilkinson HS, Gatehouse JA. 1997. Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochemistry and Molecular Biology* **27**: 625–638.

Boyko A, Blevins T, Yao Y, Golubov A, Bilichak A, Ilnytskyy Y, Hollunder J, Meins F Jr, Kovalchuk I. 2010. Transgenerational adaptation of *Arabidopsis* to stress requires DNA methylation and the function of Dicer-like proteins. *PLoS ONE* **5**: e9514.

Boyko A, Kathiria P, Zemp FJ, Yao Y, Pogribny I, Kovalchuk I. 2007. Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). *Nucleic Acids Research* **35**: 1714–1725.

Boyko A, Kovalchuk I. 2011. Genetic and epigenetic effects of plant-pathogen interactions: an evolutionary perspective. *Molecular Plant* **4**: 1014–1023.

Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Reforgiato-Recupero G, Martin C. 2012. Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. *Plant Cell* **24**: 1242–1255.

Castillo-Gonzalez C, Liu X, Huang C, Zhao C, Ma Z, Hu T, Sun F, Zhou Y, Zhou X, Wang XJ *et al.* 2015. Geminivirus-encoded TrAP suppressor inhibits the histone methyltransferase SUVH4/KYP to counter host defense. *eLife* **4**: e06671.

Catoni M, Tsang JM, Greco AP, Zabet NR. 2018. DMRcaller: a versatile R/Bioconductor package for detection and visualization of differentially methylated regions in CpG and non-CpG contexts. *Nucleic Acids Research* **46**: e114.

Cavrak VV, Lettner N, Jamge S, Kosarewicz A, Bayer LM, Scheid OM. 2014. How a retrotransposon exploits the plant's heat stress response for its activation. *PLoS Genetics* **10**: e1004115.

Couldridge C, Newbury HJ, Ford-Lloyd B, Bale J, Pritchard J. 2007. Exploring plant responses to aphid feeding using a full Arabidopsis microarray reveals a small number of genes with significantly altered expression. *Bulletin of Entomological Research* **97**: 523–532.

Darwin C. 1899. *The various contrivances by which orchids are fertilized by insects.* London: John Murray.

De Vos M, Jander G. 2009. *Myzus persicae* (green peach aphid) salivary components induce defence responses in *Arabidopsis thaliana*. *Plant, Cell & Environment* **32**: 1548–1560.

De Vos M, Van Oosten VR, Van Poecke RM, Van Pelt JA, Pozo MJ, Mueller MJ, Buchala AJ, Metraux JP, Van Loon LC, Dicke M *et al.* 2005. Signal signature and transcriptome changes of *Arabidopsis* during pathogen and insect attack. *Molecular Plant–Microbe Interactions* **18**: 923–937.

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

Dooner HK, Robbins TP, Jorgensen RA. 1991. Genetic and developmental control of anthocyanin biosynthesis. *Annual Review of Genetics* **25**: 173–199.

Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR. 2012. Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences, USA* **109**: E2183–E2191.

Du J, Johnson LM, Jacobsen SE, Patel DJ. 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nature Reviews Molecular Cell Biology* **16**: 519–532.

Duan CG, Zhang H, Tang K, Zhu X, Qian W, Hou YJ, Wang B, Lang Z, Zhao Y, Wang X *et al.* 2015. Specific but interdependent functions for *Arabidopsis* AGO4 and AGO6 in RNA-directed DNA methylation. *EMBO Journal* **34**: 581–592.

Dunoyer P, Schott G, Himber C, Meyer D, Takeda A, Carrington JC, Voinnet O. 2010. Small RNA duplexes function as mobile silencing signals between plant cells. *Science* **328**: 912–916.

Foyer CH, Verrall SR, Hancock RD. 2015. Systematic analysis of phloem-feeding insect-induced transcriptional reprogramming in *Arabidopsis* highlights common features and reveals distinct responses to specialist and generalist insects. *Journal of Experimental Botany* **66**: 495–512.

Gao LL, Kamphuis LG, Kakar K, Edwards OR, Udvardi MK, Singh KB. 2010. Identification of potential early regulators of aphid resistance in *Medicago truncatula* via transcription factor expression profiling. *New Phytologist* **186**: 980–994.

Gutzat R, Mittelsten SO. 2012. Epigenetic responses to stress: triple defense? *Current Opinion in Plant Biology* **15**: 568–573.

Ito H, Kim JM, Matsunaga W, Saze H, Matsui A, Endo TA, Harukawa Y, Takagi H, Yaegashi H, Masuta Y *et al.* 2016. A stress-activated transposon in *Arabidopsis* induces transgenerational abscisic acid insensitivity. *Scientific Reports* **6**: 23181.

Jackson JP, Lindroth AM, Cao X, Jacobsen SE. 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416**: 556–560.

Joo S, Liu Y, Lueth A, Zhang S. 2008. MAPK phosphorylation-induced stabilization of ACS6 protein is mediated by the non-catalytic C-terminal domain, which also contains the cis-determinant for rapid degradation by the 26S proteasome pathway. *The Plant Journal* **54**: 129–140.

Kathiria P, Sidler C, Golubov A, Kalischuk M, Kawchuk LM, Kovalchuk I. 2010. Tobacco mosaic virus infection results in an increase in recombination frequency and resistance to viral, bacterial, and fungal pathogens in the progeny of infected tobacco plants. *Plant Physiology* **153**: 1859–1870.

Kidwell MG, Lisch D. 1997. Transposable elements as sources of variation in animals and plants. *Proceedings of the National Academy of Sciences, USA* **94**: 7704–7711.

Kloth KJ, Wiegers GL, Busscher-Lange J, van Haarst JC, Kruijer W, Bouwmeester HJ, Dicke M, Jongsma MA. 2016. AtWRKY22 promotes susceptibility to aphids and modulates salicylic acid and jasmonic acid signalling. *Journal of Experimental Botany* **67**: 3383–3396.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**: 1571–1572.

Kusnierczyk A, Winge P, Midelfart H, Armbruster WS, Rossiter JT, Bones AM. 2007. Transcriptional responses of *Arabidopsis thaliana* ecotypes with different glucosinolate profiles after attack by polyphagous *Myzus persicae* and oligophagous *Brevicoryne brassicae*. *Journal of Experimental Botany* **58**: 2537–2552.

Lamke J, Baurle I. 2017. Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants. *Genome Biology* **18**: 124.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**: R25.

Law JA, Jacobsen SE. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics* **11**: 204–220.

Lermontova I, Sandmann M, Mascher M, Schmit AC, Chaboute ME. 2015. Centromeric chromatin and its dynamics in plants. *The Plant Journal* **83**: 4–17.

Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, Lavine K, Mittal V, May B, Kasschau KD *et al.* 2004. Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**: 471–476.

Lippman Z, May B, Yordan C, Singer T, Martienssen R. 2003. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biology* **1**: E67.

Lisch D. 2009. Epigenetic regulation of transposable elements in plants. *Annual Review of Plant Biology* **60**: 43–66.

Lopez A, Ramirez V, Garcia-Andrade J, Flors V, Vera P. 2011. The RNA silencing enzyme RNA polymerase v is required for plant immunity. *PLoS Genetics* **7**: e1002434.

Louis J, Shah J. 2013. *Arabidopsis thaliana-Myzus persicae* interaction: Shaping the understanding of plant defense against phloem-feeding aphids. *Frontiers in Plant Science* **4**: 213.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Luna E, Bruce TJ, Roberts MR, Flors V, Ton J. 2012. Next-generation systemic acquired resistance. *Plant Physiology* **158**: 844–853.

Luna E, Lopez A, Kooiman J, Ton J. 2014. Role of NPR1 and KYP in long-lasting induced resistance by beta-aminobutyric acid. *Frontiers in Plant Science* **5**: 184.

Luna E, Ton J. 2012. The epigenetic machinery controlling transgenerational systemic acquired resistance. *Plant Signaling & Behavior* **7**: 615–618.

Luo C, Sidote DJ, Zhang Y, Kerstetter RA, Michael TP, Lam E. 2013. Integrative analysis of chromatin states in *Arabidopsis* identified potential regulatory mechanisms for natural antisense transcript production. *The Plant Journal* **73**: 77–90.

Matsunaga W, Ohama N, Tanabe N, Masuta Y, Masuda S, Mitani N, Yamaguchi-Shinozaki K, Ma JF, Kato A, Ito H. 2015. A small RNA mediated regulation of a stress-activated retrotransposon and the tissue specific transposition during the reproductive period in *Arabidopsis*. *Frontiers in Plant Science* **6**: 48.

Matzke MA, Mosher RA. 2014. RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nature Reviews Genetics* **15**: 394–408.

McCue AD, Nuthikattu S, Reeder SH, Slotkin RK. 2012. Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA. *PLOS Genetics* **8**: e1002474.

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. 2003. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* **15**: 809–834.

Molnar A, Melnyk CW, Bassett A, Hardcastle TJ, Dunn R, Baulcombe DC. 2010. Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. *Science* **328**: 872–875.

Moran PJ, Cheng Y, Cassell JL, Thompson GA. 2002. Gene expression profiling of *Arabidopsis thaliana* in compatible plant-aphid interactions. *Archives of Insect Biochemistry and Physiology* **51**: 182–203.

Mozgova I, Mikulski P, Pecinka A, Farrona S. 2019. Epigenetic Mechanisms of Abiotic Stress Response and Memory in Plants. In: Alvarez-Venegas R, De-la-Peña C, Casas-Mollano JA, eds. *Epigenetics in plants of agronomic importance: fundamentals and applications: transcriptional regulation and chromatin remodelling in plants*. Cham, Switzerland: Springer International, 1–64.

Nakano T, Suzuki K, Fujimura T, Shinshi H. 2006. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiology* **140**: 411–432.

Ninkovic V, Olsson U, Pettersson J. 2002. Mixing barley cultivars affects aphid host plant acceptance in field experiments. *Entomologia Experimentalis et Applicata* **102**: 177–182.

Pacis A, Mailhot-Leonard F, Tailleux L, Randolph HE, Yotova V, Dumaine A, Grenier JC, Barreiro LB. 2019. Gene activation precedes DNA demethylation in response to infection in human dendritic cells. *Proceedings of the National Academy of Sciences, USA* **116**: 6938–6943.

Panda K, Ji L, Neumann DA, Daron J, Schmitz RJ, Slotkin RK. 2016. Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biology* **17**: 170.

Prado E, Tjallingii WF. 1997. Effects of previous plant infestation on sieve element acceptance by two aphids. *Entomologia Experimentalis et Applicata* **82**: 189–200.

Raja P, Sanville BC, Buchmann RC, Bisaro DM. 2008. Viral genome methylation as an epigenetic defense against geminiviruses. *Journal of Virology* **82**: 8997–9007.

Rasmann S, De Vos M, Casteel CL, Tian D, Halitschke R, Sun JY, Agrawal AA, Felton GW, Jander G. 2012. Herbivory in the previous generation primes plants for enhanced insect resistance. *Plant Physiology* **158**: 854–863.

Renner T, Specht CD. 2013. Inside the trap: gland morphologies, digestive enzymes, and the evolution of plant carnivory in the *Caryophyllales*. *Current Opinion in Plant Biology* **16**: 436–442.

Rowley MJ, Rothi MH, Bohmdorfer G, Kucinski J, Wierzbicki AT. 2017. Long-range control of gene expression via RNA-directed DNA methylation. *PLoS Genetics* **13**: e1006749.

Schoonhoven LM, van Loon JJA, Dicke M. 2005. *Insect-plant biology*. Oxford, UK: OUP Oxford.

Schweighofer A, Kazanaviciute V, Scheikl E, Teige M, Doczi R, Hirt H, Schwanninger M, Kant M, Schuurink R, Mauch F *et al.* 2007. The PP2C-type phosphatase AP2C1, which negatively regulates MPK4 and MPK6, modulates innate immunity, jasmonic acid, and ethylene levels in *Arabidopsis*. *Plant Cell* **19**: 2213–2224.

Secco D, Wang C, Shou HX, Schultz MD, Chiarenza S, Nussaume L, Ecker JR, Whelan J, Lister R. 2015. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *eLife* **4**: e09343.

Sharma M, Banday ZZ, Shukla BN, Laxmi A. 2019. Glucose-regulated HLP1 acts as a key molecule in governing thermomemory. *Plant Physiology* **180**: 1081–1100.

Slaughter A, Daniel X, Flors V, Luna E, Hohn B, Mauch-Mani B. 2012. Descendants of primed *Arabidopsis* plants exhibit resistance to biotic stress. *Plant Physiology* **158**: 835–843.

Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA, Martienssen RA. 2009. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136**: 461–472.

Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. 2013. Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**: 352–364.

Sun YW, Tee CS, Ma YH, Wang G, Yao XM, Ye J. 2015. Attenuation of histone methyltransferase KRYPTONITE-mediated transcriptional gene silencing by Geminivirus. *Scientific Reports* **5**: 16476.

Tanaka Y, Asano T, Kanemitsu Y, Goto T, Yoshida Y, Yasuba K, Misawa Y, Nakatani S, Kobata K. 2019. Positional differences of intronic transposons in pAMT affect the pungency level in chili pepper through altered splicing efficiency. *The Plant Journal* **100**: 693–705.

Thieme CJ, Rojas-Triana M, Stecyk E, Schudoma C, Zhang W, Yang L, Minambres M, Walther D, Schulze WX, Paz-Ares J *et al.* 2015. Endogenous *Arabidopsis* messenger RNAs transported to distant tissues. *Nature Plants* **1**: 15025.

Tsuchiya T, Eulgem T. 2013. An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication. *Proceedings of the National Academy of Sciences, USA* **110**: E3535–E3543.

Underwood CJ, Henderson IR, Martienssen RA. 2017. Genetic and epigenetic variation of transposable elements in *Arabidopsis*. *Current Opinion in Plant Biology* **36**: 135–141.

Waghmare S, Lileikyte E, Karnik R, Goodman JK, Blatt MR, Jones AME. 2018. SNAREs SYP121 and SYP122 mediate the secretion of distinct cargo subsets. *Plant Physiology* **178**: 1679–1688.

Wang X, Weigel D, Smith LM. 2013. Transposon variants and their effects on gene expression in *Arabidopsis*. *PLoS Genetics* **9**: e1003255.

Warman C, Panda K, Vejlupkova Z, Hokin S, Unger-Wallace E, Cole RA, Chettoor AM, Jiang D, Vollbrecht E, Evans MMS *et al.* 2020. High expression in maize pollen correlates with genetic contributions to pollen fitness as well as with coordinated transcription from neighboring transposable elements. *PLoS Genetics* **16**: e1008462.

Wickham H. 2009. *ggplot2: elegant graphics for data analysis*. New York, NY, USA: Springer.

Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. 2009. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nature Genetics* **41**: 630–634.

Wu Z, Tang J, Zhuo J, Tian Y, Zhao F, Li Z, Yan Y, Yang R. 2019. Chromatin signature and transcription factor binding provide a predictive basis for understanding plant gene expression. *Plant and Cell Physiology* **60**: 1471–1486.

Yu A, Lepere G, Jay F, Wang J, Bapaume L, Wang Y, Abraham AL, Penterman J, Fischer RL, Voinnet O *et al.* 2013. Dynamics and biological relevance of DNA demethylation in *Arabidopsis* antibacterial defense. *Proceedings of the National Academy of Sciences, USA* **110**: 2389–2394.

Zervudacki J, Yu A, Amesefe D, Wang J, Drouaud J, Navarro L, Deleris A. 2018. Transcriptional control and exploitation of an immune-responsive family of plant retrotransposons. *EMBO Journal* **37**: e98482.

Zhang H, Lang Z, Zhu JK. 2018. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology* **19**: 489–506.

Zhang L, Hu J, Han X, Li J, Gao Y, Richards CM, Zhang C, Tian Y, Liu G, Gul H *et al.* 2019. A high-quality apple genome assembly reveals the association of a retrotransposon and red fruit colour. *Nature Communications* **10**: 1494.

Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. 2009. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biology* **10**: R62.

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE. 2007. Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biology* **5**: e129.

Zhou J, Wang X, He K, Charron J-BF, Elling AA, Deng XW. 2010. Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in *Arabidopsis* reveals correlation between multiple histone marks and gene expression. *Plant Molecular Biology* **72**: 585–595.

Zhu Y, Rowley MJ, Bohmdorfer G, Wierzbicki AT. 2013. A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. *Molecular Cell* **49**: 298–309.

Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. 2007. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genetics* **39**: 61–69.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Expression of TEs represented in the ATH1 microarray in *Myzus persicae*-infested samples.

Fig. S2 Analysis of differential gene expression in mRNA sequencing libraries.

Fig. S3 Analysis of differential TE expression in mRNA sequencing libraries and global characterization of sRNA sequencing libraries.

Fig. S4 Characterization of DMR regions and genome-wide bisulfite DNA libraries.

Fig. S5 Examples of genes associated with DMRs.

Fig. S6 Characterization of differentially expressed genes and TFs associated with DMRs.

Fig. S7 Phenotypic and epigenetic characterization of mutants analyzed in the aphid settling test.

Table S1 RT-qPCR primers used in this study.

Table S2 Public data used in this analysis.

Table S3 High-throughput sequencing data produced in this analysis.

Table S4 Significantly differentially expressed genes upon aphid infestation.

Table S5 Significantly differentially expressed TEs upon aphid infestation.

Table S6 Differentially methylated regions identified.

Table S7 Genes identified in the proximity of DMRs.

Table S8 Differentially expressed genes associated with hypomethylated DMRs.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.