Genome-wide association study suggests potential effects of plant apparency on the initial colonization of aphids in *Arabidopsis thaliana*

**Chongmeng Xu**1†, **Yasuhiro Sato**1,2†\*, **Misako Yamazaki**1, **Marcel Brasser**1, **Matthew A. Barbour**1,3, **Jordi Bascompte**1, **Kentaro K. Shimizu**1,4\*

1 Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland  
2 Research Institute for Food and Agriculture, Ryukoku University, Yokotani 1-5, Seta Oe-cho, Otsu, Shiga 520-2194, Japan  
3 Départemente de Biologie, Université de Sherbrooke, 2500 boulevard de l’Université, Sherbrooke, QC, J1K 2R1, Canada  
4 Kihara Institute for Biological Research, Yokohama City University, Maioka 641-12, Totsuka-ward, Yokohama 244-0813, Japan  
† *These authors equally contributed to this study*  
\* *Co-correspondence; Y.S. (*[*yasuhiro.sato@uzh.ch*](mailto:yasuhiro.sato@uzh.ch)*); K.K.S. (*[*kentaro.shimizu@uzh.ch*](mailto:kentaro.shimizu@uzh.ch)*)*

# Abstract

Plant size, height, and other architectural traits, collectively called ‘plant apparency’, determine plant’s visibility and consequent probability of being colonized by herbivores. Although plant genes involved in physical and chemical defense are well studied in herbivory, less is known about the genetic basis underlying effects of plant apparency on herbivore colonization. Here, we conducted a genome-wide association study (GWAS) of aphid abundance in a field population of *Arabidopsis thaliana*. This GWAS of aphid abundance detected a rare but significant SNP variant on the third chromosome of *A. thaliana*, which was also partly associated with the presence or absence of inflorescence. Out of candidate genes near this significant SNP, a mutant of a ribosomal gene (AT3G13882) exhibited slower growth and later flowering than a wild type under laboratory conditions. A no-choice assay with the turnip aphid, *Lipaphis erysimi*, found that aphids were unable to successfully establish on the mutant. These findings suggest that genes relevant to plant apparency might underpin aphid colonization in the field.

# Introduction

Plants are attacked by herbivores across their life cycle in natural environments. While chemical and physical traits have long been the main focus of anti-herbivore resistance [1–4], plant life-history traits also account for herbivory variation in field environments [5,6]. For example, plant size, height, and the other architectural traits often shape plants’ visibility for herbivores [7,8]. These sets of visible traits, collectively called plant apparency [9], changes across ontogeny from vegetative to reproductive stage [6] and thereby affect the probability of being harbored by herbivores [10,11]. By focusing on genetic variation within a plant species, several studies have shown that plant architectural traits shape heritable variation in herbivore abundance and community composition [4,12,13]. Yet, less is known about the underlying genetic basis of the plant apparency on herbivore colonization.

Genome-wide association study (GWAS) is an effective way to dissect the genetic architecture of ecologically important traits [14,15]. Through associations between single nucleotide polymorphisms (SNPs) and traits, GWAS provides a hypothesis-free approach to identify novel genes from natural phenotypic variation [16,17]. Recent studies showed that controlled laboratory conditions are unlikely to reflect outdoor environments where interspecific interactions typically occur [18,19]. This fact emphasizes the importance of *in natura* study of functional genes [20–24]. To achieve this goal, it is important to conduct GWAS under field conditions.

*Arabidopsis thaliana* is a model plant species distributed and naturalized around the world. While *A. thaliana* usually blooms in spring after over-wintering, some cohorts have overlapping life cycles from spring to autumn [24–26]. When *A. thaliana* plants emerge from late spring to early summer, they are threatened by various herbivores [4,27]. Of the diverse insect herbivores, aphids are known to be major herbivores occurring across the natural distribution range of *A. thaliana* [28]. Because aphids often suck phloem sap from leaf veins and flowering stems, we hypothesized that plant apparency may play a key role in harboring aphids under field conditions.

To reveal the genetic architecture of aphid abundance, we combined GWAS with a mutant analysis in *A. thaliana*. We first conducted GWAS of aphid abundance on 196 *A. thaliana* accessions grown in a field site in Zurich, Switzerland. To further test candidate genes, we then cultivated and released the turnip aphid *Lipaphis erysimi* on *A. thaliana* mutants.

# Materials & Methods

## Field GWAS

### Plant genotypes

We obtained *Arabidopsis thaliana* genotypes that were selfed and maintained as inbred lines, called “accessions”, from the Arabidopsis Biological Resource Center (<https://abrc.osu.edu/>). We used the same set of 196 *A. thaliana* accessions as a previous study [29] except for two trichome mutants and an ungenotyped accession. All of these accessions were genotyped in the RegMap [30] and 1001 Genomes [31] projects. The list of accessions and phenotypes measured in this study is available in Table S1.

### Field experiment

To observe aphids in a simulated late cohort, we exposed *A. thaliana* accessions to a field environment from 4 to 25 July 2018. This field experiment was conducted at Zurich, Switzerland, to use a field site within a native distribution range of *A. thaliana*. To keep all the accessions in the rosette stage at the start of the field experiment, we initially cultivated *A. thaliana* in a laboratory under a short-day condition (8h light/16h dark cycle at 20C). Seeds were sown on 33-mm diameter Jiffy-seven(R) pots and stratified under a constant dark 4C condition for a week. Seedlings were cultivated in a growth chamber for 6 weeks under the short-day condition. Plants grown on the Jiffy-seven pots were then planted in a new plastic pot filled with mixed soils of agricultural composts (Profi Substrat Classic CL ED73, Einheitserde Co.) and perlites with compost to perlite ratio of 3:1 litter volume. Eight replicates of the 196 accessions were then transferred to the outdoor garden at the University of Zurich-Irchel (47 23’N, 8 33’E). Aphids were identified and counted by a single observer (Y. Sato) every two or three days. The two species of specialist aphids, *Lipaphis erysimi* and *Brevicoryne brassicae*, were identified based on the presence or absence of waxy compounds on their abdomen. These two species can be distinguished from the generalist aphid *Myzus persicae* based on the length of cornicules, though *M. persicae* did not occur during the present field experiment. To examine whether the aphid abundance differed between plants with and without flowering stems, we also recorded the presence or absence of bolting two weeks after the start of field experiment.

### Data analysis

All GWAS analyses were performed using the GWA-portal (<https://gwas.gmi.oeaw.ac.at>) [32]. The imputed full-sequence dataset [32,33] was used as SNP data for the 196 accessions, which provided combined SNP data imputed between 250k SNP chip genotyping by the RegMap project [34] and high-throughput sequencing by the 1001 Genome Project [35] Pseudo-heritability [32] was calculated for the target phenotype before association mapping. Accelerated mixed models [32] were used for association mapping with a correction of kinship structure. After the association mapping, candidate genes were searched within ca. 5 kb near a focal SNP. To inspect organ-specific expression levels of candidate genes, we referred to Klepikova Arabidopsis Atlas [36] via The Arabidopsis Information Resource (<https://www.arabidopsis.org/>).

We analyzed aphid abundance as well as the other three traits as the target phenotype in the GWA-portal. The aphid abundance was quantified as the maximum number of aphids, which included *Lipaphis erysimi* and *Brevicoryne brassicae* (see Results), observed on a plant during the experiment. The number of aphids was ln()-transformed to improve normality.

## Mutant analysis

### *Arabidopsis thaliana* mutants

T-DNA sequence-indexed lines of *A. thaliana* were obtained from the Nottingham Arabidopsis Stock Centre (NASC) (<https://arabidopsis.info/>). In addition to Columbia-0 (Col-0, NASC Accession ID: N70000) wild type, we ordered four mutant lines for a ribosomal gene (AT3G13882) (Table S2). These original mutants were back-crossed to the Col-0 wild type three times.

Following the instruction [37], we examined the insertion site by polymerase chain reaction (PCR) amplification and Sanger sequencing; and gene expression levels by Semi-quantitative reverse transcription and PCR (sqRT-PCR). To confirm the T-DNA insertion site of SALK\_039481, we extracted DNA from leaves using the CTAB method. We then amplified the DNA by polymerase chain reaction (PCR) as follows: 2 min at 95C; 35 cycles of 15 s at 95C, 30 s at 55C, 1.5 min at 72C; and a final extension step of 3 min at 72C. The PCR product was finally sequenced by Sanger sequencing to confirm the insertion site.

To perform sqRT-PCR, we extracted the RNA from leaves using RNeasy kit (Qiagen: Catalogue Number: 74181) and purified the RNA with DNA-free kit (Ambion: Cat. No. AM1906). RNA concentration was measured by Qubit spectrophotometer (Invitrogen: Cat. No. Q10211). Then we obtained the cDNA using High-Capacity RNA-to-cDNA kit (Applied Biosystems: Cat. No. 4387406) from 500 ng of the total RNA. The cDNA was amplified by PCR as follows: 3 min at 95C; 28 cycles of 15 s at 95C, 30 s at 55C, 1 min at 72C; and a final extension step of 5 min at 72C. A gel electrophoresis was performed using a 1% agarose with 120 V for 60 min. The PCR products were finally visualized an UV trans-illuminator system.

We found that one of the four lines, the SALK\_039481 (NASC Accession ID: N670586), indeed had a T-DNA insertion on an exon of one of two splice variants (Fig. [S1](#fig:tDNA)) and reduced expression level of AT3G13882 (Fig. [S2](#fig:sqRT_PCR)), suggesting that the insertion disrupted the gene. In the other three lines, the insertion was not found or low germination rate prevented further experiments.

### Laboratory experiments

To observe plant growth, we cultivated ten replicates of the ribosomal gene mutant and the Col-0 wild type under a long-day condition (16h light/8h dark cycle at 22C/20C). Seeds were sown on a 294 cm (= 7 7 6 cm) pot filled with the agricultural composts (Profi Substrat Classic CL ED73), and stratified at 4C under a constant dark condition for a week. The stratified seeds were then transferred to the long day condition. Seedlings were grown for 20 days. The rosette diameter (cm) was recorded as an index of plant size before aphids were released, as described next.

To test whether aphids could establish a colony on the mutant plants, we released the turnip aphid *L. erysimi* on the wild type and the mutant of *A. thaliana*. The potted plants were separately enclosed with a mesh net. Five wingless females of adult aphids were released on each plant. The experimental aphids were derived from a source population established by a previous study [38]. The enclosed plants were incubated under the long-day condition. The number of aphids per plant was counted by eye 3, 7, 10, and 14 days after the release of aphids. We did not count aphids that escaped outside the area of a plant. Flowering time was defined as the number of days to flowering, and recorded during the aphid experiment.

### Data analysis

We used linear mixed models (LMMs) or generalized linear mixed models (GLMMs) to test phenotypic differences between the mutant and the Col-0 wild type. The plant size and flowering time were analyzed using LMMs that assumed Gaussian errors. The number of aphids i.e., the count response was analyzed using GLMMs with Poisson error structure and log link function. Paired positions of a mutant and wild type plant were incorporated as a random effect to consider spatial heterogeneity within a growth chamber environment. An analysis of deviance with a -test was used to test the significance of the mutant vs. wild type (df1 = 1) against phenotypic variation within the random effect of the paired positions (df2 = 10 pairs - 1 fixed effect = 9). All statistical analyses were performed using R version 4.0.3 [39]. For LMM and GLMM, we used the lmer and glmer function implemented in the lme4 package [40].

# Results

## Field GWAS of the aphid abundance

To monitor aphid abundance as well as visible plant traits, we transplanted 196 *A. thaliana* accessions in the field in Zurich within a native distribution range of *A. thaliana*. At the transplantation, all plants were at the rosette stage, i.e., no bolting occurred. After two weeks, 38% of individual plants initiated bolting, i.e., a stem was observed. The main herbivores were the two species of specialist aphids, *Lipaphis erysimi* and *Brevicoryne brassicae*. The aphid abundance was higher on bolted accessions than on non-bolted accessions (non-bolted and bolted plants = average 0.59 and 2.07 aphids, respectively; Welch’s -test, ), suggesting that plant life cycle might be associated with plants’ capacity to avoid aphids. In addition, we also distinguished the abundance of winged and wingless aphids to infer the colonization process of aphids on *A. thaliana*. Winged and wingless aphids were observed at the rosette stage at the first monitoring after the transplant, but many of these aphids did not establish a colony in subsequent monitoring (the days between 07 and 10 July 2018: Fig. [S3](#fig:aphid_day)). This observation suggests that colonized aphids do not always establish a colony and thereby the success of the colony establishment also depends on the presence of stem after colonization.

To reveal genetic architecture underlying variation in aphid abundance, we calculated heritability and then performed association mapping. Aphid abundance had high heritability among the plant accessions (), indicating the genetic control of this trait. Our mapping also detected a significant SNP in an intergenic region above the genome-wide Bonferroni threshold (chr3-4579292, , MAF=0.026: Fig. [1](#fig:ManPlot); see also Fig. [S4](#fig:qqplot) for quantile-quantile plots). This top-scoring SNP was also associated with the bolting to non-significant but strong extent (-log10(p)=5.26; Fig. SX). The significance of chr3-4579292 turned smaller but remained at -log10(p)=6.40 when we adjusted the effects of bolting on aphid abundance, indicating the partial contribution of the bolting to shaping the significant association between chr3-4579292 SNP and aphid abundance. In contrast, chr3-4579292 remained significant (-log10(p)=8.77) when we adjusted the effects of initial plant size on aphid abundance. This weak influence of the initial size was plausible because plants were grown in a laboratory before field experiment to minimize variation in their initial size. The line of GWAS analyses suggests that the SNP at chr3-4579292 was significantly associated with aphid abundance through its potential influence on flowering.

To narrow down candidate genes, we further focused on genomic region near the significant SNP at chr3-4579292. Five out of the 196 accessions carried a rare allele increasing aphid abundance, while the other accessions have major alleles (Figure SXa). Genome sequences of four (line V, X, Y, and Z) of the five rare accessions are available in the 1001 Genome Project [35], where three (line V, X, and Y) of the four available accessions shared similar patterns near chr3-4579292 but differed from major accessions (Figure SXb). Three candidate genes were located nearest to the chr3-4579292, such as a putative ribosomal gene (AT3G13882) that is homologous to a ribosome protein L34 gene (RPL34) [41], *EPIDERMAL PATTERNING FACTOR LIKE 3* (*EPFL3*: AT3G13898), and *MYB26*. Out of these three genes, the ribosomal gene (AT3G13882) is known to be highly expressed in vegetative organs such as leaves [36]. The other two genes, *EPFL3* and *MYB26*, are highly expressed only in reproductive organs such as anthers or pistils [36]. Because aphids were unlikely to suck saps from anthers and pistils, we focused on the ribosomal gene (AT3G13882) for further investigation.

## Mutant plant growth and aphid colony establishment in the laboratory

To examine visible phenotypes of the ribosomal gene mutant (AT3G13882), we compared growth and flowering time of this mutant with the Col-0 wild type. After 20 days of growth, the AT3G13882 mutant had a significantly smaller size than the wild type (: Fig. [2](#fig:mutant)A,B). The flowering time of the AT3G13882 mutant was also significantly later than the wild type (: Fig. [2](#fig:mutant)A,C). The slower growth and delayed flowering of the ribosomal gene mutant led us to test whether the delayed growth could prevent the establishment of aphid colonies after colonization.

Then, to examine colony establishment after aphid colonization, we released wingless individuals of *Lipaphis erysimi* on rosette plants of the ribosomal gene mutant (AT3G13882) and the wild type. We observed a reduced number of aphids on the AT3G13882 mutant compared to the wild type at 7, 10, 14 days after the release of aphids ( at 7 days; Fig. [2](#fig:mutant)D: see also Fig. [S5](#fig:aphid_stat) for results at 10 and 14 days), suggesting that the delayed growth of the host negatively affected aphid colony establishment. We also incorporated the plant size or flowering time as an offset term in GLMMs to examine their confounding influence on aphid abundance. When the plant size was offset, the number of aphids less significantly differed between the wild type and mutant (Table SX). When the flowering time was offset, the number of aphids more significantly differed between the wild type and mutant (Table SX). These additional analyses support the observational evidence that delayed growth and flowering of the mutant plant results in less successful establishment of aphid colony than the wild type.

# Discussion

Guided by the field GWAS of aphid abundance, we discovered a mutant plant that exhibited slower growth and harbored fewer aphids in *A. thaliana*. While ribosomal genes have long been considered housekeeping genes of the protein synthesis machinery, mutants of ribosome-related genes exhibit a wide variety of growth and reproductive phenotypes. For example, previous studies reported the reduction of leaf cell number [42], reduced root length [43], and the reduction of pollen number [17,44] regarding ribosomal gene mutations. In this context, our findings from the ribosomal gene AT3G13882 add insights into biotic interactions besides growth deterioration.

We should note, however, that further study on natural variants responsible for the delayed growth and reduced aphid abundance would be necessary to validate its importance in the field. Multiple alleles on the AT3G13882 gene are also needed to provide strong evidence of those phenotypes. In the studies of a ribosomal gene *REDUCED POLLEN NUMBER1* (*RDP1*), null mutants showed a pleiotropic effect on plant growth and pollen number in *A. thaliana* [17,44]. Natural alleles of *RDP1* could alleviate pleiotropic growth defects [17]. In our study, the other growth-related genes or other mutations of AT3G13882 might have reduced the aphid abundance. Because transgenic approaches may not be effective to identify mutation sites affecting quantitative traits, natural alleles needs to be created using genome editing technology and then exchanged between natural accessions by crossing [17], Such quantitative complementation tests [17,44] will be needed to identify natural causal variants that alter aphid abundance through delayed growth.

Discovery of flowering-associated loci is sometimes of main interests in GWAS, but other times conceals net genetic effects on other complex traits of interests [45,46]. Our GWAS of aphid abundance unexpectedly led us to detect a significant rare variant that was also partially associated with the bolting. By ruling out the plant size and flowering time as confounding factors, we were able to detect substantial contributions of plant apparency to aphid abundance. This finding does not necessarily indicates the irrelevance of chemical and physical defense traits to herbivore colonization. Although our top-scoring SNP did not match known GWAS peaks of glucosinolates [47], it was still possible that secondary metabolites and volatile-mediated defense signaling were more influential on aphid colonization. While plant apparency can alter the initial colonization of aphids in *A. thaliana*, its long-term consequence on plant defense and resultant fitness remains unknown. Further studies on fitness consequence besides the identification of natural causal variants would be needed to verify ecological impacts of a single-gene on plant-herbivore interactions.

Overall, our study suggests the potential importance of genetically based plant apparency in altering herbivore abundance. Previous field studies used mutant plants to illustrate *in natura* roles of well-studied functional genes in chemical defense (e.g., *LOX*s in *Nicotiana attenuata* [2,3]) and physical defense (*GLABRA1* in *A. thaliana* [4]) against herbivores. In contrast, field GWAS offered a hypothesis-free approach to screen candidate genes responsible for herbivore abundance [15]. These lines of studies suggest that plant genetic variation governs herbivore abundance and communities [4,12,13,15], but a keystone gene shaping the ecological communities was not identified until recently [38]. Barbour et al. [38] have experimentally shown that pleiotropic effects of a glucosinolate biosynthesis gene *AOP2* on plant growth alter *A. thaliana*’s capacity to harbor aphids and their parasitoids. Our candidate genes might also exhibit pleiotropy on plant growth and secondary metabolism, but our findings support the notion that genes associated with plant growth can structure populations or communities of associated organisms. Since aphids and aphidophagous insects are widespread across terrestrial ecosystems [48,49], future studies may reveal cascading effects of delayed plant growth on food webs.

# Data availability

The data and codes are available at the GitHub repository (<https://github.com/yassato/AraAphidGWAS>) and its published version is deposited on Dryad (<doi::xxxxxx>).

# Authors’ contributions

C.X.: laboratory investigation, project administration, data curation, draft writing, review and editing; Y.S.: conceptualization, funding acquisition, project administration, supervision, field investigation, data curation, formal analysis, draft writing, review and editing; M.Y.: laboratory investigation, methodology (molecular), project administration, review and editing; M.B.: laboratory investigation, methodology (molecular); M.A.B: resources (insects), review and editing; J.B.: conceptualization, funding acquisition, supervision, review and editing; K.K.S.: conceptualization, funding acquisition, project administration, supervision, draft writing, review and editing.

# Competing interests

The authors declare no conflicts of interests concerning this study.

# Acknowledgements

The authors thank L. Mohn, K.K. Thomsen, and all members of Shimizu group for their help with the establishment of field plots; and R. Hostettler for assistance with the molecular experiments.

# Funding

This study was supported by the University of Zurich through the University Research Priority Program for “Global Change and Biodiversity”; Swiss National Science Foundation grant (31003A\_182318 to K.K.S.); Japan Science and Technology Agency (Grant numbers JPMJCR16O3 to K.K.S. and JPMJPR17Q4 to Y.S.); and Japan Society for the Promotion of Science, Grant-in-Aid for Transformative Research Areas (22H05179 to K.K.S.).

# Figures

![Figure 1. Genome-wide association study of aphid abundance on 196 Arabidopsis thaliana accessions grown in the field. (A) A genomic region within ca. 5 kbp from the top-scoring SNP at Chr3-4579292 displays the position of candidate genes. Only the longest splice variant (black horizontal arrow) is shown for each gene. (B) Manhattan plot shows the association score of -log10(p) across five chromosomes of A. thaliana with MAF cut-off at 0.025. A horizontal dashed line indicates the genome-wide Bonferroni threshold at p=0.05.](data:application/pdf;base64,)

Figure 1. Genome-wide association study of aphid abundance on 196 *Arabidopsis thaliana* accessions grown in the field. (A) A genomic region within ca. 5 kbp from the top-scoring SNP at Chr3-4579292 displays the position of candidate genes. Only the longest splice variant (black horizontal arrow) is shown for each gene. (B) Manhattan plot shows the association score of -log10() across five chromosomes of *A. thaliana* with MAF cut-off at 0.025. A horizontal dashed line indicates the genome-wide Bonferroni threshold at .

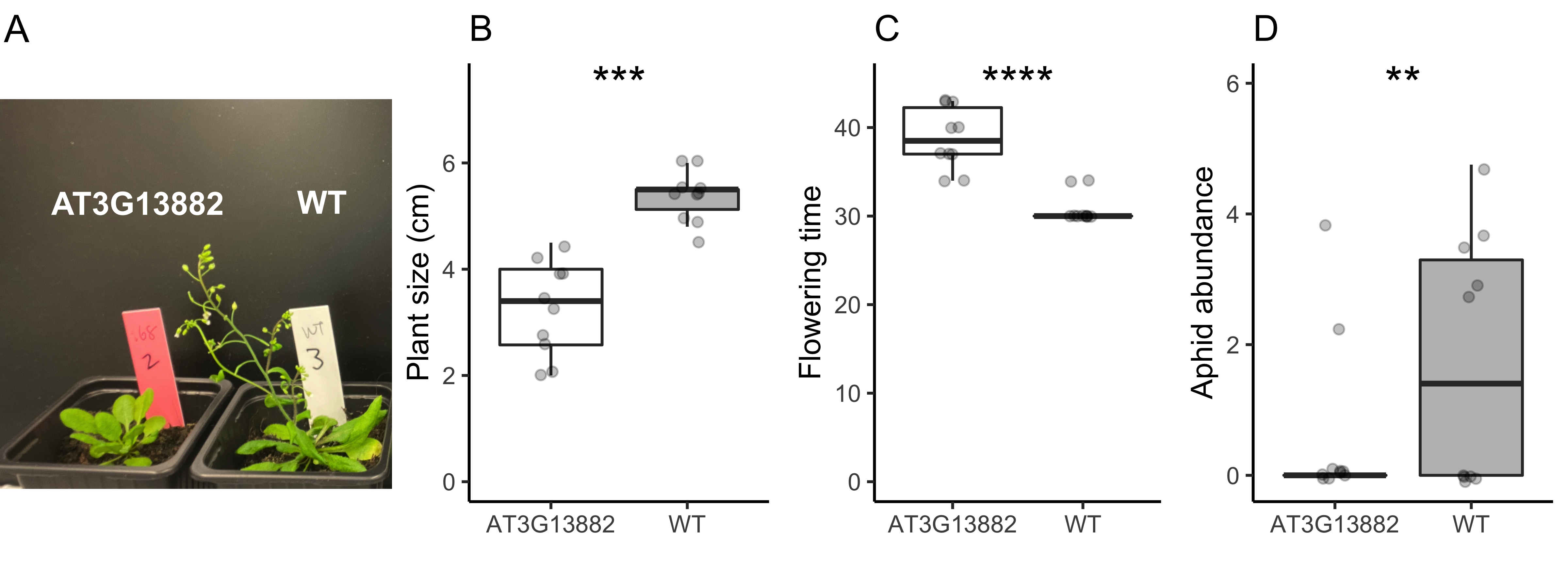


Figure 2. The photograph (A), plant size (B), flowering time (C), and aphid abundance (D) of the Col-0 wild type (WT) and the ribosomal gene mutant (AT3G13882) of *Arabidopsis thaliana* under laboratory conditions. Flowering time and aphid abundance represents the number of days to flowering and log2(no. of aphids + 1), respectively. Asterisks indicate the statistical significance by generalized linear mixed models; \*\* ; \*\*\* ; \*\*\*\* . Boxes: median with upper and lower quartile; Whiskers: 1.5 inter-quartile range.

# Supplementary Materials

![Figure S1. Schematic structure of the ribosomal gene AT3G13882. Green boxes: mRNA for two splice variants; triangle: the position of Transfer-DNA (T-DNA) insersion; red arrrows: primers for checking the T-DNA insertion. The primer IDs shown by four digits (oXXXX) correspond to those listed in Table S3. We downloaded the annotations from Genbank (Gene ID: 3768810), and visualized the position of the gene and primers using QIAGEN CLC Main Workbench.](data:application/pdf;base64,)

Figure S1. Schematic structure of the ribosomal gene AT3G13882. Green boxes: mRNA for two splice variants; triangle: the position of Transfer-DNA (T-DNA) insersion; red arrrows: primers for checking the T-DNA insertion. The primer IDs shown by four digits (oXXXX) correspond to those listed in Table S3. We downloaded the annotations from Genbank (Gene ID: 3768810), and visualized the position of the gene and primers using QIAGEN CLC Main Workbench.

Figure S2. Semi-quantitative reverse transcription and polymerase chain reaction (sqRT-PCR) of the ribosomal gene mutant mutant AT3G13882 (SALK_039481). Two primer sets A-D were used for AT3G13882; and E for an internal control gene ETL1 (see Table S4 for the primer information). This figure shows that the gene expression of AT3G13882 is suppressed in the mutant line SALK_039481.

Figure S2. Semi-quantitative reverse transcription and polymerase chain reaction (sqRT-PCR) of the ribosomal gene mutant mutant AT3G13882 (SALK\_039481). Two primer sets A-D were used for AT3G13882; and E for an internal control gene *ETL1* (see Table S4 for the primer information). This figure shows that the gene expression of AT3G13882 is suppressed in the mutant line SALK\_039481.

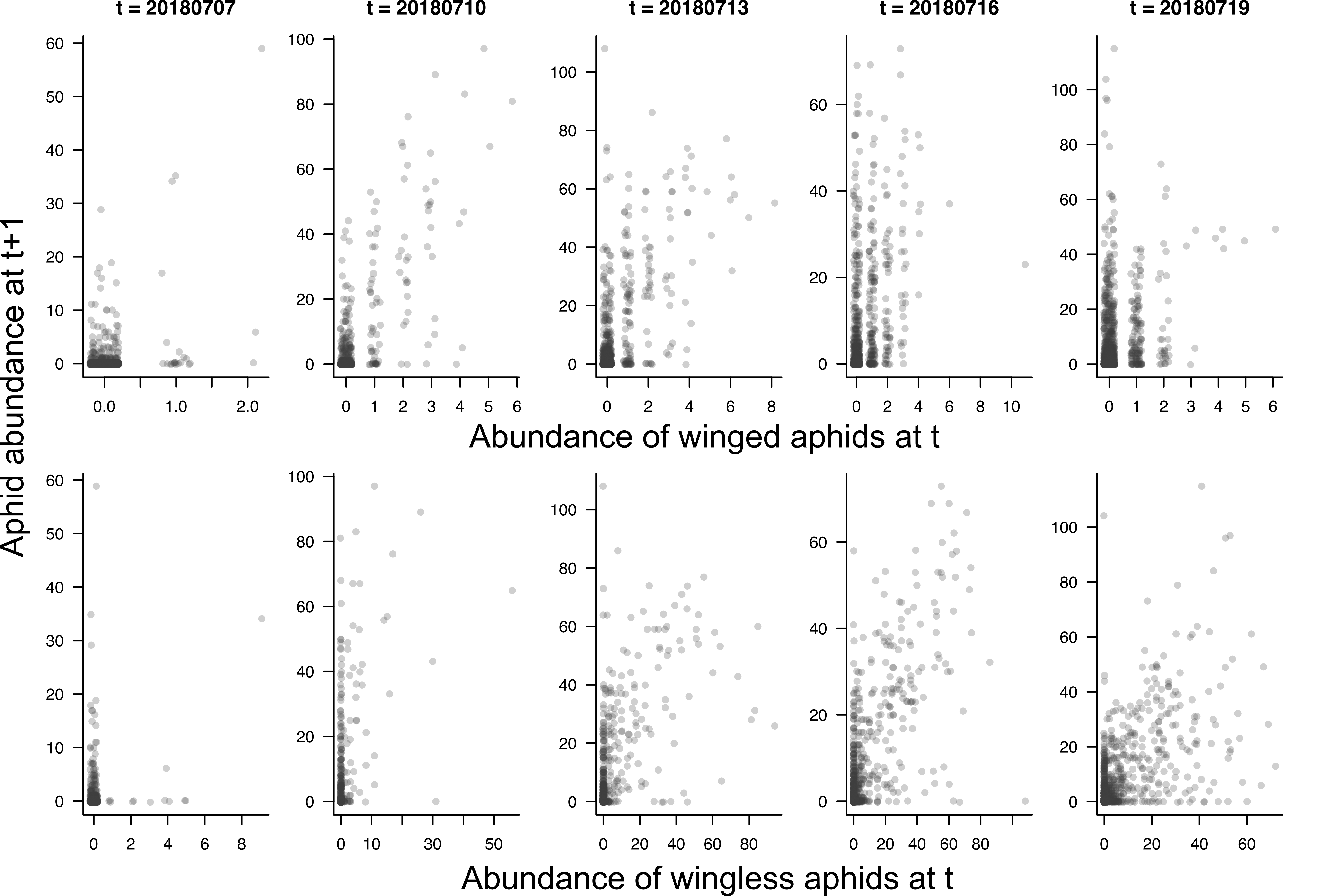


Figure S3. Temporal patterns of the emergence of wingless and winged aphids during the field GWAS experiment. X-axes show the number of winged (top row) or wingless (bottom row) aphids at current monitoring day (t=year-month-day). Y-axes show the total number of both winged and wingless aphids at the next monitoring date (t+1). A single point corresponds to an individual plant. The number of aphids represents the total number of individuals of both *Lipaphis erysimi* and *Brevicoryne brassicae*.

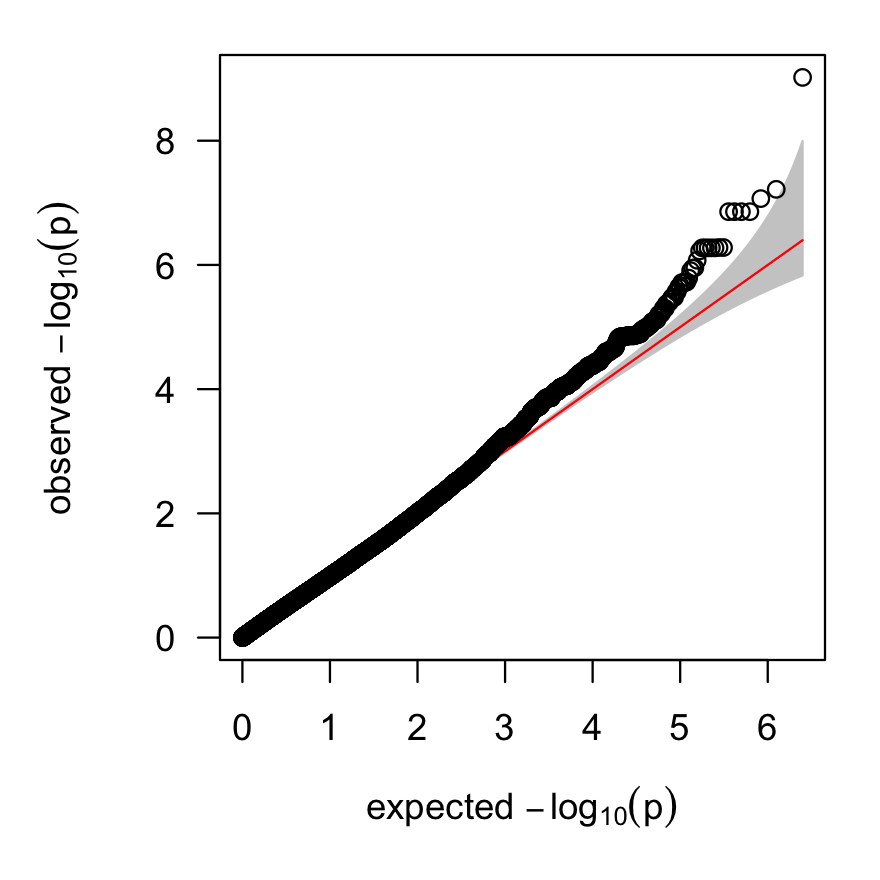


Figure S4. A quantile-quantile (QQ) plot showing relationships between the observed and expected association score of -log10(). A red solid line indicates randomly expected association scores abd the shaded area corresponds to its 95% confidence intervals. This figure shows that the top-scoring SNP at Chr3-4579292 is larger than the upper 95% confidence interval.

![Figure S5. Aphid abundance during the later period of incubation. Aphid abundance represents log2(no. of aphids + 1). Same as the result at 7 days (Fig. 2D), both the left and right panels have significant differences of the aphid abundance between the mutant (AT3G13882) and wild type (WT) (F_{1,9}=56.96, p<10^{-4} for 10 days; F_{1,9}=131.3, p<10^{-5} for 14 days).](data:application/pdf;base64,)

Figure S5. Aphid abundance during the later period of incubation. Aphid abundance represents log2(no. of aphids + 1). Same as the result at 7 days (Fig. [2](#fig:mutant)D), both the left and right panels have significant differences of the aphid abundance between the mutant (AT3G13882) and wild type (WT) ( for 10 days; for 14 days).

Table S1. List of GWAS accessions and phenotypes.

Table S2. List of mutant lines examined in this study.

Table S3. List of primers to check the T-DNA insertion or gene expression.

Table S4. Primer set information for sqRT-PCR results shown in Figure [S2](#fig:sqRT_PCR).

# References

1. Schoonhoven LM, Van Loon B, Loon JJ van, Dicke M. 2005 *Insect-plant biology*. Oxford University Press.

2. Kessler A, Halitschke R, Baldwin IT. 2004 Silencing the jasmonate cascade: Induced plant defenses and insect populations. *Science* **305**, 665–668.

3. Schuman MC, Allmann S, Baldwin IT. 2015 Plant defense phenotypes determine the consequences of volatile emission for individuals and neighbors. *Elife* **4**, e04490.

4. Sato Y, Shimizu-Inatsugi R, Yamazaki M, Shimizu KK, Nagano AJ. 2019 Plant trichomes and a single gene *GLABRA1* contribute to insect community composition on field-grown *Arabidopsis thaliana*. *BMC Plant Biology* **19**, 1–12.

5. Carmona D, Lajeunesse MJ, Johnson MT. 2011 Plant traits that predict resistance to herbivores. *Functional Ecology* **25**, 358–367.

6. Barton KE, Boege K. 2017 Future directions in the ontogeny of plant defence: Understanding the evolutionary causes and consequences. *Ecology Letters* **20**, 403–411.

7. Zverev V, Zvereva EL, Kozlov MV. 2017 Ontogenetic changes in insect herbivory in birch (*Betula pubesecens*): The importance of plant apparency. *Functional Ecology* **31**, 2224–2232.

8. Higuchi Y, Kawakita A. 2019 Leaf shape deters plant processing by an herbivorous weevil. *Nature Plants* **5**, 959–964.

9. Feeny P. 1976 Plant apparency and chemical defense. In *Biochemical interaction between plants and insects* (eds JW Wallace, RL Mansell), pp. 1–40. Boston, MA: Springer US. (doi:[10.1007/978-1-4684-2646-5\_1](https://doi.org/10.1007/978-1-4684-2646-5_1))

10. Kawagoe T, Kudoh H. 2010 Escape from floral herbivory by early flowering in *Arabidopsis halleri* subsp. *gemmifera*. *Oecologia* **164**, 713–720.

11. Marquis RJ, Moura RF. 2021 Escape as a mechanism of plant resistance against herbivores. In *Plant-animal interactions*, pp. 39–57. Springer.

12. Johnson MT, Agrawal AA, Maron JL, Salminen J-P. 2009 Heritability, covariation and natural selection on 24 traits of common evening primrose (*Oenothera biennis*) from a field experiment. *Journal of Evolutionary Biology* **22**, 1295–1307.

13. Barbour MA, Rodriguez-Cabal MA, Wu ET, Julkunen-Tiitto R, Ritland CE, Miscampbell AE, Jules ES, Crutsinger GM. 2015 Multiple plant traits shape the genetic basis of herbivore community assembly. *Functional Ecology* **29**, 995–1006.

14. Santure AW, Garant D. 2018 Wild GWAS—association mapping in natural populations. *Molecular Ecology Resources* **18**, 729–738.

15. Barker HL, Riehl JF, Bernhardsson C, Rubert-Nason KF, Holeski LM, Ingvarsson PK, Lindroth RL. 2019 Linking plant genes to insect communities: Identifying the genetic bases of plant traits and community composition. *Molecular ecology* **28**, 4404–4421.

16. Fujii S *et al.* 2019 A stigmatic gene confers interspecies incompatibility in the Brassicaceae. *Nature Plants* **5**, 731–741.

17. Tsuchimatsu T *et al.* 2020 Adaptive reduction of male gamete number in the selfing plant *Arabidopsis thaliana*. *Nature Communications* **11**, 1–9.

18. Honjo MN, Emura N, Kawagoe T, Sugisaka J, Kamitani M, Nagano AJ, Kudoh H. 2020 Seasonality of interactions between a plant virus and its host during persistent infection in a natural environment. *The ISME Journal* **14**, 506–518.

19. Sato Y, Tezuka A, Kashima M, Deguchi A, Shimizu-Inatsugi R, Yamazaki M, Shimizu KK, Nagano AJ. 2019 Transcriptional variation in glucosinolate biosynthetic genes and inducible responses to aphid herbivory on field-grown *Arabidopsis thaliana*. *Frontiers in Genetics*, 787.

20. Shimizu KK, Kudoh H, Kobayashi MJ. 2011 Plant sexual reproduction during climate change: Gene function *in natura* studied by ecological and evolutionary systems biology. *Annals of Botany* **108**, 777–787.

21. Kudoh H. 2016 Molecular phenology in plants: In natura systems biology for the comprehensive understanding of seasonal responses under natural environments. *New Phytologist* **210**, 399–412.

22. Yamasaki E *et al.* 2017 Genomics meets remote sensing in global change studies: Monitoring and predicting phenology, evolution and biodiversity. *Current Opinion in Environmental Sustainability* **29**, 177–186.

23. Zaidem ML, Groen SC, Purugganan MD. 2019 Evolutionary and ecological functional genomics, from lab to the wild. *The Plant Journal* **97**, 40–55.

24. Stockenhuber R *et al.* 2021 The *UV RESISTANCE LOCUS 8*-mediated UV-b response is required alongside *CRYPTOCHROME1* for plant survival under sunlight in the field. *bioRxiv* (doi:[10.1101/2021.12.08.471623](https://doi.org/10.1101/2021.12.08.471623))

25. Thompson L. 1994 The spatiotemporal effects of nitrogen and litter on the population dynamics of *Arabidopsis thaliana*. *Journal of Ecology*, 63–68.

26. Taylor MA, Cooper MD, Sellamuthu R, Braun P, Migneault A, Browning A, Perry E, Schmitt J. 2017 Interacting effects of genetic variation for seed dormancy and flowering time on phenology, life history, and fitness of experimental *Arabidopsis thaliana* populations over multiple generations in the field. *New Phytologist* **216**, 291–302.

27. Mosleh Arany A, Jong TJ de, Meijden E van der. 2009 Herbivory and local genetic differentiation in natural populations of *Arabidopsis thaliana* (Brassicaceae). *Plant Ecology* **201**, 651–659.

28. Züst T, Heichinger C, Grossniklaus U, Harrington R, Kliebenstein DJ, Turnbull LA. 2012 Natural enemies drive geographic variation in plant defenses. *Science* **338**, 116–119.

29. Sato Y, Yamamoto E, Shimizu KK, Nagano AJ. 2021 Neighbor GWAS: Incorporating neighbor genotypic identity into genome-wide association studies of field herbivory. *Heredity* **126**, 597–614.

30. Horton MW *et al.* 2012 Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature Genetics* **44**, 212–216.

31. Alonso-Blanco C *et al.* 2016 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 481–491.

32. Seren Ü. 2018 GWA-portal: Genome-wide association studies made easy. In *Root development*, pp. 303–319. Springer.

33. Togninalli M, Seren Ü, Meng D, Fitz J, Nordborg M, Weigel D, Borgwardt K, Korte A, Grimm DG. 2018 The AraGWAS Catalog: A curated and standardized Arabidopsis thaliana GWAS catalog. *Nucleic Acids Research* **46**, D1150–D1156. (doi:[10.1093/nar/gkx954](https://doi.org/10.1093/nar/gkx954))

34. Horton MW *et al.* 2012 Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature Genetics* **44**, 212–216. (doi:[10.1038/ng.1042](https://doi.org/10.1038/ng.1042))

35. Alonso-Blanco C *et al.* 2016 1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. *Cell* **166**, 481–491. (doi:[10.1016/j.cell.2016.05.063](https://doi.org/10.1016/j.cell.2016.05.063))

36. Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA. 2016 A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *The Plant Journal* **88**, 1058–1070.

37. O’Malley RC, Barragan CC, Ecker JR. 2015 A user’s guide to the *Arabidopsis* t-DNA insertion mutant collections. In *Plant functional genomics*, pp. 323–342. Springer.

38. Barbour MA, Kliebenstein DJ, Bascompte J. 2022 A keystone gene underlies the persistence of an experimental food web. *Science* **376**, 70–73.

39. R Core Team. 2020 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <https://www.R-project.org/>.

40. Bates D, Mächler M, Bolker B, Walker S. 2015 Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48. (doi:[10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01))

41. Cheng C-Y, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. 2017 Araport11: A complete reannotation of the *Arabidopsis thaliana* reference genome. *The Plant Journal* **89**, 789–804.

42. Fujikura U, Horiguchi G, Ponce MR, Micol JL, Tsukaya H. 2009 Coordination of cell proliferation and cell expansion mediated by ribosome-related processes in the leaves of *Arabidopsis thaliana*. *The Plant Journal* **59**, 499–508.

43. Creff A, Sormani R, Desnos T. 2010 The two *Arabidopsis* *RPS6* genes, encoding for cytoplasmic ribosomal proteins S6, are functionally equivalent. *Plant Molecular Biology* **73**, 533–546.

44. Kakui H, Tsuchimatsu T, Yamazaki M, Hatakeyama M, Shimizu KK. 2022 Pollen number and ribosome gene expression altered in a genome-editing mutant of *REDUCED POLLEN NUMBER1* gene. *Frontiers in Plant Science*, 3159.

45. Hori K, Kataoka T, Miura K, Yamaguchi M, Saka N, Nakahara T, Sunohara Y, Ebana K, Yano M. 2012 Variation in heading date conceals quantitative trait loci for other traits of importance in breeding selection of rice. *Breeding science* **62**, 223–234.

46. Onogi A, Ideta O, Yoshioka T, Ebana K, Yamasaki M, Iwata H. 2016 Uncovering a nuisance influence of a phenological trait of plants using a nonlinear structural equation: Application to days to heading and culm length in asian cultivated rice (oryza sativa l.). *PLoS One* **11**, e0148609.

47. Brachi B, Meyer CG, Villoutreix R, Platt A, Morton TC, Roux F, Bergelson J. 2015 Coselected genes determine adaptive variation in herbivore resistance throughout the native range of arabidopsis thaliana. *Proceedings of the National Academy of Sciences* **112**, 4032–4037.

48. Dixon A. 1977 Aphid ecology: Life cycles, polymorphism, and population regulation. *Annual Review of Ecology and Systematics* **8**, 329–353.

49. Snyder WE, Ives AR. 2003 Interactions between specialist and generalist natural enemies: Parasitoids, predators, and pea aphid biocontrol. *Ecology* **84**, 91–107.