# Genome-wide association study highlights escape by delayed growth from aphid herbivory on field-grown *Arabidopsis thaliana*

**Author 1**1  
**Author 2**1,2  
**Author 3**1,3

1 University of XX  
2 YY University  
3 ZZ University

# Abstract (should be revised once the main text is fixed)

Growing number of ecological studies have shown that plant phenological and architectural traits greatly contribute to shaping field herbivory, but their key genes remain unknown. Here, we conducted a genome-wide association study (GWAS) of aphid abundance in a field population of *Arabidopsis thaliana*. Near a single significant peak on the third chromosome, we observed known growth-related genes, such as *ROOT HAIR DEFECTIVE3*, and candidate genes with unknown functions. Out of the unknown genes, we discovered a new locus *NAME OF THE GENE* (*NOG*) that encoded a putative ribosomal gene (AT3G13882). Knockout mutants of *nog* exhibited slower growth and ca. x-days later flowering than Col-0 under a long-day condition in a laboratory. Our laboratory bioassay further showed that a colony of the mustard aphid *Lipaphis erysimi* was difficult to establish on *nog* due to its small rosette size. The field-based detection of a known and novel growth-related gene suggests that side effects of life-history traits on biotic interaction play a more critical role *in natura* than currently appreciated.

# Introduction

Plants are attacked by herbivores across their life cycles in natural environments. While chemical and physical traits have long been a main focus of anti-herbivore defense [1], plant life-history traits also account for herbivory in field environments [2, 3]. For example, phenological changes can allow plants to escape from seasonal herbivory [4, 5]. Plants’ visibility for herbivores, namely plant apparency [6], also changes from vegetative to reproductive phase, which alters the risk of herbivore attacks across plant ontogeny [3, 7]. By focusing on intraspecific variation within a plant species, several studies have shown that plant phenological and architectural traits shape heritable variation in herbivory among plant genotypes [8–10], though key genes remain unknown.

Genome-wide association study (GWAS) is increasingly recognized as a powerful approach to dissect the genetic architecture of ecologically important traits [11, 12] and identify novel genes from natural phenotypic variation [13, 14]. However, field studies have suggested that controlled laboratory conditions unlikely reflect outdoor environments where interspecific interactions typically occur [15, 16], highlighting the importance of *in natura* study on gene functions [17–20]. For *in natura* understanding of interspecific interactions, it is necessary to conduct GWAS under field conditions.

*Arabidopsis thaliana* is the model plant species distributed in Europe and naturalized around the world. While *A. thaliana* usually bloom in spring after over-wintering, some cohorts have overlapped life cycles from spring to autumn [20–22]. When plants emerge from late spring to early summer, they are threatened by various herbivores, such as aphids and beetles [10, 23]. Of the diverse insect herbivores, aphids are a major herbivore occurring across a natural distribution range of *A. thaliana* [24]. Because aphids often suck phloem saps from flowering stems, we hypothesized that plant life-history traits may play a key role in aphid colonization in the field.

To reveal the genetic architecture of aphid herbivory, we combined GWAS and mutant analysis in *A. thaliana*. We first conducted GWAS of aphid abundance on 196 *A. thaliana* accessions grown in a field site of Zurich, Switzerland. Near a GWAS peak, we discovered a new locus *NAME OF THE GENE* (*NOG*) as well as several known genes involved in delayed growth. To further validate the candidate genes, we then cultivated and released the mustard aphid *Lipaphis erysimi* on *A. thaliana* mutants. The field-based discovery of *NOG* will gain genetic insights into plant apparency on insect herbivory *in natura*.

# Materials & Methods

## Field GWAS (*Yasu will fill in this section*).

### Plant accessions

We selected 196 *A. thaliana* accessions from RegMap [25] and 1001 Genomes [26] projects, most of which were overlapped with previous GWAS of biotic interaction [27]. The list of accessions is available in a supplementary material.

### Field experiment and trait survey

To observe aphids on a simulated late cohort, we exposed *A. thaliana* to field environment from 4 to 25 July in 2018 at a field site within Europe. We initially cultivated eight replicates of the 196 accessions in a laboratory. Seeds are sown on 33-mm diameter Jiffy-seven(R) pots and stratified under a constant dark 4C condition for a week. Seedlings were grown for 6 weeks under a short-day condition (8L:16D cycle with 20C air temperature). Plants were then potted in a plastic pot filled with agricultural soils (3:1L volume ratio of xxxx soils:perlites), and then transferred to the outdoor garden at the University of Zurich-Irchel (Zurich, Switzerland; XXxx’N, YYyy’E). Aphids were counted by a single observer every two or three days. 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

GWAS was performed using the GWA-portal website (<https://gwas.gmi.oeaw.ac.at>) [28]. Target phenotypes were the maximum number of aphids per plant throughout the experiment. The imputed fullsequence dataset were chosen as SNP data for the 196 accessions. Pseudo-heritability was calculated for each target phenotype before association mapping. Accelerated mixed models were used for association mapping with a correction of kinship structure. The genome-wide significance level was given at with Bonferroni correction of multiple testing. The number of aphids was log()-transformed to improve normality. Input phenotype data are available as a supplementary material. After the association mapping, candidate genes were searched within ca. 10 kb near a focal SNP. To estimate gene functions from gene expression information, we used the Arabidopsis eFP browser [29], 1001 transcriptome data [30], and ATTED-II coexpression database [31] (<https://atted.jp/>).

To estimate gene functions associated with the aphid abundance, we conducted gene ontology (GO) enrichment analysis for the output of GWA-portal. To deal with statistical non-independence of SNPs, we employed an unbiased GO analysis implemented in the Gowinda program [32]. The cut-off value of minor allele frequency (MAF) was set at 0.025 for the entire SNP set, leaving 2,500,979 SNPs in total. SNPs having the top 1% value of -log10() were tested against the entire set. The latest gene feature format (GFF) and GO slim annotation of *A. thaliana* were obtained from The Arabidopsis Information Resource (TAIR) (<https://www.arabidopsis.org/>). The Gowinda program was run with the key arguments of –simulations 1000000 –min-significance 1 –gene-definition updownstream1000 –mode gene –min-genes 2.

## Mutant analysis (*Please fill in this section*)

### Plants

T-DNA insertion lines of *A. thaliana* were obtained from the Nottingham Arabidopsis Stock Centre (NASC) (<https://arabidopsis.info/>).

Please describe  
- which lines did you order?  
- how did you confirm the knockout mutations?  
- how many times did you backcross?

### Bioassay

Please describe  
- what species of aphids did you use?  
- how did you cultivate plants?  
- how did you construct the experimental arena?  
- how did you record the number of aphids?

# Results

## Field GWAS of the aphid abundance (*Yasu will fill in this section*)

We observed the two species of specialist aphids, *Lipaphis erysimi* and *Brevicoryne brassicae*, occurring on the 196 accessions of *A. thaliana* during early summer in Zurich. These aphids more likely colonized on bolted accessions than non-bolted accessions (non-bolted and bolted plants = 0 and 7 aphids in median, respectively; Mann-Whitney’s -test, ), indicating that the presence of flowering stem was highly significantly associated with the aphid abundance. The pseudo-heritability of the aphid number was 0.7. The large heritability suggests that the observed trait is likely under genetic control and thus deserves further association mapping.

We then performed GWAS using the accelerated mixed model that corrects a kinship structure. Regarding the aphid number, we observed a significant SNP above the genome-wide Bonferroni threshold (chr3-4579292, , MAF=0.026: Fig. [1](#fig:ManPlot)). When comparing trait values between two alleles on the significant SNP marker, three of five accessions shared similar haplotypes from AT3G13870 to AT3G13890 locus (Fig. [S1](#fig:1001browser)), spanning a X kbp region. Several known genes were located near this genomic region: AT3G13870 locus, also known as *ROOT HAIR DEFECTIVE3* (*RHD3*), is known to regulate root hair developments [33] and thereby results in delayed growth [34]. AT3G13890 locus is known to encode MYB26 transcription factor responsible for the anther dehiscence and male sterility [35]. AT3G13880 (*OTP72*) locus itself has no visible phenotype, but one allele *otp72-2* is known to affect gene the expression level of *RHD3* [36]. Known functions of these candidate genes led us to hypothesize that genes involved in growth or reproduction have side effects on aphid colonization on a plant stem.

To estimate known gene functions associated with the aphid numbers, we also performed GO enrichment analysis for SNPs having the top 1% score of -log10(). We detected two significant GOs, “DNA-binding transcription factor activity” and “positive regulation of transcription, DNA-templated” (FDR-adjusted ). Although the other GOs were not significant after the FDR correction, the top 22 categories included GOs involving defense, growth, and reproduction, such as “positive regulation of salicylic acid mediated signaling pathway”, “positive regulation of leaf senescence”, “pollen tube development” (non-adjusted : Fig. [S2](#fig:gomap)). The results of GO enrichment analysis support the relevance of plant life-history traits as well as defense signaling to the aphid numbers.

## Mutant analysis in a laboratory (*Please fill in this section*)

### Growth and flowering of *A. thaliana* mutants

Two small unknown genes (AT3G13882 and AT3G13898) as well as *MYB26* were located nearby the significant SNP (chr3-4579292). Thus, we cultivated single-gene mutants of these three genes to examine their visible phenotypes. **[Describe the results of growth and flowering time].**

### Aphid attachment experiments

To test whether the delayed growth could prevent aphid colonization, we then released *Lipaphis erysimi*, one of the two aphid species observed in the field, on mutant plants. **[Describe the results of aphid numbers].**

# Discussion (not yet written, combined Results & Discussion possible)

possible topics  
- ecological discussion on the genetic basis of phenological escape from herbivory  
- side effects of ribosomal genes on plant growth (Takashi & Hiro’s paper)  
- caveats on other candidates near the GWAS peak  
- summary or concluding paragraph

# Tables & Figures

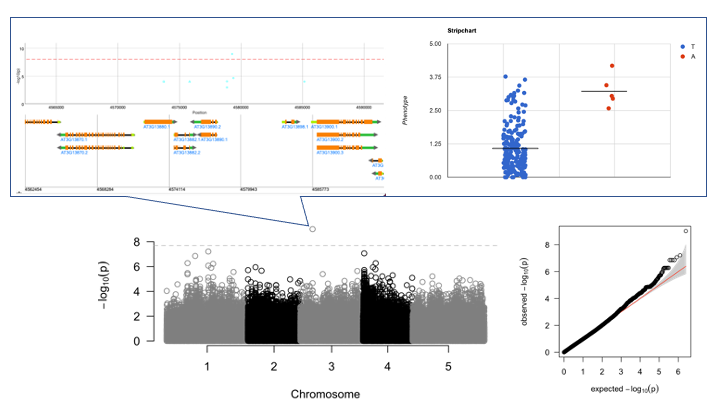


Figure 1. GWAS of the aphid abundance on 196 *A. thaliana* accessions grown in the field. Manhattan plot shows the associaiton score of -log10() against five choromosomes of *A. thaliana* at MAF cut-off = 0.025, where a horizontal dashed line indicates the genome-wide Bonferroni threshold at . QQ-plot shows relationships between the observed and expected -log10() values, where a solid line indicates randomly expected -log10() and the shaded area corresponds to its 95% confidence intervals. The upper inset focuses on the top-scroing SNP at Chr3-4579292, and desplays the position of candidate genes (left) or compares phenotype values i.e., log(no. of aphids + 1) between the two SNP alleles (right).

# Supplementary Materials

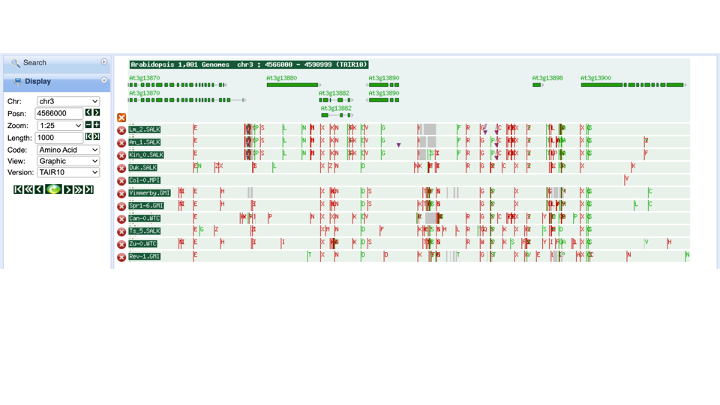


Figure S1. A snapshot of the genomic region near Chr3-4579292 in the 1001 Genome Browser (<http://signal.salk.edu/atg1001/3.0/gebrowser.php>). The upper four accessions carried a rare allele susceptible to aphids (inset of Fig. [1](#fig:ManPlot)). The accessions listed below Col-0 carried a major allele and harboured no aphids in the field GWAS.

![Figure S2. REVIGO treemap summarizing suggestive GOs detected by the Gowinda program. The panel size corresponds to -log10(p). Colors indicate similar GO families summarized by REVIGO.](data:application/pdf;base64,)

Figure S2. REVIGO treemap summarizing suggestive GOs detected by the Gowinda program. The panel size corresponds to -log10(). Colors indicate similar GO families summarized by REVIGO.

# References

1. Schoonhoven LM, Van Loon B, Loon JJ van, Dicke M: *Insect-Plant Biology*. Oxford University Press on Demand; 2005.

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

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

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

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

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

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

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

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

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

11. Anderson JT, Mitchell-Olds T: **Ecological genetics and genomics of plant defences: Evidence and approaches**. *Functional Ecology* 2011, **25**:312–324.

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

13. Fujii S, Tsuchimatsu T, Kimura Y, Ishida S, Tangpranomkorn S, Shimosato-Asano H, Iwano M, Furukawa S, Itoyama W, Wada Y, et al.: **A stigmatic gene confers interspecies incompatibility in the Brassicaceae**. *Nature Plants* 2019, **5**:731–741.

14. Tsuchimatsu T, Kakui H, Yamazaki M, Marona C, Tsutsui H, Hedhly A, Meng D, Sato Y, Städler T, Grossniklaus U, et al.: **Adaptive reduction of male gamete number in the selfing plant *Arabidopsis thaliana***. *Nature Communications* 2020, **11**:1–9.

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

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

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

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

19. Yamasaki E, Altermatt F, Cavender-Bares J, Schuman MC, Zuppinger-Dingley D, Garonna I, Schneider FD, Guillén-Escribà C, Moorsel SJ van, Hahl T, et al.: **Genomics meets remote sensing in global change studies: Monitoring and predicting phenology, evolution and biodiversity**. *Current Opinion in Environmental Sustainability* 2017, **29**:177–186.

20. Stockenhuber R, Akiyama R, Tissot N, Yamazaki M, Wyler M, Arongaus AB, Podolec R, Sato Y, Milosavljevic S, Widmer A, et al.: **The *UV RESISTANCE LOCUS 8*-mediated UV-b response is required alongside *CRYPTOCHROME1* for plant survival under sunlight in the field**. *bioRxiv* 2021.

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

22. Taylor MA, Cooper MD, Sellamuthu R, Braun P, Migneault A, Browning A, Perry E, Schmitt J: **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* 2017, **216**:291–302.

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

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

25. Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Muliyati NW, Platt A, Sperone FG, Vilhjálmsson BJ, et al.: **Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel**. *Nature Genetics* 2012, **44**:212–216.

26. Alonso-Blanco C, Andrade J, Becker C, Bemm F, Bergelson J, Borgwardt KM, Cao J, Chae E, Dezwaan TM, Ding W, et al.: **1,135 genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana***. *Cell* 2016, **166**:481–491.

27. Horton MW, Bodenhausen N, Beilsmith K, Meng D, Muegge BD, Subramanian S, Vetter MM, Vilhjálmsson BJ, Nordborg M, Gordon JI, et al.: **Genome-wide association study of *Arabidopsis thaliana* leaf microbial community**. *Nature Communications* 2014, **5**:1–7.

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

29. Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ: **An “electronic fluorescent pictograph” browser for exploring and analyzing large-scale biological data sets**. *PloS One* 2007, **2**:e718.

30. Kawakatsu T, Huang SC, Jupe F, Sasaki E, Schmitz RJ, Urich MA, Castanon R, Nery JR, Barragan C, He Y, et al.: **Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions**. *Cell* 2016, **166**:492–505.

31. Obayashi T, Aoki Y, Tadaka S, Kagaya Y, Kinoshita K: **ATTED-II in 2018: A plant coexpression database based on investigation of the statistical property of the mutual rank index**. *Plant and Cell Physiology* 2018, **59**:e3–e3.

32. Kofler R, Schlötterer C: **Gowinda: Unbiased analysis of gene set enrichment for genome-wide association studies**. *Bioinformatics* 2012, **28**:2084–2085.

33. Schiefelbein JW, Somerville C: **Genetic control of root hair development in *Arabidopsis thaliana*.** *The Plant Cell* 1990, **2**:235–243.

34. Zhang M, Wu F, Shi J, Zhu Y, Zhu Z, Gong Q, Hu J: ***ROOT HAIR DEFECTIVE3* family of dynamin-like GTPases mediates homotypic endoplasmic reticulum fusion and is essential for *Arabidopsis* development**. *Plant Physiology* 2013, **163**:713–720.

35. Mitsuda N, Hiratsu K, Todaka D, Nakashima K, Yamaguchi-Shinozaki K, Ohme-Takagi M: **Efficient production of male and female sterile plants by expression of a chimeric repressor in *Arabidopsis* and rice**. *Plant Biotechnology Journal* 2006, **4**:325–332.

36. Chateigner-Boutin A-L, Colas des Francs-Small C, Fujii S, Okuda K, Tanz SK, Small I: **The e domains of pentatricopeptide repeat proteins from different organelles are not functionally equivalent for RNA editing**. *The Plant Journal* 2013, **74**:935–945.