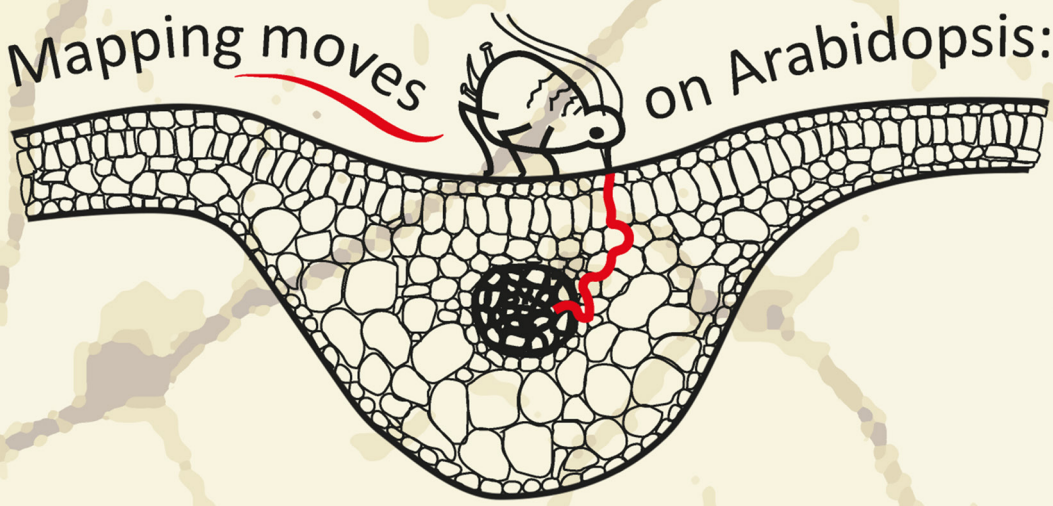



# From natural variation to single genes affecting aphid behaviour

Karen J. Kloth

# Mapping moves on Arabidopsis:

## From natural variation to single genes affecting aphid behaviour

Karen J. Kloth

**Thesis committee**

**Promotors**
Prof. Dr Marcel Dicke
Professor of Entomology
Wageningen University

Prof. Dr Harro J. Bouwmeester
Professor of Plant Physiology
Wageningen University

**Co-promotor**
Dr Maarten A. Jongsma
Senior Scientist, Business Unit PRI Bioscience
Wageningen University and Research Centre

**Other members**
Prof. Dr René A.A. van der Vlugt, Wageningen University
Prof. Dr Joost J.B. Keurentjes, University of Amsterdam, Wageningen University
Dr Christa S. Testerink, University of Amsterdam
Dr Sjoukje Heimovaara, Royal van Zanten B.V., Rijsenhout

This research was conducted under the auspices of
the Graduate School of Experimental Plant Sciences (EPS)

# Mapping moves on Arabidopsis: From natural variation to single genes affecting aphid behaviour

Karen J. Kloth

**Thesis**
submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 11 March 2016
at 4 p.m. in the Aula.

Karen J. Kloth

Mapping moves on Arabidopsis:

From natural variation to single genes affecting aphid behaviour
270 pages.

PhD thesis, Wageningen University, Wageningen, NL (2016)
With references, with a summary in English

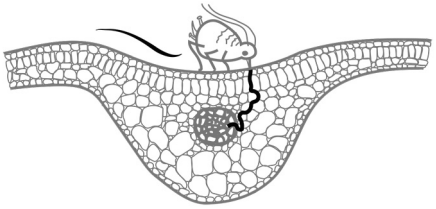
ISBN 978-94-6257-648-3

# Contents

*voor Mariëlle*

# Chapter 1

## General introduction



Karen J. Kloth

## Aphid pests

Aphids (Hemiptera: Aphididae) are worldwide pest insects in many crops. They are phloem-feeding insects that retard plant growth and development and can transmit multiple plant viruses (Dixon, 1998; Girousse *et al.*, 2003; Minks and Harrewijn, 1989). Aphids have the highest economic importance in temperate regions, where 26% of the major pest insects on the main food crops (maize, wheat, potatoes, sugar beet, barley and tomatoes) belong to the Aphididae family (Hill, 1987). Aphids remove photoassimilates from the host plant, excrete honeydew droplets which are a substrate for fungal growth, and, most importantly, have been described as a vector of almost 300 plant viruses (Dedryver *et al.*, 2010; Nault, 1997). As a consequence, aphid-infested crops suffer from yield loss and unmarketable products. The economic impact of aphids reaches up to 150 million US dollars per year for soybean and more than 10 million US dollars per year for conifer plantations (FAO, 2015; Kim *et al.*, 2008a). Between 1989 and 2002, aphid-transmitted viruses resulted in a mean yield loss of 20% in barley (Dedryver *et al.*, 2010). The most common pest control strategy against aphids is the application of insecticides, such as neonicotinoids, synthetic pyrethroids or carbamates, often combined with mineral oils (Asjes, 2000; Asjes and Blom-Barnhoorn, 2001; Martín-López *et al.*, 2006). These chemicals prevent aphids to feed and settle on a host plant, lower the transmission of viruses, and can reduce aphid reproduction rate and longevity. In 2012, approximately 80,000 kg of active insecticidal compounds were applied in the Netherlands on an estimated surface of 3000 $km^2$ (Centraal Bureau voor de Statistiek, 2015). The downside of these practices is the accumulation of harmful chemicals in the environment (Beketov *et al.*, 2013). There are several options to reduce the excessive application of insecticides. Growing crops in polycultures instead of monocultures and the use of biological control agents, such as ladybirds (Coccinellidae) and entomopathogenic fungi, can reduce the occurrence and impact of pest outbreaks (Obrycki *et al.*, 2009; Schoonhoven *et al.*, 2005; Shah and Pell, 2003). One of the cheapest and most effective pest management strategies is to use resistant crops (Russell, 1978). After generations of domestication, many cultivated plants have lost important resistance traits (Kollner *et al.*, 2008; Wink, 1988), making them vulnerable targets for herbivorous insects. Wild crop relatives are, therefore, a valuable resource for plant breeding programmes that aim to improve crop innate immunity to pests and pathogens. Essential in these programmes is the identification of resistant plant lines and development of genetic markers for resistance. Robust assessment of plant resistance to piercing-sucking insects is, however, a major bottleneck. Large-scale screening of hundreds of plants usually precludes to have controlled conditions and delicate scoring of insect performance. As a consequence, natural host-plant resistance mechanisms to aphids are largely elusive and hardly exploited in crops.

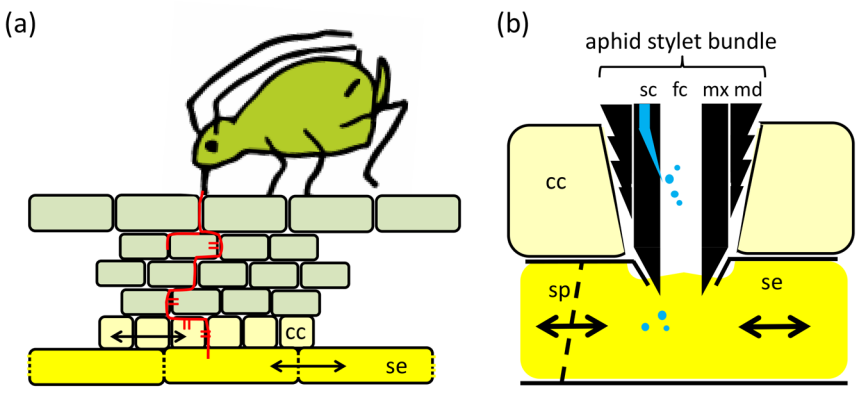



**Figure 1. Schematic overview of a probing aphid. a.** The aphid penetrates the plant intercellularly via a multibranched pathway (red). The majority of the cells is punctured (double red lines), in particular the companion cells (cc), until they reach a sieve element (se) where the aphids can ingest phloem sap. **b.** An aphid stylet tip in a sieve element. The hollow, needle-like structure of the aphid stylet bundle is shaped by mandibular stylets (md) and maxillary stylets (mx). The stylet bundle has a separate food channel (fc) and salivary channel (sc). After penetration of a sieve element (sp= sieve plate), watery saliva is secreted in the entrance of the food channel, and disperses most likely both into the sieve element and partially inwards into the food channel (dimensions are not drawn to scale).

## Aphid feeding behaviour

Aphids are highly specialised herbivorous insects that feed almost exclusively on the phloem. Although they are attracted to plant volatiles and visual plant cues, there is discussion on whether aphids can fly and navigate well enough to locate host plants from a long distance (Dixon, 1998; Kennedy *et al.*, 1961; Pickett *et al.*, 1992; Verheggen *et al.*, 2013). Their piercing-sucking mouthparts, or "stylet bundle", consists of specialised mandibles that form a hollow, needle-like structure containing a salivary channel and a food channel (Dixon, 1998). The outside of the stylet tip is covered with tactile- and chemoreceptive sensilla (Tjallingii, 1978). After exploration of the host plant surface with their proboscis, aphids manoeuvre their flexible stylet bundle intercellularly through the epidermis and mesophyll (**Figure 1**). While probing through the apoplast, they secrete gelling saliva. From microscopic studies it has been inferred that almost all cells along the trajectory are punctured (Tjallingii and Hogen Esch, 1993). In some cases these punctures only involve the cell wall, but often also the plasma membrane is punctured and small amounts of cell content are ingested without disrupting cell organelles. The companion cells surrounding the phloem sieve elements are usually punctured multiple times (Tjallingii and Hogen Esch, 1993). This behaviour is considered to play a role

in gustatory assessment of host plant quality and localisation of the phloem vessels (Powell *et al.*, 2006). Aphids require between 15 and 45 minutes of probing to reach the vascular bundle (Prado and Tjallingii, 2007; Tjallingii, 1994; van Helden and Tjallingii, 1993). Once the stylet bundle arrives at a sieve element, aphids inject watery saliva into the vessel before they ingest phloem (Prado and Tjallingii, 1994). Aphids can feed for hours or even days from a single sieve element (Tjallingii, 1995). In general, they feed passively. Phloem in intact plants has considerable turgor pressure, and aphids regulate the ingestion rate with an adjustable piston valve in the food channel (Dixon, 1998). Frequently, aphids ingest xylem sap, which has a putative function in osmoregulation (Pompon *et al.*, 2010).

## Aphid life-history characteristics

Phloem primarily consists of sugars. It is, therefore, not surprising that nitrogen and amino acids are growth-limiting factors for aphids. The aphid life cycle seems to be perfectly adjusted to this diet. Aphids spend most of their life feeding and process large quantities of phloem sap. Adults ingest at least their own weight of sap each day, and even more when nitrogen content of the phloem sap is low (Dixon, 1998). The excess of sugars is excreted via honeydew droplets. Aphids are characterised by polyphenism, i.e. the occurrence of multiple discrete phenotypes in genetically identical individuals. Among the several thousand aphid species, the vast majority is known to appear in both sexual and asexual morphs, and the latter occur with and without wings (Moran, 1992). Sexual morphs occur generally in autumn. During the largest part of the year,

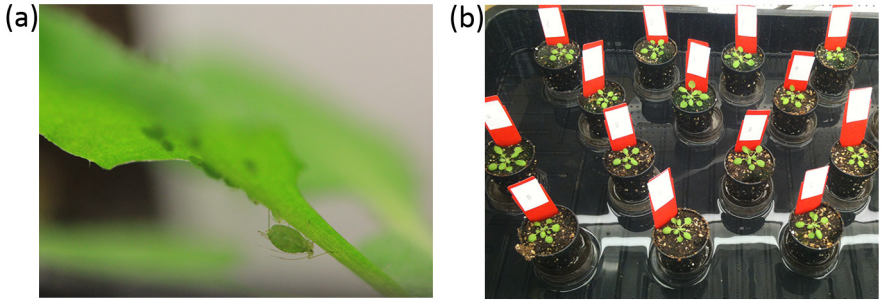



**Figure 2. Aphid population development assays. a.** *Myzus persicae* adult with offspring on an Arabidopsis plant. **b.** Aphid-infested *Arabidopsis thaliana* plants with a soap-diluted water barrier to prevent aphids from moving between plants.

however, parthenogenesis takes place and no time and energy is spent on finding a partner. Asexual morphs do usually not produce protein-rich eggs, but are viviparous. They have telescoping generations, implying that embryo development already starts before the mother is born. Aphid population size can therefore increase exponentially (Dixon, 1998). For example, one neonate *M. persicae* aphid on an Arabidopsis plant can establish a population of more than 40 individuals within two weeks (Kloth *et al.*, 2015), and an infestation with 5 neonate *Brevicoryne brassicae* aphids on a cabbage plant can result in more than 300 aphids within 30 days (Broekgaarden *et al.*, 2008). Since one aphid is enough for the (re-)establishment of a large aphid population, aphid pest management is a challenge.

## Host-plant resistance

Host-plant resistance to an herbivorous insect species can be defined as a heritable plant trait which reduces the degree of damage imposed by that insect species. Plants can defend themselves via several strategies. Antixenosis involves the repellence or deterrence of insects with, for example, volatiles or trichomes (Panda and Khush, 1995). After settlement of the insect on the host, antibiosis mechanisms can reduce insect population growth via detrimental proteins, such as protease inhibitors, or toxic metabolites, such as glucosinolates. Alternatively, some plants tolerate insect herbivory without any fitness effects. Finally, constitutive and induced resistance mechanisms are discriminated. The first occur irrespective of infestation, such as an epicutilar wax layer, while the latter are activated upon infestation, such as herbivore-induced plant volatiles (Dicke and Baldwin, 2010; Schoonhoven *et al.*, 2005).

## Plant resistance to aphids

With respect to aphids, there are few examples of absolute plant resistance. One of the most extreme resistance effects has been found in lettuce, where *Nasonovia ribisnigri* aphids could reach sieve elements but were unable to ingest phloem (Eenink *et al.*, 1982; ten Broeke *et al.*, 2013a). This *Nr*-gene-mediated resistance has been postulated to involve phloem proteins or other compounds that plug the aphid's food channel and block phloem ingestion (Martin *et al.*, 2003; ten Broeke *et al.*, 2013a). An alternative hypothesis presumes that protein aggregation and callose deposition in phloem sieve elements inhibit food ingestion by reduced phloem flow (Will *et al.*, 2013). Empirical data to test these hypotheses are however still lacking. Two other cases of near-complete resistance involve coiled-coil-nucleotide-binding-site-leucine-rich repeat (CC-NBS-LRR) proteins, which play a crucial role in the recognition of aphid infestation and

the induction of defence responses. In tomato, the CC-NBS-LRR protein Mi-1 severely affected the survival of the potato aphid *Macrosiphum euphorbiae* via a salicylic acid-dependent signalling pathway (Kaloshian, 2004; Li *et al.*, 2006; Rossi *et al.*, 1998). In melon, the CC-NBS-LRR protein Vat mediates resistance to *Aphis gossypii* and the viruses transmitted by this aphid (Dogimont *et al.*, 2014). The underlying mechanism that makes these plants more resistant to aphids remains, however, unclear. In contrast to the above mentioned examples of near-complete resistance, most studies reveal that plant resistance to aphids is a complex trait, involving multiple mechanisms, each with a small or moderate effect on aphid behaviour or population development. These quantitative resistance mechanisms include glucosinolates (Levy *et al.*, 2005; Mewis *et al.*, 2005; Pfalz *et al.*, 2009), camalexin (Kettles *et al.*, 2013), premature senescence (Pegadaraju *et al.*, 2005), lectins (War *et al.*, 2012), glandular trichomes (Alvarez *et al.*, 2006), cell wall modifications (Divol *et al.*, 2007; Dreyer and Campbell, 1984), and epicuticular wax components (Eigenbrode and Espelie, 1995). Key players in aphid-induced plant responses are the phytohormones jasmonic acid (JA), ethylene (ET) and salicylic acid (SA) (De Vos *et al.*, 2005; Kusnierczyk *et al.*, 2008; Kusnierczyk *et al.*, 2007; Thompson and Goggin, 2006). JA and ET are mainly associated with plant responses to necrotrophic pathogens and herbivores that cause substantial injury to the plant, such as caterpillars and thrips, and to a lesser extent also aphids (De Vos *et al.*, 2005). JA and ET influence the accumulation of secondary metabolites, such as glucosinolates, and digestibility reducers, such as proteinase inhibitors (Howe and Jander, 2008). In contrast to most other insects, aphids also induce SA, which is furthermore induced by e.g. biotrophic pathogens and insect eggs, and has antimicrobial properties associated with the accumulation of reactive oxygen species (De Vos *et al.*, 2005; Hilfiker *et al.*, 2014; Vlot *et al.*, 2009). SA and JA mainly act antagonistically, but both induce effective plant defence mechanisms against aphids.

## Genome-wide association mapping

To identify genes involved in plant resistance of aphids, quantitative trait locus (QTL) mapping and genome-wide association (GWA) mapping can be performed. Both methods employ natural genetic variation within plant populations, to find genomic regions and ultimately genes that are correlated to plant resistance. QTL mapping is most commonly applied in plant resistance studies (Young, 1996) and in general uses segregating generations derived from a biparental cross (Jansen, 1993). GWA mapping is, however, new in the field of plant-insect studies, and employs genetic variation in natural populations (Hirschhorn and Daly, 2005). By testing single nucleotide polymorphisms (SNPs) for their association with the phenotype, for example plant

resistance to aphids, genomic regions can be identified which are putatively involved in the regulation of the trait. One of the advantages of GWA mapping is, that it uses a large gene pool of several hundred genotypes harbouring ancient recombination events that may have been subjected to natural selection (Bergelson and Roux, 2010; Mitchell-Olds, 2010). In comparison to a QTL mapping population, GWA mapping populations contain more recombination events and have consequently smaller genomic regions linked to independent mutations. Therefore, fine mapping is usually not required to identify the gene of interest (Bergelson and Roux, 2010). Since GWA mapping only

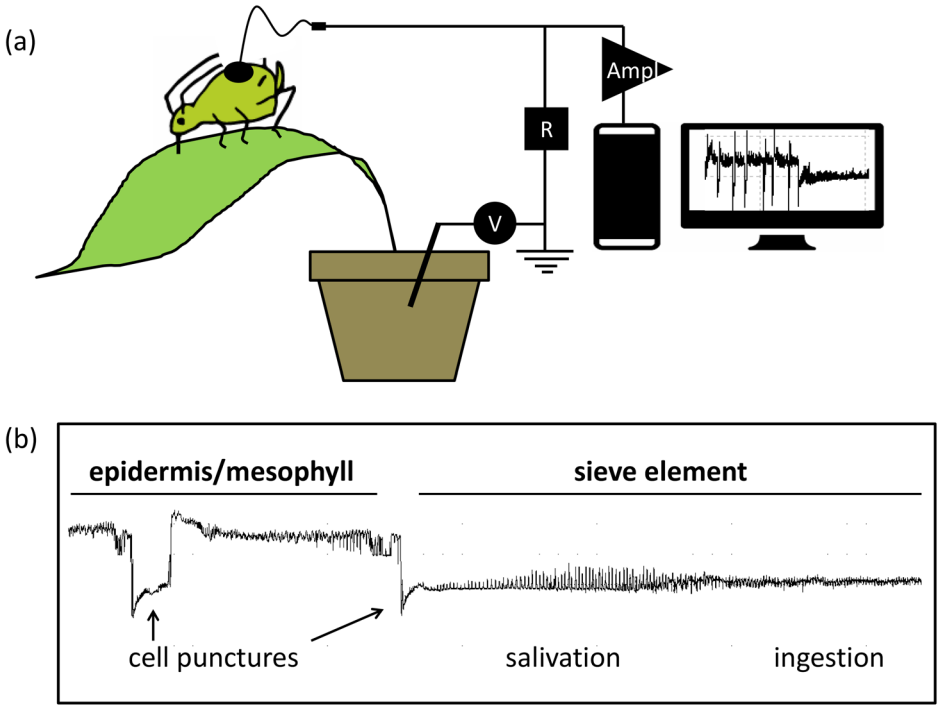



**Figure 3. Electrical Penetration Graph (EPG) setup. a.** The aphid and the plant are made part of an electrical circuit by an electrode in the potting soil and a golden wire glued to the aphid's abdomen. Electrical waveforms are recorded with EPG software (R=electrical resistor, V= voltage source, Ampl= signal amplifier connected to a computer, picture adjusted from www.epgsystems.eu). **b.** Electrical waveforms during 1.5 minutes of probing by a *M. persicae* aphid on an Arabidopsis plant. The aphid starts probing the epidermis/mesophyll cell wall, punctures the membrane of an epidermal/mesophyll cell during three seconds (potential drop), continues probing through the cell wall, and subsequently reaches a sieve element (again potential drop), where it salivates for approximately 25 seconds before it starts ingesting phloem.

delivers correlative information, a causal relationship between a candidate gene and the trait still needs to be validated in follow-up experiments. A common approach is to use knockout mutants or overexpression lines. In addition, allelic complementation and QTL mapping on recombinant inbred, near-isogenic, or double haploid lines can be employed to test for haplotype effects and epistatic interactions (Keurentjes *et al.*, 2007).

## Phenotyping techniques

There are several approaches to measure plant resistance to aphids. Since aphids do usually not inflict visual plant damage, not the plant but the aphids themselves are the benchmark for quantitative differences in plant resistance. Aphid population development is a robust estimate of plant resistance, and involves the assessment of the pre-reproductive period of aphids and the number of offspring within a time frame of several weeks (Wyatt and White, 1977) (**Figure 2**). The slower aphid development and the fewer offspring, the more resistant a host plant is. To elucidate the underlying mechanisms of plant resistance in more detail, characterisation of aphid behaviour is required. Choice assays can reveal differences in host plant preference due to e.g. plant volatiles, visual or gustatory cues (Powell *et al.*, 2006). For further in-depth profiling of feeding behaviour of piercing-sucking insects, a delicate technique has been developed, called Electrical Penetration Graph (EPG) recording (McLean and Kinsey, 1964; Tjallingii, 1988). In the EPG setup, the plant and the aphid are connected via an electrical circuit by inserting an electrode in the potting soil and gently attaching a golden wire to the aphid's abdomen (**Figure 3**). As soon as the aphid penetrates the plant cuticle, the electrical circuit is completed. Potential fluctuations in this circuit are generated by electrical resistance and electromotive force. Electrical resistance is affected by the valve in the aphid's food channel and the stylet tip and its location in the plant. The electromotive force fluctuates as a consequence of membrane potentials in plant cells and streaming potentials in the aphid's food channel (Tjallingii, 1988). These signals are amplified and recorded with EPG software (www.epgsystems.eu). Microscopic studies have characterised correlations between changes in the electrical signal (voltage level, amplitude and frequency), and the aphid stylet activity and position, such as penetration of epidermis/mesophyll tissue, salivation in the phloem, phloem ingestion, xylem ingestion and penetration difficulties (Tjallingii, 1985). EPG recording thus delivers detailed information on the duration and frequency of stylet activities in the plant. To assess the volumetric amount of food uptake by aphids, other approaches are required. Stylectomy is a technique that involves the collection of phloem sap exudates from

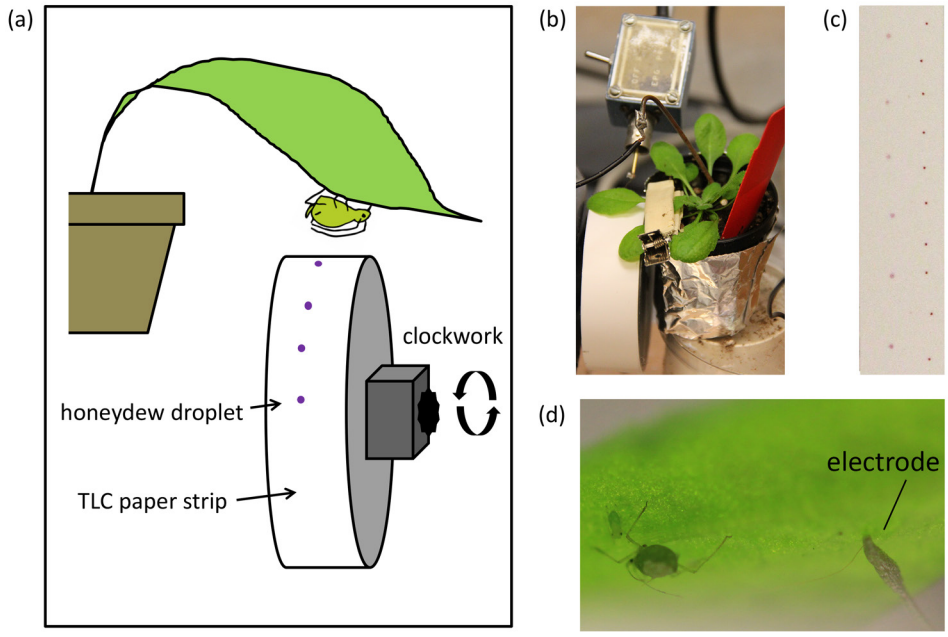



**Figure 4. Honeydew droplet collection with a honeydew clock. a.** Honeydew clock setup. **b.** Honeydew droplet collection from Myzus persicae aphids during EPG recording on Arabidopsis thaliana. **c.** TLC paper strip with honeydew droplets. Droplets were visualised with ninhydrin, a stain for amino acids. **d.** An adult M. persicae aphid was attached to the EPG electrode with a thin golden wire.

severed aphid stylets (Fisher and Frame, 1984; Kennedy and Mittler, 1953). In addition, food ingestion can indirectly be determined by quantification of honeydew excretion. This can simply be achieved by collecting honeydew droplets on a paper sheet underneath an infested plant. To assess the excretion rate, however, honeydew clocks are required (**Figure 4**). By collecting honeydew droplets on a rotating disc, the interval between droplet excretions can be determined, and with simultaneous EPG recording, this can be linked to the duration of phloem ingestion (Tjallingii, 1995). Although the above described techniques are indispensable for studying plant-aphid interactions, they are time- and labour-consuming. For large assays, such as the screening of GWA mapping populations consisting of several hundred plant lines, the availability of more efficient phenotyping techniques would be an asset.

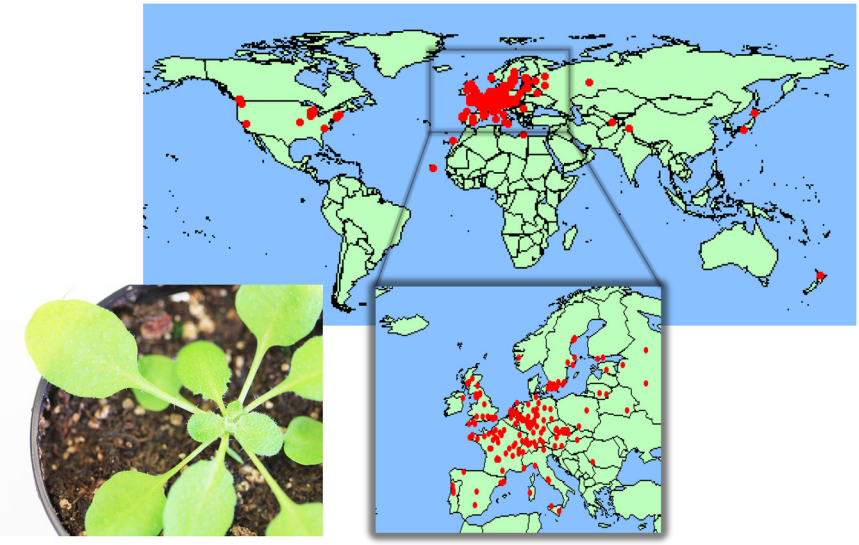


**Figure 5. Geographical distribution of the 350 natural accessions of *Arabidopsis thaliana* that were tested for resistance to *Myzus persicae* aphids.** Left bottom corner: A 4-week-old Arabidopsis plant (accession PHW-14, London, UK) in its vegetative state.

## Study system

In this thesis I study the genetic and physiological mechanisms underlying *Arabidopsis thaliana* (L.) Heynh. resistance to the green peach aphid, *Myzus persicae* (Sulzer). Among the several thousand species of Aphididae, *M. persicae* is considered to be the most polyphagous of all, occurring on plants from more than 40 plant families in the northern hemisphere, including the Brassicaceae (Blackman and Eastop, 2006). In addition, it is a common pest in greenhouses and is renowned as a vector of many plant viruses. To identify novel aphid-resistance genes, GWA mapping is performed on a natural population of 350 Arabidopsis accessions (**Figure 5**). Arabidopsis (Brassicaceae) is native to Eurasia and occurs in a variety of habitats throughout the northern hemisphere (Price *et al.*, 1994). Arabidopsis has been shown to display natural variation in abiotic stress responses and developmental traits (Alonso-Blanco *et al.*, 2009; Alonso-Blanco and Koornneef, 2000), as well as in defence mechanisms against insects (Kliebenstein, 2014). Although Arabidopsis is not a major host for aphids, aphid infestation of Arabidopsis in the field has been described (Arany, 2006; Mauricio and Rausher, 1997).

Arabidopsis was the first plant species to be sequenced (Arabidopsis Genome Initiative, 2000), and has widely been used as a model plant for a multitude of biological processes (Koornneef and Meinke, 2010). Some of the advantages of Arabidopsis are that it has a short generation cycle, a small diploid genome of five chromosomes, and is mainly a self-pollinating species with a high level of homozygosity (Abbott and Gomes, 1989). The availability of dense genetic maps for hundreds of accessions and transfer (T)-DNA insertion mutants for virtually every gene is unique to Arabidopsis, and gives the opportunity for unravelling the genetic architecture of plant resistance mechanisms. For GWA mapping one of the assets of Arabidopsis is, that the populations are highly segregated and that SNPs are generally in linkage with genomic regions of less than 50 kb (Nordborg *et al.*, 2005; Nordborg and Weigel, 2008).

## Scope of this thesis

This thesis is supported by The Netherlands Organisation for Scientific Research (NWO) through the Technology Foundation Perspective Programme 'Learning from Nature' (STW10989). The project objectives that I address are: (1) the development of a video-tracking platform to screen plants for resistance to *M. persicae* aphids, (2) the identification and characterisation of novel aphid-resistance genes and mechanisms in Arabidopsis, and (3) the utilisation of the first two deliverables to screen lily lines for resistance to *M. persicae* aphids. This thesis presents the results of the first two objectives.

**Chapter 2** is a literature review of the application of GWA mapping in plant-insect studies. GWA mapping has so far hardly been used to find genes involved in plant resistance to herbivorous insects. In this review the potentials and pitfalls of GWA mapping are discussed, including the challenges for phenotyping different aspects of plant resistance to herbivorous insects in several hundred plant lines.

In **Chapter 3**, a novel high-throughput methodology is presented to screen large plant populations for resistance to aphids. With an automated video-tracking platform detailed body movements of aphids are recorded and correlated to feeding behaviour. The efficiency and accuracy of the platform is tested with Arabidopsis and lettuce plants and two different aphid species, *M. persicae* and *N. ribisnigri*.

**Chapter 4** presents a multivariate GWA mapping analysis of 30 abiotic and biotic stress responses in 350 natural Arabidopsis accessions. To explore the common genetic basis of different plant resistance mechanisms, the phenotypic data of all projects in the 'Learning from Nature' programme are combined. This results in one of the first multi-trait GWA mapping studies that integrate such a broad variety of plant stress responses.

**Chapters 5** and **6** describe two novel aphid-resistance genes. 350 Natural Arabidopsis accessions are screened with the video-tracking platform, described in Chapter 3. Subsequently GWA mapping is performed on different aspects of *M. persicae* feeding behaviour to unravel the genetic architecture underlying plant resistance. **Chapter 5** presents the identification of *PHLO*, a small heat-shock-like gene with a previously unknown function. Characterisation of near-isogenic lines and T-DNA insertion mutants shows that *PHLO* is involved in the restriction of phloem ingestion by aphids and phloem responses to moderate heat stress. **Chapter 6** describes the role of the transcription factor WRKY22 in Arabidopsis resistance to aphids and its involvement in the modulation of the phytohormones SA and JA.

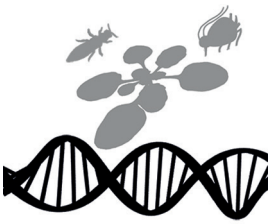
Finally, **Chapter 7** comprises a general discussion about the relationship between plants and aphids and the challenges in elucidating the genetic and physiological mechanisms underlying their interactions.

## Acknowledgements

I would like to thank Maarten Jongsma, Harro Bouwmeester and Marcel Dicke for constructive comments that helped to improve this chapter.

# Chapter 2

## Association mapping of plant resistance to insects



Karen J. Kloth*, Manus P.M. Thoen*,
Harro J. Bouwmeester, Maarten A. Jongsma and Marcel Dicke

* These authors contributed equally to this review.

*Trends in Plant Science, 2012, 17, 311-319*

## Abstract

Association mapping is rapidly becoming an important method to explore the genetic architecture of complex traits in plants, and offers unique opportunities for studying resistance to insect herbivores. Recent studies indicate that there is a trade-off between resistance against generalist and specialist insects. Most studies, however, use a targeted approach that will easily miss important components of insect resistance. Genome-wide association mapping provides a comprehensive approach to explore the whole array of plant defense mechanisms in the context of the generalist-specialist paradigm. As association mapping involves the screening of large numbers of plant lines, specific and accurate High-Throughput Phenotyping (HTP) methods are needed. Here, we discuss the prospects of association mapping for insect resistance and HTP requirements.

## Enhancing host-plant resistance against generalist and specialist insects

Host-plant resistance is one of the cornerstones of environmentally benign pest management systems (Panda and Khush, 1995; Schoonhoven *et al.*, 2005). Devastating pests and diseases only rarely occur in nature, which is due to the tremendous degree of natural variation in plant defense mechanisms (Alonso-Blanco and Koornneef, 2000; Anderson and Mitchell-Olds, 2011). Only a relatively small degree of such variation is contained in cultivated crop populations (Gols *et al.*, 2008), but wild populations provide ample opportunities for discovering novel mechanisms responsible for resistance to insects. A wide range of resistance mechanisms against herbivorous insects has been described (Panda and Khush, 1995; Schoonhoven *et al.*, 2005), and the impact of mechanisms depends on the characteristics of the herbivore, such as insect diet breadth (Mewis *et al.*, 2006; Rohr *et al.*, 2011). Although specialist insects, feeding on one or a few plant species within one family, are considered to be resistant to toxic compounds of their host (Karban and Agrawal, 2002), generalist insects are thought to thrive on a wider range of hosts with relatively low levels of allelochemicals (Loxdale *et al.*, 2011; Price *et al.*, 2011). Toxins, however, affect the performance of specialists as well (Vandenborre *et al.*, 2010), and generalists can cope with variable levels of secondary metabolites (Loxdale *et al.*, 2011), implying a more complex relationship between insect host range and plant defense. More insight into plant defenses against specialist and generalist insects is needed to understand how plants deal with herbivorous insects that differ in the degree of specialization and to improve host-plant resistance of economically important crops against insect pests. Most studies have addressed this topic with a targeted approach, focusing on only one or a few types of secondary

**Glossary box**

- **Association mapping**: a population based method of mapping quantitative trait loci (QTLs) that takes advantage of historic linkage disequilibrium to link phenotypes to genotypes (also known as "linkage disequilibrium mapping").
- **Candidate gene**: a gene, located in a chromosome region suspected of being involved in the expression of a trait of interest.
- **Confounding effect**: an extraneous variable in a statistical model that correlates (positively or negatively) with both the dependent and independent variable.
- **Genome-wide association (GWA) mapping**: comprehensive approach to systematically search the genome for causal genetic variation, using a large number of markers, by association between genotypes at each locus and a given phenotype.
- **High-Throughput Phenotyping (HTP)**: experimental set-up in which large amounts of specimens can be phenotypically screened, preferably automatic, fast, accurate, and with low costs.
- **Linkage disequilibrium**: two loci that are in linkage disequilibrium (LD) are inherited together more often or less often than would be expected by chance.
- **QTL**: Quantitative Trait Locus; a region in the genome that is responsible for variation in the quantitative trait of interest.
- **QTL mapping**: a family based mapping method using well known pedigrees to generate F2 crosses in which the genetic architecture of traits can be explored (also known as traditional linkage mapping).
- **Quantitative genetics**: the study of the heritability of quantitative traits, which are the products of two or more genes.



metabolites and a restricted amount of natural variation therein. In order to unravel the paradigm about resistance against specialists and generalists and to identify new plant defense mechanisms, comprehensive technologies are needed that can explore the apparent natural variation in multiple resistance mechanisms at the level of the genotype and phenotype.

Association mapping (see glossary) allows to screen many different wild and cultivated populations for genes involved in complex plant traits. Although association mapping has hardly been used in plant–insect studies thus far, it has the potential to allow new developments in eco-genomic studies of plant–insect interactions. One of the major prospects is the possibility to do genome-wide association (GWA) mapping in order to retrieve functional genetic loci involved in plant defenses against herbivorous insects in an untargeted way. GWA mapping involves the screening of large numbers of plant lines, which is currently a bottleneck because of the costs involved in this time- and

labor-intensive methodology. The large number of plant lines to be screened in insect resistance studies will require High-Throughput Phenotyping (HTP) techniques that succeed in accurately identifying different resistance traits. Particularly in view of the high diversity in insect-resistance mechanisms and their degree of specificity towards their enemies, this will pose some challenges. In this review, we discuss the perspectives of GWA mapping and HTP techniques in the context of insect resistance, with special reference to strategies against specialist and generalist insects.

## Association studies and linkage mapping

Understanding the genetic basis of phenotypic variation is one of the key goals in evolutionary biology. Family based QTL mapping (which uses well-characterized pedigrees (Balasubramanian *et al.*, 2009; Brotman *et al.*, 2011; Dobón *et al.*, 2011)) and association mapping (which uses linkage disequilibrium among numerous individuals of different populations (Atwell *et al.*, 2010; Ingvarsson and Street, 2011)) are the most commonly used tools for dissecting the genetic basis of phenotypic trait variation. In QTL mapping only a limited number of recombination events that have occurred within families and pedigrees can be studied, whereas with association mapping the recombination events that have accumulated over thousands of generations can be exploited (Zhu *et al.*, 2008). Since the 1980's, QTL mapping has been used most frequently, but association mapping is a promising alternative method for dissecting complex traits (Chan *et al.*, 2011; Chan *et al.*, 2010). Increased mapping resolution, reduced research time, and larger allele numbers have been put forward as main advantages over traditional QTL mapping (Yu and Buckler, 2006; Zhu *et al.*, 2008). Association studies can be divided into two broad categories: (i) candidate-gene association mapping, in which variation in a gene of interest is tested for correlation with the phenotypic trait of interest, and (ii) Genome-Wide Association (GWA) mapping, where genetic variation is explored within the whole genome, aiming to find signals of association with the complex trait (Zhu *et al.*, 2008) (see **Table 1** for an overview). Because GWA mapping is less dependent on prior information about candidate genes than QTL mapping and candidate-gene association mapping, this is a promising method to identify novel loci involved in complex phenotypic traits. However, GWA mapping should not be regarded as a replacement of traditional QTL mapping. In fact, GWA mapping and QTL mapping have complementary advantages and disadvantages, that can lead to a better understanding of causal genetic polymorphism when these approaches are combined (Chan *et al.*, 2010; Mitchell-Olds, 2010).

**Table 1.** Comparison of family based (QTL) and population based (association mapping) methods that aim to unravel the genetic basis of complex traits in plants.

| | QTL mapping | Candidate gene association mapping | Genome wide association mapping |
|---|---|---|---|
| **Main advantages** | - No population structure effects<br>- Identification of rare alleles<br>- Few genetic markers required | - Allows fine mapping<br>- Relatively low costs | - Allows untargeted fine mapping (blind approach)<br>- Detection of common alleles |
| **Main disadvantages** | - Limited genetic diversity<br>- Not always possible to create crosses<br>- Cannot distinguish between pleiotropic and physically close genes | - Detailed functional knowledge of trait is required<br>- No novel traits will be found | - Confounding effects due to population structure<br>- Will miss rare and weak effect alleles |
| **General requirements** | - Small 'original population size', low number of genetic markers, many replicates needed<br>- Generated mapping material (eg. F2 population, (AI-)RILs, MAGIC lines, NILs, HIFs etc.) | - Large population size, small number of genetic markers, the bigger the population size, the less replicates needed<br>- Prior genetic and biochemical knowledge on trait of interest<br>- Prior knowledge on LD, nucleotide-polymorphism, breeding system and population structure | - Large population size, many genetic markers, the bigger the population size, the fewer replicates needed<br>- Prior knowledge on LD, nucleotide-polymorphism, breeding system and population structure |
| **Recent case study in Arabidopsis** | QTL mapping with AI-RILs on flowering time (Balasubramanian *et al.*, 2009)<br>- 2 AI-RIL populationS (approximately 280 individuals each)<br>- 181 and 224 markers<br>- 12 to 70 replicates | Candidate gene approach on flowering time (Ehrenreich *et al.*, 2009)<br>- 251 accessions<br>- 51 SNPs<br>- 10 replicates per accession | Whole genome approach on multiple phenotypic traits (Atwell *et al.*, 2010)<br>- 199 accessions in total<br>- 216.150 SNPs<br>- 4 replicates in general |

Abbreviations: QTL; Quantitative Trait Locus, RIL, AI-RIL, Advanced Intercross-Recombinant Inbred Line; MAGIC, Multiparent Advanced Generation InterCross; NIL, Near-Isogenic Line; HIFs, Heterogeneous Inbred Family; LD, Linkage Disequilibrium; SNPs, Single Nucleotide Polymorphisms.
Combinations of these three approaches can allow the identification of false positives and negatives, but is much more laborious: a recent dual QTL mapping-GWA study (Brachi *et al.*, 2010) involved phenotyping nearly 20,000 individual plants, including 184 worldwide natural accessions genotyped for 216,509 SNPs and 4,366 RILs derived from 13 independent crosses. See Bergelson and Roux (2010) for an overview of different linkage mapping populations mentioned in this table.

## Association mapping in plant sciences

In the last decade, GWA mapping has emerged as a tool for studying the genetics of natural variation and economically important traits in plants (Atwell *et al.*, 2010). Flowering time, chemical composition, disease resistance, taste and many other economically and evolutionarily important traits have been studied in crop species (see (Zhu *et al.*, 2008) for an overview). Apart from agriculturally relevant crops, the model plant Arabidopsis (*Arabidopsis thaliana*) is of great value for understanding complex traits using GWA mapping (**Box 1**).

The presence of recombination events that have accumulated in plants over thousands of generations, is both an advantage as well as a potential pitfall of GWA mapping, because functional QTLs that are correlated with population structure can result in many false positives (Mitchell-Olds, 2010). Several statistical methods have been developed that use neutral genotypic information to account for confounding effects of population structure in GWA studies (Price *et al.*, 2006; Yu *et al.*, 2006; Zhao *et al.*, 2007). However, inadequate use of these models can lead to over-correction, resulting in false negatives which are equally problematic (Mitchell-Olds, 2010). Studies that have combined GWA- and QTL mapping strategies (dual linkage-association mapping) pointed out a false-positive rate of 40% and a false-negative rate of 24% in assays that solely involved GWA mapping (Bergelson and Roux, 2010). A major drawback of such a dual linkage-association mapping, however, is that it requires phenotyping of several thousands of individual plants, and the genesis of numerous linkage mapping populations (Brachi *et al.*, 2010). GWA mapping in regional mapping populations (instead of GWA mapping at the species scale) is an alternative approach to reduce confounding due to population structure (Bergelson and Roux, 2010).

Another major impediment in GWA studies is the phenomenon of missing heritability. Often, the associated QTLs can explain very little of the phenotypic variation, even after accounting for the effects of population structure. This phenomenon is attributed to several factors, including a scattered signal across numerous QTLs, each contributing to only a marginal proportion of the phenotype. Complex traits, such as insect resistance, are likely to encounter this problem (Myles *et al.*, 2009; Visscher, 2008). Integrating association mapping with transcriptional network analysis can decrease high false-positive rates and increase the resolution in scattered associations (Chan *et al.*, 2011). The scattering of genotype–phenotype associations can also be reduced by phenotyping multiple component traits instead of one multifactorial trait, as will be further discussed in the paragraph 'Requirements for phenotyping'.

**Box 1. Arabidopsis-insect interactions as a model for GWA studies**

The model species, Arabidopsis (*Arabidopsis thaliana*), is often used in plant-insect studies for obvious reasons, such as the availability of extensive information about genetic variation and physiology, and numerous mutants. Even though Arabidopsis is not a crop, there are numerous devastating crop pest insects (such as the generalist insect herbivores *Frankliniella occidentalis* and *Myzus persicae* and the specialist insect herbivores *Pieris rapae, Plutella xylostella* and *Brevicoryne brassicae*) that readily feed on Arabidopsis (Abe *et al.*, 2008; Bidart-Bouzat and Kliebenstein, 2011; Bruessow *et al.*, 2010; De Vos and Jander, 2009; De Vos *et al.*, 2005). However, one disadvantage in the light of insect-plant biology is that many accessions of Arabidopsis are winter annuals, so the life cycle of Arabidopsis does not temporally overlap with the life cycle of many herbivorous insects. It is known that herbivore performance (quantified in terms of mortality and developmental time) is commonly better on plants with such a 'pausing' strategy, indicating that such plants may invest less in defense traits (Van Poecke, 2007). This has likely influenced the evolution of signaling pathways in Arabidopsis, because the main biotic stresses likely comprise pathogens such as oomycetes, bacteria and fungi. Still, Arabidopsis is of great interest for studying insect resistance, since many insect defense mechanisms have been evolved within the Brassicaceae family, such as glucosinolates (Mewis *et al.*, 2005; Mewis *et al.*, 2006; Rohr *et al.*, 2011), and, many defense mechanisms against pathogens are also effective against herbivorous insects. Leaf toughness is for example effective against both microbial pathogens and insects (Schoonhoven *et al.*, 2005), and salicylic acid-, jasmonic acid- and ethylene-regulated defenses are involved in defenses against both pathogen and insect infestations (De Vos *et al.*, 2005; Mewis *et al.*, 2006; Pieterse *et al.*, 2009; Verhage *et al.*, 2011).

## Association mapping of plant-insect interactions

The complexity in the orchestration of insect resistance and its evolution in plants, makes it a difficult trait to study in a genomic context (Anderson and Mitchell-Olds, 2011). So far, only few GWA studies have been reported that deal explicitly with plant-defense mechanisms against herbivorous insects (see (Atwell *et al.*, 2010) for an example on aphids). One such study on glucosinolates (GSL) - secondary defense metabolites within the Brassicaceae family involved in resistance against herbivorous insects (Mewis *et al.*, 2005; Mewis *et al.*, 2006; Rohr *et al.*, 2011) - was conducted using 96 Arabidopsis accessions exhibiting 43 distinct GSL phenotypes and 230.000 SNPs (Chan *et al.*, 2010). In this study, GWA analysis successfully identified two major polymorphic loci controlling GSL variation in natural populations, but variation in resistance to specialist and generalist insects remains to be investigated for these accessions. This would require an experimental setup in which GWA mapping and HTP of insect resistance are integrated (**Figure 1**). GWA mapping of insect resistance will likely encounter similar obstacles as recognized in other GWA studies. Because insect resistance is generally

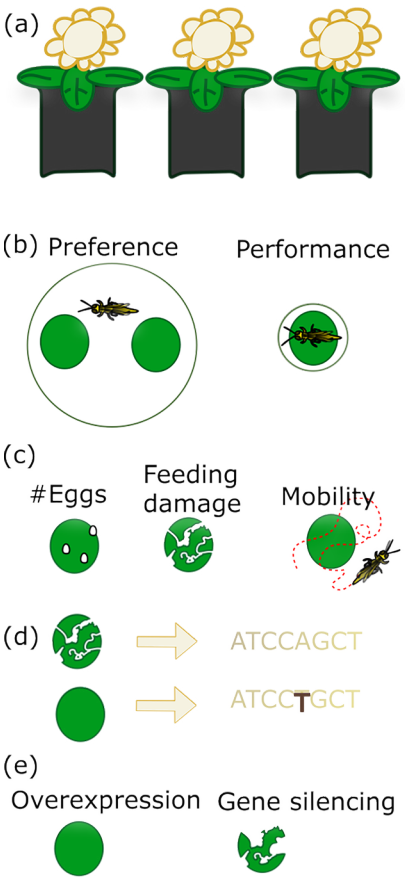



**Figure 1. Screening plants for insect resistance through GWA mapping.** This simplified overview shows how the genetic architecture underlying insect resistance can be determined in five steps, using GWA mapping. **(a)** Genotype SNPs for numerous accessions of the plant of interest; **(b)** Develop HTP choice and no-choice experiments to screen for insect preference and performance (using leaf discs in this example); **(c)** Screen for relevant insect-resistance parameters; **(d)** Find the genetic basis of phenotypic differences, using GWA mapping; **(e)** Validate candidate genes with reverse genetic tools, like overexpression and gene-silencing.

under strong positive selection pressure, GWA mapping of insect resistance might, however, unlike GWA studies of human diseases (Ingvarsson and Street, 2011; Myles *et al.*, 2009), be less affected by rare alleles that are not included in the haplotype map. Nevertheless, a good representation of all (sub)populations is indispensable for detecting variation in host-plant resistance and preventing them from having a too low allele frequency in the experimental set up. Particularly, the confounding effects of population structure can have a large effect on the success of GWA studies of host-plant resistance, because resistance against specific insects could have evolved independently and be based on different mechanisms in different populations and habitats (Poelman *et al.*, 2008b). Moreover, confounding effects due to strong population differences can be severe, when an intense evolutionary arms race between plant and herbivore has occurred as may be the case for specialist herbivorous insects and their host plants (Becerra, 2007; Poelman *et al.*, 2008b; Thompson, 2005; Vermeer *et al.*, 2011). This will require statistical correction of population structure, which can enhance the chance of false negatives due to over-correction. This problem is expected to be less evident with generalists, because they lack a reciprocal evolutionary interaction with specific plants (Price *et al.*, 2011).

## Resistance against specialist versus generalist insect herbivores

Specialist and generalist insect herbivores have different ways to deal with the defensive mechanisms of their host plants, and this is expected to result in different associations. Besides morphological and structural aspects, chemical defenses involving secondary metabolites play a major role in plant defense against insects (Schoonhoven *et al.*, 2005). Secondary metabolites can be divided into two broad functional categories, based on their modes of actions: qualitative compounds, which can be interpreted as toxins, and quantitative defensive compounds, with a dose-dependent effect, such as digestibility reducers (Price *et al.*, 2011). Recent studies show that qualitative compounds (e.g. GSL and alkaloids) often fail to affect specialist insects, because specialist insects evolved ways to detoxify or tolerate these compounds (Price *et al.*, 2011). In other words, if secondary metabolites play a role in defense against specialist insects, predominantly quantitative defensive compounds that reduce the digestibility are expected to be functional, whereas defense against polyphagous insects is mainly achieved by qualitative compounds. Toxins are even used by specialist insects to locate their host plants, or sequester these toxins for their own defense (Panda and Khush, 1995; Schoonhoven *et al.*, 2005). Thus, plants have to 'choose' between investing in substantial concentrations of qualitative compounds to deter polyphagous insects, or marginal concentrations of the same compounds to decrease preference by specialist insects (Poelman *et al.*, 2010; Van der Meijden, 1996). The evolution of defensive traits against generalists could, therefore, lead to an increased host-plant preference by specialists and vice versa. This trade-off between resistance to specialists and generalists is expected to be reflected in genotype-phenotype associations of the host plant. There are, however, many examples that do not support the qualitative-quantitative dichotomy. The generalist aphid *Myzus persicae* feeds on herbaceous plants in over 40 plant families, including families such as the Solanaceae that are well-known producers of toxic alkaloids (Blackman and Eastop, 2006; Loxdale *et al.*, 2011). Moreover, specific toxins do affect specialist herbivores. For instance, silencing nicotine production in tobacco (*Nicotiana tabacum*) results in improved performance of the specialist herbivore *Manduca sexta* (Steppuhn *et al.*, 2004) and overexpression of the lectin agglutinin in tobacco negatively affected the larval performance of *M. sexta* (Vandenborre *et al.*, 2010). Isothiocyanates, breakdown products of GLS, negatively affect the performance of the specialist herbivore *Pieris rapae* (Agrawal and Kurashige, 2003). The performance of *P. rapae* on the *coi1* mutants of Arabidopsis, that is compromised in the JA signal-transduction pathway, is significantly improved in comparison to wild-type plants, showing that even a specialist is affected by inducible plant defenses (Reymond *et al.*, 2004). Interestingly, the effects of quantitative and qualitative defenses may interact: nicotine prevents a compensatory response of the

generalist herbivore *Spodoptera exigua* to proteinase inhibitors and thus counters an insect adaptation to a qualitative defense (Steppuhn and Baldwin, 2007).

The main deficiency in addressing the defense mechanisms of plants against specialist and generalist insect herbivores, is that most studies have used a targeted approach, focusing on only one or a few types of secondary metabolites in a limited number of plant lines. Because resistance and tolerance are likely to be phenotypic traits that are composed of multiple factors, a targeted approach will easily miss important components. This is true for resistance to both generalists and specialists, but comparing the components and their relative strength of resistance to specialists and generalists may reveal how these traits are balanced.

A more comprehensive approach is, for example, taken in transcript profiling studies, where gene-expression signatures of infested plants and/or herbivorous insects are analyzed in different treatments (Bidart-Bouzat and Kliebenstein, 2011; Reymond *et al.*, 2004). Although several studies did not find a different plant response to specialist and generalist insects (Bidart-Bouzat and Kliebenstein, 2011; Govind *et al.*, 2010; Reymond *et al.*, 2004), one study (Agrawal and Kurashige, 2003) found a differential response in the insects that foraged on wild-type and mutant *Nicotiana attenuata*. The specialist *M. sexta* showed diet-specific alterations in gene expression, whereas the generalist *Heliothis virescens* regulated similar transcripts over different diets, indicating that the specialist is better adapted to both qualitative (nicotine), and quantitative (trypsin protease inhibitor) compounds of the host (Govind *et al.*, 2010). Another explorative approach is taken in a recent study, where metabolite fingerprints of *Plantago lanceolata* leaves differed after they were attacked by specialist or generalist herbivores, and by insects belonging to different taxa (Sutter and Muller, 2011). These examples show that untargeted approaches, such as transcript profiling, metabolic fingerprinting, and GWA mapping, allow to explore a large array of plant defense mechanisms in many plant lines.

## Requirements for phenotyping

Phenotyping is a prime factor in GWA mapping of host plant resistance. Among vast numbers of genome-wide markers, the aim is to achieve significant statistical power for only those molecular markers that are located close to the genes that influence the phenotypic trait of interest. In reality, functional associations between phenotype and molecular markers are often confounded, both in association and QTL mapping studies (Aranzana *et al.*, 2005; Chan *et al.*, 2010; Nemri *et al.*, 2010).

In the discussion about missing heritability of associations, where the identified genetic loci explain only little of the phenotypic variation, little attention has been paid to the role of phenotypes and phenotyping techniques. Some association studies of

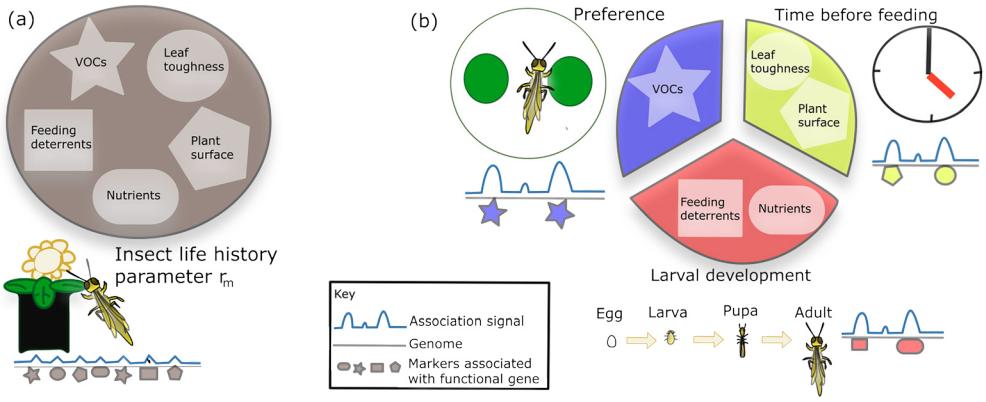



**Figure 2. Dissecting insect resistance into component traits.** Association mapping of a complex trait such as insect resistance can result in numerous associations with low statistical power. This is illustrated in **(a)** where the life history parameter $r_m$ of the insect is associated with many genetic loci. One approach to improve resolution in genotype-phenotype associations, is to dissect the complex phenotype into component traits **(b)**, e.g. insect preference (detection of repellent VOCs), time before the insect starts feeding (screening for the influence of leaf toughness and deterrent structures on the plant surface), and larval development (detection of e.g. feeding deterrents, toxins and nutrient content). Whereas the genetic architecture can overlap to some degree due to similar underlying processes, mapping these component traits will result in fewer genotype-phenotype associations with larger statistical power, and a higher proportion of functional associations. Genotype-phenotype associations can be further elucidated with, for example, metabolite fingerprints of VOCs, plant tissues or epicuticular waxes.

crop yield, for example, resulted in the characterization of numerous minor functional genes (Schon *et al.*, 2004). This confirms the infinitesimal model of Fischer (Fischer, 1918), which assumes a very large number of loci to be involved in quantitative genetics, each with a marginal effect on the phenotype. It is to be expected, however, that the number of functional (low-effect) QTLs involved is trait-specific. A complex trait is generally the result of numerous processes, which will result in a scattered association across multiple genetic loci: numerous QTLs are involved, that have a reduced statistical significance and each contribute to only a marginal proportion of the effect size of the phenotypic variation (**Figure 2**). Although a multifactorial character is inherent to complex traits, the efficiency of association mapping can be optimized by dissecting the phenotype into quantitative components with a minimum expected number of responsible mechanisms (Li *et al.*, 2011). A genome-wide screening within the scope of only a few mechanisms attributing to the trait of interest will increase the success of finding novel functional genes. A drawback is that it narrows the scope of a genome-wide

**Box 2. Plant resistance to herbivorous insects**

Host-plant resistance against herbivorous insects is generally defined as "the relative amount of heritable qualities possessed by the plant which influence the ultimate degree of damage done by the insect in the field" (Panda and Khush, 1995). Herbivorous insects use host plants for oviposition, feeding and shelter. Plants can achieve protection against herbivorous insects by both indirect defense, i.e. the attraction and facilitation of natural enemies of the insect herbivore, and direct defense against the pest insect (Dicke, 1999; Dicke and Baldwin, 2010; Heil, 2008). Three main categories of resistance against insect herbivores are (i) antixenosis, (ii) antibiosis, and (iii) tolerance (Panda and Khush, 1995).

Antixenosis mechanisms deter the insect, or, after the insect has arrived on the plant, prevent it from settling. Generally, the insect 'decides' not to colonize the plant due to the absence or low availability of an attractant, or the presence or quantity of a deterrent. A wide range of components can act as attractants or deterrents: Volatile Organic Compounds (VOCs), color, topology of the plant, chemicals and morphology of the plant surface (e.g. trichomes, epicuticular waxes, substrate texture), and physical and chemical characteristics of internal plant tissues (e.g. secondary metabolites, nutrient content, toughness of the cell wall) (Schoonhoven *et al.*, 2005). Herbivorous insects use olfactory and visual cues in the pre-alighting stage, and assess olfactory, visual, tactile, and gustatory traits after arriving on the host plant. Plants that exhibit antixenosis have a reduced number of initial colonizers and a relatively small population of herbivorous insects.

After the insect has 'decided' to utilize the host, antibiosis mechanisms of the host can affect insect performance (e.g. growth, development, reproduction, and survival) by toxins released after tissue damaging, feeding deterrents (e.g. protease inhibitors), nutritional imbalance or tissue toughness. Antibiosis causes a decrease in the insect population size (Panda and Khush, 1995). Plants can display antixenosis and antibiosis mechanisms constitutively, or after induction by e.g. herbivory or egg deposition (Dicke, 1999; Hilker and Meiners, 2006).

Finally, tolerance represents the plant's ability to compensate insect damage by increased growth, reproduction or repair of the damage. In contrast to antixenosis and antibiosis, tolerance does not severely affect the insect herbivore, but rather minimizes the impact of herbivory on the performance of the plant itself (Panda and Khush, 1995; Schoonhoven *et al.*, 2005).

survey. Complex traits are generally based on gene networks; therefore the assessment of individual components will likely overlook interactions between components, and the network as a whole and its environment (Benfey and Mitchell-Olds, 2008; Hammer *et al.*, 2006).

In insect resistance studies, typically multiple traits are phenotyped and reduced to one resistance variable, R. Most often, the total of life-history parameters of the insect are summarized in the variable $r_m$, the intrinsic rate of population increase (Awmack and Leather, 2002; Krips *et al.*, 1998). This summary statistic is an accurate parameter of the effect of resistance mechanisms on the herbivorous insects. However, insect

performance is typically dependent on multiple plant traits (e.g. nutritional components of the host plant and multiple resistance mechanisms of the plant (Awmack and Leather, 2002)). Hence, $r_m$ may lack resolution in association studies and using this parameter may result in a high proportion of missing heritability due to scattered signals (**Figure 2**). We expect that dissecting the complex parameter in multiple specific phenotypic components, e.g. host preference, time interval before the insect starts feeding, reproduction, larval development time, and mortality, will contribute to solving the problem of missing heritability and will help to identify multiple underlying mechanisms (**Figure 2**). The combination of these individual mechanisms will ultimately allow plant breeding to achieve sustainable host-plant resistance in crops. Indeed, multiparameter approaches, using a combination of phenotypic traits, for example both concentration of secondary metabolites and insect performance, have been postulated to deliver more significant relations to functional genetic data (Benfey and Mitchell-Olds, 2008; Eberius and Lima-Guerra, 2009). Apart from the parameterization of the phenotype(s), increasing the number of plant lines is of major importance for the statistical support of relevant associations (Ingvarsson and Street, 2011; Myles *et al.*, 2009). So far, most studies have used sample sizes of approximately 100 to 500 plant lines, but more genetic lines will increase the number and frequency of functional alleles and thereby improve the statistical power to detect them (Aranzana *et al.*, 2005; Chan *et al.*, 2010; Kang *et al.*, 2008; Li *et al.*, 2010; Zhao *et al.*, 2007). Secondly, a larger number of replicates within plant lines will increase the accuracy of the phenotype and the statistical support of genotype-phenotype associations. Particularly phenotyping insect resistance, involving the interaction among two or more organisms and species, is sensitive to stochastic errors and could result in relatively high levels of missing heritability. Although there is an example of successful GWA mapping by assessing aphid offspring in four replicates on 96 Arabidopsis lines (Atwell *et al.*, 2010), more replicates will reduce confounding effects. Moreover, the quality of phenotypic data can be improved by eliminating noise induced by the environment (Benfey and Mitchell-Olds, 2008; Hall *et al.*, 2010). Many studies have shown that insect resistance is an adaptive response to several biotic and abiotic factors (**Box 2**) (Ballaré, 2009; Chan *et al.*, 2011; Holopainen and Gershenzon, 2010; Poelman *et al.*, 2010). For example, it has been shown that the developmental stage of the plant altered the outcome of the GWA analysis, resulting in the identification of different functional genetic loci in different developmental stages (Chan *et al.*, 2011). This underlines the need for an experimental setup with uniform conditions among the genetic lines (**Figure 1**). Some noise will be inevitable for plant species harbouring a high diversity of ecotypes that differ in optimal growth conditions and development time. On the other hand, 'uniform' laboratory assays can deliver functional associations different from field conditions (Atwell *et al.*, 2010) due to genotype-by-environment

interactions (Hammer *et al.*, 2006). Including several (a)biotic treatments or an additional field assay could yield more field-predictive outcomes.

## High-throughput phenotyping

For quantitative traits such as insect resistance, reliable phenotyping requires a substantial amount of space, time and manpower, and this will be increasingly so in the context of association studies that require large sample sizes. There is, thus, a need for HTP methods that are accurate and yet predictive of field performance. Particularly, in view of the differential impact of mechanisms to specialist and generalist insects as discussed earlier, insect and plant performance are not necessarily correlated with each other, as high levels of deterrent compounds do not always negatively affect the performance of herbivorous insects (Poelman *et al.*, 2010; Van der Meijden, 1996), and good insect performance does not always result in reduced plant performance (**Box 2**). Therefore, both plant and insect traits are relevant for assessing the underlying mechanisms of insect resistance. Because association mapping requires at least hundreds of plant lines to be screened, it poses some challenges to the phenotyping efforts. Below some potential HTP techniques for assessing insect resistance are discussed.

## High-throughput phenotyping of plant defense

In the last decades, plant phenotyping techniques have gone through major developments (Fernie and Schauer, 2009; Montes *et al.*, 2007). Several of these methods can be applied to detect antixenosis, antibiosis or tolerance against insects and the benefits and costs involved for the plant (**Box 2**). Metabolite profiling techniques, like mass spectrometry and nuclear magnetic resonance, are the most obvious methods for screening primary and secondary proteins and metabolites in large-scale experiments (Fernie and Schauer, 2009). However, image processing techniques are also highly suitable for HTP platforms (Fernie and Schauer, 2009; Kokorian *et al.*, 2010; Montes *et al.*, 2007). These techniques translate changes in the spectral signature of a plant to quantify characteristics concerning plant growth, yield and (a)biotic stress. In the visible spectrum it is possible to detect damage caused by leaf chewing insects or for example silver damage due to thrips feeding (Abe *et al.*, 2008). Multi-colour fluorescence imaging has been used to assess feeding damage of mites and stylet penetrations of whiteflies (Buschmann and Lichtenthaler, 1998). In the near-infrared spectrum, stress-related changes in plants and changes in organic compounds can be detected (Chaerle and Van der Straeten, 2000; Cozzolino, 2009; Kramer *et al.*, 2000; Rutherford and vanStaden, 1996).

## High-throughput phenotyping of insect performance and preference

Assessing insect performance rather than that of the host plant, delivers the opportunity to study the direct and indirect impacts of plant nutritional quality and defense mechanisms on the dynamics of the herbivore population (Awmack and Leather, 2002). Although a wide variety of insect phenotyping techniques is available, only a marginal portion of these techniques is translated into high-throughput devices. This field in particular faces some challenges in developing methodologies that have low demands in terms of space, time and labor but are yet accurate and predictive of field performance. Most insect studies have focused on insect performance; e.g. population density, insect growth, development rate, fecundity, survival and the intrinsic rate of population increase ($r_m$) (Awmack and Leather, 2002). These parameters are correlated to both antixenosis and antibiosis. Assessing insect performance can be time consuming, depending on the generation time and life cycle of the insect, and is usually done in a non-automated way (Krips *et al.*, 1998; Poelman *et al.*, 2008a). Image analysis of photographs or videos represents potential for automated indexing of insect performance parameters (e.g. the number of eggs, larvae and surviving adults).

A behavioral assay can, in contrast to just monitoring insect performance or plant traits, result in a detailed chronological dataset of the process of host selection and food uptake. An additional advantage is that a behavioral assay can potentially be much shorter than an end-point measurement of reproduction and survival (Foster *et al.*, 2005; Hardie *et al.*, 1992). Food uptake is an important aspect of insect behavior, related to insect performance and host–plant resistance (Awmack and Leather, 2002). Electronic monitoring of probing behavior in piercing-sucking insects has proven to be successful in finding feeding deterrents (Backus and Bennett, 2009; Tjallingii and Hogen Esch, 1993), but is hardly feasible in large-scale experiments necessary for association mapping. Alternatively, automated tracking of insect behavior allows to measure multiple factors involved in host selection: e.g. host preference, mobility of the insect, and the timing and duration of food uptake. An additional advantage is that it allows to screen the behavior of multiple individual insects and multiple arenas simultaneously (Allemand *et al.*, 1994; Beeuwkes *et al.*, 2008; Lacey and Carde, 2011; Noldus *et al.*, 2002; Pistori *et al.*, 2010; Reynolds and Riley, 2002). The major challenge of high-throughput video-monitoring of insect behavior is to realize two- or three-dimensional arenas predictive of field performance (Prasifka *et al.*, 2010). In large field trials a mark-release-recapture technique can be a cost-effective method to assess host preference and population growth of insects (Hagler and Jackson, 2001). Ultimately, the choice of a phenotyping technique will largely depend on the study system and research focus.

## Future perspectives

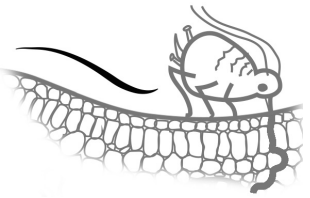
The development of accurate and field-predictive HTP will allow GWA mapping to increase insight into the genetic architecture of plant resistance to generalist versus specialist insects that will contribute to the development of host-plant resistance in crops. 'Blind' screening, unbiased by parental phenotypes and candidate genes, is the basis of this method and opens the opportunity to analyze the full scope of existing natural variation in resistance mechanisms. Although current studies mainly focus on one or a few candidate mechanisms, the untargeted nature of GWA mapping will include multiple factors that contribute to resistance against generalist and specialist herbivores. We expect that the current assumptions about differential resistance mechanisms against specialists and generalists can be addressed more comprehensively using such an unbiased approach. A further step forward will be the integration of association mapping with transcriptomics, proteomics and metabolomics, to assess insect resistance at the levels of the genotype, gene expression, and metabolite and protein networks (Chan *et al.*, 2011; DellaPenna and Last, 2008; Ingvarsson and Street, 2011; Keurentjes *et al.*, 2011; Keurentjes *et al.*, 2006; Macel *et al.*, 2010; Myles *et al.*, 2009). However, a major determinant of finding phenotype–genotype associations is imposed by the plant species itself. At present, Next Generation Sequencing technologies result in an increasing amount of sequenced plant species and lines within a species, so that the scope of plant-insect association studies will be expanded to additional biological systems with a wider array of plant–insect interactions and resistance mechanisms. In the near future, also the genomes and genetic variation of an increasing number of insect herbivores will become available (Whiteman and Jander, 2010). Comparing functional mechanisms in insect and plant populations at the genomic level, will allow the development of ecological insights in the evolution of plant-herbivore interactions and will take host-plant resistance studies to a next level.

## Acknowledgments

We thank Fred van Eeuwijk and three anonymous reviewers for stimulating discussions and comments on a previous version of the manuscript. The Netherlands Organization for Scientific Research (NWO) is acknowledged for funding through the Technology Foundation (grant STW10989, Perspectief Programme 'Learning from Nature').

# Chapter 3

## High-throughput phenotyping of plant resistance to aphids by automated video tracking



Karen J. Kloth, Cindy J.M. ten Broeke, Manus P.M. Thoen, Marianne Hanhart-van den Brink, Gerrie L. Wiegers, Olga E. Krips, Lucas P.J.J. Noldus, Marcel Dicke and Maarten A. Jongsma

*Plant Methods, 2015, 11, 4*

## Abstract

Piercing-sucking insects are major vectors of plant viruses causing significant yield losses in crops. Functional genomics of plant resistance to these insects would greatly benefit from the availability of high-throughput, quantitative phenotyping methods. We have developed an automated video tracking platform that quantifies aphid feeding behaviour on leaf discs to assess the level of plant resistance. Through the analysis of aphid movement, the start and duration of plant penetrations by aphids were estimated. As a case study, video tracking confirmed the near-complete resistance of lettuce cultivar 'Corbana' against *Nasonovia ribisnigri* (Mosely), biotype Nr:0, and revealed quantitative resistance in Arabidopsis accession Co-2 against *Myzus persicae* (Sulzer). The video tracking platform was benchmarked against Electrical Penetration Graph (EPG) recordings and aphid population development assays. The use of leaf discs instead of intact plants reduced the intensity of the resistance effect in video tracking, but sufficiently replicated experiments resulted in similar conclusions as EPG recordings and aphid population assays. One video tracking platform could screen 100 samples in parallel. Automated video tracking can be used to screen large plant populations for resistance to aphids and other piercing-sucking insects.

## Background

More than 100 aphid species (Aphididae) are economically significant pest insects and most crops are host to at least one species (Blackman and Eastop, 2006). Aphids feed on phloem sap, and to reach the phloem they move their stylets between plant cells towards a sieve element, making short punctures in cells along the way. Most probes are prematurely interrupted in the epidermis and mesophyll. When, however, a phloem vessel is reached, aphids can ingest phloem sap continuously for many hours or even days (Tjallingii, 1995). Although aphids inflict little tissue damage, they transmit plant viruses and deplete host plants of photoassimilates and free amino acids (De Vos *et al.*, 2005; Minks and Harrewijn, 1989). In wild plant populations aphids rarely constitute pests due to effective natural defence strategies, such as epicuticular waxes, protease inhibitors, and induced production of secondary metabolites (Ceci *et al.*, 2003; De Vos *et al.*, 2005; Du *et al.*, 1998; Eigenbrode and Espelie, 1995; Mewis *et al.*, 2005; Pegadaraju *et al.*, 2007; Smith and Boyko, 2007; Wang *et al.*, 2001; Will and van Bel, 2006). After generations of domestication many of these defence traits have been diminished or lost in cultivated plants, making them vulnerable targets of herbivorous insects (Kollner *et al.*, 2008; Wink, 1988). The genetic backgrounds of resistance mechanisms still remain

largely elusive and genomics studies strongly depend on the capacity for phenotyping large panels of plants. Few high-throughput methods have been established for assessing plant resistance to insect herbivores, such as aphids or other piercing-sucking insects (Chen *et al.*, 2012; Eenink and Dieleman, 1977; Kloth *et al.*, 2012; Smith *et al.*, 1994; Smyrnioudis *et al.*, 2002; Stelinski and Tiwari, 2013). Generally, two approaches are used to quantify the level of plant defence against aphids; either assessment of aphid population development or investigation of aphid feeding behaviour. Aphid population assays are generally the most demanding in terms of time and space, since they require the availability of a climate-controlled compartment for 1 or 2 weeks and extensive manual work (Broekgaarden *et al.*, 2008; Mewis *et al.*, 2006; Moran and Thompson, 2001). On the contrary, aphid feeding behaviour can be measured within a couple of hours via the Electrical Penetration Graph (EPG) technique. EPG recording delivers electrical waveforms comprising information on the plant tissue that is penetrated (phloem vessel, xylem vessel or other cells) and the stylet penetration activity (cell puncture, salivation, ingestion, penetration difficulties) (McLean and Kinsey, 1964; Tjallingii, 1988). EPG studies have shown that aphids prolong phloem ingestion on suitable host plants and delay and reduce feeding on resistant or non-host plants (Alvarez *et al.*, 2006; Boquel *et al.*, 2011; Nalam *et al.*, 2012b; ten Broeke *et al.*, 2013b; Tjallingii, 2006; van Emden and Harrington, 2007; van Helden and Tjallingii, 1993; Will *et al.*, 2007). The high specificity of the information about plant tissue and key components of aphid behaviour, makes this methodology appealing for exploring defence mechanisms. A drawback of EPG is, however, the restricted capacity, generally 8 plants per setup (ten Broeke *et al.*, 2013a), and the labour-intensive nature of wiring aphids and annotating electrical signals.

Here, we present the methodology and validation of image-based tracking of aphid feeding behaviour. Automated video tracking was introduced in the early 1990s and has since been used in many animal behaviour studies (Bakchine *et al.*, 1990; Dell *et al.*, 2014; Moreno-Delafuente *et al.*, 2013; Noldus *et al.*, 2002; Pinkiewicz *et al.*, 2011; Shcherbakov *et al.*, 2010; Spitzen *et al.*, 2013). Video tracking involves software-engineered pattern analysis of grids of pixels in order to quantify the location and movement of subjects over time. In this study, we used movement patterns of the central body point of aphids to estimate the duration of plant penetrations made by the aphid's mouth parts. Previous EPG studies showed that probes shorter than approximately 3 minutes represent penetrations in the epidermis and/or mesophyll (van Emden and Harrington, 2007), and that probes involving phloem uptake last on average at least 25 min (Prado and Tjallingii, 2007; Tjallingii, 1994; van Helden and Tjallingii, 1993). This allowed us to discriminate test probes from putative phloem uptake events in video observations in order to identify plants that are resistant to aphids. We benchmarked the performance of the high- throughput video tracking platform against EPG recordings and aphid

population development assays, using natural accessions of *Arabidopsis thaliana*, and lettuce cultivars, *Lactuca sativa*, in combination with the green peach aphid, *Myzus persicae* (Sulzer), and the black-currant lettuce aphid, *Nasonovia ribisnigri* (Mosely) (Hemiptera: Aphididae), respectively.

## Results

### Tracking aphid feeding behaviour

Automated video tracking of aphid feeding behaviour was performed using video tracking software and a stationary camera mounted above 20 no-choice arenas. We introduced one aphid onto each arena, consisting of an agar substrate almost completely covered by a leaf disc, and recorded 20 arenas simultaneously with a frame rate of 25 frames $s^{-1}$ (**Figure 1**, **Figure S1**). Because the aphid's mouthparts were not visible in the multi-arena setup, we made the assumption that when the aphid's centre point was located on the leaf disc and did not move, the aphid was penetrating the leaf tissue with its stylets. By assessing video images by eye, we defined velocity thresholds for the start and end of probing events of two aphid species, *M. persicae* and *N. ribisnigri* (**Figure 2**, **Figure S2**). According to our observations, the software was more vulnerable to premature probe endings of *N. ribisnigri* due to body movements during probing (such as event γ in **Figure 2**). As this aphid species is somewhat larger (± 1.9 mm body length, versus ± 1.7 mm for *M. persicae*), movements around the fixated mouth resulted in a higher tangential velocity, and therefore required a higher velocity threshold.

### Accuracy

To test the accuracy of the platform, we performed automated video tracking and human observations simultaneously. A camera was attached to a stereo microscope to deliver a side-view on the arena for manual scoring of probes (**Additional file 2**). Among a total of 139 probes of 16 different *M. persicae* aphids scored by hand, 88% was detected with video tracking (**Figure 3a**). Undetected and false positive probes involved only short events (< 3 min). Of the detected probes, 19% was either underrated (multiple 'true' probes were considered as one probe), or overrated (one 'true' probe was translated into multiple probes by the software). Underrated samples were caused by undetected probe stops due to slow movements below the velocity threshold. Overrated samples were caused by false probe stops when, for example, the aphid was immobile on the edge of the leaf disc and the assigned position continuously switched between an "on the leaf disc" and "off the leaf disc" status (**Figure 3b**). Three times this incident occurred, leading to 17 redundant probes of which 10 were filtered out automatically (see Materials and methods, section Software settings). Other reasons

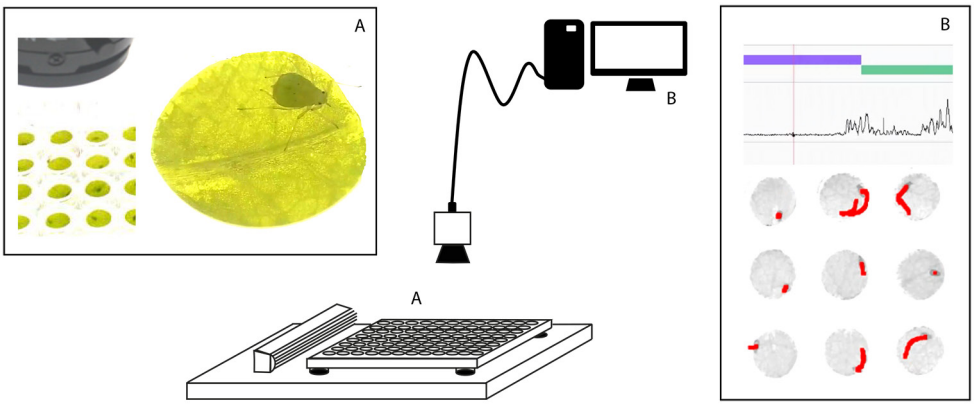



**Figure 1. Video tracking platform.** A stationary camera is mounted above a microtitre plate which is placed on top of a backlight unit with ventilation at the left. Wells in the microtitre plate contained a leaf disc and an aphid (A). Cling film was wrapped around the plate to prevent aphids from escaping. The camera was connected to a computer with EthoVision XT video tracking software (B). Movements of the aphid's centre point were automatically tracked (red track shows movements across 30 seconds). With this information the software calculated aphid velocity (line graph) and estimated probing (purple bar) and non-probing events (green bar).



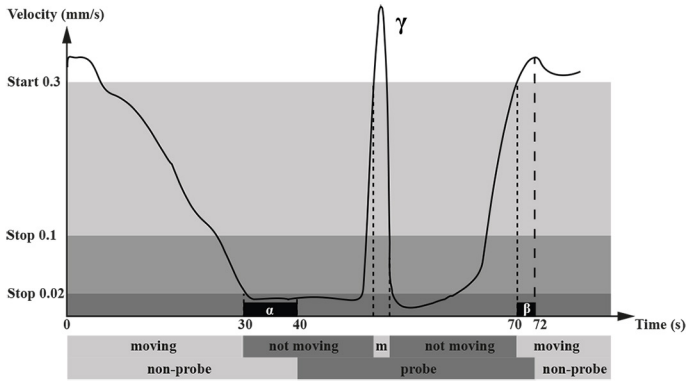



**Figure 2. Velocity thresholds for registration of probes.** An example of how aphid feeding behaviour was measured using a resolution of 275 pixels per mm$^2$. Subject states can be defined as 'moving' or 'not moving' by means of two thresholds: the start velocity at which the subject begins to move, and the stop velocity at which the state changes from moving to not moving. Probe starts were recorded if the velocity of the aphid's centre point dropped below 0.02 mm/s for at least 10 seconds (α). Probe stops were recorded if the velocity of *M. persicae* aphids exceeded 0. 3 mm/s for at least 2 seconds (β), or 0.35 mm/s for at least 2 seconds in the case of winged *N. ribisnigri* aphids. To avoid premature probe endings due to short movements during probing (event γ), probe stops were only recorded when velocity increased above 0. 1 mm/s for more than 2 seconds. Figure adjusted from the EthoVision XT Reference Manual (version 8) (Grieco *et al.*, 2010).

for premature probe stops were abdominal movements during probing related to e.g. reproduction or honeydew excretion. The longer probes lasted, the higher the risk was of encountering such incidents. Indeed automatically tracked probes were in general biased to end 73 to 12 seconds too early (**Figure 3c**), and the total duration of probing was underestimated (on average 46 min ± 2.5 min standard error, versus 50 min ± 1.9, P=0.01, Mann-Whitney U test, total observation duration: 55min). Nevertheless, the recorded number and duration of probes were highly correlated to human observations (**Figure 4**, average $r^2$=0.7 with 275 pixels per $mm^2$). Other parameters, such as distance moved, were also highly correlated with feeding behaviour in general, but were less informative with regard to long probes (**Figure 4l**). Although automated video tracking did not achieve a precision as high as manual scoring, it enabled observing multiple arenas simultaneously. In the above described tests, we used 275 pixels per $mm^2$, equal to a coverage of 20 arenas with our 768 x 576 pixels camera. To determine whether the capacity could be increased, we repeated the experiment with only 155 pixels per $mm^2$, equal to a coverage of 35 arenas, but found that reduced resolution resulted in decreased correlations with human observations (average $r^2$<0.5).

### Benchmarking against EPG recording with Arabidopsis

To validate whether automated video tracking delivered a reliable proxy for plant resistance, feeding behaviour of *M. persicae* was measured during 8 hours continuous recording on two natural accessions of Arabidopsis, Co-2 and Sanna-2 (**Additional file 3**). These accessions were selected from a population of hundreds of accessions based on preliminary video tracking data. Automated video tracking showed that *M. persicae* aphids walked larger distances on Co-2 and reduced the duration of long probes (**Table 1**). EPG recordings on intact plants confirmed shorter durations of (sustained) phloem ingestion, and additionally revealed more short probes, non-probing behaviour and a delayed phloem uptake on Co-2 (**Table 1**). This behaviour is an indication of both epidermis/mesophyll-located and phloem-located resistance in Co-2 against *M. persicae*. All aphids ingested phloem, but quantitative differences in feeding behaviour between aphids on Co-2 and on Sanna-2 were already apparent in the first hour (**Figure 5**). A reproduction assay on intact plants confirmed that Co-2 was indeed more resistant than Sanna-2, although the resistance was not absolute. Depending on plant age, aphids either started reproduction later or produced fewer offspring (**Figure 6**). Although we had been able to correctly identify a quantitative difference in resistance with automated video tracking, the effects were smaller than in EPG recordings on intact plants. To verify whether the plant line effects in the video tracking assay were attenuated due to the use of excised plant tissue, the EPG experiment was repeated with leaf discs. Particularly for the resistant accession, aphid feeding behaviour was different and involved more

phloem uptake and fewer short probes on leaf discs compared to intact plants (**Table S2**). The only significant difference between the accessions that remained was a reduced duration of phloem uptake events on Co-2 (**Table 1**). In addition, contribution of salivation to the phloem phase, required to suppress (callose-mediated) sieve-plate occlusion (Will *et al.*, 2009), was equal on leaf discs but higher on intact plants of Co-2 (**Figure 7**). This indicates that the resistance mechanisms in intact plants were partially lost in leaf discs.

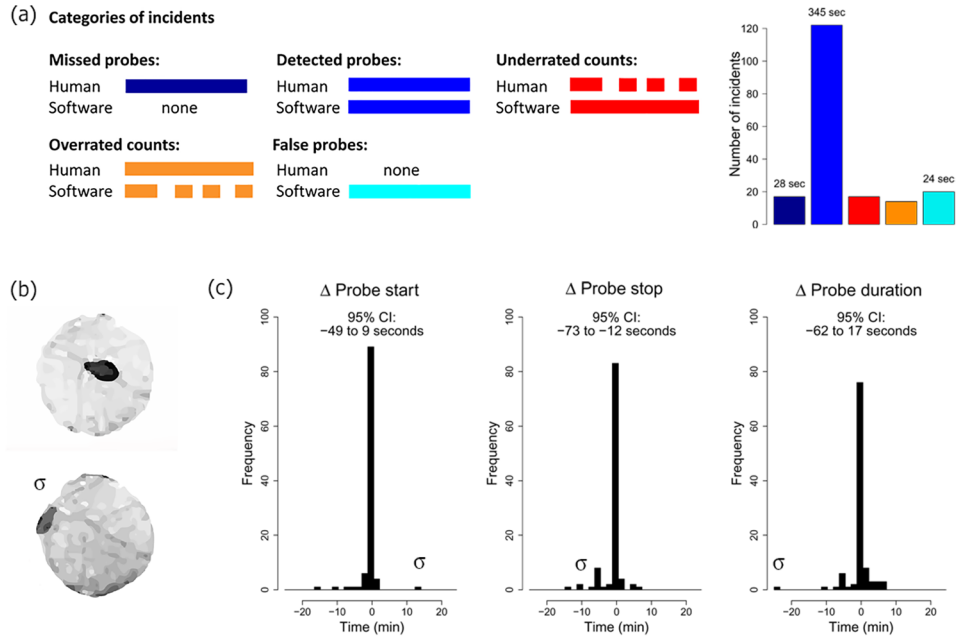



**Figure 3. Accuracy of automated tracking in comparison to human observations.** *M. persicae* feeding behaviour was measured on Arabidopsis leaf discs by automated video tracking and human observations simultaneously. **(a)** Out of 139 probes of 16 aphids scored by hand, 88% was detected by automated video tracking. Probes were considered a match when their duration overlapped at least partially. Some of the detected probes were matched by too few (underrated) or too many (overrated) probes. For these situations, the amount of missed or redundant probes is shown. 17 Probes went undetected and 20 false probes were recorded. Mean duration per probe is shown above the bars. **(b)** Screenshots of the top-view video used for automated tracking. The lower image (σ) shows an aphid positioned on the edge of the leaf disc for more than 20 min, causing overrated probe counts by the software due to continuous switching between an "on the leaf disc" and "off the leaf disc" status. **(c)** Differences between software and human observations per matched probe. 95% Confidence Intervals are shown above the histograms. Negative values correspond to too early probe starts, too early probe endings, resp. too short duration of probes compared to the human observations. In case of overrated probe counts, the probe with the most similar duration as the manually scored probe was included. The outlier caused by the example in (b) is annotated with σ.

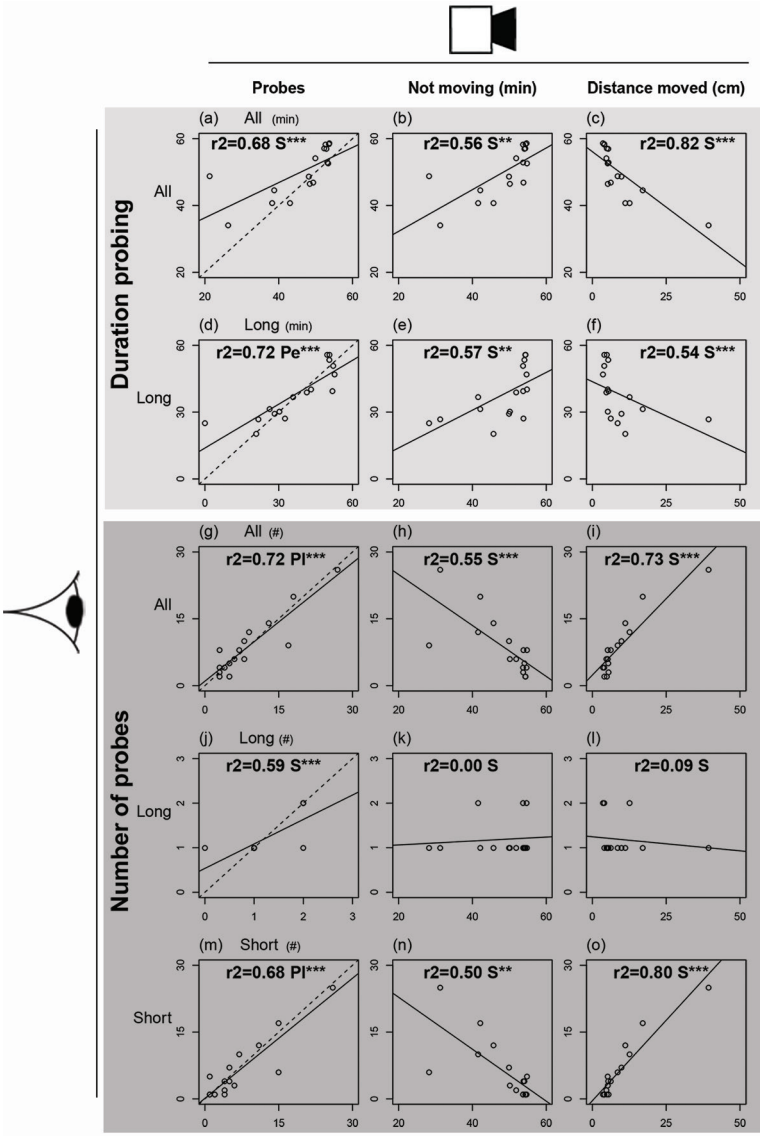



**Figure 4. Correlation between automated video tracking and human observations.** *M. persicae* behaviour was measured by automated video tracking (x-axes) and human observations simultaneously (y-axes). Three categories of probes were distinguished: All probes, Long probes (> 15 min), and Short probes (< 3 min). The duration (min) and number of probes measured by human observations were compared to: **(a,d,g,j,m)** the duration (min) and number of probes (all, long, and short probes) measured by video tracking, **(b,e,h,k,n)** the total time not moving (min), and **(c,f,i,l,o)** the distance moved by the aphids (cm) (*P<0.05; **P<0.01; ***P<0.001, Pe = Pearson correlation test, Pl = Pearson correlation test on log transformed data, S = Spearman correlation test, dashed lines represent a hypothetical $r^2$=1, n=16 recordings of 1 aphid for 55 min, 275 pixels per $mm^2$).

**Table 1. Feeding behaviour of *M. persicae* on two Arabidopsis accessions and *N. ribisnigri* on two lettuce cultivars.** Within each EPG and video tracking experiment it was tested whether aphid behaviour differed between the two plant lines. The mean duration and latency were only calculated if the corresponding events did occur (all samples of *M. persicae*; *N. ribsinigri*: EPG n=2 Corbana and n=20 Terlana, video tracking n=21 Corbana and n=26 Terlana).

| | | Arabidopsis – *M. persicae* | | Lettuce – *N. ribisnigri* | |
|---|---|---|---|---|---|
| | | Co-2 (R) | Sanna-2 (S) | Corbana (R) | Terlana (S) |
| EPG intact plants | Total duration non-penetration | 139 ± 15 | 60.2 ± 8.1 Ts*** | 121 ± 16 | 29.2 ± 5.8 Tl*** |
| | Total duration phloem feeding | 110 ± 13 | 296 ± 31 M*** | 2.34 ± 1.68 | 399 ± 17 M*** |
| | Total duration phloem feeding (> 10 min) | 27.7 ± 7.9 | 259 ± 32 M*** | 0.00 ± 0.00 | 399 ± 17 M*** |
| | Total duration other penetration activities | 231 ± 15.8 | 124 ± 28 M** | 357 ± 16 | 51.5 ± 14 T*** |
| | Number of non-penetrations | 57.7 ± 6.2 | 22.4 ± 3.1 M*** | 27.8 ± 7.58 | 7.58 ± 1.33 T*** |
| | Number of short probes (< 3 min) | 49.8 ± 5.8 | 18.4 ± 2.4 M*** | 13.7 ± 2.0 | 4.53 ± 1.17 M*** |
| | Number of phloem feeding events | 6.37 ± 0.55 | 4.28 ± 0.70 T* | 0.11 ± 0.07 | 1.16 ± 0.12 M*** |
| | Mean duration of phloem feeding events | 17.6 ± 1.4 | 117 ± 26 M*** | 22.2 ± 6.16 | 379 ± 26 M* |
| | Latency to first phloem feeding event | 188 ± 25 | 111 ± 26 M* | 228 ± 94 | 106 ± 17 M |
| | Contribution salivation to phloem phase (%) | 13.3 ± 2.2 | 2.45 ± 1.3 M*** | 98.7 ± 1.0 | 1.15 ± 0.3 M*** |
| EPG leaf discs | Total duration non-penetration | 164 ± 48 | 173 ± 59 M | - | - |
| | Total duration phloem feeding | 220 ± 42 | 201 ± 48 T | - | - |
| | Total duration phloem feeding (> 10 min) | 158 ± 46 | 177 ± 47 T | - | - |
| | Total duration other penetration activities | 95.3 ± 25.7 | 106 ± 17 T | - | - |
| | Number of non-penetrations | 19.1 ± 5.7 | 14.4 ± 3.0 Ts | - | - |
| | Number of short probes (< 3 min) | 16.6 ± 4.8 | 8.63 ± 2.0 Ts | - | - |
| | Number of phloem feeding events | 8.11 ± 1.53 | 2.63 ± 0.84 M* | - | - |
| | Mean duration of phloem feeding events | 34.2 ± 10.8 | 99.2 ± 31.7 Ts* | - | - |
| | Latency to first phloem feeding event | 144 ± 51 | 250 ± 53 T | - | - |
| | Contribution salivation to phloem phase (%) | 3.03 ± 0.6 | 3.63 ± 2.2 M | - | - |
| Video leaf discs | Total duration non-probing | 61.5 ± 9.7 | 42.6 ± 9.4 T | 243 ± 27 | 148 ± 25 T* |
| | Total duration long probes (> 25 min) | 338 ± 20 | 377 ± 17 T | 153 ± 26 | 283 ± 25 T*** |
| | Total duration sust. probes (> 35 min) | 276 ± 27 | 353 ± 19 M* | 132 ± 24 | 260 ± 25 T*** |
| | Total duration other probes (<= 25 min) | 80.6 ± 12.4 | 60.5 ± 9.9 T | 84.1 ± 7.0 | 48.4 ± 9.3 M*** |
| | Number of non-penetrations | 33.2 ± 3.9 | 22.1 ± 3.8 T | 31.1 ± 3.1 | 23.0 ± 2.6 T |
| | Number of short probes (< 3 min) | 20.3 ± 3.1 | 12.4 ± 3.4 T | 22.4 ± 3.1 | 15.6 ± 2.3 M |
| | Number of long probes (>= 25 min) | 5.95 ± 0.36 | 4.71 ± 0.36 M* | 2.19 ± 0.33 | 3.25 ± 0.30 T* |
| | Mean duration of long probes (>= 25 min) | 62.8 ± 7.0 | 90.0 ± 9.8 M* | 72.4 ± 9.5 | 99.8 ± 11.2 M* |
| | Latency to first long probe (>= 25 min) | 18.5 ± 5.8 | 27.1 ± 7.1 M | 142 ± 27 | 65.1 ± 14.2 M* |
| | Total duration not moving (min) | 445 ± 7.3 | 465 ± 3.5 M | 257 ± 27 | 350 ± 25 T* |
| | Total distance moved (cm) | 46.9 ± 4.6 | 34.0 ± 4.6 T* | 204 ± 47 | 103 ± 23 T |
| | Max velocity (mm/s) | 0.72 ± 0.12 | 0.53 ± 0.05 T | 2.78 ± 0.54 | 1.88 ± 0.20 M |

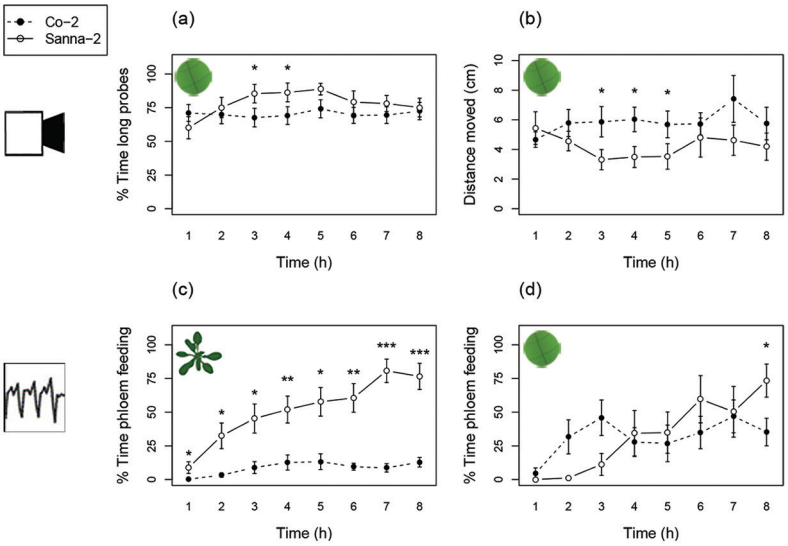



**Figure 5. Behavioural parameters of** *M. persicae* **on two natural Arabidopsis accessions, Co-2 (resistant) and Sanna-2 (susceptible). (a)** Percentage of the time spent on long probes (> 25 min), and **(b)** distance moved (cm) were measured by automated video tracking. Percentage of the time spent on phloem feeding (waveform 5) were measured by **(c)** EPGs on intact plants, and **(d)** EPGs on leaf discs (Mann-Whitney U test, *P<0.05; **P<0.01; ***P<0.001, video tracking: Co-2 n=20, Sanna-2 n=17, EPG recording intact plants: n=19, EPG recording leaf discs: Co-2 n=9, Sanna-2 n=8, error bars represent standard error).

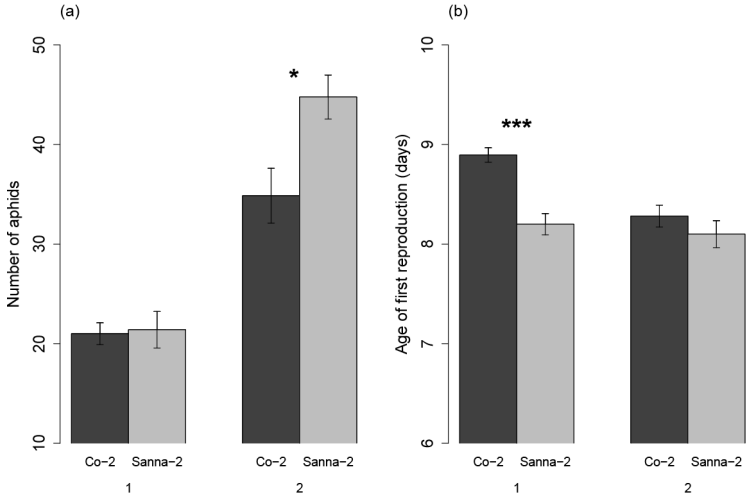



**Figure 6. Reproduction of** *M. persicae* **on two Arabidopsis accessions.** One neonate aphid was introduced to a 2.5-week-old plant (assay 1) or a 3.5-week-old plant (assay 2). **(a)** Total number of aphids per plant 2 weeks after infestation. **(b)** Days until the first nymph was produced by the aphid (Mann-Whitney U test, *P<0.05, ***P<0.001, assay 1: Co-2 n=19, Sanna-2 n=15, assay 2: Co-2 n=14, Sanna-2 n=13, error bars represent standard error).

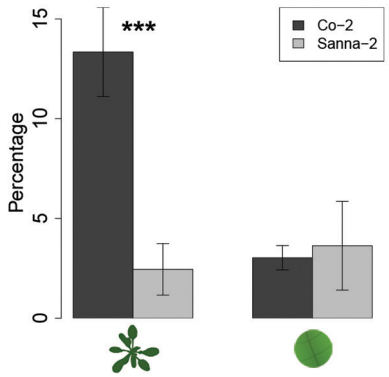



**Figure 7. Contribution of salivation to phloem ingestion.** Percentage of time spent salivating in the phloem compared to the total phloem phase (salivation + ingestion) of *M. persicae* aphids on Arabidopsis accessions Co-2 (resistant) and Sanna-2 (susceptible) (Mann-Whitney U test, *P<0.05; **P<0.01; ***P<0.001, left bars: EPG recording intact plants: n=19, right bars: EPG recording leaf discs: Co-2 n=9, Sanna-2 n=8, error bars represent standard error).



## Benchmarking against EPG recording with lettuce

Apart from a study system with partial resistance, an example of near-complete resistance was tested with the video tracking platform. The behaviour of black-currant lettuce aphids, *N. ribisnigri*, biotype Nr:0 was recorded on two near-isogenic lettuce cultivars, the resistant 'Corbana' and susceptible 'Terlana'. Previous studies showed that the *Nr* gene is responsible for near-complete resistance in Corbana against this biotype of aphids, mainly due to a phloem-located mechanism (Reinink and Dieleman, 1989; ten Broeke *et al.*, 2013a). Our video tracking observations on leaf discs were compared to EPG recording data by ten Broeke et al. (2013). Seven out of nine video tracking variables confirmed that cultivar Corbana was more resistant than cultivar Terlana (**Table 1**). Aphids on Corbana spent less time on long probes and more time on shorter probes and other activities. In addition, aphids increased their walking activity over time on both cultivars, but generally covered larger distances on Corbana leaf discs (mixed linear model: time effect: P=0.00, cultivar effect: P=0.03, time x cultivar interaction: P=0.77, **Figure 8**). Yet, the resistance effect was less pronounced in video tracking compared to EPG recording on intact plants: only 11% of the aphids in EPG recordings showed phloem ingestion on Corbana plants, while 78% of the aphids in the video assay performed long probes on Corbana. These long probing events could include other activities, such as water ingestion from xylem vessels, since EPGs showed that on Corbana plants more aphids penetrated xylem sieve elements (12 aphids on Corbana versus 2 aphids on Terlana).

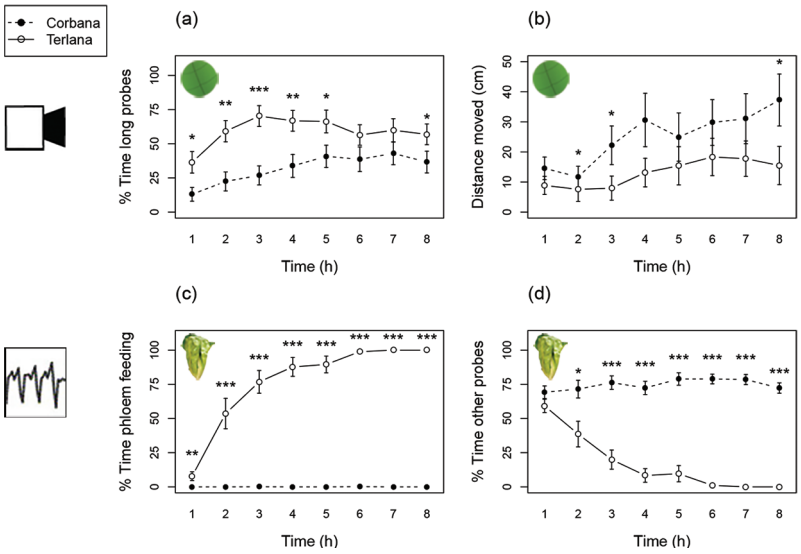



**Figure 8. Behavioural parameters of** *N. ribisnigri* **on two lettuce cultivars, Corbana (resistant) and Terlana (susceptible). (a)** Percentage of the time spent on long probes (> 25 min), and **(b)** distance moved (cm) were measured by automated video tracking on leaf discs. **(c)** Percentage of the time spent on phloem feeding (waveform 5), and **(d)** percentage of time spent on other probes (pathway, phloem salivation and xylem feeding) were measured by EPGs on intact plants (Mann-Whitney U test per time bin, *P<0.05; **P<0.01; ***P<0.001, video tracking: Corbana n=27, Terlana n=28, EPG recording: n=19).

**Table 2. Required video duration and number of replicates for identifying a significant effect.** Student's t-tests have been applied to subsamples of two simulated data sets with a similar mean and standard deviation as two response variables from the video tracking assays. The percentage of tests with a significant outcome in at least one of the two variables is represented as the detection rate (Bonferroni correction: P<0.025).

| Plant-aphid system | Duration | Replicates | Detection rate |
|---|---|---|---|
| Arabidopsis - *M. persicae* | 8 h | 20 | >80% |
| | 6 h | 25 | >80% |
| Lettuce – *N. ribisnigri* | 8 h | 20 | >80% |
| | 4 h | 15 | >80% |

**Table 3. Comparison of automated video tracking and EPG recording characteristics.** Minimum duration of observations and minimum sample sizes were estimated with simulations. Preparation time per replicate reflects a rough estimation of time required for an experienced person to prepare one arena, resp. plant-aphid individual, for a recording.

| | EPG | Video |
|---|---|---|
| Plant material | intact plants | leaf discs |
| Min. observation duration | 4-8 h | 4-8 h |
| Min. number of replicates (identification rate ± 80%) | 3-5 | 15-20 |
| Maximum sample size per set up | ± 8 | ± 100 |
| Preparation time per sample | ± 5 min | ± 2 min |
| Annotation of electrical patterns/video images | ± 15 min | automated |



## Throughput

Using simulated data with a similar plant line effect as the data sets from the plant-aphid assays described here, we assessed the required sample size and recording duration for automated video tracking (**Table 2**). With 20 replicates of 8-hour observations, significant resistance was detected in more than 80% of the cases for the Arabidopsis plant line effect on *M. persicae* (2 response variables tested per simulated data set, Bonferroni correction: $P<0.025$). The near-complete resistance of Corbana lettuce against N. ribisnigri biotype Nr:0 was detected in more than 80% of the cases with 15 replicates of 4 hours of video observations. Subtle differences in resistance in Arabidopsis were more difficult to detect when video observations were shorter than 8 hours (**Table 2**). On the other hand, reducing the video duration to the first 4 hours improved the detection of near-complete resistance, as with *N. ribisnigri* biotype Nr:0 on the Corbana lettuce cultivar. Apparently, in this case, the precision of video tracking decreased over time. While the EPG recording with lettuce did not reveal an increase in aphid activities in the xylem or mesophyll over time (**Figure 8**), the last stretch of the video observation was likely confounded by sessile behaviour other than probing. The risk of falsely rejecting the null hypothesis was limited to 1% (*M. persicae* on Arabidopsis accession Col-0). Overall, video tracking required similar observation durations as EPG recording, but a larger sample size to detect significant plant effects (**Table 3**). The required amount of replicates was, however, compensated by screening many samples simultaneously and automated data annotation.

## Discussion

### Leaf discs

The effect sizes measured in video tracking with leaf discs were substantially smaller compared to EPG recording on intact plants. EPG recording on leaf discs confirmed that the application of excised plant tissue partially impaired plant resistance (Liu and McCreight, 2006; ten Broeke, 2013), possibly due to the interrupted supply of ions and metabolites in the phloem, or due to the interference by jasmonic acid and ethylene mediated wound responses (Leon *et al.*, 2001). Furthermore, aphids can be disturbed by the decrease in pressure in the sieve elements of excised plant tissue, although they are well capable of active uptake of sap (Mittler, 1988; Will *et al.*, 2008). The increase of coagulating proteins and cellular debris in the phloem after plant wounding may plug sieve plates and the aphid's food canal in the stylets (Martin *et al.*, 2003; Tjallingii, 2006; Turgeon, 2010). To prevent such potential clogging of sieve elements, aphids might increase the injection of watery saliva into the phloem or shorten their feeding events, but neither of these effects were observed consistently. To maintain turgor better the use of leaves still connected to intact plants would be favourable, but this is currently not feasible in view of poor detection of aphids in more complex environments. Arenas designed to hold entire detached leaves or seedlings on agar could, however, be a feasible alternative to leaf discs.

### Application

High-throughput phenotyping techniques of sucking insect species are urgently needed in view of functional genomics studies aiming to find subtle allelic differences in plant populations measuring many hundreds of plants. Conventional methods, like EPG and population studies, are less scalable for this purpose and carry much higher investments in terms of time (labour, duration) and costs (equipment, greenhouses). In this study, automated video tracking was used to study aphid feeding behaviour, but it could as well be applied to track the behaviour of other piercing-sucking insects. We recommend to validate the velocity thresholds for each species first, by checking several video files by hand. As shown here with two aphid species, size and velocity can differ and will affect the accuracy of probe estimations. When studying plants with thick or dark leaves, increased resolution, better (macro) lenses, and lateral light sources instead of backlight can help to improve the detection of insects. We worked with EthoVision XT video tracking and analysis software, but other programs or programming environments, such as MatLab and ImageJ, could as well serve as robust video tracking tools (Dell *et al.*, 2014; Husson *et al.*, September 10, 2012; Pinkiewicz *et al.*, 2011).

## Conclusions

The aim of this study was to develop a high-throughput method to screen large plant populations for resistance to aphids and other piercing-sucking insects. For the first time it is shown that automated video tracking of aphid body movement can be used to estimate how often the insects are penetrating plant tissue and are reaching the vascular bundle. The use of leaf discs instead of intact plants enhanced the throughput of the video tracking platform, but EPG recording illustrated that resistance effects were partially lost in leaf discs. Nevertheless, we could identify both intermediate and extreme levels of resistance with video tracking. In Arabidopsis accession Co-2, we found a quantitative resistance level. This was confirmed in additional bioassays, suggesting the involvement of constitutive or rapidly activated resistance mechanisms in both epidermis/mesophyll and phloem, resulting in a small detrimental effect on the aphid population. The video tracking platform also confirmed the near-complete resistance of the lettuce cultivar Corbana to *N. ribisnigri* biotype Nr:0. Although video tracking requires more replicates to identify resistant plants than the conventional EPG technique, it can screen many samples simultaneously in a confined space. In addition, computerized data acquisition reduces laborious exercises, such as annotation of electrical patterns or counting of aphid populations, and only little plant material is required which can be advantageous when studying segregating populations with only one plant per genotype. These features make automated video tracking a valuable phenotyping method for screening large plant populations for resistance to piercing-sucking insects that are serious pests in our crops.



## Materials and methods

### Plants and insects

Arabidopsis, *Arabidopsis thaliana* (L.) Heynh., plants were grown for 4-5 weeks in pots (5 cm diameter) with pasteurized potting soil (4 h at 80°C; Lentse potgrond, Lent, The Netherlands) in a climate room at 21±1 °C, 50-70% relative humidity, an 8/16 h day/night cycle, and a light intensity of 200 μmol $m^{-2}s^{-1}$. Four natural accessions of Arabidopsis were used throughout this study: Col-0 (CS76113), Van-0 (CS76297), Co-2 (CS28163) and Sanna-2 (CS76223). Seeds were acquired from the European Arabidopsis Stock Centre and propagated by the Laboratory of Genetics, Wageningen University.
Lettuce, *Lactuca sativa* (L.), cultivars Corbana (resistant) and Terlana (susceptible) were grown for 3 to 4 weeks in a greenhouse compartment at a temperature of 20±3°C during the day and 18±3°C during the night, 50-70% relative humidity and a 14/10 h

day/night cycle using artificial lighting. Seeds were acquired from Enza Zaden bv. *Myzus persicae* (Sulzer) aphids were reared in a climate room on radish plants at 19°C, 50-70% relative humidity and a 16/8 h day/night cycle. *Nasonovia ribisnigri* (Mosely) biotype Nr:0 aphids were reared on the susceptible lettuce cultivar Terlana in a greenhouse compartment at a temperature of 20±3°C during the day and 18±3°C during the night, 50-70% relative humidity, and a 14/10 h day/night cycle.

### Video tracking platform

Aphid behaviour was recorded with an analogue, monochrome camera (Ikegami, model: I CD-49E, type: REV, 768 x 576 pixels) with a varifocal lens (Computar H3Z4512 CS-IR, 4.5-12.5 mm F1.2) mounted above the arenas (**Figure 1**). An arena consisted of a well in a 96-well microtitre plate, having a 6.5 mm inner diameter (Sarstedt, sterile flat bottom suspension cells. No. 831835500), containing a leaf disc with the abaxial side up on a substrate of 1% agar (technical agar no.3, Oxoid). One aphid was introduced per arena and cling film was tightly wrapped around the plate to prevent aphids from escaping. The microtitre plate was placed on a platform, 1 cm above a backlight unit (FL tubes, 5000 K). A fan was attached between the platform and backlight unit to prevent water condensation inside the arenas. Room temperature was controlled at 21-22°C.

### Software settings

EthoVision® XT 8.5 video tracking and analysis software (Noldus Information Technology bv, Wageningen, The Netherlands) was used for automated video tracking of aphid feeding behaviour in multiple arenas simultaneously (Noldus *et al.*, 2001; Noldus *et al.*, 2002). Subject detection was achieved with grey scaling (**Additional file 1**, **Table S1**). Arenas contained two zones: the leaf disc (zone 1) and the space surrounding the leaf disc (zone 2) (**Figure S1**). Zone 1 had a diameter of approximately 5 mm, excluding the outer edges of the leaf disc to prevent aphids on the arena wall to be falsely assigned to the leaf disc. Because zone 1 and zone 2 required different grey scale thresholds, optimal thresholds for zone 1, the leaf disc, were chosen. Consequently, only behavioural data acquired in zone 1 were used throughout this study. Velocity and time thresholds appropriate to starting and ending a probe were fine-tuned using simultaneous observations of the top-view camera (275 pixels per $mm^2$) and a side-view camera attached to a stereo microscope (20-40x magnification), capturing close-up recording of proboscis and antennae movements of *M. persicae* aphids (**Additional file 2**). A probe start was automatically recorded when the aphid was positioned on the leaf disc and its velocity dropped below 0.02 mm/s and did not exceed 0.3 mm/s for at least 10 seconds (**Figure 2**, **Figure S2**). A probe stop was recorded when aphid velocity

exceeded 0.3 mm/s for the relatively small wingless *M. persicae* or 0.35 mm/s for the larger winged *N. ribisnigri* and did not decrease below 0.1 mm/s for at least 2 seconds. Confounding movements during probing were generally characterized by a repetitive pattern of short movements. The 2 seconds time delay prevented that these movements resulted in false probe stops. Zone-transition problems, which occurred when aphids were positioned exactly on the edge of zone 1 and zone 2, were filtered from the data set after acquisition in EthoVision XT, with the statistical computing program R (**Additional file 4**). These incidences, characterized by a train of consecutive short probes in the output, were filtered out by excluding probes with a duration of less than 3 seconds that were preceded by a very short non-probe bout of maximally 15 seconds. These thresholds were selected by hand using some examples of zone transition problems in this study.



### Video recording versus human observations

To validate automated tracking of probes with manual scoring, we used a camera mounted on a stereo microscope (20-40x) with a side-view on a single arena (n=16) (**Additional file 2**). Each replicate consisted of a 55 min continuous recording of one arena with a single adult *M. persicae* aphid and an Arabidopsis leaf disc, by both the mounted and side-view camera. Aphids were starved between 30 minutes and three hours before the experiment. Recordings with the mounted camera were performed at two distances: capturing 20 arenas with 275 pixels per $mm^2$, and capturing 35 arenas with 155 pixels per $mm^2$. Leaf discs of 6 mm in diameter were cut just below the leaf apex of 4-5 week old Col-0 and Van-0 plants. The Observer® XT 10 software (Noldus Information Technology bv, Wageningen, The Netherlands) was used for manual scoring of probes. Probe starts were manually recorded when body movement stopped, the proboscis was touching the leaf and antennae moved backwards. If the aphid's proboscis was obscured, body arrestment on the leaf disc with subsequent backward movement of antennae was defined as a probe start according to Hardie *et al.* (1992) and Hardie and Powell (2000). Probe endings were manually recorded when antennae moved upward and the aphid removed its proboscis from the leaf, or, when the latter was not visible, when the antennae moved upwards followed by locomotion. Apart from probe estimations, we also tracked the "total time not moving" across the whole observation, using a start velocity of 0.3 mm/s and a stop velocity of 0.02 mm/s. Velocities were averaged across 5 frames, using a sample rate of 5 frames per second.

### Video tracking assays

In each recording twenty arenas were tracked simultaneously for 8 hours, with a frame rate of 25 $s^{-1}$, and a resolution of 275 pixels per $mm^2$ (**Additional file 3**). All arenas consisted of a different plant and aphid individual and within each recording the 2 involved plant lines were equally represented. For Arabidopsis accessions Co-2 and Sanna-2, automated video tracking was performed with 7 to 8 day old wingless *M. persicae* aphids (Co-2 n=20, Sanna-2 n=17). Leaf discs of 6 mm in diameter were made just below the apex of intermediately aged leaves. Aphid survival was checked the day after recording. Subject detection was checked after data acquisition on 6 time points across the video. Three samples with no or low quality detection were excluded from the analysis. Video tracking of winged *N. ribisnigri* biotype Nr:0 on lettuce cultivars Terlana and Corbana was performed with 4 mm leaf discs (Corbana n=27, Terlana n=28). In view of the large contour of winged *N. ribisnigri* aphids, we used arenas with leaf discs of 4 mm diameter and a 3-4 mm leaf edge-to-wall distance in order to have a clear distinction between aphids on the leaf disc and aphids on agar or the arena wall. Leaf discs were made near the leaf base of the third oldest leaf, next to the mid vein. None of the aphids had died the day after recording. Five samples with no or low quality detection were excluded from the analysis. The response variable "duration not moving" was measured using a start velocity of 0.3 mm/s and a stop velocity of 0.02 mm/s. Velocities were averaged across 5 frames, using a sample rate of 5 frames per second.

### EPG recording

Feeding behaviour of the green peach aphid, *M. persicae*, was analysed with EPG recording on two natural accessions of Arabidopsis, Co-2 and Sanna-2, during 8-hour observations. EPG recording was made on both intact plants (Co-2 n=19, Sanna-2 n=18) and leaf discs (Co-2 n=9, Sanna-2 n=8), using direct currents (DC) according to the methodology of ten Broeke et al. (ten Broeke *et al.*, 2013a). An electrode was inserted in the potting soil or agar respectively, and a thin gold wire (1.5 cm length for intact plants, 1 cm length for leaf discs) was gently attached to the dorsum of 8 to 11 day old wingless aphids with silver glue. The electrical circuit was completed when the aphid's piercing-sucking mouthparts penetrated the plant cuticle and the electrical signals, correlated to stylet activities, were recorded instantly (Tjallingii, 1988). Each replicate consisted of a different aphid and plant individual, employing one leaf disc per plant. Leaf discs of 9 mm in diameter were processed just below the apex of intermediately aged Arabidopsis leaves and placed abaxial side up in a Petri dish on a 1% agar substrate. A transparent plastic sheet covered the agar surrounding the leaf

disc to prevent aphids to get stuck or make probes in the agar. Aphids that did not start probing within the first 3 hours of the observation were excluded from the analysis. EPG recording of winged *N. ribisnigri* biotype Nr:0 on lettuce cultivars Corbana and Terlana has been made in a previous study by ten Broeke et al. (ten Broeke, 2013) (8-hour recording, n=19).

### Aphid population development

One *M. persicae* neonate (0 to 24 h old) was transferred to each Arabidopsis plant in a climate chamber (21±1°C, 50-70% relative humidity, an 8/16h day/night cycle, light intensity of 200 µmol $m^{-2}s^{-1}$). In the first assay 2.5-week-old plants were infested, in the second assay 3.5-week-old plants. A soap-diluted water barrier prevented aphids from moving between plants. Six, seven, and eight days after introduction the presence of the aphid and its offspring was checked. None of the aphids developed wings. 14 Days after infestation the number of aphids was counted per plant. Plants without an adult aphid 8 days after introduction, and plants without any adults or neonates 14 days after introduction were excluded from the analysis (assay 1: Co-2 n=19, Sanna-2 n=15; assay 2: Co-2 n=14, Sanna-2 n=13).



### Simulations

In simulations, $10^4$ random draws were taken from a normal distribution with the mean and standard deviation of a response variable of the Arabidopsis-*M. persicae* and lettuce-*N. ribisnigri* data sets (**Table S3**). For video observations data was simulated with two probing variables: the mean duration of long probes and the total duration of sustained probes. For EPGs the total duration of phloem ingestion was simulated. Random draws were excluded when values were below zero, below the minimum duration of the probe category, or above the maximum recording duration. The generated data sets were subsampled with 1000 iterations without replacement for several replicate levels (n=10, 15, 20, 25, 30, 35, 40). Student's t-tests were executed for each iteration and the percentage of significant p-values per replicate level was calculated. Video tracking simulation tests were defined significant if they had a P-value below $\alpha=0.025$ for at least one of the two probing variables (Bonferroni correction: $\alpha=0.05/2$). For EPG simulations one variable and P-values below $\alpha=0.05$ already delivered maximum detection rates. This process was performed on complete data sets of EPG and video recording (8 h observations) and on data sets rescaled to shorter durations (6 and 4 hour observations). The proportion of tests where the null hypothesis is incorrectly rejected, was calculated with simulations based on a data set of 8-hour video recording of *M. persicae* on Arabidopsis accession Col-0 (data set n=53, replicate levels n=15 and n=20, two variables, $P<0.025$, **Table S3**).

### Statistical analysis

An R script was written to calculate response variables of video tracking, such as the total number and total duration of short and long probes in each observation and for each hour (**Additional file 4**). For EPG recording, the start time and duration of waveforms were analysed with the EPG PROBE 3.0 software (EPG-Systems, Wageningen, The Netherlands). Further calculations and analyses of EPG data were performed with the statistical computing program R. The duration of phloem ingestion events in EPG recording were calculated as the sum of three subsequent waveforms: (a) inter- and intracellular penetrations followed by (b) phloem salivation and (c) phloem ingestion. Bar graphs were produced with the R package sciplot version 1.1-0 (Morales 2012) (R-Core-Team, 2013). Data distributions and homogeneity of variances were tested with a Shapiro test and a Levene's test. In case data transformations (square root, log, logit, arcsine) did not result in a distribution that approaches a normal distribution, non-parametric tests were applied. Human observations were compared to video tracking parameters with a paired t-test or, when data were not normally distributed with a Wilcoxon signed ranks matched pairs test. Correlations were tested with a Pearson correlation test or, when data were not normally distributed, with a Spearman correlation test. For benchmarking of video tracking against EPGs with susceptible and resistant Arabidopsis and lettuce lines and for the reproduction assay, response variables were tested with a Student's t-test, or when the data were not normally distributed with a Mann-Whitney U test. Walking activity of aphids was tested across 8 time bins of 1 hour. The distance moved was not normally distributed and, therefore, transformed to ranks ranging from the lowest to highest value within the complete data set. A mixed linear model was applied on the ranks, using plant line, time bin, and plant line x time bin interaction as fixed effects and plant/aphid individual as a random effect.

## Acknowledgements

This work was supported by The Netherlands Organization for Scientific Research (NWO) through the Technology Foundation Perspective Programme 'Learning from Nature' [STW10989]. We would like to thank Lia Hemerik and Freddy Tjallingii for their valuable advice during this study.

## Supporting Information

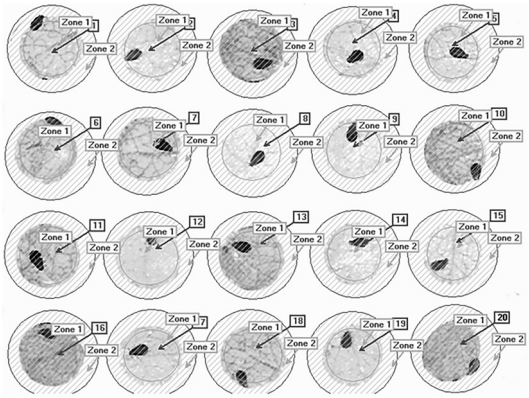



**Figure S1.** Arena settings in EthoVision XT.

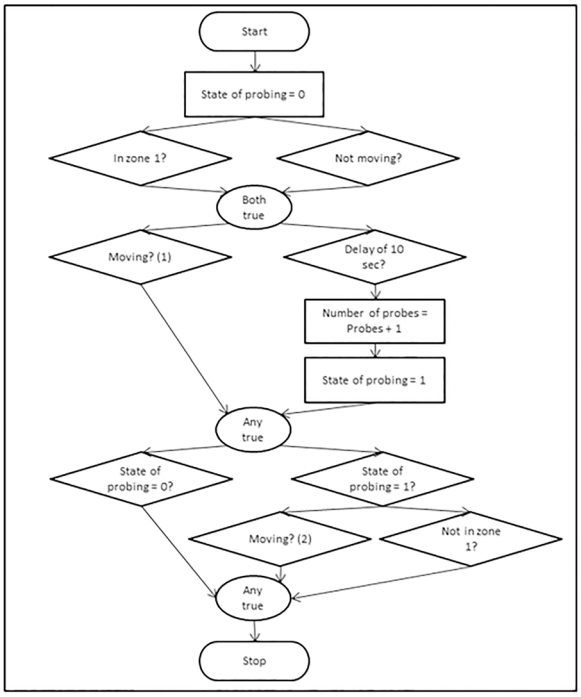



**Figure S2.** Trial Control Settings in EthoVision XT for measuring the start time and duration of probes. Zone 1 is the leaf disc. 'Not moving' and 'Moving (1)' were defined by a start velocity of 0.3 mm/s and a stop velocity of 0.02 mm/s, averaged over 5 samples (0.2 seconds). The condition 'Moving (2)' was defined by a start velocity of 0.3 mm/s for *M. persicae*, resp. 0.35 mm/s for *N. ribsinigri*, a stop velocity of 0.1 mm/s, and a minimum duration of 2 seconds. This procedure was repeated indefinitely during data acquisition. Velocities were calculated by taking the average velocity of 5 frames, with a frame rate of 5 frames per second.

**Table S1. Subject detection settings in EthoVision XT for *M. persicae*.**

| Parameters | Settings |
|---|---|
| Detection method | Grey scaling, centre point detection |
| Grey scale range | 6-200 |
| Subject size | 85-600 pixels |
| Contour erosion (first) | 2 pixels |
| Contour dilation | 2 pixels |
| Sample rate | 5 samples per second |
| Pixel smoothing | Medium |
| Track noise reduction | Off |
| Track smoothing profile | Lowess, 10 samples before and after |

## Online files (file 2 - 4)

http://plantmethods.biomedcentral.com/articles/10.1186/s13007-015-0044-z#Sec22

**- Additional file 2: (.mp4)**

Movie of aphid feeding behaviour measured by manual annotations and automated video tracking simultaneously. Events were estimated where the *M. persicae* aphid penetrated the Arabidopsis leaf disc with its piercing-sucking mouth parts. Manual scoring of probes was achieved with a side-view camera mounted on a stereo microscope and The Observer XT software. Automated video tracking was performed with a top view on the arena (275 pixels/mm) and EthoVision XT software (Noldus Information Technology bv, Wageningen, The Netherlands). EthoVision XT recorded a probe when the aphid's centre point was located on the leaf disc and did not move. Other motion parameters, such as the distance moved (cm) and velocity (cm/s), were also acquired with automated tracking. The movie is displayed at double speed.

**- Additional file 3: (.mp4)**

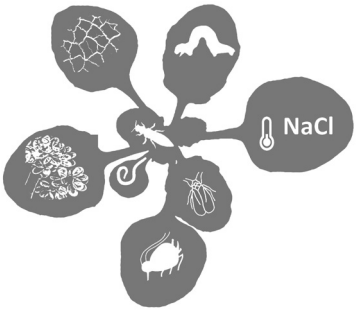
Movie of automated video tracking of 20 aphids simultaneously. *M. persicae* aphids were tracked on Arabidopsis leaf discs using EthoVision XT video tracking software (Noldus Information Technology bv, Wageningen, The Netherlands). The aphid's centre point (red) was automatically detected and tracks of the previous 30 seconds are visualised in red. Arenas are numbered from 1 (top left) to 20 (bottom right). The movie is displayed at 16x speed.

**- Additional file 4: (.txt)**

R script for feature extraction from EthoVision XT output.

# Chapter 4

## Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping




Manus P.M. Thoen*, Nelson H. Davila Olivas*, Karen J. Kloth*, Silvia Coolen*, Ping-Ping Huang*, Mark G.M. Aarts, Johanna A. Bac-Molenaar, Jaap Bakker, Harro J. Bouwmeester, Colette Broekgaarden, Johan Bucher, Jacqueline Busscher-Lange, Xi Cheng, Emilie F. Fradin, Maarten A. Jongsma, Magdalena M. Julkowska, Joost J.B. Keurentjes, Wilco Ligterink, Corné M.J. Pieterse, Carolien Ruyter-Spira, Geert Smant, Christa Testerink, Björn Usadel, Joop. J.A. van Loon, Johan A. van Pelt, Casper C. van Schaik, Saskia C.M. van Wees, Richard G.F. Visser, Roeland Voorrips, Ben Vosman, Dick Vreugdenhil, Sonja Warmerdam, Gerrie L. Wiegers, Joost van Heerwaarden, Willem Kruijer, Fred A. van Eeuwijk and Marcel Dicke

* These authors contributed equally to this work

*under review*

## Abstract

Plants are exposed to combinations of various biotic and abiotic stresses, but stress responses are usually investigated for single stresses only. Here we investigated the genetic architecture underlying plant responses to 11 single stresses and several of their combinations by phenotyping 350 *Arabidopsis thaliana* accessions. A set of 214k SNPs was screened for marker-trait associations in Genome-Wide Association analyses using tailored multi-trait mixed models. Stress responses that share phytohormonal signaling pathways also share genetic architecture underlying these responses. For the 30 most significant SNPs, average QTL-effect-sizes were stronger for dual stresses than single stresses. Plants appear to deploy broad-spectrum defensive mechanisms influencing multiple traits in response to combined stresses. Association analyses identified QTLs with contrasting and with similar responses to (a) biotic versus abiotic stresses and (b) belowground versus aboveground stresses. Our approach allowed for an unprecedented comprehensive genetic analysis of how plants deal with a wide spectrum of stress conditions.

## Background

In nature, plants face variable environments that impose a wide range of biotic and abiotic stresses. These include e.g. belowground and aboveground stresses, stresses imposed by unicellular and multicellular organisms, short and long-lasting stresses. Thus, plants are under strong selection to adapt to local conditions and have evolved sophisticated mechanisms to withstand multiple adverse environmental conditions (Brachi *et al.*, 2015; Howe and Jander, 2008; Julkowska and Testerink, 2015; Kerwin *et al.*, 2015; Pieterse *et al.*, 2012; Stam *et al.*, 2014). Yet, investigating this experimentally is a major challenge due to the complexity of multiple stress exposure. To gain insight into the adaptation of plants to the wide variety of stress-inducing conditions they face, genetic variation and mechanisms underlying stress resistance should be studied (Alonso-Blanco *et al.*, 2009; Brachi *et al.*, 2015; Kerwin *et al.*, 2015). The responses of plants to stresses have traditionally been investigated for individual stresses (Howe and Jander, 2008), but research focus is currently shifting towards plant responses to combinations of stresses (Holopainen and Gershenzon, 2010; Kissoudis *et al.*, 2015; Pierik and Testerink, 2014; Stam *et al.*, 2014; Suzuki *et al.*, 2014). The emerging picture is that responses to stress combinations cannot be predicted reliably from the responses to individual stresses (Makumburage *et al.*, 2013). For instance, the majority of transcriptional responses of Arabidopsis to combinations of two abiotic stresses could not be predicted from

responses to the individual stresses (Rasmussen *et al.*, 2013). Moreover, phenotype expression in response to two biotic stresses could not be predicted on the basis of existing information on interactions between underlying signaling pathways (De Vos *et al.*, 2006). Phytohormones are major players in a signaling network, mediating responses to both biotic and abiotic stresses (Pieterse *et al.*, 2009). For instance, chewing insect herbivores elicit especially the jasmonic acid (JA), abscisic acid (ABA) and ethylene (ET) signaling pathways, phloem-sucking insects and biotrophic microbial pathogens elicit especially the salicylic acid (SA) pathway, and drought elicits the abscisic acid (ABA) pathway (Pieterse *et al.*, 2009). The phytohormonal responses exhibit extensive crosstalk, resulting in specific changes in plant phenotype in response to individual stresses (De Vos *et al.*, 2005; Pieterse *et al.*, 2012).

Most studies that examined plant responses to multiple stresses included only one or a few genotypes (Holopainen and Gershenzon, 2010; Kissoudis *et al.*, 2015; Pierik and Testerink, 2014; Rasmussen *et al.*, 2013; Stam *et al.*, 2014; Suzuki *et al.*, 2014). To obtain a further understanding of the genetic architecture of complex traits such as plant adaptation to a diversity of stresses, extensive study of the natural genetic variation within a species is instrumental. Genome-wide association (GWA) analysis is an important tool for this, requiring a large number of well-genotyped plant accessions. Yet, although the interest in natural variation and GWA mapping is rapidly increasing (Ogura and Busch, 2015; Wijnen and Keurentjes, 2014), a large-scale evaluation of natural genetic variation for resistance of plants to the diversity of stresses that they are exposed to, including pathogens, herbivores and abiotic stresses and their interactions, has not been made to date. To elucidate the genetic architecture of plant stress resistance, an integrated approach is needed that models the genetics of responses to a range of single and combined stresses, including the interaction between those responses. Here, we have taken a comprehensive and integrated approach to investigate the genetics underlying plant responses to 15 single stresses or stress combinations (**Table 1**), making use of a global population of 350 Arabidopsis accessions that have been genotyped for 214k SNPs (Baxter *et al.*, 2010; Li *et al.*, 2010). We developed a tailored multi-trait GWA analysis that allowed the identification of candidate genes associated with adaptive plant responses to multiple stresses that were validated by gene expression and mutant analyses.



## Results

The phenotypic response of a population of 350 Arabidopsis accessions to an extensive set of stress-inducing conditions was quantified relative to the respective control treatments. Thirty traits, including e.g. root length, number of damaged leaves,

or number of pathogen-inflicted spreading lesions (**Table 1**) were quantified when the plants were exposed to 15 different stresses, i.e. four abiotic stresses (drought, salt stress, osmotic stress and heat), seven biotic stresses (parasitic plant, phloem-feeding aphid, phloem-feeding whitefly, cell-content feeding thrips, leaf-chewing caterpillar, root-feeding nematode, and necrotrophic fungus) and four stress combinations (fungus and caterpillar, drought and fungus, drought and caterpillar, caterpillar and osmotic stress). For detailed information on the stress treatments and the trait definitions see Supplementary Materials and methods (Sections 2 and 3).`

## Heritability of responses to biotic and abiotic stresses

The phenotypic analysis resulted in a wide range of marker-based narrow sense heritability (Kruijer *et al.*, 2015) estimates with 15 traits of low ($h^2<0.2$), 10 of moderate ($0.2<h^2<0.5$) and 5 of high ($h^2>0.5$) heritability (**Figure S1**). The number of abiotic stress traits per heritability category was similar, while the number of traits related to biotic and combined stresses decreased with increasing heritability class. The most heritable traits were responses to feeding damage by thrips (Thrips_1; $h^2=0.8$), and nematodes ($h^2=0.7$), and responses to salt (Salt_1 and Salt_3; resp. $h^2=0.6$ and $h^2=0.7$) and heat (Heat; $h^2=0.6$) (**Table S1**). The traits related to combined stresses have predominantly low heritabilities.

## Genetic commonality underlying responses to different stresses

To analyze the phenotypic variation between Arabidopsis accessions as a function of molecular marker variation, we used various mixed model approaches (see Supplementary Materials and methods, section 4). We estimated marker-based genetic correlations, i.e. correlations based on the genome-wide commonality of SNP effects underlying pairs of traits (see Materials and methods), to investigate the magnitude of genetic commonality underlying resistance mechanisms in response to a range of biotic and abiotic stresses. For brevity, we will refer to these marker-based genetic correlations as genetic correlations. Such genetic correlations can be interpreted as upper boundaries to the joint determination of pairs of traits by genetic factors. Genetic correlation

**Table 1. Phenotypes assessed.** The dataset contains three plant stress categories; abiotic stress, biotic stress and combinations of both abiotic and biotic stress. Phenotype assessments that were performed under similar environmental conditions have similar background shading (light and dark grey). 'Phenotype' refers to different phenotypic assessments (in some cases the first principal component of a group of phenotypes). 'Treatment' refers to the sort of stress that was applied. Additional information on traits can be found in Supplementary methods (section 3). ►

| | Stress | Trait name | Trait phenotype | Treatment |
|---|---|---|---|---|
| Abiotic stresses | Salt | Salt_1 | Main root length, number of lateral roots and straightness | 75 mM NaCl |
| | | Salt_2 | Main root length | 125 mM NaCl |
| | | Salt_3 | Number of lateral roots | 125 mM NaCl |
| | | Salt_4 | Main root angle | 125 mM NaCl |
| | | Salt_5 | Biomass | 25 mM NaCl |
| | Drought | Drought_1 | Biomass | Drought |
| | | Drought_2 | Biomass | Drought |
| | Osmotic | Osmotic | Biomass | PEG8000 |
| | Heat | Heat | Number of siliques | 35 °C |
| Biotic stresses | Parasitic plant | Parasitic plant | Attachments | *Phelipanche ramosa* |
| | Nematode | Nematode | Offspring, eggmass | *Meloidogyne incognita* |
| | Whitefly | Whitefly_1 | Survival, whiteflies | *Aleyrodes proletella* |
| | | Whitefly_2 | Reproduction, eggs | *A. proletella* |
| | Aphid | Aphid_1 | Behavior T1, probing | *Myzus persicae* |
| | | Aphid_2 | Behavior T2, probing | *M. persicae* |
| | | Aphid_3 | Offspring, aphids | *M. persicae* |
| | Thrips | Thrips_1 | Feeding damage | *Frankliniella occidentalis* |
| | | Thrips_2 | Behavior T1 | *F. occidentalis* |
| | | Thrips_3 | Behavior T2 | *F. occidentalis* |
| | Caterpillar | Caterpillar_1 | Leaf area consumed | *Pieris rapae* |
| | | Caterpillar_2 | Biomass | *P. rapae* |
| | | Caterpillar_3 | Number of damaged leaves and feeding sites | *P. rapae* |
| | Fungus | Fungus | Number of spreading lesions | *Botrytis cinerea* |
| Abiotic and biotic stress | Double stress | Fungus and caterpillar_1 | Biomass | *B. cinerea* and *P. rapae* |
| | | Fungus and caterpillar_2 | Number of damaged leaves and feeding sites | *B. cinerea and P. rapae* |
| | | Caterpillar and fungus | Number of spreading lesions | *P. rapae* and *B. cinerea* |
| | | Drought and fungus | Number of spreading lesions | Drought and *B. cinerea* |
| | | Drought and caterpillar | Number of damaged leaves and feeding sites | Drought and *P. rapae* |
| | Double stress | Caterpillar and osmotic_1 | Projected leaf area | *P. rapae* and PEG8000 |
| | | Caterpillar and osmotic_2 | Biomass | *P. rapae* and PEG8000 |

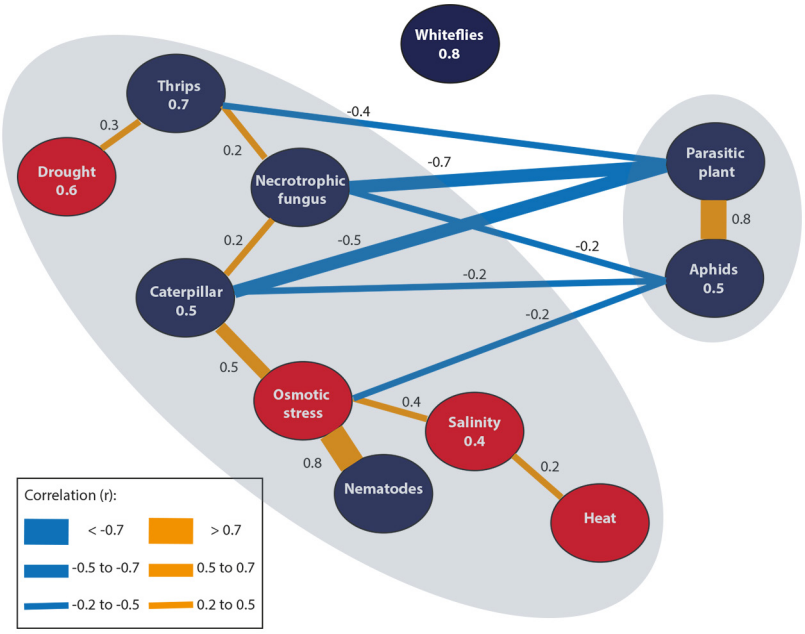



**Figure 1. Mean genetic correlations between responses to abiotic (red) and biotic (dark blue) plant stresses.** Thickness of lines represents the strength of mean genome-wide correlations, annotated with r values (orange=positive, blue=negative correlation). The more shared genetic associations between stresses, the higher the absolute genetic correlation. Correlations are negative when alleles have opposite effects, i.e. resulting in increased resistance to one stress, but decreased resistance to the other stress. Values in balloons represent mean within-group correlation (not shown for groups consisting of a single trait). Mean between-group correlations are not shown if they are below an absolute value of r=0.2. Two clusters can be distinguished: (1) parasitic plants and aphids and (2) the other stresses, except whiteflies.

analysis revealed a strong connection between the responses to parasitic plants and to aphids (r=0.8), which were both negatively associated with other stress responses (**Figure 1**). Parasitic plants and aphids have in common that they target phloem and xylem tissue (Dorr and Kollmann, 1995; Tjallingii and Hogen Esch, 1993), and induce the SA phytohormonal pathway (De Vos *et al.*, 2005; Runyon *et al.*, 2008). In contrast, the biotic stress responses that were negatively associated with the responses to parasitic plants and aphids, i.e. responses to necrotrophic fungi, caterpillars, and thrips, represent JA-inducing stresses (De Vos *et al.*, 2005; Pieterse *et al.*, 2009; Pieterse *et al.*, 2012). Because the SA and JA pathways predominantly interact through negative crosstalk (Pieterse *et al.*, 2009), the two main clusters resulting from the genetic correlation analysis represent different phytohormonal signaling response mechanisms. We also observed a strong genetic correlation between plant responses to osmotic stress and

root-feeding nematodes. This supports the notion that root-knot nematodes trigger a differentiation of root cells to multinucleate giant cells with severely altered water potential and osmotic pressure (Baldacci-Cresp *et al.*, 2015). While the correlations between traits at the phenotypic level were generally rather low, the genetic correlation analysis revealed a common genetic architecture underlying the responses to sets of single and combined stresses (**Figure S2**).

### Candidate genes underlying responses to stresses

To identify individual candidate genes that contributed most to the pattern of genetic correlations, we fitted multi-trait QTL mixed models (MTMMs) to the total set of 30 traits, using a 214k SNP set that is commonly used for GWA studies in Arabidopsis (Atwell *et al.*, 2010; Bac-Molenaar *et al.*, 2015a; Horton *et al.*, 2012; Kim *et al.*, 2007; Li *et al.*, 2010). Our multi-trait GWA approach closely follows the modeling framework developed by Zhou and Stephens (2014) and generalizes the use of MTMMs as described previously (Alimi *et al.*, 2013; Boer *et al.*, 2007; Malosetti *et al.*, 2008) for classical biparental offspring populations to association panels. This GWA analysis identified 30 chromosome regions with multiple, significant SNP-trait associations. From each of those regions, the most significant SNP was chosen to represent the locus (**Figure 2**; **Table S2**). Clustering of stresses by estimated SNP-effect profiles (**Figure 2**) revealed that multiple SNPs were associated with response to more than one stress. Stress combinations induced large QTL allele substitution effects in the MTMM mapping (**Figure 2** and **Table S2**), indicating that combinations of stresses trigger broad-spectrum defensive mechanisms. A total of 125 genes were in linkage disequilibrium (LD) with the 30 most significant SNPs from the GWA analysis. Twenty of these genes were stress-related according to gene ontology (GO) annotation data (**Table S3**). Of these 20 genes, six have been functionally characterized by at least one study (**Table 2a**). For these six genes, we explored expression data to evaluate the biological relevance of these genes in stress-responsive mechanisms of Arabidopsis (**Figure S3**). Of special interest were SNPs chr5.7493620, chr5.22041081 and chr4.6805259, that were in LD with *WRKY38* (encoding a WRKY transcription factor involved in SA-dependent disease resistance) (Kim *et al.*, 2008c), *AtCNGC4* (involved in pathogen resistance) (Chin *et al.*, 2013) and *RMG1* (coding for disease resistance protein) (Yu *et al.*, 2013) respectively.

### Phytohormonal signaling underlying contrasts in stress responses

The MTMM framework allowed imposing constraints on the values of the estimated QTL effects (see Materials and methods). In this way specific hypotheses can be tested about stresses sharing a common QTL effect or having opposite QTL effects. We

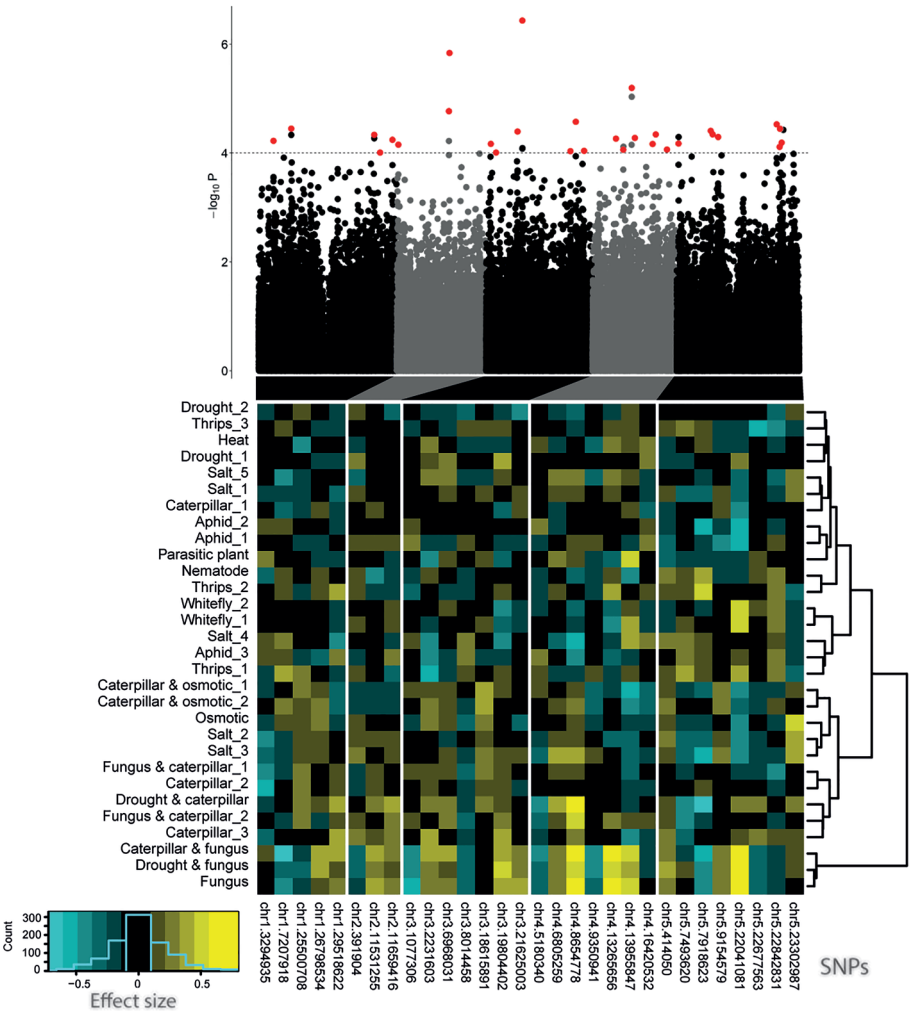



**Figure 2. Multi-trait mixed-model (MTMM) GWA mapping with 30 different stress responses of Arabidopsis.** The top panel shows the 214k SNPs with their corresponding -$\log_{10}(P)$ values for the five chromosomes. The lower panel depicts the trait-specific effect size of the rare allele for SNPs with a LOD score higher than 4 (effects were estimated from the full MTMM). When several SNPs were located within a 20 kb half-window of each other, only the SNP with the highest absolute cumulative effect size was included (red-flagged in the Manhattan plot). SNPs are named by chromosome number and position on the chromosome. Negative effect sizes (blue) correspond to reduced plant resistance due to the rare allele, positive effect sizes (yellow) to increased resistance due to the rare allele. Stress responses were clustered hierarchically according to their effect size, using Ward's minimum variance method. The key shows the frequency distribution of SNPs across effect sizes.

**Table 2.** Candidate genes resulting from **a.** MTMM analysis of all 30 stress responses as presented in Figure 2 and **b.** contrast-specific analysis with MTMM for contrasting effects of biotic and abiotic stresses as presented in Figure 3.

**Table 2a**

| Marker* | Gene in LD | Name | Gene description** | References |
|---|---|---|---|---|
| chr2. 11659416 | *AT2G27250* | *CLV3* | One of the three *CLAVATA* genes controlling the size of the shoot apical meristem (SAM) in Arabidopsis | (Clark *et al.*, 1996; Fletcher *et al.*, 1999; Shinohara and Matsubayashi, 2010) |
| chr3. 19804402 | *AT3G53420* | *PIP2* | A member of the plasma membrane intrinsic protein subfamily PIP2. | (Martiniere *et al.*, 2012; Peret *et al.*, 2012; Sanchez-Romera *et al.*, 2014) |
| chr4. 6805259 | *AT4G11170* | *RMG1* | Encodes RMG1 (Resistance Methylated Gene 1), an NB-LRR disease resistance protein with a Toll/interleukin-1 receptor (TIR) domain at its N terminus. | (Yu *et al.*, 2013) |
| chr5. 7493620 | *AT5G22570* | *WRKY38* | Member of WRKY Transcription Factor; Group III | (Kim *et al.*, 2008c; Mare *et al.*, 2004) |
| chr5. 22041081 | *AT5G54250* | *CNGC4* | Member of Cyclic Nucleotide Gated Channel family, a downstream component of the signaling pathways leading to hypersensitive response (HR) resistance. Mutant plants exhibit gene-for-gene disease resistance against avirulent *Pseudomonas syringae* despite the near-complete absence of the HR. Salicylic acid accumulation in *dnd2* mutants is completely *PAD4*-independent. | (Chin *et al.*, 2013; Jurkowski *et al.*, 2004; Keisa *et al.*, 2011) |
| chr5. 23302987 | *AT5G57560* | *TCH4* | Encodes a cell wall modifying enzyme, rapidly upregulated in response to environmental stimuli | (Braam and Davis, 1990; Iliev *et al.*, 2002; Purugganan *et al.*, 1997; Xu *et al.*, 1996) |

* markers derived from MTMM analysis (see Figure 2), ** based on information on http://www.arabidopsis.org/tools/bulk/go/index.jsp

**Table 2b** See legend previous page.

| Marker | Gene in LD | Name | Gene description* | Reference |
|---|---|---|---|---|
| chr1. 30381439 | *AT1G80820* | *CCR2* | *CINNAMOYL COA REDUCTASE*. Encodes a cinnamoyl CoA reductase isoform. Involved in lignin biosynthesis. | (Lauvergeat *et al.*, 2001; Luderitz and Grisebach, 1981; Zhou *et al.*, 2010) |
| chr1. 30381439 | *AT1G80840* | *WRKY40* | Pathogen-induced transcription factor. Binds W-box sequences in vitro. Forms protein complexes with itself and with WRKY60. Co-expression with *WRKY18* or *WRKY60* made plants more susceptible to both *P. syringae* and *Botrytis*. | (Chen *et al.*, 2010a; Liu *et al.*, 2012; Pandey *et al.*, 2010) |
| chr1. 6038270 | *AT1G17610* | *CHS1* | *CHILLING SENSITIVE 1*, mutant accumulates steryl-esters at low temperature. | (Wang *et al.*, 2013; Zbierzak *et al.*, 2013) |
| chr5. 171177 | *AT5G17640* | *ASG1* | *ABIOTIC STRESS GENE 1*; Expression of this gene is induced by abscisic acid and salt stress. | (Batelli *et al.*, 2012; Coste *et al.*, 2008) |
| chr5. 23247572 | *AT5G57380* | *VIN3* | Encodes a plant homeodomain protein VERNALIZATION INSENSITIVE 3 (VIN3). In planta VIN3 and VRN2, VERNALIZATION 2, are part of a large protein complex that can include the polycomb group (PcG) proteins FERTILIZATION INDEPENDENT ENDOSPERM (FIE), CURLY LEAF (CLF), and SWINGER (SWN or EZA1). The complex has a role in establishing FLC (FLOWERING LOCUS C) repression during vernalization. | (Bond *et al.*, 2009; Finnegan *et al.*, 2011; Sung *et al.*, 2007) |
| chr5. 23293119 | *AT5G57560* | *TCH4* | Encodes a cell wall-modifying enzyme | (Braam and Davis, 1990; Iliev *et al.*, 2002; Purugganan *et al.*, 1997; Xu *et al.*, 1996) |
| chr5. 23293870 | *AT5G57490* | *VDAC4* | Encodes a voltage-dependent anion channel (VDAC: AT3G01280/VDAC1) | (Lee *et al.*, 2009; Tateda *et al.*, 2011) |
| chr5. 23366252 | *AT5G57685* | *GDU3* | Encodes a member of the GDU (glutamine dumper) family proteins involved in amino acid export: At4g31730 (GDU1) | (Chen *et al.*, 2010b) |

* based on information on http://www.arabidopsis.org/tools/bulk/go/index.jsp

investigated whether polymorphisms for genes involved in SA and JA biosynthesis or genes responsive to signals from these pathways were the cause of the negative genetic correlations between the groups of traits sharing one or the other phytohormonal signaling pathway. To this end, we performed a multi-trait GWA mapping to test the contrast between: (1) parasitic plant and aphid response, versus (2) the most negatively correlated traits, i.e. fungus, caterpillar, thrips and drought response (**Figure 1**). Fifteen SNPs were significantly associated with contrasting effects between the two trait clusters (**Figure S4**). Seven of these SNPs, were in LD with one or more genes known to be involved in JA-, SA- or resistance-related signal transduction (**Table S4**). Among these genes are *LOX5*, whose product is involved in facilitating aphid feeding, *MYB107* encoding a transcription factor responsive to SA, the JA-inducible genes *TPS02* and *TPS03* encoding terpene synthases and *MES16*, encoding a methyl jasmonate esterase. In addition to screening for SNPs with contrasting effects, we screened for SNPs with a similar effect across the above-mentioned trait clusters (**Figure S5**) and found candidate genes involved in oxidative stress and plant responses to salinity and pathogens (**Table S5**).



### QTLs underlying contrasts in responses to biotic and abiotic stresses

We expected a negative correlation between the responses to biotic and abiotic stresses. Testing for this contrast within the GWA analysis using our MTMM approach significantly identified 43 SNPs with a QTL effect size that changed sign between biotic and abiotic conditions. Traits were then ordered by a cluster analysis on estimated SNP effects across significant SNPs, while SNPs were ordered by clustering their effects across traits. Figure 3 shows the SNPs with the strongest overall effects, identified in 18 LD intervals. The minor alleles of nine of these SNPs displayed a positive effect on biotic stress response traits and a negative effect on abiotic response traits. The remaining nine SNPs displayed the opposite effect (**Figure 3**). Several candidate genes were identified in LD with the SNPs that are specific for plant responses to either abiotic or biotic stresses (**Table 2b**), such as *TCH4* (encoding a cell-wall modifying enzyme), *AtCCR2* (involvement in lignin biosynthesis) and *ASG1* (a gene induced by ABA and salt stress), were identified. Transcription data (**Figure S6**) support the notion that these genes play a contrasting role in responses to abiotic and biotic stresses. A screen for QTLs with similar effects on resistance to biotic and abiotic stress (**Figure S7**) identified three genes annotated to be responsive to stress stimuli. Transcriptional data show that these genes respond differentially to different (a)biotic stresses and phytohormones (**Figure S8**). Genes like *ARGAH2* (involved in JA-mediated resistance to necrotrophic fungus) and *PKS1* (involved in light responses) are promising candidates for consistent effects across biotic and abiotic stresses (**Table S6**).

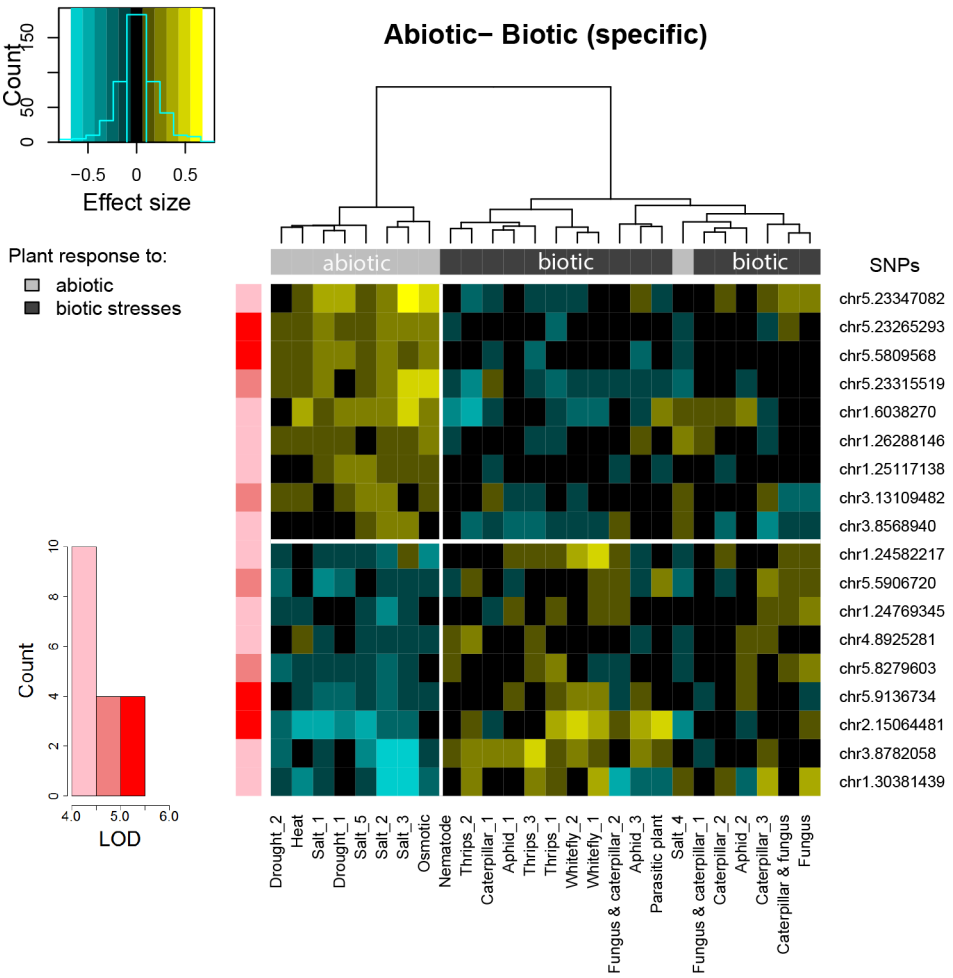



**Figure 3. Genetic associations specific for contrasting plant responses to either abiotic or biotic stress.** Genetic associations were estimated with a contrast-specific analysis using MTMM. SNPs with a significance above LOD score 4 ($P \leq 10^{-4}$) for the biotic-abiotic contrast are clustered according to trait-specific effect size estimated from the full MTMM. If there was another SNP in LD that had a higher effect size, this SNP was used a representative for the LD block. Negative effect sizes (blue) were cases where the rare allele was associated with a detrimental effect on the plants, positive effect sizes (yellow) were cases where the rare allele was associated with increased resistance to the stress. The rare alleles of the top 9 SNPs are associated with enhanced resistance to abiotic stresses and reduced resistance to biotic stresses; the bottom 9 SNPs show the inverse. Stresses are clustered according to effect size, using Ward's minimum variance method. If SNPs were located within a 20 kb half-window of each other, only the SNP with the highest absolute cumulative effect size was included. The key shows the frequency distribution of SNPs across effect sizes.

### QTLs underlying contrasts in responses to below- and aboveground stresses

We expected a negative correlation between responses to below- and aboveground stresses. A strong QTL signal was found on chromosome 1 for this contrasting response (**Figure S9**). The associated marker (chr1. 13729757) had 12 genes in LD with it, of which 11 are annotated as pseudogenes. Transcriptional data on abiotic stresses for the only protein coding gene (*AT1G36510*) shows an upregulation in above tissues, yet a downregulation in the root tissues. Marker chr5.16012837 showed the strongest signal for similar effects on responses to below- and aboveground stresses (**Figure S10**) for which the *pathogenesis-related thaumatin superfamily protein* (*AT5G40020*) is the most promising candidate gene.

### Validation of identified QTLs

To obtain experimental support for the most interesting QTLs resulting from the MTMM, we tested homozygous T-DNA insertion lines for candidate genes *RMG1* and *WRKY38* (both resulting from the MTMM analysis), and *TCH4* (from MTMM analysis on biotic versus abiotic contrast) for several of the stresses addressed in this study. Two independent *rmg1* T-DNA insertion lines showed a phenotype that was different from the wild type (Col-0) for some of the stress conditions (**Figure 4**, Supplementary Materials and methods section 3.11), being more resistant to caterpillar feeding and osmotic stress (**Figure 4**). *RMG1* (*AT4G11170*) encodes an NB-LRR disease resistance protein, and transcription is highly induced by the bacterial peptide flg22 (Yu *et al.*, 2013). The rare allele of the corresponding marker chr4.6805259 is associated with enhanced resistance to salt stress and the combined stresses 'caterpillar and drought' and 'caterpillar and fungus' and with enhanced susceptibility to drought and thrips stress. Gene expression data show that *RMG1* is upregulated by several abiotic and biotic stresses (**Figure 4**). T-DNA insertion lines for *TCH4* and *WRKY38* did not show a phenotype different from the wild type (Col-0) for any of the tested stress conditions. Whether this is dependent on the genetic background used, remains to be investigated.

Summarizing, our multi-trait GWA methodology facilitated a detailed analysis of the genetic architecture of resistance in Arabidopsis to a wide diversity of biotic and abiotic stresses. Application of this methodology revealed novel candidate genes associated with multiple stress responses, where specific contrasts were identified with some genes positively associated with the resistance to one set of stresses while being negatively associated with another set of stresses. In plant breeding (Ballesteros *et al.*, 2015; Brady *et al.*, 2005), such genes are classified as adaptive. Alternatively, other genes were identified with consistent effects across a wide spectrum of stress conditions. Such genes are labelled as constitutive in the plant breeding literature (Ballesteros *et al.*,

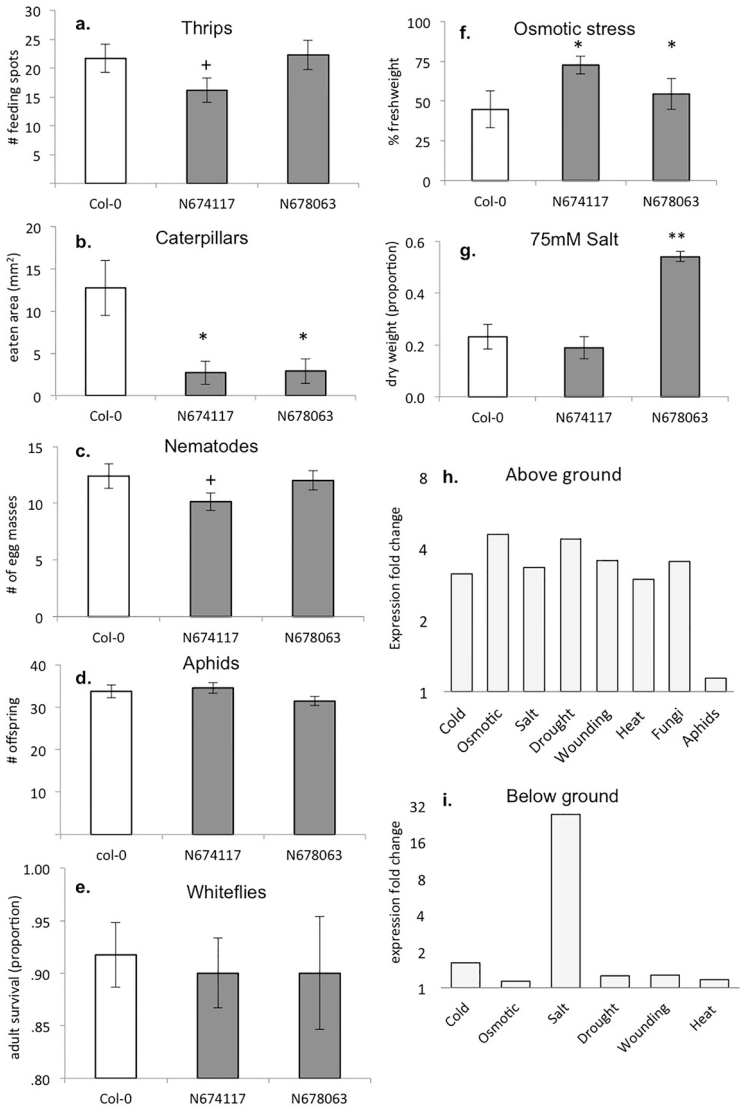



**Figure 4. Phenotypes of** *RMG1* **T-DNA mutant screenings.** Phenotypes are given for two T-DNA lines in the *RMG1* gene and for Col-0 as control. **a.** Number of thrips feeding spots on a detached leaf, 6 days post infestation (n=24); **b.** Leaf area consumed by *P. rapae* caterpillars (n=6); **c.** Number of nematode egg masses (n=23); **d.** Number of *M. persicae* aphid offspring (n=10-17); **e.** Percent survival of adult whiteflies (*A. proletella*) (n=10); **f.** Plant fresh weight after osmotic treatment in comparison to control (% relative to control) (n=4); **g.** Plant dry weight after 75mM salt treatment in comparison to control (ratio)(n=7-10) ( mean ± SE, +: $P < 0.10$, *: $P < 0.05$, **: $P < 0.01$, difference in comparison to Col-0); **h.** Relative expression fold change for *RMG1* compared to untreated control plants in aboveground and; **i.** Belowground tissue. Expression data from Arabidopsis eFP browser (http://bbc.botany.utoronto.ca).

2015; Brady *et al.*, 2005). Both adaptive and constitutive QTLs are important factors to contribute to improved stress resistance and tolerance in commercial crop species (Ballesteros *et al.*, 2015; Brady *et al.*, 2005).

## Discussion

Using a novel mixed-model approach to multi-trait GWA mapping, the genetic architecture of Arabidopsis underlying a total of 30 stress response traits was analyzed. A special feature of our statistical approach is that the GWA analysis accounted simultaneously for dependencies between genotypes and between traits. Through this approach, candidate genes for adaptive stress responses were identified that are involved in contrasting responses when comparing biotic and abiotic stresses, above- and belowground stresses, and attack by phloem feeders compared with other biotic stresses. Among these genes many are involved in phytohormone-mediated processes, supporting the notion that the phytohormonal regulatory network plays an important role in plant stress responses (Pieterse *et al.*, 2012). The MTMM approach further showed that certain SNPs were associated to multiple stress responses and that transcriptional patterns of genes to which the SNPs were linked, as well as the phenotype expressed upon knocking out one of these genes, matched with the observed stress responses of the plants. The *RMG1* gene that was identified through this procedure has relevant effects on plant phenotype in the context of responses to individual stresses. *RMG1* is a bacterium-inducible resistance gene whose activity is modulated by the plant through RNA-directed DNA methylation (RdDM) (Yu *et al.*, 2013). *RMG1* expression activates the SA pathway (Yu *et al.*, 2013). Thus, the increased resistance against caterpillars in *rmg1* mutants may be the result of elimination of SA-mediated interference with JA-induced resistance to caterpillars (Pieterse *et al.*, 2012). *RMG1* appears to be inducible by several stresses and deserves further in-depth analysis for its role in plant response to multiple stresses.

Our data show that for the 30 most significant SNPs resulting from the MTMM analysis, the average absolute effect size for double stresses is on average higher than that for single stresses ($P < 0.007$, **Table S2**). This indicates that resistance mechanisms involved in countering dual stresses are of a more general nature, in contrast to the rather specific resistance mechanisms involved in single stress responses.

The MTMM framework that we used for GWA mapping provides unbiased estimates for QTL allele substitution effects together with correct standard errors for these effects. Within the same framework we developed unique facilities to test hypotheses on QTL-by-stress interactions in multi-trait models, which are not available in competing meta-analysis approaches (Zhu *et al.*, 2015). The variance-covariance

structure that we used for the polygenic term protects against inflated type I error, i.e. too many false positive SNP-trait associations, as a consequence of population structure and kinship on the genotypic side and genetic correlations between traits on the trait side. The inclusion of trait correlations will, for most QTLs, improve the power of detection in comparison to single-trait GWA mapping. Our choice for the variance-covariance structure of the polygenic term as a Kronecker product of a compressed kinship on the genotypes with an approximated unstructured variance-covariance model on the environments is sometimes used in plant breeding for genomic prediction models (Burgueno *et al.*, 2012). However, implementation of such models in GWA mapping and especially on the scale that we present here, with 30 traits, is unprecedented and is practically far from straightforward. It required substantial work on preparatory phenotypic analyses as well as fine-tuning of the genotypic and trait variance-covariance structures to achieve convergence of the mixed models.

The MTMM analyses identified candidate genes associated with contrasting responses to biotic and abiotic stresses. Stress combinations appeared to have a strong influence on the MTMM outcome, indicative for significant interactions between different stresses when occurring simultaneously, and underlining the importance of studying the resistance of plants to combinations of stress. Transcriptional data and phenotyping of mutants provide initial support for the role of several of the candidate genes identified Studies of plant responses to a diverse set of biotic stresses show that the transcriptional pattern is stress-specific and that phytohormonal signaling pathways can explain up to 70% of the induced gene regulation (De Vos *et al.*, 2005). Taking the outcome of the MTMM analyses to investigate the involvement of identified candidate genes in the resistance of plants to several stresses, not only in Arabidopsis but also in related crop species such as e.g. *Brassica* species will be valuable in the breeding by design of future crops to protect them against combinations of stresses, including biotic and abiotic stresses. This will be of great value for next generation crops.

## Acknowledgements

This research is supported by the Perspective Programme 'Learning from Nature' of the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs.

## Supplementary information



## Materials and methods

### 1. *Arabidopsis thaliana* population

In this study we included 350 *Arabidopsis thaliana* (L.) Heynh. accessions from the Hapmap population (http://naturalvariation.org/hapmap). The Hapmap population has been genotyped for 250K bi-allelic SNPs (Baxter *et al.*, 2010; Chao *et al.*, 2012; Platt *et al.*, 2010) and after quality control and imputation this SNP-set was reduced to a set of 214,051 SNPs.

### 2. Definition of the target traits

For every experiment, the target traits were derived from the individual plant data using the following strategy. First, when traits followed a non-normal distribution, a logarithmic, arcsine or square root transformation was applied to the original observations. Second, genotypic (accession) means for each treatment were calculated using a mixed model to account for design effects. Different mixed models were used in the experiments, reflecting the different designs. In all cases, accession effects were modelled as fixed, and the accession means were the best linear unbiased estimator (BLUE) of these effects. Third, for traits measured in treatment and control conditions, differences or residuals (when regressing treatment on control values) were defined, in order to obtain a measure of stress tolerance independent of the trait under control conditions. Finally, within each experiment, the traits were replaced by the first principal component if the latter explained more than half of the variation in all traits in this experiment; in all other cases the original traits were retained. An overview of final traits and their corresponding sections in this supplementary methods file can be found in **Methods Table M1**. In case of replacement by the first principal component, original traits and the variance explained by the first principal component are listed (**Methods Tables M2-M6**). In total, phenotypic data for 73 individual traits were obtained by 10 different groups. All calculation were performed in R, unless stated otherwise. Mixed model analysis was performed with the R-package asreml (Butler *et al.*, 2009). In all equations the term $E$ denotes residual error. All other terms represent fixed effects unless stated otherwise. A colon (:) is used to define interactions between terms.

**Methods Table M1.** Trait list of tested phenotypes and section.

| | Stress | Trait name | Section | Trait phenotype | Treatment |
|---|---|---|---|---|---|
| Abiotic stresses | Salt | Salt_1 | 3.1 | Main root length, number of lateral roots and straightness | 75 mM NaCl |
| | | Salt_2 | 3.1 | Main root length | 125 mM NaCl |
| | | Salt_3 | 3.1 | Number of lateral roots | 125 mM NaCl |
| | | Salt_4 | 3.1 | Main root angle | 125 mM NaCl |
| | | Salt_5 | 3.2 | Biomass | 25 mM NaCl |
| | Drought | Drought_1 | 3.2 | Biomass | Drought |
| | | Drought_2 | 3.9 | Biomass | Drought |
| | Osmotic | Osmotic | 3.2 | Biomass | PEG8000 |
| | Heat | Heat | 3.2 | Number of siliques | 35°C |
| Biotic stresses | Parasitic plant | Parasitic plant | 3.4 | Attachments | *Phelipanche ramosa* |
| | Nematode | Nematode | 3.3 | Offspring, eggmass | *Meloidogyne incognita* |
| | Whitefly | Whitefly_1 | 3.5 | Survival, whiteflies | *Aleyrodes proletella* |
| | | Whitefly_2 | 3.5 | Reproduction, eggs | *A. proletella* |
| | Aphid | Aphid_1 | 3.6 | Behavior T1, probing | *Myzus persicae* |
| | | Aphid_2 | 3.6 | Behavior T2, probing | *M. persicae* |
| | | Aphid_3 | 3.6 | Offspring, aphids | *M. persicae* |
| | Thrips | Thrips_1 | 3.7 | Feeding damage | *Frankliniella occidentalis* |
| | | Thrips_2 | 3.7 | Behavior T1 | *F. occidentalis* |
| | | Thrips_3 | 3.7 | Behavior T2 | *F. occidentalis* |
| | Caterpillar | Caterpillar_1 | 3.8 | Leaf area consumed | *Pieris rapae* |
| | | Caterpillar_2 | 3.9 | Biomass | *P. rapae* |
| | | Caterpillar_3 | 3.9 | Number of damaged leaves and feeding sites | *P. rapae* |
| | Fungus | Fungus | 3.10 | Number of spreading lesions | *Botrytis cinerea* |
| Abiotic and biotic stress | Double stress | Fungus and caterpillar_1 | 3.9 | Biomass | *B. cinerea* and *P. rapae* |
| | | Fungus and caterpillar_2 | 3.9 | Number of damaged leaves and feeding sites | *B. cinerea and P. rapae* |
| | | Caterpillar and fungus | 3.10 | Number of spreading lesions | *P. rapae* and *B. cinerea* |
| | | Drought and fungus | 3.10 | Number of spreading lesions | Drought and *B. cinerea* |
| | | Drought and caterpillar | 3.9 | Number of damaged leaves and feeding sites | Drought and *P. rapae* |
| | Double stress | Caterpillar and osmotic_1 | 3.8 | Projected leaf area | *P. rapae* and PEG8000 |
| | | Caterpillar and osmotic_2 | 3.8 | Biomass | *P. rapae* and PEG8000 |

## 3. Phenotyping the HapMap population

### 3.1 Salt

#### Traits

Salt_1: root response to mild salt stress (75 mM NaCl), in terms of a combination of main root vector length (MRVL), number of lateral roots per main root (noLR) and straightness (main root length divided by MRVL)
Salt_2: root response to severe salt stress (125 mM NaCl), in terms of MRVL
Salt_3: root response to severe salt stress (125 mM NaCl), in terms of noLR
Salt_4: root response to severe salt stress (125 mM NaCl), in terms of straightness.

**Methods Table M2.** Salt trait reduction overview.

| Original traits[1] | Variance explained | Trait |
|---|---|---|
| Main root vector length at 75 mM NaCl<br>Number of lateral roots per main root at 75 mM NaCl<br>Straightness at 75 mM NaCl | 0.585977 | Salt_1 |
| Main root vector length 125 mM NaCl | 1 | Salt_2 |
| Number of lateral roots per main root at 125 mM NaCl | 1 | Salt_3 |
| Straightness at 125 | 1 | Salt_4 |

[1] Residuals from control are taken for all original traits.

#### Growing conditions

Seeds were stratified at 4°C for 3 days. Seeds germinated on square agar plates positioned under an angle of 70 degrees containing half strength Murashi-Skoog medium (MS), 1% Dashin agar, 0.5% sucrose, 0.1% M.E.S. buffer, pH5.8 (KOH). 4-Day-old seedlings were transferred into agar plates containing different salt concentrations (0, 75 and 125mM). After transfer, plants were grown for 8 days at the same conditions as those to which they were exposed when they were germinated, and scanned every second day. The Root System Architecture was determined of 8-day-old plants in control conditions and 12-day-old plants in both salt stress conditions.
Four-day-old seedlings were transferred to plates containing 0, 75 or 125 mM NaCl (control condition, mild stress, severe stress, respectively). Phenotypes were measured on 8–day-old plants in control conditions and 12-day-old plants in both salt stress conditions. Of each plant Root System Architecture was determined using EZ-Rhizo software (Armengaud *et al.*, 2009).

#### Experimental design

Plants were screened in 7 rounds (experiments), each containing a maximum of 106 accessions. Most of the accessions (198) were present in only one round; Col-0 was present in all rounds. In each round, at least four plants were included per accession-treatment combination. All three treatments (0, 75 or 125 mM NaCl) were screened simultaneously. Plants were allocated to plates, each plate containing 2 plants of 2 accessions. The within-plate average of each accession was

the basis for subsequent analysis. The position of every plate in the racks was recorded. The growth chamber contained 6 racks, each holding 64 plates. Positions of racks were also recorded.

### Genotypic means

For each of the traits genotypic means were calculated. We obtained BLUEs (best linear unbiased estimator) for all genotype-treatment combinations using the following model:

$$Y = \mu + EXP + TRT + GEN + GEN{:}TRT + EXP{:}TRT + EXP{:}RCK + EXP{:}DIST + GEN{:}EXP + EXP{:}RCK{:}PLT + GEN{:}EXP{:}TRT + E,$$

where EXP is experiment, TRT is treatment, GEN is genotype, RCK is rack, PLT is plate and DIST is distance to the wall. The terms EXP, TRT, GEN, GEN:TRT and EXP:TRT were modeled as fixed effects and all other terms as random.

### Definition of target traits

Stress response was defined as the residual obtained from the regression of the genotypic means for salt stress (either mild or severe) on the values for control conditions. Salt_1 was defined as the first principal component of the response to mild stress of MRVL, noLR and straightness. Salt_2, Salt_3 and Salt_4 were defined as the severe stress of MRVL, noLR and straightness individually.

## 3.2 Abiotic

### Traits

Salt_5: plant response to mild salt stress (25 mM NaCl), in terms of plant fresh weight, dry weight and water content

Drought_1: plant response to drought stress (0.22 g $H_2O$/ g soil at day 14), in terms of plant fresh weight and water content

Osmotic: plant response to osmotic stress (10% of PEG8000 from day 8 until 18), in terms of fresh weight, dry weight, water content and rosette area

Heat: plant response to heat stress (1 day, 35 °C), in terms of number and length of siliques

### Growing conditions

Four types of stress were studied in four different experiments. Seeds were sown in Petri dishes on wet filter paper. After 4 days of cold treatment, they were placed in the light at room temperature for 1.5 day to germinate. Germinated seeds were placed on rockwool blocks saturated with nutrient solution (Hyponex, 1mM N, 1.1 mM P, 5.9 mM K). For the stress treatments control plants received nutrient solution only. Salt treatment contained Hyponex + 25 mM NaCl. The plants of the salt experiments were automatically watered by a flooding system three times a week. For the osmotic treatment, the plants received nutrient solution containing 0.1 g/ml PEG8000 on day 8, 11, 13 and 15. For the other experiments plants automatically watered by a flooding system for approximately 5 minutes, three times a week. All experiments were performed under 125 μmol $m^{-2}$ $s^{-1}$ light, 16h/8h light/dark schedule, 20/18°C and 70% humidity.

**Methods Table M3.** Abiotic trait reduction overview.

| Original traits | Variance explained | Trait |
|---|---|---|
| Fresh Weight of the Rosette at day 28<br>Dry Weight of Rosette at day 28<br>Water Content of Rosette at day 28<br>Dry weight of largest leaf at day 24 | 0.789968 | Salt_5 |
| Fresh Weight of Rosette at day 28<br>Fresh Weight of largest leaf at day 24<br>Rehydrated Weight of largest leaf at day 24<br>Water Content of the largest leaf at day 24 | 0.541105 | Drought_1 |
| Dry Weight of Rosette at day 28[2]<br>Fresh Weight of Rosette at day 28[2]<br>Water Content of Rosette at day 28[2]<br>Rosette Area at day 28[2]<br>Rosette Area at day 28 without bolting plants.[2] | 0.679514 | Osmotic |
| Number of aborted siliques along the inflorescence<br>Number of silique (<5mm) along the inflorescence<br>Number of silique (<5mm) in the region -10 until 20. [1]<br>Average of the length of all siliques along the inflorescence<br>Average of the length of siliques 0 until 10[1]<br>Average of the length of siliques 0 until 20[1]<br>Average of the length of siliques 10 until 20[1]<br>Average of the length of siliques 20 until 30[1]<br>Average of the length of siliques -10 until 0[1] | 0.647069 | Heat |

[1] Silique zero belongs to the flower that opened first on the day of the treatment. [2] Stress did not disappear when watering with PEG-containing nutrient solution stopped, because PEG is not evaporating.



## Experimental design

In the salt experiment 3 blocks received control treatment and 3 blocks received treatment conditions. For the drought experiments the Phenopsis phenotyping platform was used, preventing position related differences in plant growth within the climate chamber (Bac-Molenaar *et al.*, 2015b; Granier *et al.*, 2006). The plants were grown in 4 rounds of 84 accessions. Four of these 84 were used for reference accessions, which were grown in each round. Each round contained 3 blocks and all 84 accessions were present in each block. Plants for the PEG experiment were grown in 6 blocks. Each accession was present in each block. Within the block the plants were grown in 9 trays each containing 40 plants. Within a tray the plants had a fixed position. The 9 trays were positioned randomly within the block. 3 blocks received PEG treatment and 3 blocks received control treatment.

For the heat experiment, 8 plants of each accession were grown in controlled conditions (Bac-Molenaar *et al.*, 2015b). Five replicates received a heat treatment and 3 replicates served as controls. One to 2 weeks after the first replicate of each accession started to flower, the heat treatment was applied. A small number of accessions received the treatment outside this window. The first flower that opened first on the day of the treatment was tagged with a thread. Three replicates per accession were kept in the climate room as controls. Five replicates per

accession were transferred to a climate cabinet where they received heat treatment. At the start of the day, the temperature was raised from 20°C to 35°C within two hours. The temperature was kept at 35°C for 13.5 hours. At the end of the light period, the temperature was decreased again to 20°C in two hours. The day after the treatment the plants returned to the climate room.

#### Genotypic means and definition of target traits

All data in Drought_1 are log-transformed. In the Salt and PEG experiments, genotypic means were calculated using a mixed model containing random block effects and genotypic fixed effects. For the Heat experiment, the model included fixed effects for treatment, genotype and genotype x treatment interaction. For the Drought experiment, we fitted the mixed model used in (Bac-Molenaar *et al.*, 2015b). Next, for all traits in the four experiments, the impact of the stress was quantified using the residuals from the regression of genotypic means under stress on those under control conditions, except for the heat traits 'Rosette Area at day 28', where no control was available. Finally, the four target traits were defined as the first principal component of all traits (residuals) from the corresponding experiments.

### 3.3 Nematodes

#### Trait

Nematode: Plant response to nematodes (*Melodoigyne incognita*, 180 stage-2 juveniles), in terms of number of *M. incognita* egg masses per plant.

#### Growing conditions

Seeds were vapor-sterilized for 5 hours and transferred to a 6-well plate with MS20 (5% gelrite). After 4 nights in the dark at 4°C the plates were transferred to 24°C in 12 h light. At the age of 1 week the seedlings were transferred individually to a well of a 6-wells plate. *Melodoigyne incognita* infection was induced with 180 juveniles stage-2 added to 2-week-old seedlings. 6-Well plates with nematodes and seedlings were incubated in the dark at 24°C for 6 weeks. Plants were grown for 2 weeks : 24°C, 12 h. light,/12 h. dark , then 6 weeks 24°C, dark. Egg masses were quantified manually.

#### Experimental design

Plants were screened in rounds of 20 accessions. Each round included a 6-well plate with 1 Col-0 plant as reference.

#### Genotypic means

'Nematode' was defined as the number of egg masses, after arcsine-square root transformation. Genotypic means were calculated using the following mixed model:

$Y = \mu + GEN + RND + E,$

where GEN is genotype (accession) and RND is a random effect for round.

### 3.4 Parasitic plants

#### Trait

Parasitic plant: Plant response to parasitic plant (*Phellipanche ramose*), in terms of the total number of parasitic plant organ attachments onto the host root.

#### Growing conditions

Arabidopsis seeds were put on filter paper in the dark at 4°C for 2 days. Then, Arabidopsis seeds were sown on river sand (with a thin layer of soil on the top of river sand). Arabidopsis plants were grown for 2 weeks on river sand at 21°C, 60% RH, 100 μmol $m^{-2}$ $s^{-1}$ light intensity, 12h:12h L:D photoperiod. After 2 weeks, Arabidopsis seedlings were surface-sterilized with 70% ethanol for 5 seconds and washed with sterile demi-water. The rhizotron system was prepared by cutting a hole at the side of 14.5 cm diameter round Petri dish, putting successively a piece of round rock wool slice (14.5 cm diameter, 1.5 cm in thickness) at the bottom of Petri dish, a piece of 12 cm diameter glass-fibre filter discs and a piece of 14.5 cm diameter nylon mesh on top. The rhizotron system was supplied with sterile ½ Hoagland liquid medium. Sterile seedlings were then moved to prepared rhizotron system by fitting the plant in the hole of the Petri dish. Leaves and shoots of the seedlings were kept outside of Petri dishes. The roots were carefully separated and organized on the top of nylon mesh by forceps. Arabidopsis seedling were grown in rhizotron system at 21°C, 60% RH, 100 μmol $m^{-2}$ $s^{-1}$ light intensity, 12h:12h L:D photoperiod for another 2weeks.

Sterile *Phellipanche ramosa* seeds were spread on 5 cm diameter glass-fiber filter discs (Whatman GF/A paper) which were wetted with 0.8 ml sterilized demi-water and placed in 9 cm diameter Petri dishes. The Petri dishes were sealed with parafilm and then kept in dark in a growth chamber at 20°C for a 12 days precondition period. Preconditioned seeds on a glass-fiber filter disc were dried and treated with 0.8 ml strigolactone analog GR24 at the concentration of $3.3x10^{-3}$μM for 1 day under dark at 25°C. GR24 treatment triggered the initial germination of *P. ramosa*. After 1 day, GR24 was immediately washed off the *P. ramosa* seeds by sterile demi-water.

Pre-germinated *P. ramosa* seeds were spread along 4-week-old Arabidopsis seedlings in the rhizotron system with painting brushes. The rhizotron petri dish were sealed with tape and covered by aluminium foil. Plant were grown at the same condition for the following 4 weeks. Pictures of *P. ramosa*-infested roots in the rhizotron system were taken 4 weeks after infection with Canon camera EOS 60D DSLR (with EF-S 18-135mm IS Lens).

#### Experimental design

The 359 accessions were screened in 2 rounds (the first 200 accessions, the second with 160 accessions, 2 accessions were used for control in both rounds). Rhizotron petri dishes were randomly arranged in trays. Positions of trays and Petri dishes were rearranged randomly every 3 days. Pictures of rhizotrons were taken after 4 weeks. Image analysis was done with the ImageJ software (Schneider *et al.*, 2012). The number of attachment organs was counted. The total number of pre-germinated *P. ramosa* seeds was recorded as a co-variable.

### Genotypic means and definition of target traits

Values for diameter of attachment organs, number of attachment organs and number of pre-germinated seeds were log-transformed for normality and were averaged over technical replicates where present. Since there was significant correlation between the two variables of interest and the number of pre-germinated *P.ramosa* seeds, we used the residuals from the regression on the number of pre-germinated seeds for further analysis.

## 3.5 Whiteflies

### Traits

Whitefly_1: Plant response to whitefly (*Aleyrodes proletella*, 5 females), in terms of whitefly survival
Whitefly_2: Plant response to whitefly (*Aleyrodes proletella*, 5 females), in terms of number of eggs

### Growing conditions

Plants were grown for 5 weeks at 20°C, 70% RH, 100 μmol $m^{-2}$ $s^{-1}$ light intensity and 10h:14h L:D photoperiod. One leaf of each accession was infested with 5 female whiteflies (placed in clip cages) that were allowed to feed and oviposit. Seven days after infestation, the number of living and dead females was counted as well as the number of eggs. From this, we calculated the survival (number of living flies divided by the total number of flies) and oviposition rate (eggs laid per female per day).

### Experimental design

Accessions were screened in 3 blocks of 120 accessions with 5 reference accessions (Col-0, Ler-1, WS-0, Cvi-0, Kin-0) in each block. The whole experiment was repeated 5 times.

### Genotypic means

Genotypic means were calculated with Genstat 15th edition (Payne, 2009), using the following mixed model:

$Y = \mu + REP + GEN + REP:BLOCK + E,$

where *GEN* is genotype (accession), *REP* denotes complete replicates (experiments with 3 blocks) and REP:BLOCK is a random effect for incomplete blocks within replicates.

## 3.6 Aphids

### Traits

Aphid_1: Plant response to aphids (*Myzus persicae*), in terms of behavior at t1
Aphid_2: Plant response to aphids (*M. persicae*), in terms of behavior at t2
Aphid_3: Plant response to aphids (*M. persicae*), in terms of aphid reproduction

**Methods Table M4.** Aphid trait reduction overview.

| Original traits | Variance explained | Trait |
|---|---|---|
| Total duration probing (logit) [1]<br>Total duration of short probes (< 3 min , arcsine) [1]<br>Total duration of intermediate probes (< 15 min, arcsine) [1] | 0.74997 | Aphid_1 |
| Total duration probing (logit) [2]<br>% of aphids making long probes (>= 15 min) [2]<br>Total duration of intermediate probes (< 15 min, arcsine)[2] | 0.655366 | Aphid_2 |
| Number of aphids per plant [3] | 1 | Aphid_3 |

[1] 0h after inoculation
[2] 4.5h after inoculation
[3] 2 weeks after inoculation

## Growing conditions

Plants were grown for 4 to 5 weeks at 23°C, 70% RH, 200 μmol $m^{-2}$ $s^{-1}$ light intensity and 8h:16h L:D photoperiod. Green peach aphids, *M. persicae*, were reared on radish, *Raphanus sativus*, at 19°C, 50-70% relative humidity and a 16h day and 8h night cycle. Behaviour of the green peach aphid was screened by automated video-tracking. One leaf disc was collected per plant from an intermediately aged leaf and placed abaxial side up on a 1% agar substrate in a well of a 96-well microtitre plate. One 7- to 8-day-old wingless aphid was released on the leaf disc and cling film was used to cover the arena. 20 Arenas were recorded simultaneously with a mounted camera. Aphids were observed for 85 minutes on 2 time points: (1) immediately after introducing the aphids into the arenas, and (2) 4.5 hours after the start of the first observation. Motion analysis was performed with EthoVision XT® 8.5 software (Noldus Information Technology bv, Wageningen, The Netherlands). Aphids are phloem-feeding insects and probe with their piercing mouthparts between plant cells to feed from the plant sap. Start time and duration of probes were registered with automated video-tracking (Kloth *et al.*, 2015). Aphid survival was checked 24 hours after recording. Subject detection was checked on 4 time points within each movie. Samples with no survival, low subject detection or with less than 5 replicates were excluded from analysis. Probes were categorized into short (< 3 min) probes and intermediate (< 15 min) probes, both associated with penetration of the plant epidermis or mesophyll, and long (>=15 min) probes, putatively associated with phloem uptake. Response variables expressed in seconds were arcsin or logit transformed to approach a normal distribution.

Aphid reproduction was measured in a whole-plant assay. Each 2-to-3-week-old plant was inoculated with one 0-to-24-hour-old nymph. Two weeks after infestation, aphid population size was measured per plant. Plants were placed in a Petri dish in trays with a water barrier to prevent aphids to move between plants. Each tray contained 20 plants, none of the aphids developed wings.

### Experimental design

Automated video tracking of aphid behavior was performed in an incomplete block design with each complete replicate consisting of 18 incomplete blocks of 20 accessions. One replicate of the complete Hapmap collection was acquired in 6 days, 60 plants were screened each day across 3 batches. An alpha design was generated with Gendex (http://designcomputing.net/gendex/) to assign accessions to blocks. For each accession 5 to 6 replicates were acquired.
Phenotyping of aphid reproduction was performed in an incomplete block design with 7 incomplete blocks. Blocks were defined according to the position in the climate cell. Each replicate consisted of 3 to 4 blocks and plant genotypes were randomized across blocks between replicates. For each genotype 2 to 3 replicates were acquired.

### Genotypic means and definition of target traits

Genotypic means were calculated using the following linear mixed model:
$Y = \mu + REP + GEN + REP:BLOCK + E$,
where REP denotes complete replicate and REP:BLOCK is a random term for block nested within replicate.

## 3.7 Thrips

### Traits

Thrips_1: Plant response to thrips (*Frankliniella occidentalis*, 3 juveniles, 6 days), in terms of feeding damage on detached leaf.
Thrips_2: Plant response to thrips (*F. occidentalis* 1 adult), in terms of behavior/ preference at t1 (0 hpi) in two choice leaf disc assay
Thrips_3: Plant response to thrips (*F. occidentalis*, 1 adult), in terms of behavior/ preference at t2 (4 hpi) in two choice leaf disc assay

### Growing conditions

Plants were grown for 5 weeks at 23°C, 70% RH, 200 μmol $m^{-2}$ $s^{-1}$ light intensity and a 8h:16h L:D photoperiod.
For trait 'Thrips_1', feeding damage on detached leaves was scored. Leaves were cut from plants, and kept turgid in Petri dishes with a diameter of 5 cm (BD falcon, Product Number: 351006) containing a film of 1% technical agar. The amount of feeding damage on one leaf was manually scored after 6 days of exposure to 3 juvenile thrips.
For traits 'Thrips_2' and 'Thrips_3', thrips preference was phenotyped with an automated video tracking setup. Thrips behavior was tracked in 2-choice arenas using 96-well plates, consisting of halved leaf discs from Col-0 and one of the HapMap accessions. Position bias was corrected for, by alternating the Col-0 leaf disc position (left or right) every row. 20 plants were screened in one recording. Thrips position was automatically monitored for 40 minutes (Thrips_2), and once more for 40 minutes after 4 hours (Thrips_3). The ratio of time spent on Col-0 was used for Thrips_2 and Thrips_3. Video tracking was performed with EthoVision XT 8.5 software

(Noldus Information Technology bv, Wageningen, The Netherlands).

### Experimental design

Plants were screened in 5 rounds (complete replicates) of 360 accessions, using an incomplete block (alpha) design. Within each round plants were randomly allocated to 18 blocks of 20 accessions, the blocks representing plants being screened in one recording. One sampling day consisted of 5 blocks (100 accessions), with the exception of the last day (3 blocks, 60 accessions).

### Genotypic means

Genotypic means were calculated using the following linear mixed model:

$Y = \mu + REP + GEN + REP:BLOCK + E,$

where REP denotes complete replicate and REP:BLOCK is a random term for blocks nested within replicate.

## 3.8 Drought – combinatory stress

### Traits

Caterpillar_1: Plant response to *Pieris rapae* change in terms of rosette area

Caterpillar and osmotic_1: Plant response to *P. rapae* and osmotic stress (PEG8000), in terms of terms of rosette area

Caterpillar and osmotic_2: Plant response to *P. rapae* and osmotic stress (PEG8000), in terms of plant biomass

**Methods Table M5.** Drought – combinatory stress trait reduction overview.

| Original traits[1] | Variance explained | Trait |
|---|---|---|
| Rosette perimeter after Caterpillar treatment<br>Rosette area after Caterpillar treatment<br>Rosette ferret after Caterpillar treatment | 0.841488 | Caterpillar_1 |
| Rosette perimeter after Caterpillar/Osmotic treatment<br>Rosette area after Caterpillar/Osmotic treatment<br>Rosette ferret after Caterpillar/Osmotic treatment | 0.823736 | Caterpillar & osmotic_1 |
| Plant Fresh weight after Caterpillar/Osmotic treatment | 1 | Caterpillar & osmotic_2 |

[1]Residuals obtained from regressing treatment means on control means.

### Growing conditions

Plants were grown for 4 weeks at 21°C (day temperature) and 19°C (night temperature), 70% RH, 200 μmol $m^{-2}$ $s^{-1}$ light intensity, and 10h:14h SD photoperiod. Projected leaf area, rosette feret and rosette perimeter were measured using ImageJ software in the *P. rapae* and PEG8000 combined treatment group (Schneider *et al.*, 2012). Data were recorded at 3 time points: T1) before applying *P. rapae*; T2) before applying PEG8000 treatment; T3) after 7 days PEG8000 treatment. Rosette fresh weight was measured from both control and combinatorial stress

treatment groups at T3. We used 332 Arabidopsis accessions, grown on rock wool blocks in a climate controlled growth chamber. Plants were first treated with *P. rapae* L1 larvae for 24 hours, and then irrigated with nutrient solution that containing PEG8000 for 7 days (*P. rapae* and PEG8000 combined stress treatment). In additional, plants were grown without any stress treatment (control).

**Experimental design**

All traits were measured in a randomized complete block design with 2 complete blocks (replicates) under treatment conditions and 2 complete blocks under control conditions.

**Genotypic means**

The square root transformation was first applied to the area traits. For each treatment, genotypic means were calculated, using a linear model with a fixed effect for block.

**Definition of target traits**

We regressed the genotypic means of fresh weight at T3 after *P. rapae* and PEG8000 treatment on the means of fresh weight at T3 under control conditions; these residuals represent the effect of *P. rapae* and PEG8000 treatment on fresh weight (Caterpillar_&_osmotic_2). Similarly, we regressed each of the rosette area related traits (projected leaf area, rosette feret, and rosette perimeter) observed at T2 (only *P. rapae* treatment) on the corresponding means of these traits measured at T1 (Caterpillar_1). The resulting residuals represent the effect of *P. rapae* treatment on rosette area related traits. Finally we performed the regression of rosette area related traits at T3 on the values measured at T1 as well as T2, whose residuals represent the combined effect of *P. rapae* and PEG8000 treatment on rosette area related traits (Caterpillar_&_osmotic_1). In both the Caterpillar_1 and Caterpillar_&_osmotic_1 group, the 3 traits were replaced by the first principal component.

### 3.9 Caterpillar – combinatory stress

**Traits**

Drought_2: Plant response to drought (7 days), in terms of plant biomass after drought recovery
Caterpillar_2: Plant response to *P. rapae*, in terms of plant biomass
Fungus and caterpillar_1: Plant response to *Botrytis cinerea* and *P. rapae*, in terms of plant biomass
Caterpillar_3: Plant response to *P. rapae*, in terms of damaged leaves and feeding sites
Drought and caterpillar: Plant response to drought stress and *P. rapae*, in terms of damaged leaves and feeding sites
Fungus and caterpillar_2: Plant response to *B. cinerea* and *P. rapae*, in terms of damaged leaves and feeding sites

**Methods Table M6.** Caterpillar – combinatory stress trait reduction overview.

| Original traits | Variance explained | Trait |
|---|---|---|
| Biomass reduction (with respect to control conditions) upon drought stress following a recovery period. | 1 | Drought2 |
| Biomass reduction (with respect to control conditions) upon *P. rapae* herbivory | 1 | Caterpillar_2 |
| Biomass reduction (with respect to control conditions) upon *P. rapae* herbivory preceded by *B.cinerea*<br>Biomass reduction (with respect to *P. rapae* single stress) upon *P. rapae* herbivory preceded by *B.cinerea* | 0.910749 | Fungus & caterpillar_1 |
| Number of leaves damaged upon *P. rapae* herbivory<br>Number of feeding sites upon *P. rapae* herbivory | 0.792034 | Caterpillar_3 |
| Number of leaves damaged upon *P. rapae* herbivory preceded by drought<br>Number of feeding sites upon *P. rapae* herbivory preceded by drought | 0.792354 | Drought & caterpillar |
| Number of leaves damaged upon *P. rapae* herbivory preceded by *B.cinerea*<br>Number of feeding sites upon *P. rapae* herbivory preceded by *B.cinerea* | 0.788294 | Fungus & caterpillar_2 |



### Growing conditions

Plants were grown for 4 weeks at 23°C, 70% RH, 100 μmol $m^{-2}$ $s^{-1}$ light intensity and 8h:16h L:D photoperiod. Plants were grown under similar conditions during the first 3 weeks. Drought stress was imposed by withholding water for 7 days while the rest of plants were watered every 2 days with 1 liter of water per tray. *Botrytis* inoculation was carried out 24 h prior to *Pieris* inoculation. Plants were 4 weeks old when they were exposed to stress by *P. rapae* as single or combined stress. Plants were inoculated with 2 newly hatched L1 and the larvae were allowed to feed for 5 days until harvesting. Phenotypic measurements were taken 24 h and 5 days after inoculation with *P. rapae* as a single and combined stress. After 24 h, the number of leaves damaged and number of feeding sites was counted in plants exposed to *P. rapae* , drought and *P. rapae*, and *Botrytis* and *P. rapae.* After 5 days, fresh weight was measured for the 5 treatments.

### Experimental design

Plants were screened in rounds of 37 accessions. Three control accessions were present in all rounds (Col-0, Tsu-0, Fei-0). In each round, 6 replicates were included per accession-treatment combination. Treatments were screened simultaneously. Plants were randomly allocated in trays (28 accessions per tray). Plant positions within a tray were recorded ($X_{pos}$ and $Y_{pos}$). The chamber where the experiment were conducted consist of 6 racks**,** and each rack contained 4 shelves**.** Positions of trays within shelves within racks were also recorded.

### Genotypic means

For each of the 3 traits (shoot fresh weight, number of leaves damaged and number of feeding sites) we fitted the following mixed model:

$$Y = \mu + ROUND + RACK + SHELF + TRT + GEN + GEN{:}TRT + ROUND{:}RACK{:}SHELF + ROUND{:}RACK{:}SHELF{:}TRAY + ROUND{:}RACK{:}SHELF{:}TRAY{:}X_{POS} + ROUND{:}RACK{:}SHELF{:}TRAY{:}Y_{POS} + E,$$

where *TRT* is the treatment factor (Control and 4 treatment levels), *GEN* is genotype (accession) and *GEN:TRT* is the genotype by environment interaction. The terms GEN, TRT AND GEN*:*TRT were fitted as a fixed effect and all others as random. For each of the 3 traits, significance of each model term was assessed and only significant terms were retained for the estimations of genotypic means. Genotypic means (BLUEs) were calculated for shoot fresh weight, number of leaves damaged and number of feeding sites, for each accession-treatment combination.

### Definition of target traits

Target traits were defined based on the genotypic means for shoot fresh weight, number of leaves damaged and number of feeding sites. The traits Caterpillar_3, Drought & caterpillar and Fungus & caterpillar_2 were defined as the first principal component of the number of damaged leaves and number of feeding sites under the respective types of stress. For shoot fresh weight, stress response was defined by the residuals obtained from the regression of genotypic means under each stress condition on those for the non-stress condition (Drought2, Caterpillar_2 and Fungus & caterpillar_1). In case of Fungus & caterpillar_1, also the regression of the double stress (*P. rapae* and *B. cinerea*) on the single stress (*P. rapae*) was performed, and the trait was defined as the first principal component of the 2 residuals.

## 3.10 Fungus – combinatory stress

### Traits

Fungus: Plant response to *B. cinerea* infection ($1x10^5$ spores/ml), in terms of percentage of spreading lesions

Drought and fungus: Plant response to drought (7 days) followed by *B. cinerea* infection, in terms of percentage of spreading lesions

Caterpillar and fungus: Plant response to *P. rapae* feeding (one L1 caterpillar/ plant, 24 hours) followed by *B. cinerea* infection, in terms of percentage of spreading lesions

### Growing conditions

Seeds were sown and vernalized for two days at 4˚C on river sand supplied with half strength Hoagland medium with sequestreen. Ten-day-old seedlings were transplanted to pots containing half volume river sand and half volume sowing soil supplemented with Hoagland solution (with sequestreen). Plants were kept at ~21˚C, 70% relative humidity, 8h:16h light:dark period. At day 0 of the experiment 27-day-old plants were exposed to a period of drought stress or a normal

watering regime. At day 7 of the experiment, drought stress was stopped by re-watering the drought stressed plants. At that day 7 one first instar (L1) *P. rapae* caterpillar was put on each plant for the dual stress combination with herbivory. At day 8, *P. rapae* was taken off the plants and all the plants from the different treatments were simultaneously inoculated with *B. cinerea*. Six leaves per plant were each drop inoculated with 5 µl of $1\times10^5$ spores/ml, in half strength potato dextrose broth. Plants were kept under ~100% humidity for three days, after which the disease severity was measured on day eleven of the experiment. Severity was measured as percentage of leaves with spreading lesions caused by *B. cinerea*. In total 6 leaves per plant were scored. Lesions that did not exceed the size of the droplet, (5 µl) were scored as zero, whereas a spreading lesion was scored as a one.

### Experimental design

Plants were screened in rounds of 35 accessions. Col-0 was present in all rounds as a control. The 3 treatments were screened simultaneously.

### Genotypic means and definition of target traits

An arcsine transformation was applied to the proportion of leaves with spreading lesions, i.e. for each observed count k= 0, 1, 2, 3, 4, 5, 6, the transformed phenotype was defined as arcsin(square root(k)/6). Prior to transformation, counts equal to zero or 6 were replaced by respectively 1/4 and 5.75=6-1/4. The transformed phenotypic observations were corrected for round effects by subtracting from each observation the mean of the round it was contained in, and genotypic means were calculated based on the round corrected phenotypes. Differential sensitivity of each double stress was calculated as the residuals obtained from the linear regression of the double stress on the single stress phenotype.



## 3.11 Screening of T-DNA lines

T-DNA lines were ordered and screened for homozygosity, using primers described in **Methods Table M7**. Seeds from homozygous mutants were harvested and grown and screened individually by consortium partners (**Methods Table M8**).

**Methods Table M7.** T-DNA lines and primers.

| Mutant line | LP | TM | RP | TM | Product size |
|---|---|---|---|---|---|
| WiscDsLox489-492C21 | ATTTGGTAAACCCAAATTGGC | 59.94 | CGATGAAGGAGGATAAGAGCC | 60.18 | 1178 |
| SAIL_158_A07 | AACAAAAACCGCGTGATTTC | 59.98 | CAAGAAGACTTGCCGTTTGAC | 59.91 | 1010 |
| SAIL_422_D11 | AACAAAAACCGCGTGATTTC | 59.98 | CAAGAAGACTTGCCGTTTGAC | 59.91 | 1010 |
| SALK_023944.54.15.x | TGGTCTAATGGGCTCAATGAG | 60.08 | CATAGCCGTTGTCAATTCCAG | 60.51 | 1009 |
| SALK_007034.41.00.x | TTTAGCGGTCAACACGAAAAC | 60.16 | CCAAAATTGAAAATAGAGAACCC | 58.14 | 1196 |

**Methods Table M8.** Methodology screening of T-DNA mutants.

| Trait | Number of replicates | Method |
|---|---|---|
| Thrips | 24 | See section 3.7, Thrips_1 |
| Aphids | 10-17 | See section 3.6, Aphid_3 |
| Whitefly | 10 | See section 3.5, Whitefly_1 |
| Caterpillar | 6 | For the caterpillar treatment, each plant was exposed to 1 *Pieris rapae* 1st instar larvae for 24h, thereafter, the caterpillar was removed from the plant. Damage was assessed using ImagJ software. |
| Nematodes | 23 | See section 3.3, Nematode |
| Salt | 10 | See section 3.2, Salt_5. 75 mM Salt instead |
| Drought | 4 | Plants were irrigated with Hyponex solution containing 7.7% polyethylene glycol (PEG8000) of osmotic potential about 0.1MPa for 7 days. |

## 4. Statistics

### 4.1 Statistics: Genetic correlation networks

Pairwise genetic correlations between traits were estimated using a multi-trait mixed model (MTMM) (Korte *et al.*, 2012). Residuals were assumed uncorrelated for traits that were measured on different plants. For some pairs of traits the likelihood was monotone, which can also occur in single-trait mixed models (Kruijer *et al.*, 2015). In this case, the genetic correlation was estimated by the (Pearson) correlation between the univariate G-BLUPs (De los Campos *et al.*, 2013) estimated for these traits. A network between predefined groups of traits was constructed by connecting groups whose average genetic correlation across pairs of traits was above 0.2.

### 4.2 Statistics: Multi-trait mixed models

Following (Zhou and Stephens, 2014), we assume the MTMM, Y=XB+G+E, with Y being the genotypes by traits (n × p) matrix of phenotypic observations. The terms, XB, G and E stand for respectively the fixed effects (including trait specific intercepts and SNP-effects) and the random genetic and environmental effects. G follows a zero mean matrix-variate normal distribution with row-covariance (marker-based kinship) matrix K and column (trait) covariance matrix $V_g$.$V_g$ is a p × p matrix modeling the genetic correlations between traits. This is equivalent with g= vec(G) (the vector containing the columns of G being multivariate normal with a covariance matrix defined by the Kronecker product $V_g$ x K (Zhou and Stephens, 2014)). Similarly, vec(E) follows a zero mean normal distribution with covariance $V_e$ x $I_n$, where $V_e$ accounts for the non-genetic correlations between traits.

### 4.3 Statistics: Factor-analytic models

Since $V_g$ and $V_e$ contain a total of p(p + 1) parameters, the MTMM above becomes difficult to fit for more than 10 traits (Zhou and Stephens, 2014). For $V_g$ we therefore assumed a factor analytic

model, which is well known in the context of QTL-mapping for experimental populations with limited numbers of markers (Boer et al., 2007), but has not been used in the context of multivariate GWAS. As almost all traits were derived from measurements on different plants, a diagonal model , $V_e$= diag($\sigma_{e,1}^2$,..., $\sigma_{e,p}^2$), was chosen for the environmental covariances. For $V_g$ a first order factor analytic structure was chosen $V_g=\sigma_g^2$ ($\lambda\lambda^t$+diag($\tau_1^2$,…,$\tau_p^2$ )), where $\sigma_g^2$ represents a scale parameter, the magnitude of genetic effects, the vector $\lambda=(\lambda_1,\ldots,\lambda_p)^t$ contains the trait specific scores belonging to the factor analytic part of the model that provides a rank one variance-covariance structure between traits, and diag($\tau_1^2$,…,$\tau_p^2$ ) provides trait specific residual genetic variances (Meijer, 2009; Piepho, 1997). The model was fitted with the R-package ASRreml (Butler *et al.*, 2009).

### 4.4 Statistics: Compressed kinship

Factor analytic models have been successfully applied to experimental populations with a simple genetic relatedness structure (Alimi *et al.*, 2013; Boer *et al.*, 2007; Malosetti *et al.*, 2008), but currently available software could not perform REML-estimation for the hapmap-population. The kinship matrix was therefore replaced by a compressed kinship matrix (Bradbury *et al.*, 2007; Zhang *et al.*, 2010), modelling the genetic relatedness between a number of internally homogeneous groups. Assuming there are m such groups, containing $n_1$, ..., $n_m$ accessions each, the original kinship matrix K is replaced by $ZK_C Z^t$, where $K_c$ is the kinship matrix for the groups, and Z is the n × m incidence matrix assigning each of the n accessions to one of the m groups. The groups were created by a procedure that restricted the marker data to be linear combinations of environmental covariates representing the conditions at the place of origin of the accessions, as explained below.

Although our factor analytic model has been successfully applied to experimental populations with a simple genetic relatedness structure (Boer *et al.*, 2007; Malosetti *et al.*, 2007; Malosetti *et al.*, 2011), currently available software could not perform REML-estimation for the hapmap-population. The kinship matrix was therefore replaced by a compressed kinship matrix (Bradbury *et al.*, 2007), modeling the genetic relatedness between a number of genetically homogeneous groups. Compressed kinship was calculated as the average kinship within genetic groups. Genotypes were assigned to k genetic groups by performing Ward clustering based on the squared Euclidean distance along the first k − 1 principal components calculated from a matrix of standardized SNP scores, followed by cutting the resulting dendrogram into k distinct clusters (Odong *et al.*, 2013; van Heerwaarden *et al.*, 2012; van Heerwaarden *et al.*, 2013).

Choosing the number of groups. The use of a compressed kinship matrix requires a choice of the level of compression, as determined by the number of genetic groups over which the individual kinship is averaged. This choice needs to balance the gain in computational efficiency with model fit (Zhang *et al.*, 2010) and the ability of the compressed matrix to capture the correlation between genetic dissimilarity and phenotypic differences, which is ultimately the reason for including a kinship matrix in the association model. There are currently no standard methods to determine the optimum level of compression. We determined the appropriate level of compression for each association model based on the model likelihood, convergence and

correspondence between kinship and phenotypic and geographical similarity. The latter was quantified as the Frobenius norm of the difference between the complement of the compressed kinship matrix, expanded to a block matrix of full rank, and the Euclidean distance matrix of phenotypic traits or geographic coordinates. We considered a range of 4 to 100 groups. Correspondence with phenotypic and geographical dissimilarity increased steeply from 4 to around 35 groups, after which correspondence with geographic distance increased more slowly and the correspondence with phenotypic distance showing a local decrease until 58 groups. Model Likelihood was relatively stable above 4 groups but convergence was erratic depending on the modelled contrasts. For each model the number of groups was therefore chosen to be the minimum number of groups needed to achieve a level of correspondence approximating that found at 35 groups, under condition of model convergence.

### 4.5 Statistics: Multi-trait GWAS

Traits (columns of Y) were standardized. Along the genome, MTMMs of the type Y= XB+G+E were fitted with initially for each marker trait-specific QTL effects $\beta_1,\ldots,\beta_p$ (contained in B). To identify general QTLs with trait-specific effects, for individual markers, the null hypothesis $\beta_1=\beta_2=...\beta_p=0$ was tested by a Wald test against the alternative hypothesis that at least one of the trait specific effects was nonzero (Zhou and Stephens, 2014). To identify consistent QTLs, the null hypothesis $\beta_1=\beta_2=...\beta_p=\beta\neq 0$ was tested. To identify adaptive QTLs, contrasts defined on the trait specific QTL effects were tested. For example, suppose the first p1 of the full set of p traits represent responses measured under abiotic stresses, while the second p2 traits represent responses under biotic stresses. A contrast can now be defined to test the hypothesis whether the QTL effect for abiotic stresses differs from that for biotic stresses: $\beta_1= \beta_2= ...\beta_{p1}= \alpha_{abiotic}$; $\beta_{(p1+1)}= \beta_{(p1+2)}= \beta_p= \alpha_{biotic}$ and $H_0$: $\alpha_{abiotic}= \alpha_{biotic}$ versus $H_a$: $\alpha_{abiotic}\neq \alpha_{biotic}$. For the Wald test for the hypothesis $\beta_1 = \ldots = \beta_p$ we first fit the MTMM Y = XB + G + E with XB only containing trait specific means $\mu_1, \ldots, \mu_p$, and next test hypotheses on the marker effects. The contrast is defined through a partitioning of the traits in two groups (e.g. resistance against biotic or abiotic stress). Using the R-package asreml (Butler et al., 2009) we perform Wald tests for the following hypotheses:

$H_0 : \beta = 0$, in the constrained model $\beta_1 = \ldots = \beta_p = \beta$.

$H_0 : \alpha_1 = \alpha_2$, in the constrained model where $\alpha_1$ is the effect on all traits in the first group, and $\alpha_2$ for traits in the second group.

### Selecting candidate genes

A significance threshold of $P< 0.0001$ was chosen after implementation of genomic control (see below). For MTMM this resulted in 43 SNPs meeting this criterion. The surrounding region of interest was set to a maximum of 100kb window (50kb on both sides), where the final boundaries were determined by SNPs in LD (threshold 0.4) the furthest away of the significant SNP. This resulted in 30 genome regions, for which in each region the SNP with the lowest P value was selected as representative for the LD block. For Figures 3, S4, S5, S7, S9 and S10, however, we selected the SNP with the highest absolute effect size to maximize visual contrasts

in the figures. Contrast analyses followed the same selection procedure as MTMM.

Correcting for genomic inflation. The Wald test is known to suffer from some inflation (Zhou and Stephens, 2014), which we correct for using genomic control (GC) (Devlin and Roeder, 1999; Devlin *et al.*, 2001), which divides the observed test statistics $T_1$, …, $T_p$ by the genomic inflation factor. For both the unconstrained MTMM and the MTMM for contrasts described above, we observed inflation for small as well as large p-values (i.e. also more p-values close to one than expected). Consequently, the usual genomic control procedures based on the observed versus expected median of test statistics gave too optimistic inflation factors. We therefore applied an alternative genomic control procedure, in which we regress the observed $-\log_{10}(p)$ values on the expected ones, and correct the observed $-\log_{10}(p)$ values for the slope.

## Supplementary Tables

**Table S1.** Data overview on phenotyping the 350 *Arabidopsis thaliana* accessions of the HapMap collection.

| Trait | Section[a] | Min. | Mean | Max. | Variance | $h^2$ [b] | L 95% CI $h^2$ | R 95% CI $h^2$ | NA[c] |
|---|---|---|---|---|---|---|---|---|---|
| Salt_1 | 3.1 | -3.62 | 0.00 | 3.06 | 1.31 | 0.60 | 0.22 | 0.89 | 328 |
| Salt_2 | 3.1 | -1.34 | 0.00 | 1.29 | 0.20 | 0.43 | 0.15 | 0.77 | 323 |
| Salt_3 | 3.1 | -7.70 | -0.01 | 9.73 | 8.91 | 0.64 | 0.27 | 0.89 | 323 |
| Salt_4 | 3.1 | -0.14 | 0.00 | 0.09 | 0.00 | 0.30 | 0.08 | 0.68 | 322 |
| Fungus | 3.10 | -0.58 | 0.01 | 0.58 | 0.05 | 0.40 | 0.13 | 0.74 | 336 |
| Drought&fungus | 3.10 | -0.62 | 0.01 | 0.50 | 0.05 | 0.31 | 0.08 | 0.68 | 336 |
| Caterpillar&fungus | 3.10 | -0.72 | 0.00 | 0.45 | 0.04 | 0.17 | 0.03 | 0.55 | 336 |
| Heat | 3.2 | -6.98 | 0.00 | 5.44 | 6.31 | 0.62 | 0.25 | 0.89 | 275 |
| Osmotic | 3.2 | -4.35 | 0.00 | 10.69 | 3.62 | 0.10 | 0.004 | 0.75 | 346 |
| Drought_1 | 3.2 | -7.05 | 0.00 | 4.64 | 2.69 | 0.39 | 0.12 | 0.75 | 323 |
| Salt_5 | 3.2 | -4.47 | 0.00 | 6.37 | 2.38 | 0.15 | 0.01 | 0.76 | 334 |
| Whitefly_1 | 3.5 | -1.60 | -1.31 | -0.13 | 0.09 | 0.01 | 0.00 | 1.00 | 339 |
| Whitefly_2 | 3.5 | -0.71 | -0.24 | 1.29 | 0.05 | 0.01 | 0.00 | 1.00 | 339 |
| Aphid_1 | 3.6 | -5.11 | 0.00 | 4.14 | 2.38 | 0.10 | 0.004 | 0.76 | 341 |
| Aphid_2 | 3.6 | -4.74 | 0.00 | 4.73 | 1.94 | 0.36 | 0.08 | 0.79 | 341 |
| Aphid_3 | 3.6 | -46.28 | -27.97 | -13.12 | 31.63 | 0.19 | 0.03 | 0.66 | 337 |
| Thrips_1 | 3.7 | -56.51 | -22.29 | -2.95 | 97.60 | 0.80 | 0.37 | 0.96 | 346 |
| Thrips_2 | 3.7 | 0.14 | 0.49 | 0.83 | 0.01 | 0.14 | 0.01 | 0.66 | 347 |
| Thrips_3 | 3.7 | 0.03 | 0.47 | 1.00 | 0.03 | 0.29 | 0.06 | 0.73 | 436 |
| Caterpillar_1 | 3.8 | -4.75 | 0.00 | 6.60 | 2.53 | 0.15 | 0.01 | 0.78 | 328 |
| Caterpillar&osmotic_1 | 3.8 | -3.36 | 0.00 | 6.71 | 2.37 | 0.08 | 0.003 | 0.72 | 326 |
| Caterpillar&osmotic_2 | 3.8 | -0.15 | 0.00 | 0.45 | 0.01 | 0.08 | 0.002 | 0.82 | 324 |
| Drought_2 | 3.9 | -0.12 | 0.00 | 0.21 | 0.00 | 0.06 | 0.002 | 0.66 | 346 |
| Caterpillar_2 | 3.9 | -0.15 | 0.00 | 0.16 | 0.00 | 0.23 | 0.04 | 0.68 | 346 |
| Fungus&caterpillar_1 | 3.9 | -0.16 | 0.00 | 0.19 | 0.00 | 0.20 | 0.03 | 0.64 | 346 |
| Caterpillar_3 | 3.9 | -3.92 | 0.00 | 3.69 | 1.43 | 0.27 | 0.06 | 0.69 | 346 |
| Drought&caterpillar | 3.9 | -5.07 | 0.00 | 3.99 | 1.45 | 0.28 | 0.07 | 0.67 | 346 |
| Fungus&caterpillar_2 | 3.9 | -4.06 | 0.00 | 3.99 | 1.30 | 0.10 | 0.005 | 0.72 | 346 |
| Nematode | 3.3 | -0.50 | -0.30 | -0.15 | 0.00 | 0.72 | 0.35 | 0.93 | 313 |
| Parasitic_plant | 3.4 | -1.65 | 0.01 | 3.23 | 0.56 | 0.03 | 0.00 | 1.00 | 238 |

[a] Section in Supplementary methods where additional information on phenotyping can be found
[b] Narrow sense heritability estimated using the 'heritability' R package
[c] Number of accessions included in the analyses

**Table S2.** Summed effect sizes of 30 most significant SNPs in MTMM per trait.

| Trait | Stress | Summed absolute effect size |
|---|---|---|
| Caterpillar_2 | Single | 3.42 |
| Drought_1 | Single | 3.59 |
| Caterpillar_1 | Single | 3.81 |
| Aphid_2 | Single | 3.99 |
| Salt_1 | Single | 4.13 |
| Drought_2 | Single | 4.25 |
| Whitefly_2 | Single | 4.29 |
| Heat | Single | 4.37 |
| Thrips_3 | Single | 4.42 |
| Whitefly_1 | Single | 4.51 |
| Aphid_1 | Single | 4.54 |
| Fungus and Caterpillar_1 | Double | 4.67 |
| Salt_5 | Single | 4.99 |
| Nematode | Single | 5.09 |
| Parasitic plant | Single | 5.11 |
| Salt_2 | Single | 5.11 |
| Thrips_2 | Single | 5.19 |
| Fungus and Caterpillar_2 | Double | 5.21 |
| Osmotic | Single | 5.30 |
| Aphid_3 | Single | 5.33 |
| Caterpillar_3 | Single | 5.44 |
| Caterpillar and osmotic_2 | Double | 5.69 |
| Thrips_1 | Single | 6.03 |
| Salt_4 | Single | 6.06 |
| Caterpillar and osmotic_1 | Double | 6.17 |
| Salt_3 | Single | 6.77 |
| Drought and Caterpillar | Double | 7.42 |
| Drought and fungus | Double | 10.06 |
| Fungus | Single | 10.09 |
| Caterpillar and fungus | Double | 11.93 |

**Table S3.** 125 candidate genes derived from the Multitrait Mixed Model analysis. Stress-responsive genes are highlighted in gray.

| Significant SNP or gene in LD | Associated marker | Gene | Gene name | Gene description |
|---|---|---|---|---|
| Significant SNP | Ch1: 25500708 | *AT1G68030* | | RING/FYVE/PHD zinc finger superfamily protein |
| Significant SNP | Ch1: 26798534 | *AT1G71040* | Low Phosphate Root2 (*LPR2*) | Encodes LPR2. Function together with LPR1 (AT1G23010) and a P5-type ATPase (At5g23630/PDR2) in a common pathway that adjusts root meristem activity to inorganic phosphate availability |
| Significant SNP | Ch1: 29518622 | *AT1G78460* | | SOUL heme-binding family protein |
| Significant SNP | Ch1: 3294935 | *AT1G10090* | | Early-responsive to dehydration stress protein (ERD4) |
| Significant SNP | Ch1: 7207918 | *AT1G20750* | | RAD3-like DNA-binding helicase protein |
| Significant SNP | Ch2: 11531255 | *AT2G27020* | 20S proteasome alpha subunit G1 (*PAG1*) | Encodes 20S proteasome alpha 7 subunit PAG1 |
| Significant SNP | Ch2: 11659416 | *AT2G27240* | | Aluminium-activated malate transporter family protein |
| Significant SNP | Ch2: 391904 | *AT2G01880* | | Purple acid phosphatase 7 (PAP7) |
| Significant SNP | Ch3: 1077306 | *AT3G04110* | glutamate receptor 1.1 (*GLR1.1*) | Putative glutamate receptor (GLR1.1). Contains a functional cation - permeable pore domain. Involved in cellular cation homeostasis. |
| Significant SNP | Ch3: 18615891 | *AT3G50210* | | 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein |
| Significant SNP | Ch3: 19804402 | *AT3G53420* | plasma membrane intrinsic protein 2A (*PIP2A*) | Member of the plasma membrane intrinsic protein subfamily PIP2. Localizes to the plasma membrane and exhibits water transport activity in *Xenopus* oocyte. Expressed specifically in the vascular bundles and protein level increases slightly during leaf development. When expressed in yeast cells can conduct hydrogen peroxide into those cells |
| Significant SNP | Ch3: 21625003 | *AT3G58460* | | RHOMBOID-like protein 15 (RBL15) |
| Significant SNP | Ch3: 2231603 | *AT3G07050* | | GTP-binding family protein |
| Significant SNP | Ch3: 6968031 | *AT3G20000* | translocase of the outer mitochondrial membrane 40 (*TOM40*) | Encodes a component of the TOM receptor complex responsible for the recognition and translocation of cytosolically synthesized mitochondrial preproteins. With TOM22, functions as the transit peptide receptor at the surface of the mitochondrial outer membrane and facilitates the movement of preproteins into the translocation pore. |

**Table S3.** (continued)

| Significant SNP or gene in LD | Associated marker | Gene | Gene name | Gene description |
|---|---|---|---|---|
| Significant SNP | Ch3: 8014458 | *AT3G22640* | | PAP85 |
| Significant SNP | Ch4: 5180340 | *AT4G08200* | | Similar to unknown protein [*Arabidopsis thaliana*] (TAIR:AT1G43722.1) |
| Significant SNP | Ch4: 6805259 | *AT4G11160* | | Translation initiation factor 2, small GTP-binding protein |
| Significant SNP | Ch4: 8654778 | *AT4G15180* | | SET domain protein 2 (SDG2) |
| Significant SNP | Ch4: 9350941 | *AT4G16600* | | Nucleotide-diphospho-sugar transferases superfamily protein |
| Significant SNP | Ch4:13265656 | *AT4G26190* | | Haloacid dehalogenase-like hydrolase (HAD) superfamily protein |
| Significant SNP | Ch4:13955847 | *AT4G28080* | | Tetratricopeptide repeat (TPR)-like superfamily protein |
| Significant SNP | Ch4:16420532 | *AT4G34320* | | Protein of unknown function (DUF677) |
| Significant SNP | Ch5: 22041081 | *AT5G54280* | myosin 2 (*ATM2*) | Type VII myosin gene |
| Significant SNP | Ch5: 22677563 | *AT5G56000* | | HEAT SHOCK PROTEIN 81.4 (Hsp81.4) |
| Significant SNP | Ch5: 22842831 | *AT5G56390* | | F-box/RNI-like/FBD-like domains-containing protein |
| Significant SNP | Ch5: 23302987 | *AT5G57535* | | unknown protein |
| Significant SNP | Ch5: 414050 | *AT5G02100* | Unfertilized embryo sac 18 (*UNE18*) | Encodes a protein that binds to beta-sitosterol and localizes to the ER. The WFDE motif in ORP3a appears to be important for a direct interaction with PVA12 [Plant VAMP-Associated protein 12]. Mutation of this motif causes ORP3a to relocalize to the Golgi and cytosol. The interaction between PVA12 and ORP3a does not appear to be sterol-dependent |
| Significant SNP | Ch5: 7493620 | *AT5G22560* | | Plant protein of unknown function (DUF247) |
| Significant SNP | Ch5: 7493623 | *AT5G23480* | | SWIB/MDM2 domain |
| Significant SNP | Ch5: 9154579 | *AT5G26190* | | Cysteine/Histidine-rich C1 domain family protein |
| in_LD_with | Ch1: 25500708 | *AT1G67990* | *TSM1* | Encodes a tapetum-specific O-methyltransferase. In vitro enzyme assay indicated activity with caffeoyl-CoA, caffeoyl glucose, chlorogenic acid and polyamine conjugates. RNAi mutants had impaired silique development and seed setting. |

**Table S3.** (continued)

| Significant SNP or gene in LD | Associated marker | Gene | Gene name | Gene description |
|---|---|---|---|---|
| in_LD_with | Ch1: 25500708 | *AT1G68010* | hydroxypyruvate reductase (*HPR*) | Encodes hydroxypyruvate reductase. |
| in_LD_with | Ch1: 25500708 | *AT1G67980* | caffeoyl-CoA 3-O- methyltransferase (*CCOAMT*) | Encodes S-adenosyl-L-methionine: transcaffeoyl Coenzyme A 3-O-methyltransferase. |
| in_LD_with | Ch1: 25500708 | *AT1G67960* | | CONTAINS InterPro DOMAIN/s: Membrane protein,Tapt1/CMV receptor (InterPro:IPR008010) |
| in_LD_with | Ch1: 25500708 | *AT1G68000* | phosphatidylinositol synthase 1 (*PIS1*) | phosphatidylinositol synthase 1 |
| in_LD_with | Ch1: 25500708 | *AT1G68020* | *ATTPS6* | Encodes an enzyme putatively involved in trehalose biosynthesis. The protein has a trehalose synthase (TPS)-like domain and a trehalose phosphatase (TPP)-like domain. It can complement a yeast mutant lacking both of these activities suggesting that this is a bifunctional enzyme. |
| in_LD_with | Ch1: 25500708 | *AT1G67970* | heat shock transcription factor A8 (*HSFA8*) | member of Heat Stress Transcription Factor (Hsf) family |
| in_lD_with | Ch1: 29518622 | *AT1G78440* | *Arabidopsis thaliana* gibberellin 2-oxidase 1 (*ATGA2OX1*) | Encodes a gibberellin 2-oxidase that acts on C19 gibberellins. |
| in_lD_with | Ch1: 29518622 | *AT1G78430* | | ROP interactive partner 2 (RIP2) |
| in_lD_with | Ch1: 29518622 | *AT1G78450* | | SOUL heme-binding family protein |
| in_lD_with | Ch1: 29518622 | *AT1G78470* | | BEST *Arabidopsis thaliana* protein match is: F-box family protein (TAIR:AT1G67390.1) |
| in_LD_with | Ch1: 7207918 | *AT1G20740* | | Protein of unknown function (DUF833) |
| in_LD_with | Ch1: 7207918 | *AT1G20760* | | Calcium-binding EF hand family protein |
| in_LD_with | Ch1: 7207918 | *AT1G20780* | senescence-associated E3 ubiquitin ligase 1 (*SAUL1*) | Encodes a protein containing a U-box and an ARM domain. |
| in_LD_with | Ch1: 7207918 | *AT1G20790* | | F-box family protein |
| in_LD_with | Ch1: 7207918 | *AT1G20770* | | Unknown protein |

**Table S3.** (continued)

| Significant SNP or gene in LD | Associated marker | Gene | Gene name | Gene description |
|---|---|---|---|---|
| in_LD_with | Ch2: 11659416 | *AT2G27250* | *AtCLV3* | One of the three CLAVATA genes controlling the size of the shoot apical meristem (SAM) in Arabidopsis. Belongs to a large gene family called CLE for CLAVATA3/ESR-related. Encodes a stem cell-specific protein CLV3 presumed to be a precursor of a secreted peptide hormone. The deduced ORF encodes a 96-amino acid protein with an 18-amino acid N-terminal signal peptide. The functional form of CLV3 (MCLV3) was first reported to be a posttranscriptionally modified 12-amino acid peptide, in which two of the three prolines were modified to hydroxyproline |
| in_LD_with | Ch3: 19804402 | *AT3G53400* | | BEST *Arabidopsis thaliana* protein match is: conserved peptide upstream open reading frame 47 (TAIR:AT5G03190.1) |
| in_LD_with | Ch3: 21625003 | *AT3G53410* | | RING/U-box superfamily protein |
| in_LD_with | Ch3: 21625003 | *AT3G58490* | | Phosphatidic acid phosphatase (PAP2) family protein |
| in_LD_with | Ch3: 21625003 | *AT3G58450* | | Adenine nucleotide alpha hydrolases-like superfamily protein |
| in_LD_with | Ch3: 21625003 | *AT3G58510* | | DEA(D/H)-box RNA helicase family protein |
| in_LD_with | Ch3: 21625003 | *AT3G58440* | | TRAF-like superfamily protein |
| in_LD_with | Ch3: 21625003 | *AT3G58520* | | Ubiquitin carboxyl-terminal hydrolase family protein |
| in_LD_with | Ch3: 21625003 | *AT3G58480* | | Calmodulin-binding family protein |
| in_LD_with | Ch3: 21625003 | *AT3G58470* | | Nucleic acid binding |
| in_LD_with | Ch3: 21625003 | *AT3G58500* | | Encodes one of the isoforms of the catalytic subunit of protein phosphatase 2A: AT1G59830/PP2A-1, AT1G10430/PP2A-2, At2g42500/PP2A-3, At3g58500/PP2A-4 [Plant Molecular Biology (1993) 21:475-485 and (1994) 26:523-528 |
| in_LD_with | Ch3: 6968031 | *AT3G20010* | | SNF2 domain-containing protein / helicase domain-containing protein / zinc finger protein-related |
| in_LD_with | Ch3: 6968031 | *AT3G19990* | | Unknown protein |
| in_LD_with | Ch3: 6968031 | *AT3G19980* | | Encodes catalytic subunit of serine/threonine protein phosphatase 2A. It can associate with phytochromes A and B in vitro. Mutant plants display an accelerated flowering phenotype. |

**Table S3.** (continued)

| Significant SNP or gene in LD | Associated marker | Gene | Gene name | Gene description |
|---|---|---|---|---|
| in_LD_with | Ch3: 8014458 | *AT3G22670* | | Pentatricopeptide repeat (PPR) superfamily protein |
| in_LD_with | Ch3: 8014458 | *AT3G22680* | RNA-directed DNA methylation 1 (RDM1) | Encodes RNA-DIRECTED DNA METHYLATION 1 (RDM1), forming a complex with DMS3 (AT3G49250) and DRD1 (AT2G16390). This complex is termed DDR. The DDR complex is required for polymerase V transcripts and RNA-directed DNA methylation. |
| in_LD_with | Ch3: 8014458 | *AT3G22650* | | CEGENDUO (CEG) |
| in_LD_with | Ch3: 8014458 | *AT3G22690* | | Involved in: photosystem II assembly, regulation of chlorophyll biosynthetic process, photosystem I assembly, thylakoid membrane organization, RNA modification |
| in_LD_with | Ch3: 8014458 | *AT3G22700* | | F-box and associated interaction domains-containing protein |
| in_LD_with | Ch3: 8014458 | *AT3G22710* | | F-box family protein |
| in_LD_with | Ch3: 8014458 | *AT3G22720* | | F-box and associated interaction domains-containing protein |
| in_LD_with | Ch3: 8014458 | *AT3G22730* | | F-box and associated interaction domains-containing protein |
| in_LD_with | Ch3: 8014458 | *AT3G22740* | homocysteine S- methyltransferase 3 (*HMT3*) | Homocysteine S-methyltransferase (HMT3) |
| in_LD_with | Ch3: 8014458 | *AT3G22750* | | Protein kinase superfamily protein |
| in_LD_with | Ch3: 8014458 | *AT3G22760* | *SOL1* | CXC domain containing TSO1-like protein 1. The gene is expressed in stamens, pollen mother cells, and immature ovules. |
| in_LD_with | Ch4: 5180340 | *AT4G08190* | | P-loop containing nucleoside triphosphate hydrolases superfamily protein |
| in_LD_with | Ch4: 5180340 | *AT4G08180* | | OSBP(oxysterol binding protein)-related protein 1C (ORP1C) |
| in_LD_with | Ch4: 5180340 | *AT4G08230* | | Glycine-rich protein |
| in_LD_with | Ch4: 5180340 | *AT4G08210* | | Pentatricopeptide repeat (PPR-like) superfamily protein |
| in_LD_with | Ch4: 5180340 | *AT4G08220* | | Mutator-like transposase family, has a $5.3*10^{-67}$ P-value blast match to Q9SUF8 /145-308 Pfam PF03108 MuDR family transposase (MuDr-element domain) |

**Table S3.** (continued)

| Significant SNP or gene in LD | Associated marker | Gene | Gene name | Gene description |
|---|---|---|---|---|
| in_LD_with | Ch4: 6805259 | *AT4G11140* | cytokinin response factor 1 (*CRF1*) | Encodes a member of the ERF (ethylene response factor) subfamily B-5 of the ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 7 members in this subfamily. Also named as CRF1 (cytokinin response factor 1). |
| in_LD_with | Ch4: 6805259 | *AT4G11150* | vacuolar ATP synthase subunit E1 (*TUF*) | Encodes a vacuolar $H^+$-ATPase subunit E isoform 1 which is required for Golgi organization and vacuole function in embryogenesis. |
| in_LD_with | Ch4: 6805259 | *AT4G11170* | *RMG1* | Disease resistance protein (TIR-NB-LRR class) family |
| in_LD_with | Ch4: 8654778 | *AT4G15210* | *Arabidopsis thaliana* BETA-AMYLASE (*ATBETA-AMY*) | Cytosolic beta-amylase expressed in rosette leaves and inducible by sugar. RAM1 mutants have reduced beta amylase in leaves and stems. |
| in_LD_with | Ch4:13265656 | *AT4G26180* | | Mitochondrial substrate carrier family protein |
| in_LD_with | Ch4:13265656 | *AT4G26150* | cytokinin-responsive gata factor 1 (*CGA1*) | Encodes a member of the GATA factor family of zinc finger transcription factors. |
| in_LD_with | Ch4:13265656 | *AT4G26170* | | Molecular_function unknown |
| in_LD_with | Ch4:13265656 | *AT4G26220* | | S-adenosyl-L-methionine-dependent methyltransferases superfamily protein |
| in_LD_with | Ch4:13265656 | *AT4G26140* | beta-galactosidase 12 (*BGAL12*) | Putative beta-galactosidase |
| in_LD_with | Ch4:13265656 | *AT4G26160* | atypical CYS HIS rich thioredoxin 1 (*ACHT1*) | Encodes a member of the thioredoxin family protein. Located in the chloroplast. Shows high activity towards the chloroplast 2-Cys peroxiredoxin A, and poor activity towards the chloroplast NADP-malate dehydrogenase |
| in_LD_with | Ch4:13265656 | *AT4G26210* | | Mitochondrial ATP synthase subunit G protein |
| in_LD_with | Ch4:13955847 | *AT4G26200* | 1-amino-cyclopropane-1-carboxylate synthase 7 (*ACS7*) | Member of a family of proteins in Arabidopsis that encode 1-Amino-cyclopropane-1-carboxylate synthase, an enzyme involved in ethylene biosynthesis. Not expressed in response to IAA |
| in_LD_with | Ch4:13955847 | *AT4G28100* | | Unknown protein |
| in_LD_with | Ch4:13955847 | *AT4G28060* | | Cytochrome c oxidase, subunit Vib family protein |
| in_LD_with | Ch4:13955847 | *AT4G28070* | | AFG1-like ATPase family protein |
| in_LD_with | Ch4:13955847 | *AT4G28090* | | SKU5 similar 10 (sks10) |
| in_LD_with | Ch4:13955847 | *AT4G28085* | | Unknown protein |

**Table S3.** (continued)

| Significant SNP or gene in LD | Associated marker | Gene | Gene name | Gene description |
|---|---|---|---|---|
| in_LD_with | Ch4:13955847 | *AT4G28088* | | Low temperature and salt responsive protein family |
| in_LD_with | Ch4:13955847 | *AT4G34310* | | alpha/beta-Hydrolases superfamily protein |
| in_LD_with | Ch5: 22041081 | *AT5G54250* | cyclic nucleotide-gated cation channel 4 (*CNGC4*) | Member of Cyclic nucleotide gated channel family, downstream component of the signaling pathways leading to HR resistance. Mutant plants exhibit gene-for-gene disease resistance against avirulent *Pseudomonas syringae* despite the near-complete absence of the hypersensitive response (HR). Salicylic acid accumulation in dnd2 mutants is completely PAD4-independent. |
| in_LD_with | Ch5: 22041081 | *AT5G54260* | Meiotic recombination 11 (*MRE11*) | DNA repair and meiotic recombination protein, component of MRE11 complex with RAD50 and NBS1 |
| in_LD_with | Ch5: 22041081 | *AT5G54270* | light-harvesting chlorophyll B-binding protein 3 (*LHCB3*) | Lhcb3 protein is a component of the main light harvesting chlorophyll a/b-protein complex of Photosystem II (LHC II). |
| in_LD_with | Ch5: 22041081 | *AT5G54240* | | Protein of unknown function (DUF1223) |
| in_LD_with | Ch5: 22677563 | *AT5G55990* | calcineurin B-like protein 2 (*CBL2*) | Encodes a member of the Arabidopsis CBL (Calcineurin B-like Calcium Sensor) protein family |
| in_LD_with | Ch5: 22677563 | *AT5G55980* | | Serine-rich protein-related |
| in_LD_with | Ch5: 22677563 | *AT5G55970* | | RING/U-box superfamily protein |
| in_LD_with | Ch5: 22842831 | *AT5G56380* | | F-box/RNI-like/FBD-like domains-containing protein |
| in_LD_with | Ch5: 22842831 | *AT5G56370* | | F-box/RNI-like/FBD-like domains-containing protein |
| in_LD_with | Ch5: 22842831 | *AT5G56368* | | Encodes a defensin-like (DEFL) family protein. |
| in_LD_with | Ch5: 23302987 | *AT5G57520* | zinc finger protein 2 (*ZFP2*) | Encodes a zinc finger protein containing only a single zinc finger. |
| in_LD_with | Ch5: 23302987 | *AT5G57560* | Touch 4 (*TCH4*) | Encodes a cell wall-modifying enzyme, rapidly upregulated in response to environmental stimuli |

**Table S3.** (continued)

| Significant SNP or gene in LD | Associated marker | Gene | Gene name | Gene description |
|---|---|---|---|---|
| in_LD_with | Ch5: 23302987 | *AT5G57490* | voltage dependent anion channel 4 (*VDAC4*) | Encodes a voltage-dependent anion channel (VDAC: AT3G01280/VDAC1, AT5G67500/VDAC2, AT5G15090/VDAC3, AT5G57490/VDAC4, AT5G15090/VDAC5). VDACs are reported to be porin-type, beta-barrel diffusion pores. They are prominently localized in the outer mitochondrial membrane and are involved in metabolite exchange between the organelle and the cytosol. |
| in_LD_with | Ch5: 23302987 | *AT5G57565* | | Protein kinase superfamily protein |
| in_LD_with | Ch5: 23302987 | *AT5G57540* | | Encodes a xyloglucan endotransglucosylase/hydrolase with only only the endotransglucosylase (XET |
| in_LD_with | Ch5: 23302987 | *AT5G57550* | xyloglucan endotrans/glucosylase/hydrolase 25 (*XTH25*) | Xyloglucan endotransglycosylase-related protein (XTR3) |
| in_LD_with | Ch5: 23302987 | *AT5G57500* | | Galactosyltransferase family protein |
| in_LD_with | Ch5: 23302987 | *AT5G57530* | | Xyloglucan endotransglucosylase/hydrolase 12 (XTH12) |
| in_LD_with | Ch5: 23302987 | *AT5G57510* | | Unknown protein |
| in_LD_with | Ch5: 23302987 | *AT5G57570* | | GCK domain-containing protein |
| in_LD_with | Ch5: 23302987 | *AT5G57590* | biotin auxotroph 1 (*BIO1*) | Mutant complemented by E coli Bio A gene encoding 7,8-diaminopelargonic acid aminotransferase. |
| in_LD_with | Ch5: 23302987 | *AT5G57580* | | Calmodulin-binding protein |
| in_LD_with | Ch5: 414050 | *AT5G02110* | | CYCLIN D7 |
| in_LD_with | Ch5: 7493620 | *AT5G22550* | | Plant protein of unknown function (DUF247) |
| in_LD_with | Ch5: 7493620 | *AT5G22570* | *WRKY38* | member of WRKY Transcription Factor |
| in_LD_with | Ch5: 7493620 | *AT5G22545* | | Unknown protein |
| in_LD_with | Ch5: 7493620 | *AT5G22555* | | Unknown protein |
| in_LD_with | Ch5: 7493623 | *AT5G23510* | | Unknown protein |
| in_LD_with | Ch5: 7493623 | *AT5G23490* | | Unknown protein |

**Table S4.** Genes in linkage with SNPs with –log10(P) score above 4 (20 kb half-window size) in the contrast-specific GWA mapping of parasitic plants and aphids on the one hand versus fungus, caterpillar, thrips and drought on the other hand.

| Marker | Gene in LD | Gene name | Gene description | Reference |
|---|---|---|---|---|
| chr1.19711816 | *AT1G52900* | - | Toll-Interleukin-Resistance (TIR) domain family protein, signal transduction, defense response | (Cartieaux *et al.*, 2008) |
| chr1.24785939 | *AT1G66410* | *CAM4* | Calmodulin 4, calcium-binding EF-hand site, calcium-mediated signalling | (Zhao *et al.*, 2013) |
| chr3.672138 | *AT3G02940* | *MYB107* | Transcription factor, responsive to salicylic acid | (Stracke *et al.*, 2001) |
| chr3.7945317 | *AT3G22400* | *LOX5* | Oxidoreductase activity (9-LOX pathway), facilitates *M. persicae* aphid feeding | (Nalam *et al.*, 2012a; Nalam *et al.*, 2012b) |
| chr3.23145919 | *AT3G62610* | *MYB11* | Transcription factor, involved in production of flavonol glycosides | (Stracke *et al.*, 2007) |
| chr4.9390514 | *AT4G16730, AT4G16740,* | *TPS02, TPS03* | Terpene synthases, (*E,E*)-alpha-farnesene synthase | (Huang *et al.*, 2010) |
| | *AT4G16690* | *MES16* | Methyl jasmonate esterase | (Christ *et al.*, 2012) |
| chr5.22829754 | *AT5G56360* | *PSL4* | Calmodulin binding protein, involved in MAMP-triggered defense to bacteria | (Lu *et al.*, 2009) |

**Table S5.** Candidate genes in linkage with SNPs with $-\log_{10}(P)$ score above 4 (20 kb half-window size) that have common effects on plant response to parasitic plants and aphids on the one hand versus fungus, caterpillar, thrips and drought on the other hand.

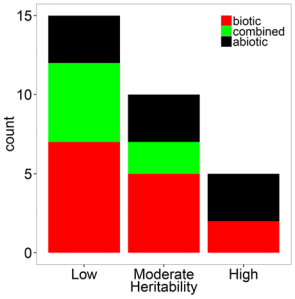
| Marker | Gene | Gene name | Description |
|---|---|---|---|
| chr2.15762021 | *AT2G37570* | *SLT1* | Encodes a protein that can complement the salt-sensitive phenotype of a calcineurin (CaN)-deficient yeast mutant. |
| in_LD_with_chr2.15762021 | *AT2G37630* | *MYB91* | Encodes a MYB-domain protein involved in specification of the leaf proximodistal axis. Also functions as a regulator of the plant immune response. |
| in_LD_with_chr3.22345759 | *AT3G60490* | - | Encodes a member of the DREB subfamily A-4 of ERF/AP2 transcription factor family. Pathogenesis-related. |
| chr4.9598560 | *AT4G17070* | - | Encodes a peptidyl-prolyl cis-trans isomerase. Involved in response to oxidative stress. |

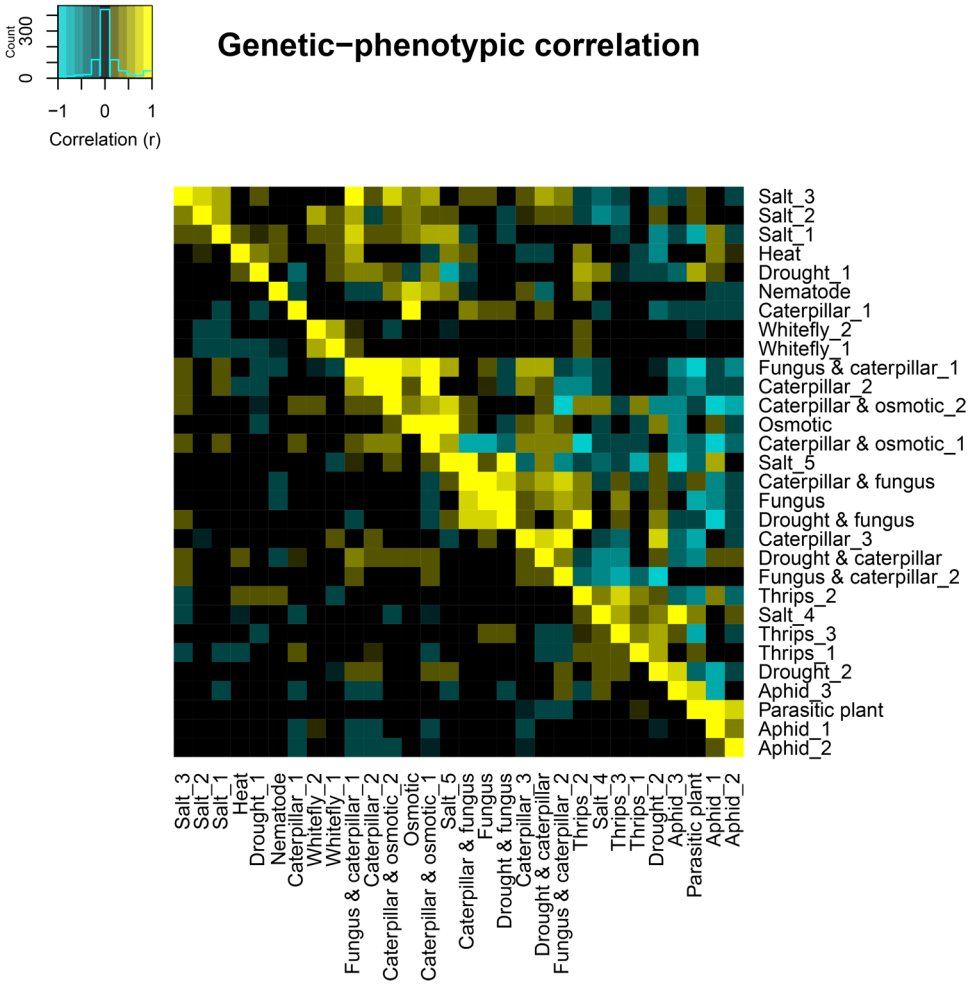
**Table S6.** Candidate genes in linkage with SNPs with $-\log_{10}(P)$ score above 4 (20 kb half-window size) that have common effects on biotic and abiotic stress responses.

| Marker | Gene | Gene name | Description |
|---|---|---|---|
| in_LD_with_ chr4.5651749 | *AT4G08870* | *ARGAH2* | Encodes one of the two arginases in the genome. Gene expression is enhanced by methyl jasmonate treatment. It is involved in the defence response to *B. cinerea*. |
| chr4. 8057710 | *AT4G13940* | *AtSAHH1* | Encodes an S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent gene silencing. |
| chr2.856085 | *AT2G02950* | *PKS1* | Encodes a basic soluble protein which can independently bind to either PHYA or PHYB, regardless of whether the phytochromes are in the Pr or Pfr state. PKS1 can be phosphorylated by oat phyA *in vitro* in a light-regulated manner. It is postulated to be a negative regulator of phyB signalling. |

## Supplementary Figures

**Figure S1.** Narrow sense heritability for *Arabidopsis thaliana* resistance to abiotic and biotic stresses. Narrow sense heritability values were estimated using the 'heritability' R package. Traits were classified in three biological categories: resistance to abiotic, biotic and double stresses. These biological categories were grouped based on their heritability in low ($h^2$<0.2), moderate (0.2<$h^2$<0.5) and high ($h^2$>0.5) heritability classes.



◄ **Figure S2.** Genetic and phenotypic correlation matrix. Heatmap displaying phenotypic correlations below the diagonal and genetic correlations above the diagonal. Phenotypic correlations were calcluated using Spearman's correlation coefficient rho, whereas the genome-wide genetic correlations were estimated bivariately and with correction for population structure (on full kinship matrix). For Whitefly_1 and Whitefly_2 the maximum likelihood estimates were not available so genetic correlations were estimated using G-BLUP. Traits were clustered according to Ward's minimum variance method for the genetic correlation coefficient values.

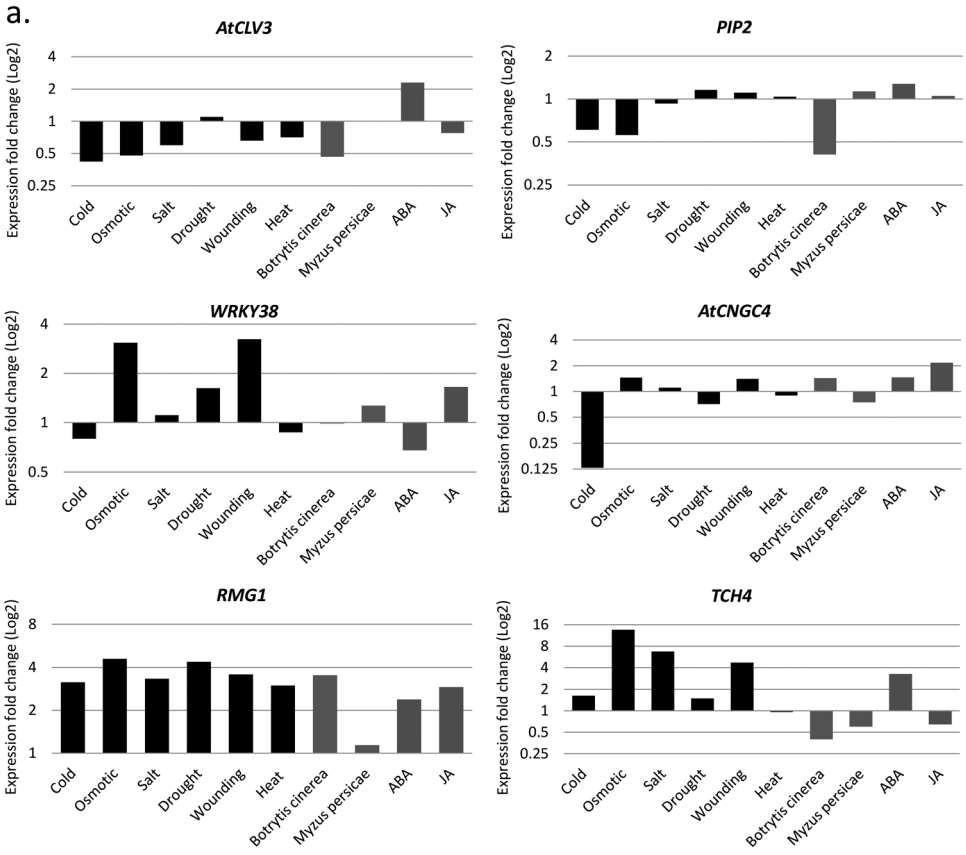



**Figure S3.** Expression data of 6 candidate genes (resulting from MTMM, see Table 2a) in plants exposed to biotic or abiotic stress factors, relative to control conditions. **a.** Shoot tissues and **b.** root tissues (next page). Expression data from Arabidopsis eFP browser (http://bbc.botany.utoronto.ca).

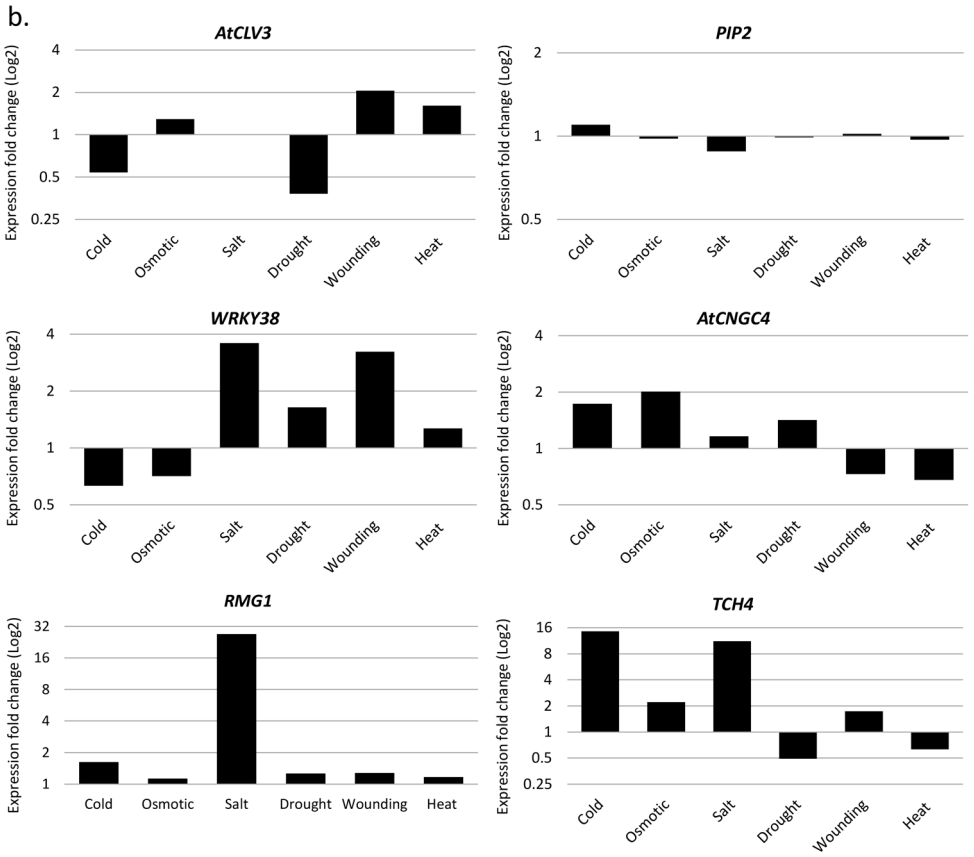



**Figure S3b** See legend previous page.

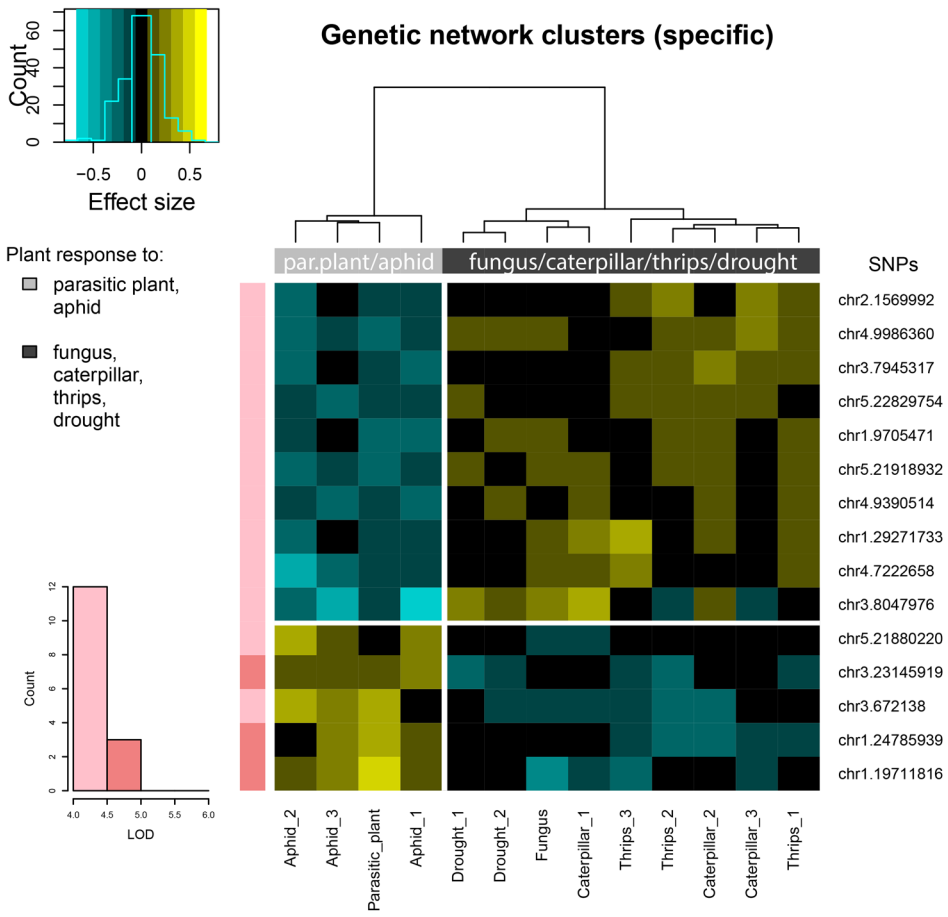



**Figure S4.** Genetic associations specific for plant responses to the main clusters of the genetic correlation network (see Figure 1): parasitic plant and aphid versus fungus, caterpillar, thrips and drought. Genetic associations were estimated with a contrast-specific analysis using MTMM. SNPs with a significance above LOD score 4 ($P \leq 10^{-4}$) for the contrast are clustered according to trait-specific effect size estimated from the full MTMM. If there was another SNP in LD that had a higher effect size, this SNP was used as representative for the LD block. Negative effect sizes (blue) were cases where the rare allele was associated with a detrimental effect on the plants, positive effect sizes (yellow) were cases where the rare allele was associated with increased resistance to the stress. The rare alleles of the top 10 SNPs are associated with enhanced resistance to fungus, caterpillar, thrips and drought stresses and reduced resistance to stresses inflicted by parasitic plants and aphids; the bottom 5 SNPs show the inverse. Stresses are clustered according to effect size, using Ward's minimum variance method. If SNPs were located within a 20 kb half-window of each other, only the SNP with the highest absolute cumulative effect size was included. The key shows the frequency distribution of SNPs across effect sizes.

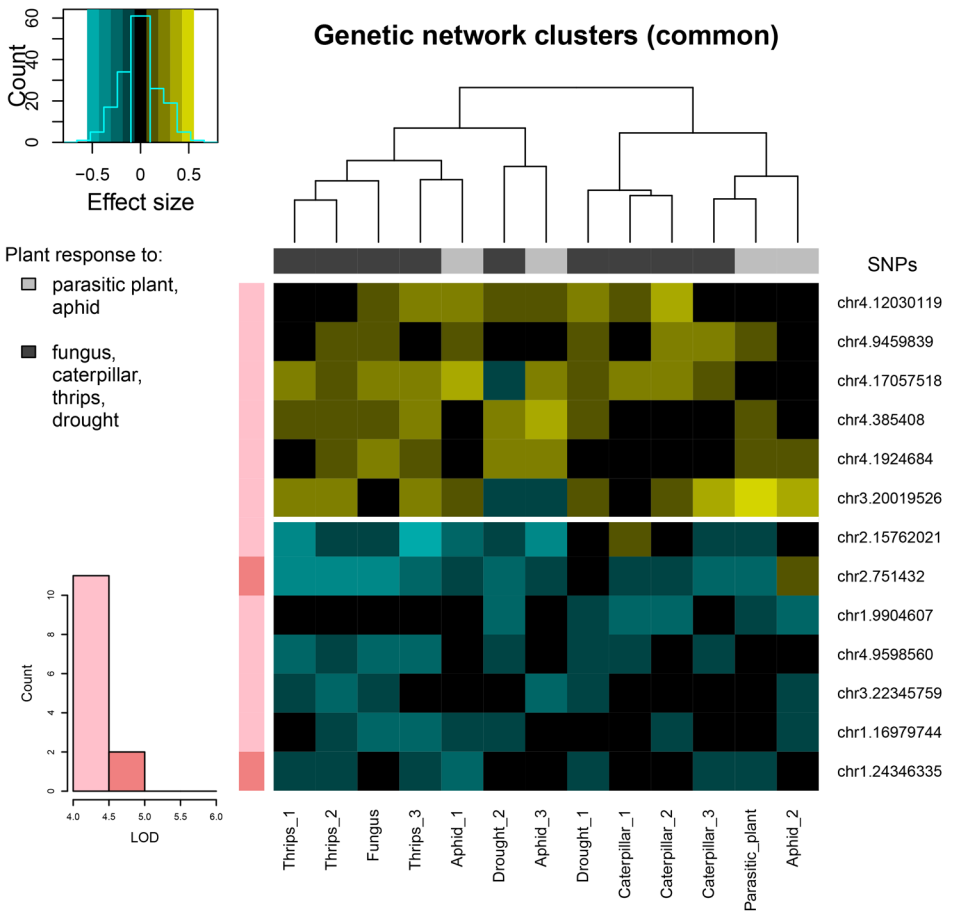



**Figure S5.** Genetic associations common for plant response to the main clusters of the genetic correlation network: parasitic plant and aphid on the one hand versus fungus, caterpillar, thrips and drought on the other hand. Genetic associations were estimated with a contrast analysis using MTMM. SNPs with a significance above LOD score 4 ($P \leq 10^{-4}$) for the common response are clustered according to trait-specific effect size estimated from the full MTMM. If there was another SNP in LD that had a higher effect size, this SNP was used as representative for the LD block. Negative effect sizes (blue) were cases where the rare allele was associated with a detrimental effect on the plants, positive effect sizes (yellow) were cases where the rare allele was associated with increased resistance to the stress. The rare alleles of the top 6 SNPs are associated with enhanced resistance to abiotic stresses and reduced resistance to biotic stresses; the bottom 7 SNPs show the inverse. Stresses are clustered according to effect size, using Ward's minimum variance method. If SNPs were located within a 20 kb half-window of each other, only the SNP with the highest absolute cumulative effect size was included. The key shows the frequency distribution of SNPs across effect sizes.

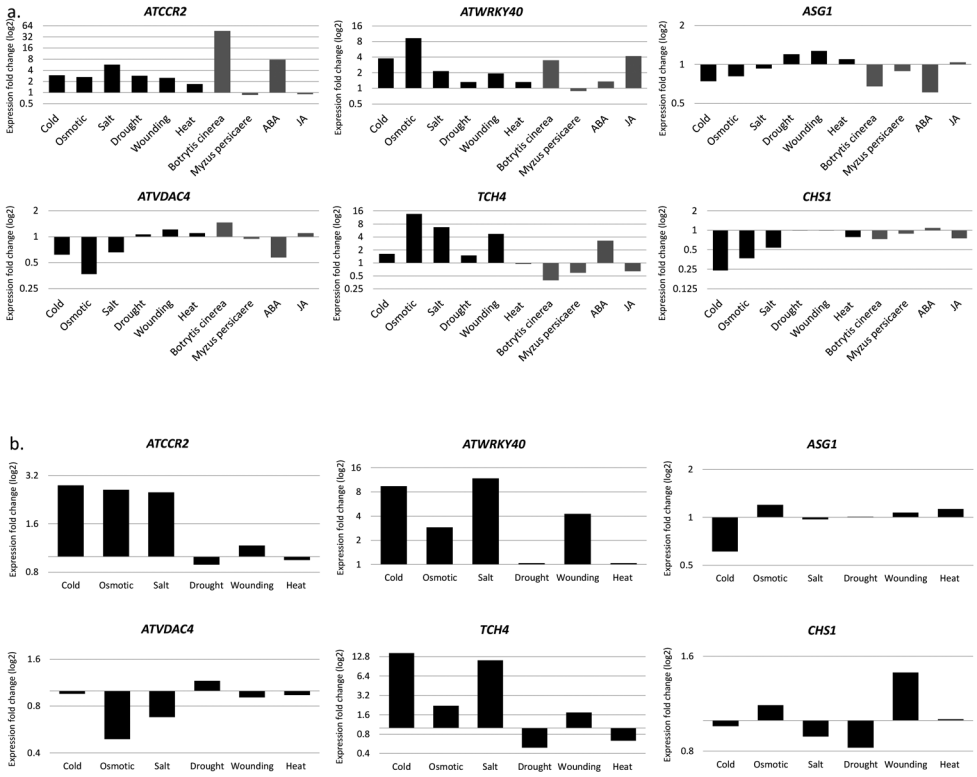



**Figure S6.** Expression data of 6 candidate genes (resulting from MTMM analysis, see Table 2b) in plants exposed to biotic or abiotic stress factors, relative to control conditions. **a.** Shoot tissues and **b.** root tissues. Expression data from Arabidopsis eFP browser (http://bbc.botany.utoronto.ca).

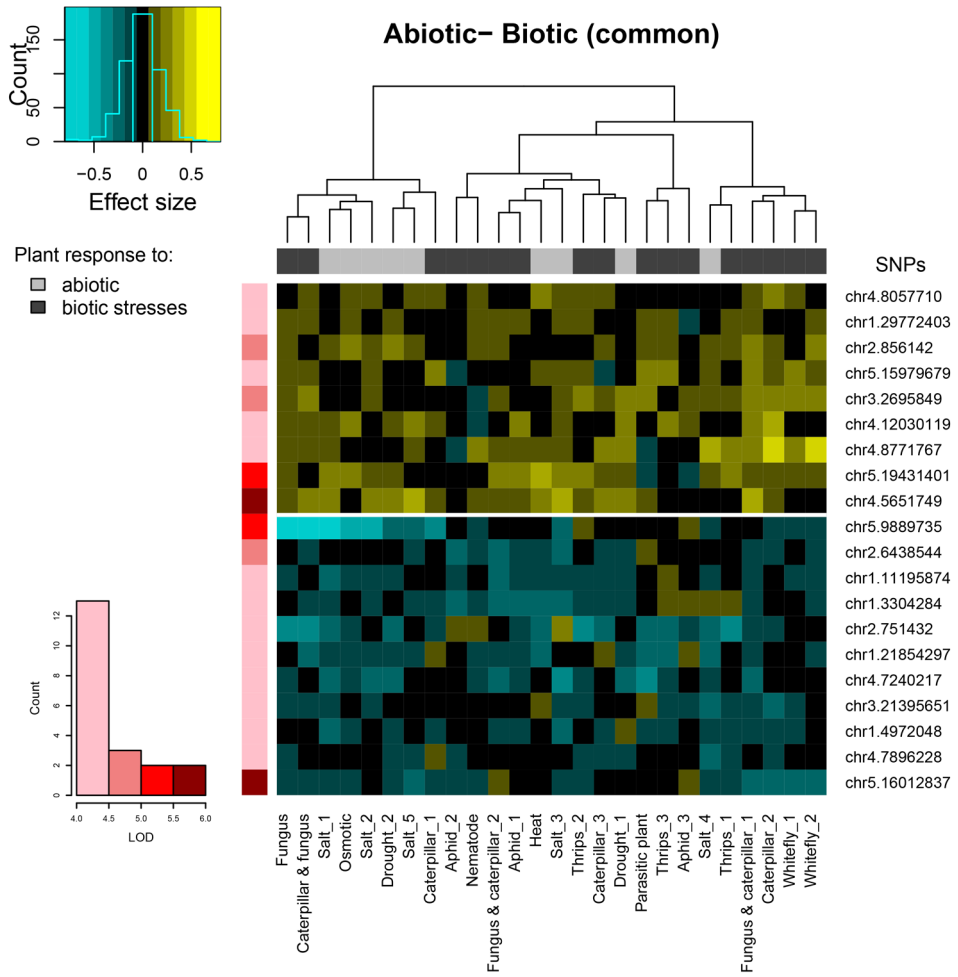



**Figure S7.** Genetic associations common for plant responses to abiotic and biotic stresses. Genetic associations were estimated with a contrast analysis using MTMM. SNPs with a significance above LOD score 4 ($P \leq 10^{-4}$) for the common response are clustered according to trait-specific effect size estimated from the full MTMM. If there was another SNP in LD that had a higher effect size, this SNP was used as representative for the LD block. Negative effect sizes (blue) were cases where the rare allele was associated with a detrimental effect on the plants, positive effect sizes (yellow) were cases where the rare allele was associated with increased resistance to the stress. The rare alleles of the top 9 SNPs are associated with enhanced resistance to abiotic and biotic stresses; the bottom 11 SNPs are associated with reduced resistance to abiotic and biotic stresses. Stresses are clustered according to effect size, using Ward's minimum variance method. If SNPs were located within a 20 kb half-window of each other, only the SNP with the highest absolute cumulative effect size was included. The key shows the frequency distribution of SNPs across effect sizes.

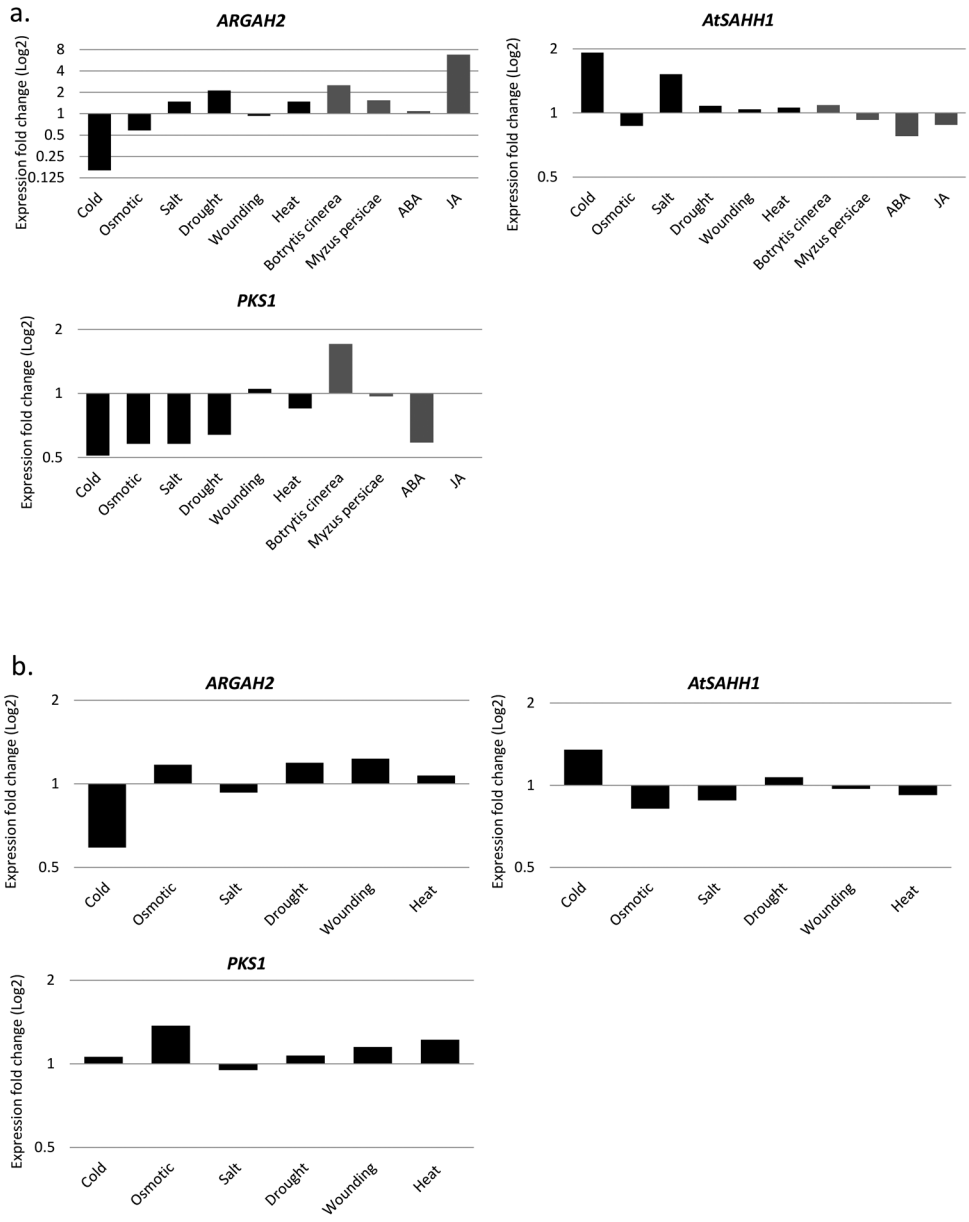



**Figure S8.** Expression data of 3 candidate genes (resulting from MTMM, see Supplementary Table S5) in plants exposed to biotic or abiotic stress factors, relative to control conditions. **a.** Shoot tissues and **b.** root tissues. Expression data from Arabidopsis eFP browser (http://bbc.botany.utoronto.ca).

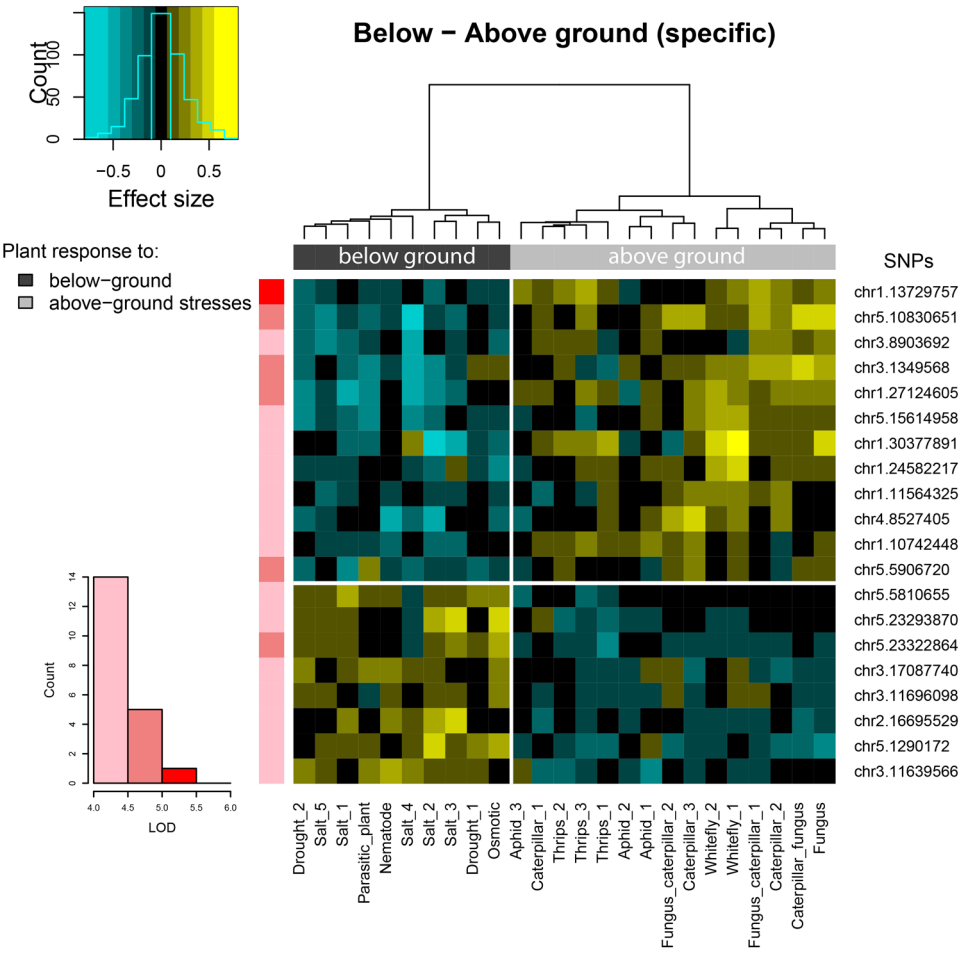



**Figure S9.** Genetic associations specific for plant responses to either below- or aboveground stress. Genetic associations were estimated with a contrast analysis using MTMM. SNPs with a significance above LOD score 4 ($P \leq 10^{-4}$) for the belowground-aboveground contrast are clustered according to trait-specific effect size estimated from the full MTMM. If there was another SNP in LD that had a higher effect size, this SNP was used as representative for the LD block. Negative effect sizes (blue) were cases where the rare allele was associated with a detrimental effect on the plants, positive effect sizes (yellow) were cases where the rare allele was associated with increased resistance to the stress. The rare alleles of the top 12 SNPs are associated with enhanced resistance to aboveground stresses and reduced resistance to belowground stresses; the bottom 8 SNPs show the inverse. Stresses are clustered according to effect size, using Ward's minimum variance method. If SNPs were located within a 20 kb half-window of each other, only the SNP with the highest absolute cumulative effect size was included. The key shows the frequency distribution of SNPs across effect sizes.

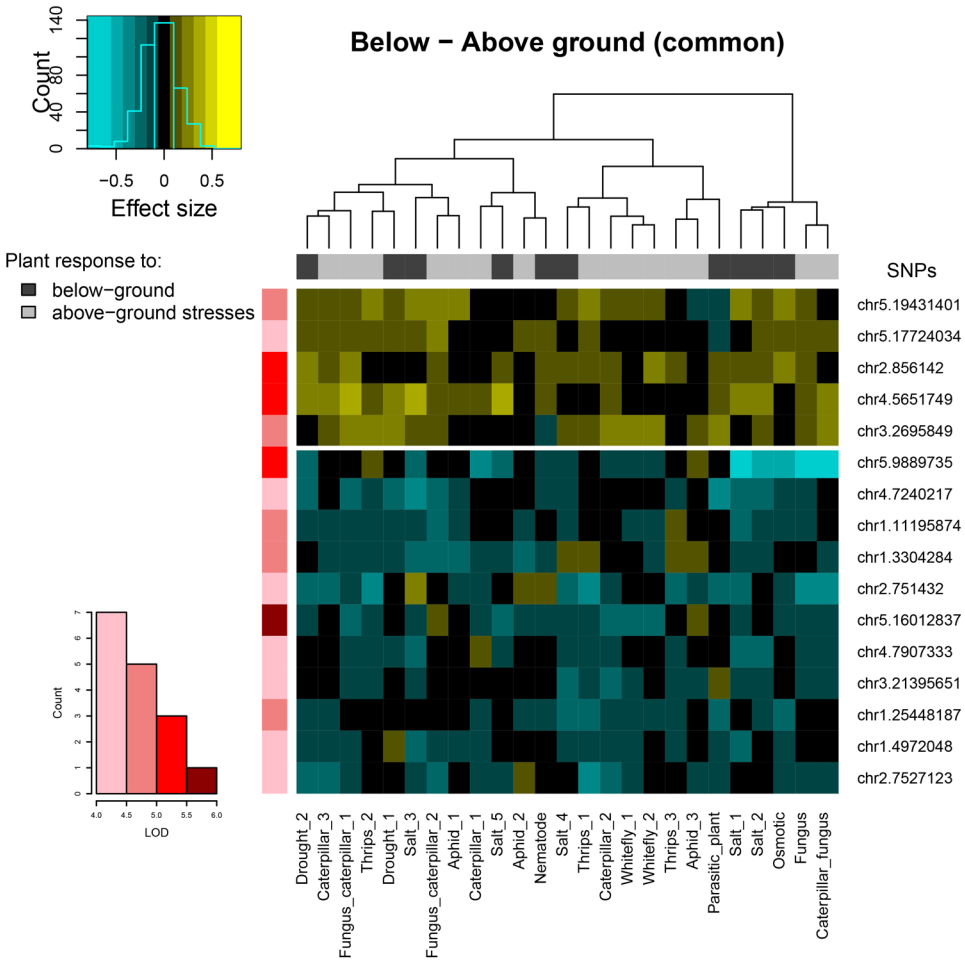
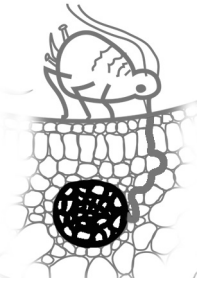



**Figure S10.** Genetic associations common for plant responses to below- and aboveground stresses. Genetic associations were estimated with a contrast analysis using MTMM. SNPs with a significance above LOD score 4 ($P \leq 10^{-4}$) for the common response are clustered according to trait-specific effect size estimated from the full MTMM. If there was another SNP in LD that had a higher effect size, this SNP was used as representative for the LD block. Negative effect sizes (blue) were cases where the rare allele was associated with a detrimental effect on the plants, positive effect sizes (yellow) were cases where the rare allele was associated with increased resistance to the stress. The rare alleles of the top 5 SNPs are associated with enhanced resistance to above- and belowground stresses; the bottom 11 SNPs are associated with reduced resistance to above- and belowground stresses. Stresses are clustered according to effect size, using Ward's minimum variance method. If SNPs were located within a 20 kb half-window of each other, only the SNP with the highest absolute cumulative effect size was included. The key shows the frequency distribution of SNPs across effect sizes.

# Chapter 5

## *PHLO*: A gene involved in restricting phloem ingestion by *Myzus persicae* aphids and phloem responses to moderate heat stress



Karen J. Kloth, Jacqueline Busscher-Lange, Gerrie L. Wiegers, Willem Kruijer, Gonda Buijs, Rhonda C. Meyer, Harro J. Bouwmeester, Marcel Dicke and Maarten A. Jongsma

*in prep.*

## Abstract

Plant resistance mechanisms to aphids are still largely elusive and hardly supported by genetic evidence. By genome-wide association mapping of *Myzus persicae* feeding behaviour on 350 natural *Arabidopsis thaliana* accessions we identified the hsp20-like *PHLOEM HEAT-RESPONSE LOCUS* (*PHLO*). Aphid behaviour was associated with polymorphisms in the intron of the *PHLO* alpha-crystallin domain. On *phlo* T-DNA insertion lines, aphids showed a strong increase in the duration of phloem ingestion, honeydew excretion rate and reproduction compared to the wild type. A similar difference in phenotype, although with less strong effects, was observed between the natural accessions Col-0 and C24, and was associated with a quantitative trait locus around *PHLO*. The effects in the T-DNA lines were more pronounced when plants were grown at 26˚C than at 20˚C. RT-qPCR revealed that *PHLO* gene expression was induced at 26˚C. Interestingly, the aphid-resistant *PHLO* haplotype occurs particularly in geographic regions with large annual temperature fluctuations. Apart from the effects on aphids, we observed a stunted growth of the inflorescence stem at 26˚C and reduced seed dormancy in *phlo-1*, indicating that PHLO might be relevant for plant fitness during high ambient temperature. *PHLO*'s closest homologue, *RESTRICTED TOBACCO ETCH POTYVIRUS (TEV) MOVEMENT 2* (*RTM2*), restricts virus transport through the vascular bundle via an unknown mechanism. Although co-expression analysis suggested a link between *PHLO* and callose-mediated sieve plate occlusion, we could not confirm a callose deposition phenotype. Yet, phloem exudation rates were lower in *phlo-1* plants at 26˚C. Based on its absence in the Poaceae family, which is devoid of phloem proteins, and its co-expression with phloem filament proteins, we postulate that PHLO might be a clamp protein that anchors phloem proteins to the lateral membrane of sieve elements and thereby obstructs the aphid's food channel and increases flow rate during moderate heat stress. Altogether, we found a gene which has substantial implications for plant resistance to aphids and requires further functional characterisation.

## Background

Aphids impose substantial economic damage to a wide variety of crops (FAO, 2015; Kim *et al.*, 2008a). They are phloem-feeding insects that have evolved a highly specialised feeding apparatus to get access to the vascular bundle of their host plants. Aphid mouthparts are called stylets, and consist of specialised mandibles that form a hollow, needle-like structure (Dixon, 1998). By manoeuvring their stylet bundle through

the cell wall matrix of the epidermis and mesophyll, they reach phloem sieve elements without inflicting much damage. Directly after penetrating a sieve element and before ingesting phloem, aphids inject watery saliva into the sieve element. The salivation duration and the occurrence of interruptive salivation events during phloem ingestion, depend on the host plant, aphid species and previous experiences of the aphid (Prado and Tjallingii, 1999; Tjallingii, 1994, Tjallingii, 2006). Injection of watery saliva has been suggested to be required for counteracting wound-responses in the phloem (Tjallingii, 2006; Will *et al.*, 2007). After salivation, aphids can feed for hours or even days from a single sieve element (Tjallingii, 1995). They are capable of active sap ingestion, but mainly feed passively. An adjustable piston valve in the stylet's food channel regulates the rate of phloem ingestion and prevents damage due to excessive turgor pressure in the phloem vessels (Dixon, 1998). For plants, a lot is at stake once insects damage or utilise the vascular network. Not only do aphids remove photoassimilates and transmit plant viruses to the vascular system, they also affect the systemic signalling network of the plant by inducing many transcriptional changes (Foyer *et al.*, 2015). To prevent deleterious effects on plant fitness, plants harbour several defence mechanisms against aphids, including the accumulation of secondary metabolites and biosynthesis of cell wall and epicuticular wax components (De Vos *et al.*, 2007; De Vos *et al.*, 2005; Levy *et al.*, 2005; Mewis *et al.*, 2005). One of the few examples of near-complete resistance to aphids occurs in lettuce and is putatively associated with a stylet-blocking mechanism in the phloem (ten Broeke *et al.*, 2013a). To further unravel the genetic and physiological background of plant resistance to aphids, investigating natural variation in plants could be a valuable approach. Wild plant populations potentially harbour resistance mechanisms that are subject to natural selection and local adaptation. Genome-wide association (GWA) mapping is a technique to mine the genomic sequences of large populations for polymorphisms associated with phenotypic traits (Hirschhorn and Daly, 2005). Due to the use of dense genotype maps in segregated populations, candidate genes can be identified with high resolution and without the need for fine-mapping (Bergelson and Roux, 2010). Nevertheless, GWA mapping has hardly been used to study the genetic architecture of plant defence mechanisms against herbivorous insects (Atwell *et al.*, 2010; Joukhadar *et al.*, 2013; Samayoa *et al.*, 2015). One of the reasons is the laborious nature of phenotyping insect performance on hundreds of plant genotypes (Goggin *et al.*, 2015; Kloth *et al.*, 2012). Now that new techniques are available to phenotype large plant collections (Chen *et al.*, 2012; Kloth *et al.*, 2015), natural resources can be exploited more efficiently to identify genes underlying plant defence.

In this study, we used a high-throughput video-tracking platform to screen the feeding behaviour of *Myzus persicae* (Sulzer) aphids on 350 natural *Arabidopsis thaliana* accessions. By GWA mapping we identified *At3g10680*, a small-heat-shock-like gene

with so far unknown function. Characterisation of aphid feeding behaviour on transfer (T)-DNA insertion lines revealed that *At3g10680* restricts phloem ingestion and reduces aphid population size. High ambient temperature induced the expression of *At3g10680* and resulted in more pronounced effects on phloem ingestion and phloem exudation rates. We, therefore, named this gene *PHLOEM HEAT-RESPONSE LOCUS* (*PHLO*).

## Results

### Identification of *PHLO* by GWA mapping

Feeding behaviour of *M. persicae* aphids on 350 natural Arabidopsis accessions was recorded by automated video tracking. By analysing aphid movements on single leaf discs we assessed the frequency and duration of tissue penetrations (Kloth *et al.*, 2015). Particularly the time that aphids spent on short (<3 min) probes showed substantial variation across the accessions (**Figure 1a**). Short probes are superficial penetrations of the epidermis and mesophyll. The more short probes aphids display, the more resistant plants generally are (van Emden and Harrington, 2007). In the Arabidopsis population, the total time aphids spent on making short probes during a timeframe of 1.5 hours, ranged from 19 seconds to 13 min with a median of 3.5 min (**Table S1**). The trait heritability was low (7%). Nevertheless, GWA mapping revealed a strong association with a region on chromosome 3 (**Figure 1b**). The most significant polymorphism (-10log(P) value=6.2, chr.3 pos. 3338114) was in linkage with a 24 kb region containing 5 genes (**Table S2**). This SNP is located in the hsp20-like gene *At3g10680* (*PHLO*) with unknown function, causing an amino acid change in the conserved alpha-crystallin domain. A haplotype analysis was performed with 173 re-sequenced accessions of the GWA mapping population (Cao *et al.*, 2011). In total, *PHLO* and its 1000 bp promoter region contained 82 SNPs, of which 18 were non-synonymous mutations in the exons (**Figure 1c-d**). Five polymorphisms, in the intron in the alpha-crystallin domain, conveyed the largest effect on the phenotype. All of them, except for the first polymorphism in the intron, were in strong linkage (>0.95) with the most significant SNP of the GWA mapping.

### Temperature-dependent phenotype

To test if *PHLO* is indeed involved in resistance to aphids, aphid feeding behaviour was characterised on two homozygous *phlo* T-DNA lines in the Col-0 background. The *phlo-1* T-DNA line (SALK_027475) had complete abolishment of *PHLO* expression,

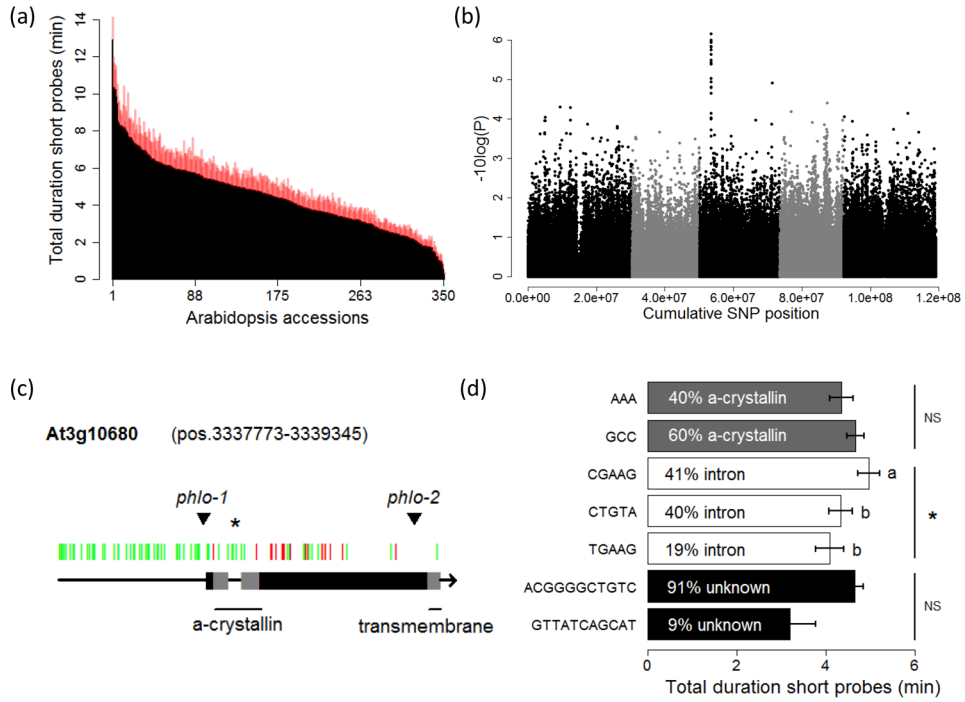



**Figure 1. Genome-wide association mapping of aphid feeding behaviour. a.** Average total time aphids were making short probes (< 3 min) on 350 Arabidopsis accessions during a 1.5-h recording (red = standard error above the mean). **b.** Genome-wide associations with 214k SNPs. A SNP in gene *At3g10680* was most significant (-10log(P)= 6.2). **c.** All SNPs in *At3g10680* and its 1000 bp promoter region of 173 resequenced Arabidopsis accessions (green= silent, red= non-synonymous compared to Col-0). Predicted gene domains are shown in gray, unknown domains in black. Triangles represent T-DNA insertions, the significant haplotype is annotated with an asterisk. **d.** For each domain haplotypes were tested for their effect on the total duration of short probes. There were no significant haplotypes in the promoter and first unknown domain (not shown). Haplotype frequency is depicted in percentage (*P<0.05, one-way ANOVA, significant differences are annotated with different letters, NS= not significant).



and the *phlo-2* T-DNA line (SAIL_1269_C01) showed a 2-fold reduction in expression (**Figure S1**). We followed the behaviour of individual aphids on whole plants for 8 hours by Electrical Penetration Graph (EPG) recording (McLean and Kinsey, 1964; Tjallingii, 1988). With this technique we assessed which tissue the aphids penetrated (epidermis/mesophyll or vascular bundle) and what stylet activities they performed (e.g. salivation, phloem ingestion, xylem ingestion). Because the 350 Arabidopsis accessions had been grown in climate compartments with the relatively high temperature of 26˚C and *PHLO* shares homology with small heat-shock proteins, we investigated a potential

| | 20˚C | | | 26˚C | | | | |
|---|---|---|---|---|---|---|---|---|
| **Variable** | **Col-0** | **phlo-1** | **phlo-2** | **Col-0** | **C24** | **Col-0 x C24** | **phlo-1** | **phlo-2** |
| Total duration NP | 195.8 ± 30.1 | 164 ± 15.4ns | 106.5 ± 21.8* | 140.9 ± 13.5ns | 98.3 ± 13.3* | 128.5 ± 18.1ns | 55.2 ± 12.2*** | 119.1 ± 24ns |
| Total duration C | 149.5 ± 20.6 | 121.3 ± 15.2ns | 91.5 ± 24.5ns | 181.4 ± 13.2ns | 175.6 ± 16.2ns | 151.6 ± 18.8ns | 69.7 ± 9*** | 93.9 ± 16.1** |
| Number of C < 3 min | 26.8 ± 4.4 | 28.5 ± 4.2ns | 26.2 ± 9.2ns | 27.4 ± 3.9ns | 29.9 ± 5ns | 27.1 ± 4.2ns | 8.4 ± 2.2*** | 21.5 ± 5.6ns |
| Total duration E1 | 6.4 ± 1.6 | 3.9 ± 0.7ns | 1.4 ± 0.3* | 9.4 ± 2.3ns | 6.1 ± 0.9ns | 4.5 ± 0.9* | 1.8 ± 0.5*** | 1.3 ± 0.3*** |
| Number of E1 | 6.5 ± 1.2 | 5.4 ± 1ns | 1.8 ± 0.3* | 9.2 ± 0.9ns | 11.1 ± 1.4* | 7.1 ± 0.9ns | 2.2 ± 0.3*** | 1.7 ± 0.4*** |
| Mean duration E1 | 1 ± 0.2 | 0.8 ± 0.1ns | 0.8 ± 0.1ns | 1 ± 0.2ns | 0.6 ± 0.1* | 0.6 ± 0.1* | 0.7 ± 0.1ns | 0.8 ± 0.1ns |
| Latency to 1st E1 | 155.7 ± 26.8 | 165.8 ± 25.6ns | 128.2 ± 22.2ns | 116.4 ± 22ns | 83.4 ± 14.2ns | 99 ± 22.6ns | 102.2 ± 16.6ns | 178.6 ± 36.1ns |
| Number of single E1 | 0.9 ± 0.4 | 0.2 ± 0.1ns | 0.2 ± 0.1ns | 0.6 ± 0.3ns | 0.2 ± 0.2ns | 0.5 ± 0.2ns | 0.3 ± 0.1ns | 0.1 ± 0.1* |
| % E1 in (E1+E2) | 11.9 ± 4.9 | 3.1 ± 0.8ns | 2 ± 1.5** | 8.5 ± 1.5ns | 3.6 ± 0.6* | 5.8 ± 2ns | 0.6 ± 0.2*** | 4.2 ± 2.4** |
| Total duration E2 | 118.4 ± 34 | 172.6 ± 20.3ns | 280.1 ± 41.2* | 137 ± 20.1ns | 197.5 ± 22.6* | 192.2 ± 34.7ns | 345.3 ± 17.9*** | 251.1 ± 41.4* |
| Number of E2 | 5.2 ± 1 | 4.9 ± 0.9ns | 1.7 ± 0.3* | 7 ± 0.9ns | 10 ± 1.2ns | 6.3 ± 1ns | 1.8 ± 0.2*** | 1.5 ± 0.3*** |
| Mean duration E2 | 54.7 ± 37.7 | 62.6 ± 13ns | 232.2 ± 43.5** | 26 ± 5.7ns | 39.5 ± 16.8ns | 45.2 ± 16ns | 254 ± 31*** | 222 ± 41.9** |
| Latency to 1st E2 | 160.9 ± 27.4 | 167.1 ± 25.7ns | 138.3 ± 28.4ns | 128.5 ± 23.3ns | 83.9 ± 14.2ns | 113.8 ± 23.7ns | 110 ± 18.6ns | 185.5 ± 38.9ns |
| Total duration E2s | 107.2 ± 33.7 | 165.8 ± 20.9* | 278.6 ± 41.5* | 119.1 ± 21ns | 170 ± 24.9ns | 173.7 ± 36.3ns | 343.7 ± 17.9*** | 249 ± 41.6* |
| Number of E2s | 2.2 ± 2.2 | 2.8 ± 2.8ns | 1.2 ± 1.2ns | 2.8 ± 2.8ns | 4.3 ± 4.3* | 2.3 ± 2.3ns | 1.4 ± 1.4*** | 0.9 ± 0.9*** |
| Mean duration E2s | 78 ± 40 | 108.4 ± 25.2ns | 269.8 ± 36.4*** | 50.6 ± 10.4ns | 59 ± 18ns | 85.5 ± 16.7* | 298.2 ± 25.9*** | 272.7 ± 39.4*** |
| % aphids E2 | 84.6 ± 10.4 | 100 ± 0ns | 92.3 ± 7.7ns | 100 ± 0ns | 100 ± 0ns | 100 ± 0ns | 100 ± 0ns | 93.8 ± 6.2ns |
| Total duration G | 9.8 ± 5.8 | 0 ± 0* | 0.5 ± 0.5ns | 1.9 ± 1.7ns | 0 ± 0ns | 0 ± 0ns | 6.6 ± 3.1ns | 2.8 ± 2.8ns |
| Number of G | 0.5 ± 0.2 | 0 ± 0* | 0.1 ± 0.1ns | 0.1 ± 0.1ns | 0 ± 0ns | 0 ± 0ns | 0.2 ± 0.1ns | 0.1 ± 0.1ns |
| Total duration F | 0 ± 0 | 18.3 ± 8.5* | 0 ± 0ns | 9.5 ± 6.9ns | 2.6 ± 2ns | 3.2 ± 3.2ns | 1.4 ± 1.4ns | 11.8 ± 5.7ns |
| Number of F | 0 ± 0 | 0.8 ± 0.3* | 0 ± 0ns | 0.2 ± 0.1ns | 0.1 ± 0.1ns | 0.1 ± 0.1ns | 0 ± 0ns | 0.8 ± 0.4ns |

Mean ± standard error, *P<0.05; **P<0.01; ***P<0.001 (Mann-Whitney U pairwise comparisons within temperature treatments: C24, Col-0xC24, *phlo-1, phlo-2* versus Col-0. Between temperature treatments: Col-0 versus Col-0). NP = non-penetration, C = probing in epidermis or mesophyll, E1 = phloem salivation, E2 = phloem ingestion, E2s = sustained phloem ingestion (> 10 min), G = xylem ingestion, F = penetration difficulties. Latency was measured from the start of the recording. Single E1s were phloem salivations that were not directly followed or preceded by phloem ingestion.

◄ **Table 1.** Aphid feeding behaviour, measured by 8-h Electrical Penetration Graph recordings on the natural Arabidopsis accessions Col-0 and C24, a near-isogenic line of Col-0 with a C24 introgression covering the region around *PHLO*, and the T-DNA lines *phlo-1* and *phlo-2* . Experiments were performed on plants grown at two different temperatures, 20˚C and 26˚C (time shown in minutes).

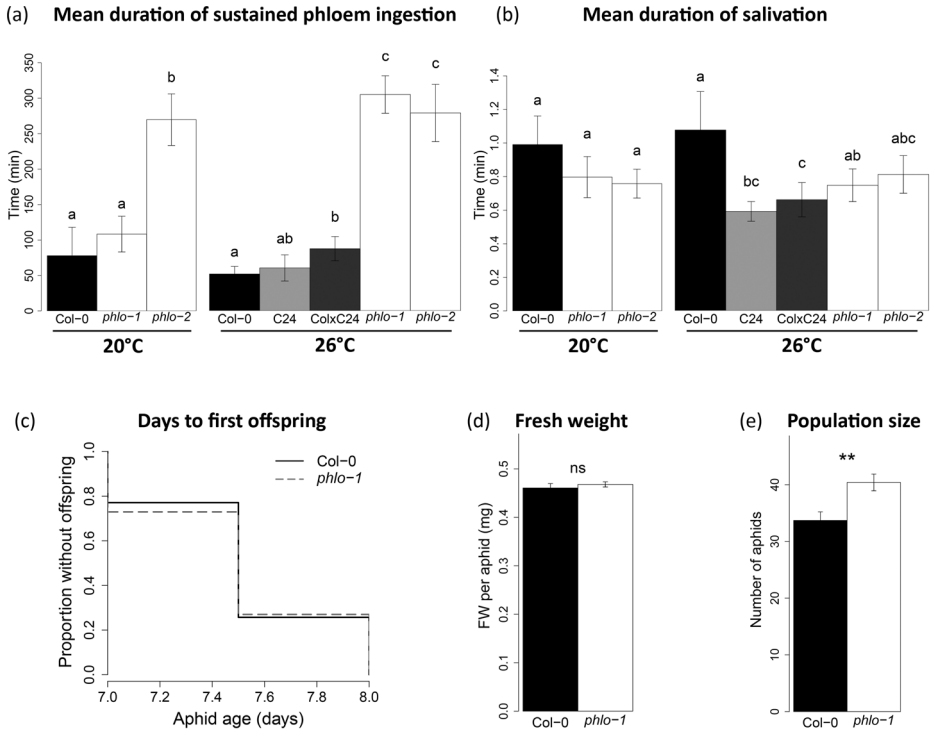



**Figure 2. Aphid performance. a.** Mean duration of sustained phloem ingestion events (> 10 min) by *M. persicae* aphids on the natural Arabidopsis accessions Col-0 and C24, a near-isogenic line of Col-0 with a C24 introgression covering the region around *PHLO*, and the T-DNA lines *phlo-1* and *phlo-2*. Plants were grown at 20˚C or 26˚C, the duration of the EPG recordings was 8 hours (Kruskal Wallis group-wise comparison within temperature treatment P<0.001, Mann-Whitney U pairwise comparisons, different letters refer to significant differences). **b.** Mean duration of salivation events in the phloem (P<0.05 Kruskal Wallis group-wise comparison, Mann-Whitney U pairwise comparisons). On wild-type and *phlo-1* plants grown at 26˚C the following life history traits were assessed: **c.** Aphid age when producing their first offspring (P=0.77, Kaplan-Meier curve of Cox proportional hazards model), **d.** fresh weight of 7-day-old aphids (P=0.51, one-way ANOVA), and **e.** aphid population size after a 2-week infestation with one neonate aphid per plant (P=0.003, one-way ANOVA).

temperature effect by including two temperature treatments. When plants were grown at 20˚C, aphids displayed a longer total duration of sustained phloem ingestion on both *phlo* T-DNA lines (**Table 1**), and a longer mean duration of sustained phloem ingestion events on *phlo-2* compared to the wild type (**Figure 2a**). The effects in *phlo-1* and *phlo-2* were, however, more pronounced when plants were grown at 26˚C. In comparison to the wild type, aphids spent less time on penetration of the epidermis and mesophyll, salivated less before and during feeding (**Table 1**), and showed a 3-fold increase in the mean duration of sustained phloem ingestion (**Figure 2a**). On wild type plants, aphid feeding behaviour was not affected by temperature. The time required to reach the phloem did not differ between *phlo* T-DNA lines and the wild type (**Table 1**), indicating that the resistance mechanism was mainly phloem-located and not associated with epidermal or mesophyll tissue. To assess the effect of the natural haplotype, aphid behaviour was recorded on accession C24, with the susceptible, non-Col-0 allele (chr. 3, pos. 3338114, **Table S3**), and a Col-0 near-isogenic line (NIL) with a 3000 kb introgression of C24 in the region covering *PHLO*. On C24 and the NIL, grown at 26˚C, aphids showed less salivation and more phloem ingestion than on Col-0 (**Figure 2a-b**, **Table 1**), confirming a phloem-located effect. To assess the long-term impact

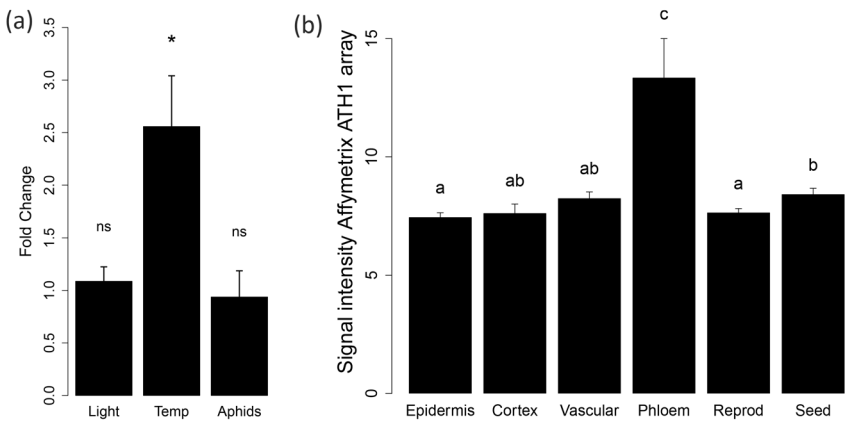



**Figure 3. Gene expression of** *PHLO*. **a.** Expression in Col-0 rosette leaves under different light intensity, temperature and aphid treatments, measured by RT-qPCR. Fold change expression is shown for plants grown at high light intensity (120 µmol m-2 s-1) relative to plants grown at low light intensity (70 µmol m-2 s-1, P=0.74, one-way ANOVA); plants grown at high temperature (26˚C) relative to low temperature (20˚C, P=0.042, one-way ANOVA), and plants infested with aphids for 6 h, relative to clean plants (P=0.57, one-way ANOVA). **b.** *PHLO* expression in different Arabidopsis tissues. Data from 105 anatomical parts were retrieved from Genevestigator (Zimmermann *et al.*, 2004), and classified in the above tissue categories (Kruskal Wallis group-wise comparison P<0.01, Mann-Whitney U pairwise comparisons, different letters refer to significant differences, bars represent mean ± standard error).

on aphid performance, several aphid life-history traits were quantified at 26˚C during a time course of two weeks on *phlo-1* and the wild type. Aphids produced significantly more offspring on *phlo-1*, although development time from neonate to adult and aphid fresh weight were unaffected (**Figure 2c-e**). RT-qPCR revealed that the expression of *PHLO* in Col-0 was higher at 26˚C compared to 20˚C, but was not upregulated by high light intensity or aphid infestation (**Figure 3a**). These results indicate that although *PHLO* exhibits substantial effects on aphid performance, it is not regulated by aphid infestation but by ambient temperature.

### Effects on callose, phloem exudation and honeydew excretion

Transcriptome repositories showed that *PHLO* expression is primarily confined to phloem tissue of root and shoot (**Figure 3b**). Its closest homologue, *RESTRICTED TOBACCO ETCH POTYVIRUS (TEV) MOVEMENT 2* (*RTM2*) is also expressed in phloem vessels, where it restricts the transport of viruses through the vascular bundle via an unknown mechanism (Bondino *et al.*, 2012; Cayla *et al.*, 2015; Chisholm *et al.*, 2001; Whitham *et al.*, 2000; Whitham *et al.*, 1999). Protein similarity between PHLO and RTM2 is mainly reflected in the alpha-crystallin and transmembrane domain (**Table**

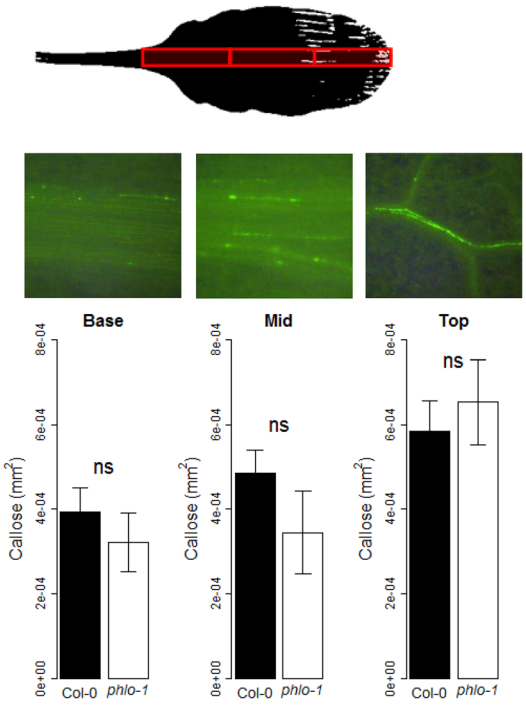



**Figure 4. Callose content in leaves of wild type and *phlo-1* plants.** Callose depositions were stained with aniline blue. The complete midvein of each leaf was imaged with a fluorescent microscope and semi-quantified with ImageJ. Images were grouped into either base, mid or top part of the leaf ($P > 0.05$, one-way ANOVA). All plants were grown at 26˚C, data represent mean ± standard error.

**Table 2.** Genes co-expressed with *At3g10680*. The first 22 genes appeared in three independent co-expression databases: (1) the top 300 genes of Atted-II version c4.1 (Obayashi *et al.*, 2014), (2) the top 100 genes of BAR Expression Angler Abiotic Stress, Pathogen, and Extended Tissue Compendium resp. (Toufighi *et al.*, 2005), and (3) the top 100 genes of Genevestigator's Pertubations data set (Zimmermann *et al.*, 2004). In total, 597 co-expressed genes were retrieved.

| AGI code | Name | Description | Number of databases |
|---|---|---|---|
| *At1g02950* | *GSTF4* | Glutathione transferase belonging to the phi class of GSTs | 3 |
| *At1g05760* | *RTM1* | *RESTRICTED TOBACCO ETCH POTYVIRUS MOVEMENT 1* | 3 |
| *At1g13920* | - | Remorin family protein | 3 |
| *At1g63310* | - | unknown protein | 3 |
| *At1g68230* | - | Reticulon family protein | 3 |
| *At1g73370* | *SUS6* | *SUCROSE SYNTHASE 6*, involved in callose biosynthesis on sieve plates | 3 |
| *At2g27140* | - | Hsp20-like chaperones superfamily protein, homologue of *RTM2* and *PHLO* | 3 |
| *At2g34530* | - | unknown protein | 3 |
| *At3g01670* | *ATSEOR2* | *SIEVE ELEMENT OCCLUSION-RELATED 2* | 3 |
| *At3g01680* | *ATSEOR1* | *SIEVE ELEMENT OCCLUSION-RELATED 1* | 3 |
| *At3g03200* | *NAC045* | NAC domain containing protein 45, DNA binding transcription factor activity | 3 |
| *At3g07420* | *NS2* | Asparaginyl-tRNA synthetase 2 | 3 |
| *At3g18260* | - | Reticulon family protein | 3 |
| *At3g20450* | - | B-cell receptor-associated protein 31-like, intracellular protein transport | 3 |
| *At3g58350* | *RTM3* | Restricted Tobacco etch potyvirus Movement 3 | 3 |
| *At4g00890* | - | Putative glycosyl hydrolase family 10 protein (xylanase) | 3 |
| *At4g24250* | *MLO13* | Homologous to barley mildew resistance (MLO) protein | 3 |
| *At4g35970* | *APX5* | Microsomal ascorbate peroxidase 5, involved in hydrogen peroxide scavenging | 3 |
| *At4g37445* | - | Putative member of Calcium-binding EF hand family protein | 3 |
| *At5g17260* | *NAC086* | NAC domain containing protein 86, DNA binding transcription factor activity | 3 |
| *At5g25370* | - | Member of C2-PLD subfamily, involved in hyperosmotic response | 3 |
| *At5g56720* | *C-NAD-MDH3* | Lactate/malate dehydrogenase family protein, located in membrane | 3 |
| *At5g37180* | *SUS5* | *SUCROSE SYNTHASE 5*, involved in callose biosynthesis on sieve plates | 2 |
| *At1g06490* | *GSL7* | *GLUCAN SYNTHASE-LIKE 7*, involved in callose biosynthesis on sieve plates | 2 |

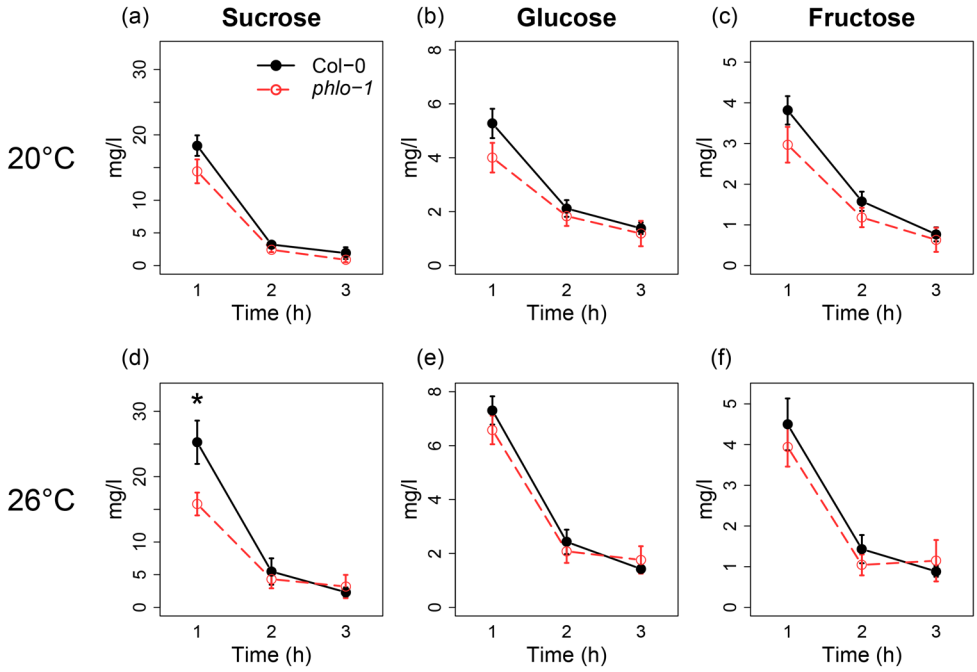



**Figure 5. Sugar content of phloem exudates of wild-type and *phlo-1* leaves. a.** Sucrose, **b.** glucose, and **c.** fructose content of phloem exudates, collected at growing conditions of 20˚C (P>0.05, one-way ANOVA). **d.** Sucrose, **e.** glucose, and **f.** fructose content of phloem exudates, collected at growing conditions of 26˚C (*P < 0.05, one-way ANOVA). Phloem was collected from 15 excised leaves per sample in a 3-hour time course and corrected for leaf fresh weight. Samples were treated with EDTA to prevent sealing of the vessels. Data represent mean ± standard error.



**S4**). Unlike RTM2, PHLO does, however, not contain a predicted coiled coil region. *PHLO* showed co-expression with several genes related to sieve element occlusion; the filamentous phloem proteins *SIEVE ELEMENT OCCLUSION-RELATED 1* (*ATSEOR1*) and *SIEVE ELEMENT OCCLUSION-RELATED 2* (*ATSEOR2*) (Ruping *et al.*, 2010), and three genes involved in callose formation on sieve plates *SUCROSE SYNTHASE 5* (*SUS5*), *SUCROSE SYNTHASE 6* (*SUS6*) and *GLUCAN SYNTHASE-LIKE* (*GSL7*) (Barratt *et al.*, 2009; Barratt *et al.*, 2011) (**Table 2**). Based on the latter, we hypothesized that *PHLO* might be required for callose-mediated plugging of sieve plates. In order to assess callose deposition in the phloem bundle, rosette leaves of *phlo-1* and the wild type were stained with aniline blue and fluorescent signals were semi-quantified along the leaf mid-vein. Since sieve elements differ in developmental stage along the mid-vein and callose content depends on the age of cell and sieve plates (Barratt *et al.*, 2011; Samuels *et al.*, 1995), callose was separately semi-quantified

for the basis, middle and top of the leaf. However, in none of these regions, nor in the complete mid-vein, did the *phlo* T-DNA line show significant differences in callose presence compared to the wild type (**Figure 4**). As an independent test for differences in sieve plate plugging, we assessed whether the phloem flow rate was different in *phlo-1*. Hereto, phloem exudates were collected from excised leaves during a 3-hour time course at both 20˚C and 26˚C. Assuming that sugar concentrations were the same in wild type and *phlo-1* phloem, we used total sugar content in EDTA-collected phloem exudates as an estimate of phloem flow rate. At 20˚C, no difference in exudation of sucrose, glucose and fructose was observed between *phlo-1* and wild type (**Figure 5a-c**). In contrast to our expectations, at 26˚C, sucrose levels were lower in *phlo-1* than in the wild type (**Figure 5d-f**), suggesting that the phloem flow rate was reduced in the T-DNA line. To assess the impact on the volume of phloem ingestion, we measured the honeydew excretion rate of aphids, which is highly correlated to the phloem ingestion rate (Tjallingii, 1995). On plants grown at 26˚C, we collected honeydew droplets with honeydew clocks and simultaneously recorded stylet activities with EPG recording. Honeydew droplets were only counted during continuous phloem ingestion events with a duration of approximately 8 hours to exclude any effects of non-feeding intervals

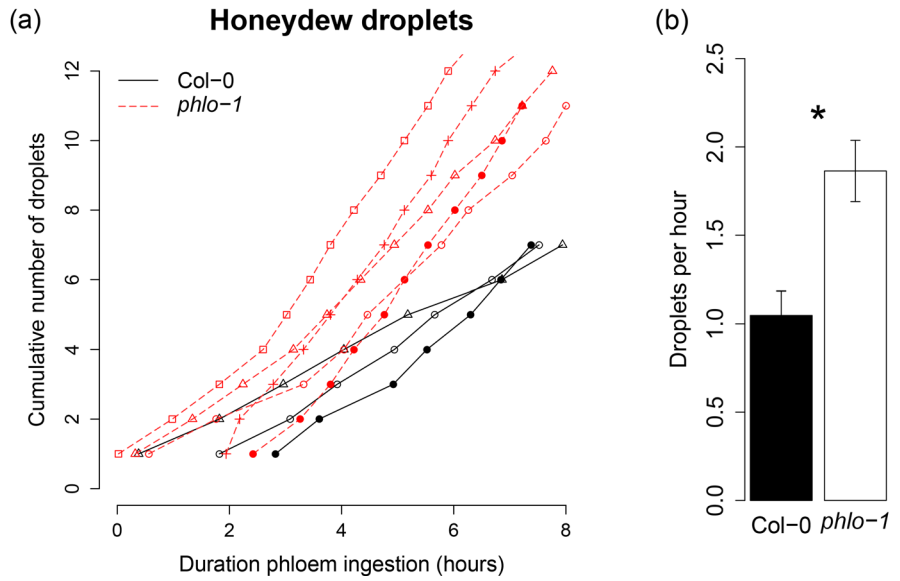



**Figure 6. Honeydew droplet excretion by** *M. persicae* **aphids on wild type and** *phlo-1* **plants. a.** Droplets were collected and counted during continuous phloem ingestion. Every line represents an aphid, each symbol represents a honeydew droplet. Plants were grown at 26˚C. **b.** Mean honeydew excretion rate (P=0.018, one-way ANOVA, data represent mean ± standard error).

or lower ingestion rates during short feeding periods. Overall, honeydew excretion occurred in regular intervals on both plant lines, but the rate of excretion was higher on *phlo-1* (**Figure 6**). Altogether, these results indicate that PHLO is involved in a callose-independent mechanism that increases the phloem translocation rate but reduces the phloem ingestion rate of aphids during high ambient temperature.

### PHLO's association with plant performance and natural environmental conditions

To assess whether PHLO affects any aspects of plant performance, several plant growth and developmental parameters were measured in *phlo-1* and the wild type. For these assays, new seeds were collected from wild type and mutant plants grown at 26˚C and subjected to 2.5 weeks of dry storage after harvest. Germination assays were performed at standard conditions of 22˚C and suboptimal conditions of 25˚C, in order to reveal subtle differences in the performance of not completely after-ripened seeds (Tamura *et al.*, 2006). Under suboptimal conditions, *phlo-1* displayed a significantly higher seed germination rate and germination percentage compared to the wild type (**Figure 7**), suggesting that *phlo-1* seeds had reduced dormancy. When seeds were cold stratified to break dormancy (Bentsink and Koornneef, 2008), the effect was not visible anymore. T-DNA plants did not display irregularities in rosette development and time to bolting when grown at 26˚C. The growth of the inflorescence stem was, however, stunted in the first phase after bolting (**Table 3**). In view of the quantitative effects of PHLO on seed dormancy and inflorescence development, we tested the geographical distribution and climate conditions for the most significant SNP of the GWA mapping analysis. The



**Table 3.** Rosette and inflorescence development of wild type and *phlo-1* plants. Seeds were cold stratified before sowing, and plants were grown at 26˚C, 8 h:16 h L:D photoperiod. Leaves were counted when they were > 1 mm in length.

| Variable | Col-0 | *phlo-1* |
|---|---|---|
| Number of leaves (15 days) | 7.8 ± 0.12 | 7.6 ± 0.14ns |
| Number of leaves (35 days) | 30.2 ± 0.59 | 29.4 ± 0.37ns |
| Rosette fresh weight (gram, 35 days) | 1.3 ± 0.06 | 1.2 ± 0.05ns |
| Time to bolting (days) | 52.8 ± 0 | 54 ± 1ns |
| Inflorescence stem height (cm, 58 days) | 8.9 ± 2 | 2.3 ± 1* |
| Inflorescence stem height (cm, 60 days) | 16.1 ± 2 | 7.6 ± 2* |
| Inflorescence stem height (cm, 63 days) | 23.3 ± 1 | 14.3 ± 3ns |

Means ± standard error, *P<0.05; **P<0.01; ***P<0.001 (one-way ANOVA, leaves at 15 days: n=13, leaves and fresh weight at 35 days: n=8, other variables: n=4-5).

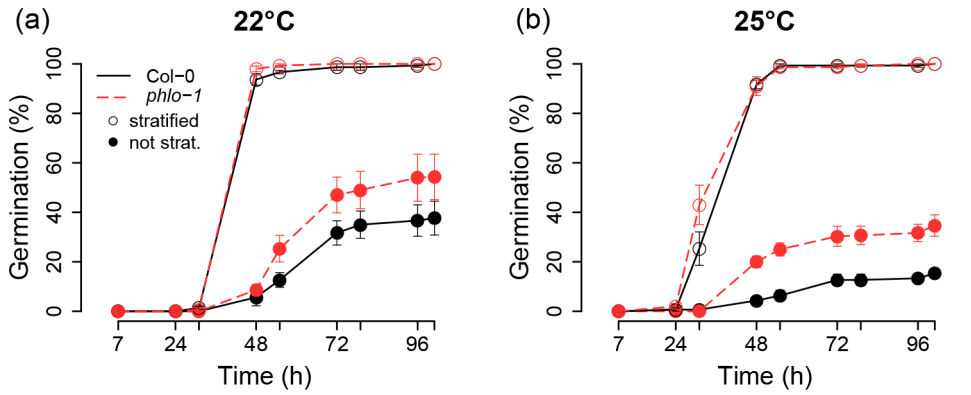



**Figure 7. Seed germination of the wild type and *phlo-1* T-DNA line. a.** Germination percentage during incubation at 22°C of stratified (open circles) and unstratified (closed circles) seeds (P>0.05, one-way ANOVA). **b.** Germination percentage during incubation at 25°C. For unstratified seeds, maximum germination percentage and area under the curve were significantly different between Col-0 and *phlo-1* (P=0.012, resp. P=0.007, one-way ANOVA).

occurrence of the resistant Col-0 haplotype was biased to Eastern Europe ($P=3.2*10^{-5}$, Mann-Whitney U test on longitude, P=0.13, Mann-Whitney U test on latitude, **Figure S2**). In addition, twelve out of 52 climate variables (**Table S5-S6**) had a significant effect on the occurrence of the haplotype (P<0.001, Mann-Whitney U test with Bonferroni correction), including nine temperature variables. Overall, the resistant allele occurred significantly more often in habitats with high temperature fluctuations throughout the year, relatively low winter temperatures and high wet season temperatures. These associations suggest that PHLO may play a role in plant responses to temperature which are relevant for plant fitness.

## Discussion

### The effect of PHLO on aphids

This is the first study to identify and validate a novel gene involved in plant resistance to aphids via GWA mapping. We used a video-tracking platform to phenotype *M. persicae* feeding behaviour on 350 natural Arabidopsis accessions. The time aphids spent on short probes showed pronounced variation between accessions. The trait heritability

was low (7%), possibly due to the use of leaf discs which is known to reduce resistance effects (ten Broeke, 2013). Nevertheless, GWA mapping revealed a strong association with a SNP in *PHLO* (*At3g10680*), encoding an hsp20-like gene. Remarkably, *PHLO* was not identified by natural variation in sustained feeding events, which was the major phenotype of the NIL and T-DNA lines. We presume, however, that the duration of the video recording was too short (1.5 h) to reveal any variation in long events. Aphid behaviour was most strongly associated with polymorphisms in the intron, which might cause alternative splicing of the alpha-crystallin domain. The involvement of PHLO in plant resistance to aphids was validated with detailed EPG recording of aphid feeding behaviour on whole plants. Testing of *phlo* T-DNA lines showed that PHLO substantially affects aphid feeding behaviour and revealed that the effect of PHLO is temperature-dependent. On *phlo* T-DNA plants grown at 26˚C, aphids showed a 3-fold increase in the mean duration of phloem ingestion events and an almost 2-fold higher rate of phloem volume ingestion compared to the wild type. In addition, aphids were salivating less before and during feeding and spent less time on penetrating the epidermis and mesophyll on plants of the T-DNA lines, most likely as a consequence of rapid phloem acceptance and subsequent feeding. When plants were grown at 20˚C, the total duration of phloem ingestion was still 1.5- to 3-fold higher on the T-DNA lines compared to the wild type, but the majority of the effects were smaller or absent. *Phlo-2* displayed a stronger effect at 20˚C than *phlo-1*. *phlo-2* Plants had a homozygous insertion in the last exon, but gene expression was not completely abolished in this T-DNA line (**Figure S1**). This may have influenced the phenotype relative to *phlo-1* in which *PHLO* expression is completely abolished due to an insertion in the promoter. Differences between natural haplotypes were small, but also here, the effects were associated with the phloem. Aphids showed less salivation in the phloem (**Figure 2b**) and prolonged phloem ingestion (**Table 1**) on the natural accession C24 and the NIL compared to Col-0. Altogether, we have found a gene with an important role in Arabidopsis resistance to *M. persicae* aphids.



### Links between PHLO and restriction of virus transport

Bondino *et al.* (2012) showed that *PHLO* is part of the phylogenetic clade UAPVI of alpha-crystallin domain genes, together with its closest homologue *RTM2*. *PHLO* and *RTM2* are considered to have evolved from a gene duplication event (Cosson *et al.*, 2012). RTM2 proteins are located at the lateral sides of phloem sieve elements as well as on sieve plates. *PHLO* shows co-expression with *RESTRICTED TEV MOVEMENT 1* (*RTM1*), *RESTRICTED TEV MOVEMENT 3* (*RTM3*), and *At2g27140*, a more distant homologue of *PHLO* and *RTM2* from clade UAPVII (Bondino *et al.*, 2012). RTM1, RTM2 and RTM3 are all required for the restriction of TEV transport through

the phloem vessels (Cayla *et al.*, 2015; Chisholm *et al.*, 2001; Cosson *et al.*, 2010; Mahajan *et al.*, 1998; Whitham *et al.*, 2000). How the *RTM* genes restrict virus transport is still unclear, but it involves multiprotein complexes of homo- and heterodimers of RTM1 and RTM3 (Cosson *et al.*, 2010). RTM2 proteins have not been found to interact with RTM1, RTM3 or itself. The co-expression analysis of *PHLO* also revealed a gene belonging to the remorin family proteins (**Table 2**), some of which have been shown to restrict viral movement through plasmodesmata (Raffaele *et al.*, 2009). Based on homology, co-expression and experimental data, it is most likely that *PHLO*, *RTM2*, and possibly also *At2g27140*, have a similar functionality in the phloem, but respond to different stimuli.

### A putative role of PHLO in sieve element occlusion-related processes

The reduced phloem exudation rate of *phlo-1* compared to the wild type and increased phloem ingestion rate by aphids, indicate that PHLO plays a role in the regulation of phloem transport. Indeed, co-expression analysis of *PHLO* reveals associations with sieve element occlusion mechanisms. *PHLO* is co-expressed with genes involved in callose-mediated sieve plate occlusion (*SUS5, SUS6, GSL7*) (Barratt *et al.*, 2009; Barratt *et al.*, 2011), and the phloem proteins *ATSEOR1* and *ATSEOR2* (**Table 2**). To counteract phloem leakage upon injury, plants are considered to employ phloem protein aggregations for rapid occlusion within 15 to 45 seconds (Furch *et al.*, 2007; Knoblauch and Mullendore, 2013), and the formation of callose collars around sieve plate pores for slower occlusion within several minutes (Thompson and Van Bel, 2013; Will and van Bel, 2006). Sieve element occlusion has also been suggested as a resistance mechanism to aphids (Dreyer and Campbell, 1987; Van Bel, 2003; Will *et al.*, 2013), and callose deposition and phloem protein dispersal have both been described during temperature fluctuations (Furch *et al.*, 2007; McNairn, 1972; McNairn and Currier, 1968). In addition, plugging of sieve plate pores and sequestering of viroid RNA by phloem proteins, as has been observed in cucumber (Gómez and Pallás, 2004; Hipper *et al.*, 2013), would both be plausible explanations for constrained long-distance movement of potyviruses by RTM2. Below, we discuss the putative role of PHLO in callose deposition and phloem protein dispersal.

### No evidence for PHLO-mediated callose deposition in the vascular bundle

Sieve plates interconnect individual sieve elements and play an important role in the regulation of phloem transport. The pore diameter in sieve plates can be dynamically altered by the deposition and removal of callose (Barratt *et al.*, 2009; Barratt *et al.*, 2011). Increased callose deposition leads to reduced phloem flow, which has been hypothesized

to reduce phloem ingestion by aphids (Dreyer and Campbell, 1987; Will and van Bel, 2006). It is, however, doubtful whether a moderate reduction in flow rate will hamper the passive ingestion of phloem by aphids. Phloem translocation rates depend on plant species, but have in general been estimated at 200–400 μm $s^{-1}$ (Knoblauch and Oparka, 2012; Windt *et al.*, 2006). Aphids generally have to reduce the excessive flow with an adjustable piston valve in the food channel (Dixon, 1998). Therefore, only large depressions in flow rate, requiring substantial callose depositions, would decrease phloem ingestion. Nonetheless, we tested the putative involvement of PHLO in callose deposition in the vascular bundle. Callose synthases are anchored in the plasma membrane and form complexes with other proteins, such as phragmoplastin and Rho-proteins, with which they synthesize callose via multiple steps (Verma and Hong, 2001). PHLO is a chaperone-like protein with a transmembrane domain, and could possibly participate in these callose-synthesis protein complexes. Due to the increased phloem ingestion by aphids, we expected a decreased callose content in the vascular bundle and a higher phloem transport velocity in *phlo-1* compared to the wild type. Aniline blue staining of *phlo-1* and wild type leaves did, however, not reveal significant differences. In addition, *phlo-1* showed a lower phloem exudation rate compared to the wild type, in contrast to what would be expected in the absence of callose plugs. In view of these findings, we consider it unlikely that PHLO is involved in callose deposition on sieve plates.

### Is PHLO involved in phloem-protein-mediated resistance to aphids?

One of the most intriguing characteristics of *PHLO* and *RTM2* is, that they do not seem to occur in the Poaceae family (Bondino *et al.*, 2012), which is known to be devoid of phloem proteins (Dinant *et al.*, 2003; Eleftheriou, 1990; Knoblauch and Mullendore, 2013). Phloem protein is a collective term for mostly uncharacterised proteins that appear in a wide variety of shapes (e.g. amorphous, crystalline, filamentous) both at the margins and in the lumen of the sieve elements (Ehlers *et al.*, 2000; Evert, 1990). Upon injury or temperature fluctuations they disperse into the sap where they can form aggregations that press against the sieve plate within seconds (Ehlers *et al.*, 2000; Furch *et al.*, 2007; Knoblauch and Van Bel, 1998). The exact function of phloem proteins is still largely unknown, but some studies showed that they reduce phloem loss (Ernst *et al.*, 2012; Knoblauch and Van Bel, 1998). Phloem proteins may also block the aphid stylet. Transverse sections of stylets during phloem ingestion revealed the accumulation of filamentous and amorphous proteins in the food channel of several aphid species (Tjallingii and Hogen Esch, 1993). The inhibition of food ingestion was indirectly illustrated by prolonged phloem exudation from cut aphid stylets in Poaceae plants compared to dicot plant species (Knoblauch and Mullendore, 2013). Aphids respond to

sieve element occlusion by the secretion of watery saliva (Furch *et al.*, 2010; Tjallingii, 2006). This watery saliva contains calcium ion-scavenging proteins that are considered to inhibit phloem protein aggregations (Will *et al.*, 2007). The phenotype of the *phlo* T-DNA lines matches the expectations of reduced phloem protein obstruction: aphids secreted less saliva into sieve elements, showed prolonged phloem feeding and ingested more volume (**Table 1**, **Figure 6**). Its absence in the Poaceae family and its co-expression with the phloem protein encoding genes, *RTM1*, *RTM3*, *ATSEOR1* and *ATSEOR2*, make *PHLO* a likely candidate for a phloem-protein-related resistance mechanism. ATSEOR1 and ATSEOR2 are both filamentous proteins that form protein meshwork and dense slime agglomerations at the margins and in the lumen of sieve elements (Anstead *et al.*, 2012; Cayla *et al.*, 2015; Froelich *et al.*, 2011; Ruping *et al.*, 2010). A previous study showed that of ATSEOR1 and ATSEOR2 were not detrimental for the development and reproduction of *M. persicae* aphids (Anstead *et al.*, 2012). Whether or not these phloem proteins have an effect on aphids might, however, depend on their state, e.g. being attached to the margins or dispersed in the sap, and forming monomers or protein aggregations. Possibly not aphids, but environmental factors such as high temperature, could determine the state and abundance of proteins such as ATSEOR1 and ATSEOR2. Based on the current insights in phloem biology, we consider two hypotheses for the functionality of PHLO: (1) it is a clamp protein which anchors phloem proteins to the lateral sides of sieve elements, and (2) PHLO is a chaperone facilitating transport of phloem proteins to the sieve elements. Some microscopic studies describe minute clamp-like structures linking sieve element components to each other and to the parietal plasma membrane (Cayla *et al.*, 2015; Ehlers *et al.*, 2000; Froelich *et al.*, 2011). As a clamp protein, PHLO could be attached to the lateral plasma membrane, endoplasmic reticulum or plastids with its C-terminal membrane domain, and anchor phloem proteins with a chaperone domain. Tjallingii and Hoogen Esch (1993) illustrated that particularly sheets of slimy, filamentous proteins along the margins of sieve elements encapsulate the stylet tip and occlude the entry of the aphid's food channel. Via the anchoring to PHLO and by the formation of homo- and heterodimers, slimy phloem protein aggregations could develop along the margins of a sieve element. Membrane-attached protein sheets could be more detrimental for aphid feeding than dispersed proteins in the sieve element lumen, because they would cover and encapsulate the stylet tip directly after penetration of the sieve element membrane (**Figure 8a**). In the absence of PHLO, phloem proteins would be dispersed in the sap and give room for passive entry of phloem into the food channel . In addition, this hypothesis would explain the reduced phloem exudation rate of *phlo* T-DNA plants, because they would have more proteinaceous obstacles in the sieve element lumen which could decrease phloem flow rate (**Figure 8b**). As an alternative hypothesis we consider the possibility

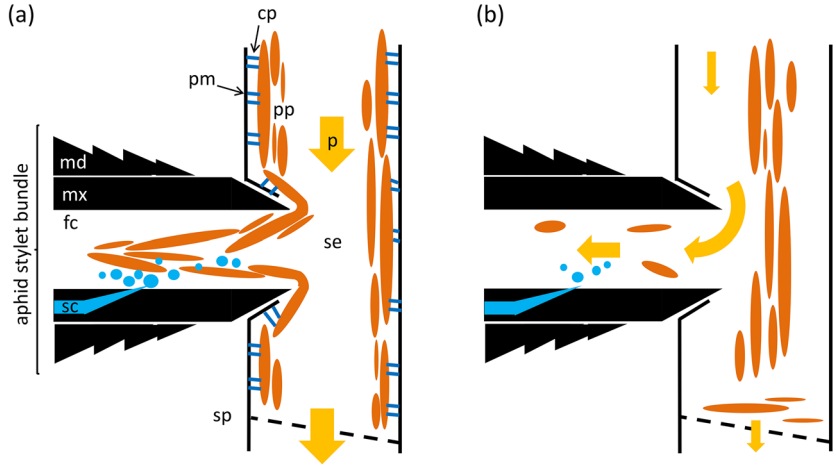



**Figure 8. Working model of PHLO as a clamp protein. a.** Aphid stylet penetration of a sieve element (se) in a wild type plant. Clamp proteins (cp) anchor phloem proteins (pp) to the plasma membrane (pm). When the stylet protrudes through the plasma membrane, sheets of protein aggregations encapsulate the stylet tip and block the aphid's food channel (fc). Aphids increase salivation to degrade the phloem proteins (more blue bubbles). **b.** In the *phlo* T-DNA line, phloem proteins are dispersed in the sieve element lumen and press against the sieve plate (sp). Phloem transport rate is decreased, but the stylet tip is free of protein agglomerations, facilitating the passive uptake of phloem sap (p) (md= mandibular stylets, mx= maxillary stylets, sc= salivary channel, dimensions are not drawn to scale).

that PHLO is a molecular chaperone involved in the transport of phloem proteins from companion cells to sieve elements (Bel *et al.*, 2002; Oparka and Turgeon, 1999) (**Figure 9**). If PHLO is such a chaperone, *phlo* T-DNA lines would lack phloem proteins in the sieve elements and allow unobstructed phloem ingestion by aphids. This scenario would, however, not explain the reduced phloem exudation rate in *phlo-1* compared to the wild type. Protein pull-down assays and imaging of individual sieve elements with fluorescently labelled PHLO and its target proteins will be instrumental in validating both hypotheses in follow-up work.



### The role PHLO and phloem proteins might play during moderate heat stress

Overall, our study shows that *PHLO* is temperature responsive and affects plant performance during moderate heat stress. Interestingly, Arabidopsis accessions with the resistant *PHLO* allele occur more often in regions with large temperature fluctuations throughout the year (**Table S6**, **Figure S2**), suggesting that natural variation in *PHLO* is relevant for plant fitness in these conditions. Seed dormancy was higher in *phlo-1*.

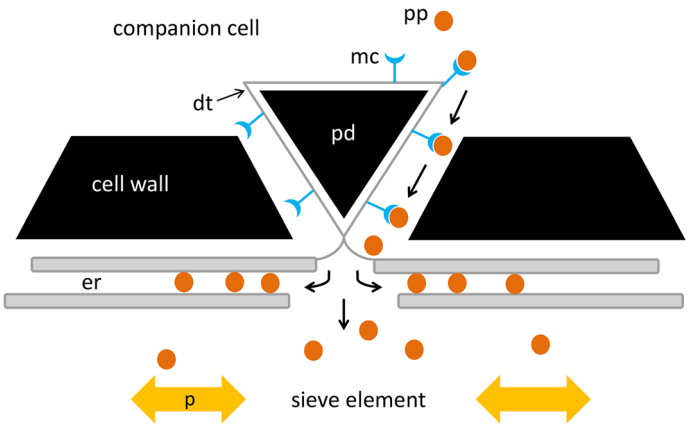



**Figure 9. Working model of PHLO as a molecular chaperone in the transport of phloem proteins from companion cells to sieve elements.** Molecular chaperones (mc) are attached to the desmotubule (dt) in plasmodesmata (pd) and transport phloem proteins (pp) to the sieve element, where they are embedded in the endoplasmatic reticulum (er) or disperse into the phloem sap (p) (picture adjusted from Oparka *et al.* (1999)).

Dormancy is known to reduce when parental plants experience high temperatures during seed maturation (He *et al.*, 2014), suggesting that T-DNA plants were more sensitive to high temperatures than Col-0. In addition, *phlo-1* plants displayed a stunted growth of the inflorescence stem at 26°C. Carbohydrate allocation is crucial during this stage of rapid elongation, and the *phlo-1* phenotype could, therefore, be caused by impaired phloem translocation. The question is, however, what type of stress high temperature imposes on the phloem. Phloem transport is affected by many temperature-sensitive processes, including photosynthesis and respiration, but most importantly, transpiration and water availability (Hölttä *et al.*, 2009; Lemoine *et al.*, 2013; Wahid *et al.*, 2007). High temperature will lead to increased transpiration and thus shortage in water availability. Since phloem translocation is driven by osmotically generated pressure-driven flow (Münch, 1926, 1927, 1930), reduced water availability will result in a reduced flow rate (Hall and Milburn, 1973; Hölttä *et al.*, 2009). Plants somehow need to maintain a constant phloem translocation rate when temperatures rise, and possibly even upregulate it due to a temperature-driven increase in metabolic activity. This might explain why we observed a higher phloem exudation rate in Col-0 at 26˚C compared to 20˚C (**Figure 5**). The next question is, what role phloem proteins might play during moderate heat stress. Previous studies have described that both heat and cooling induced changes in the state of phloem proteins (Furch *et al.*, 2007; Lemoine *et al.*, 2013; Thorpe *et*

*al.*, 2010). In an extensive literature review, Knoblauch *et al.* (2014) illustrate that most studies lack, however, empirical evidence that phloem proteins affect phloem flow rate. It was even demonstrated that the presence of ATSEOR1 and ATSEOR2 hardly reduce phloem translocation (Froelich *et al.*, 2011). Based on a phloem flow calculation model, Knoblauch *et al.* (2014) concluded that phloem protein agglomerations do not obstruct the volumetric flow rate, as long as the hydrostatic pressure increases correspondingly. Interestingly, the latter assumption is most likely violated during heat stress. In conditions of decreased water availability, the hydrostatic pressure will decrease and, according to the model, sieve element occlusion-related agglomerations could play a more dominant role in the obstruction of phloem translocation. We speculate that, to facilitate phloem translocation, it would be beneficial for plants to reduce the amount of phloem proteins in the sap or to retract them to the margins of sieve elements during heat stress (**Figure 8a**). If PHLO is anchoring phloem proteins to the margins of sieve elements, *phlo* knockout plants would have more dispersed phloem proteins (**Figure 8b**), and particularly during heat stress, experience a lower phloem flow rate than wild-type plants. Our phloem exudation results are not sufficient to validate this hypothesis, since we used detached leaves and cannot discriminate phloem exudates from carbohydrate leakage of other cells (Knoblauch *et al.*, 2014). Therefore, live tracking of phloem translocation in whole plants, for example with FRAP (fluorescence recovery after photo-bleaching) in combination with CFDA (carboxyfluorescein diacetate), is required to confirm the effect of PHLO and phloem proteins on the phloem flow rate during moderate heat stress.



## Materials and methods

### Plants and insects

A collection of 350 natural accessions of *Arabidopsis thaliana* (L.) Heynh. was obtained from the ABRC Stock Center (Baxter *et al.*, 2010). This set was selected in a previous study to represent most intraspecific genetic variation and minimal redundancy (Platt *et al.*, 2010), and was genotyped for approximately 214,000 SNPs with Atsnptile1 arrays (Atwell *et al.*, 2010; Horton *et al.*, 2012; Li *et al.*, 2010). The T-DNA lines SALK_027475 (*phlo-1*) and SAIL_1269_C01 (*phlo-2*), both in a Col-0 background, were obtained from the European Arabidopsis Stock Centre (NASC). Homozygous T-DNA plants were selected with PCR (**Table S7**) and harvested for seeds for subsequent experiments. The location of the T-DNA inserts was validated via sequencing and *At3g10670* and

*At3g10680* expression was measured with RT-qPCR (n=3, **Figure S1**). The natural accession C24 and the near-isogenic Col-0 line with a C24 introgression on chromosome 3 between position 276,917 and 3,463,232, were obtained from the Max Planck Institute of Molecular Plant Physiology in Golm, Germany (Törjék *et al.*, 2008). For all plant lines, seeds were cold stratified for 72 h at 4˚C before they were sown into pots (5 cm diameter) with pasteurized (4 h at 80˚C) Arabidopsis potting soil (Lentse potgrond, Lent, The Netherlands) in a climate room at 26 ± 1˚C, 50-70% relative humidity, an 8 h:16 h L:D photoperiod, and a light intensity of 200 µmol $m^{-2}$ $s^{-1}$. To assess temperature effects, plants were also grown in a climate cabinet at 20 ± 1˚C, 60-70% relative humidity, an 8 h:16 h L:D photoperiod, and light intensity of 120 µmol $m^{-2}$ $s^{-1}$. For assessment of light intensity effects on gene expression, light intensity in the climate cabinet was changed to 70 µmol $m^{-2}$ $s^{-1}$. Green peach aphids, *Myzus persicae* (Sulzer), were reared on radish, *Raphanus sativus* (L.), at 19˚C, 50-70% relative humidity and a 8 h:16 h L:D photoperiod. Only wingless aphids have been used throughout this study.

### Automated video tracking

Aphid behaviour was tracked on 350 natural accessions of Arabidopsis (n=3-6) according to the methodology of Kloth *et al.* (2015). One aphid was introduced into a well of a 96-well plate containing a leaf disc of 6 mm diameter, abaxial side up, on 1% agar substrate. Wells were covered with cling film, to avoid aphid escape, and 20 aphids were recorded on 20 different accessions simultaneously with a camera mounted above the plate, in a climate-controlled room at 22 ± 1˚C. EthoVision® XT 8.5 video tracking and analysis software (Noldus Information Technology bv, Wageningen, The Netherlands) was used for automated acquisition of aphid position and velocity. The number and duration of probes were subsequently calculated with the statistical computing program R (R-Core-Team, 2013). Leaf discs were made of intermediately aged leaves of 4- to 5-week-old Arabidopsis plants, one disc per plant. Aphid behaviour was recorded for 85 min, starting immediately after inoculation of the aphids. The video-tracking assay was performed in an incomplete block design with each complete replicate consisting of 18 blocks of 20 accessions. Sixty plants were screened each day across 3 blocks, and one replicate of the complete Hapmap collection was acquired in 6 days. An alpha design was generated with Gendex (http://designcomputing.net/gendex/) to assign accessions to each block.

### GWA mapping

GWA mapping was performed on the total time spent on short probes (< 3 min, data arcsine transformed) with scan_GLS (Kruijer *et al.*, 2015). Within scan_GLS, the

program asreml (VSN International Ltd.) was used to apply a mixed model to correct for population structure, which was estimated with a kinship matrix including all SNPs. SNPs with a minor allele frequency below 0.05 were excluded from analysis. Blocks and replicates were included in the model as covariates. Generalized heritability was calculated according to Oakey *et al.* (2006).

### LD, haplotypes and co-expression

Linkage disequilibrium (LD) between *At3g10680* and other genes was taken into account for polymorphisms in a 150 kb window around the highest SNP, with an LD of 0.3 as the lower threshold. LD was calculated using the re-sequenced data of more than 500 Arabidopsis accessions from the 1001 Genomes Project (www.1001genomes.org) (Cao *et al.*, 2011) using the SNP LD tool as described in Bac-Molenaar *et al.* (2015a). A haplotype analysis was based on 173 re-sequenced accessions from the GWA mapping population. For each domain, haplotypes were defined as unique SNP combinations with a frequency above 5%. A promoter region of 1000 kb was used and gene domains were obtained from Interpro (Mitchell *et al.*, 2015). In total 597 co-expressed genes of *At3g10680* were identified from 3 independent co-expression databases: (1) top 300 genes of Atted-II version c4.1 (Obayashi *et al.*, 2014), (2) top 100 genes of BAR Expression Angler Abiotic Stress, Pathogen, and Extended Tissue Compendium resp. (Toufighi *et al.*, 2005), and (3) the top 100 genes of Genevestigator's Pertubations data set (Zimmermann *et al.*, 2004).

### RT-qPCR

Two intermediately aged leaves per 4- to 5-week-old Arabidopsis plant were harvested between 12 and 3 pm. Samples were immediately frozen in liquid nitrogen, and stored at -80˚C until processing. RNA was isolated from homogenised leaf material with an InviTrap® Spin Plant RNA kit, and treated with Ambion® TURBO DNA-free™ according to the manufacturer's instructions. RNA was quantified with a NanoDrop® ND-1000 spectrophotometer, and integrity was assessed with gel electrophoresis. DNA-free RNA was converted into cDNA using the Bio-Rad iScript™ cDNA synthesis kit. Quantitative reverse transcription PCR was carried out on a Bio-Rad IQ™5 system using SYBR Green. For each primer combination (**Table S7**), RT-qPCR products were sequenced to validate the region of amplification. To test the effects of light intensity, samples were collected at 70 µmol $m^{-2}$ $s^{-1}$ (20˚C ) and 120 µmol $m^{-2}$ $s^{-1}$ (20˚C) (n=3). For temperature effects, samples were collected from plants grown at 20˚C and 26˚C (n=8). To test induction by aphids, samples with and without aphids were collected at 26˚C (n=4). Fifteen adult *M. persicae* aphids were contained on each leaf, using a Petri

dish with indentation for the petiole. Plants without aphids received an empty Petri dish. Samples were harvested 6 h post infestation.

### Electrical Penetration Graph recording

Feeding behaviour of *M. persicae* aphids was measured with EPG recording on 4- to 5-week-old plants, using direct currents (DC) according to the methodology of ten Broeke *et al.* (2013a). To adjust the radish-reared aphids to Arabidopsis, aphids were transferred to Col-0 Arabidopsis plants 24 h before the experiments. EPG recording was performed at 22 ± 2 °C and light intensity of 120 µmol $m^{-2}$ $s^{-1}$, using clean plants and one aphid per plant. An electrode was inserted in the potting soil and a thin gold wire of 1.5 cm was gently attached to the dorsum of an adult, wingless aphid with silver glue. The electrical circuit was completed when the aphid's piercing-sucking stylet mouthparts penetrated the plant cuticle. Electrical signals associated with stylet activities were recorded and annotated with EPG Stylet+ software and further processed in R (R-Core-Team, 2013; Tjallingii, 1988) (http://www.epgsystems.eu). Between 13 and 20 biological replicates were measured (20°C: Col-0 n=13, *phlo-1* n=13, *phlo-2* n=16; 26°C: Col-0 n=20, C24 n=14, Col-0xC24 n=14, *phlo-1* n=20, *phlo-2* n=16).

### Honeydew droplet collection

Aphid excretions were collected with honeydew clocks according to Tjallingii *et al.* (1995). Honeydew clocks consisted of a 12 h quartz clockwork (Hechinger, Villingen-Schwenningen, Germany) with a rotating plastic disc (20 cm diameter) attached. Honeydew droplets were collected on a 2-cm-wide TLC chromatography paper strip that completely covered the rear of the rotating disc. To visualise the droplets via amino acid staining, TLC paper was treated with ninhydrin (Acros organics ninhydrin spray, 0.5% 1-butanol solvent) and dried before the experiment. To adjust the radish-reared aphids to Arabidopsis, aphids were transferred to Col-0 Arabidopsis plants 24 h before the experiments. Adult, wingless aphids were positioned on the abaxial side of the leaves of clean plants, one aphid per plant. Honeydew clocks were placed 1 ± 0.5 cm underneath the position of the aphids. Honeydew collection and EPG recording were performed simultaneously during 12 h. Honeydew excretion rates were calculated for time intervals consisting of 8 h of continuous phloem ingestion (Col-0 n=3, *phlo-1* n=5).

### Aphid performance

To assess aphid developmental rate and population size, 3.5-week-old Arabidopsis plants were infested with 1 *M. persicae* neonate of 0- to 24-hour-old and placed in a climate

room at 26 ± 1°C, 50-70% relative humidity, an 8h:16 h L:D photoperiod, 200 µmol $m^{-2}$ $s^{-1}$ light intensity. A soap-diluted water barrier prevented aphids from moving between plants. None of the aphids developed wings. From day 7 onwards, occurrence of the first offspring was checked twice per day using 5x magnification glasses (Col-0 n=17, *phlo-1* n=20). The number of aphids per plant was counted 14 days after infestation. Plants without an adult aphid 8 days after introduction, and plants without any adults or neonates 14 days after introduction were excluded from the analysis. Aphid fresh weight was measured in a separate experiment. Five plants per plant line were infested with each 12 neonate aphids of 0- to 24-hour-old in a climate room at 26 ± 1°C, 50-70% relative humidity, an 8h:16 h L:D photoperiod, 200 µmol $m^{-2}$ $s^{-1}$ light intensity. After 8 days of infestation and before the onset of reproduction, aphids were weighed per plant, and the average fresh weight per aphid individual was calculated per plant.

### Callose measurements

Callose depositions were stained with aniline blue (Zhou, 2012). One leaf of intermediate age per 4.5-week-old plant was placed in 96% ethanol overnight, incubated in sodium phosphate buffer (70 mM, pH=9) for 1 hour and impregnated in 0.05% aniline blue (VWR Chemicals) solution of sodium phosphate buffer (70 mM, pH=9) for 2 days. Per plant line, the complete mid-vein was imaged with a fluorescent microscope, based on approximately 30 images (Nikon Instruments Europe BV, 100x, DM400 filter), for six biological replicates. Fluorescent areas were semi-quantified with ImageJ (Schneider *et al.*, 2012), counting all particles > 5 pixels using the threshold adjustment method "Triangle".

### Phloem exudation

Phloem exudates were collected according to Tetyuk *et al.* (Tetyuk *et al.*, 2013). For each sample, 15 leaves were collected from three 5.5-week-old plants grown at 20°C or 26°C. In order to chelate free calcium ions and prevent the sealing of the phloem, the fifteen leaves were stacked and the petioles were submerged in one reaction tube containing 1.4 ml 20 mM $K_2$-EDTA solution. The tubes were placed at 20 ± 2°C or 26 ± 2°C in a closed, transparent container, which was lined with wet filter paper to prevent excessive transpiration. After 1 h of incubation in the EDTA, the petioles were washed with distilled, deionized Millipore water to remove the EDTA. Immediately, the fifteen leaves were transferred into a tube with the petioles submerged in 1.4 ml autoclaved distilled, deionized Millipore water to collect phloem exudates. The tubes were placed back in the closed, transparent container at 20 ± 2°C or 26 ± 2°C. Exudates were collected for the first, second and third hour separately, by transferring the fifteen

leaves each hour into a fresh tube of water (20˚C: Col-0 n=4, *phlo-1*: n=4. 26˚C first and second hour: Col-0 n=4, *phlo-1*: n=4. 26˚C third hour: Col-0 n=2, *phlo-1*: n=2). Samples were immediately frozen in liquid nitrogen, and stored at -80°C. Sucrose, glucose and fructose content were measured by analysing 200 µl exudate with HPLC (Dionex™ ICS5000, Thermo Scientific™), using a Carbopac® PA1-column and guard column at a compartment temperature of 30˚C, and a 20-150 mM NaOH eluent gradient over 30 min and a 150 mM NaOH eluent for 10 min. The column was rinsed between runs for 5 min with 500 mM NaAc, followed by 10 min equilibration with 20 mM NaOH. Carbohydrate content was quantified with Chromeleon® 7.

### Seed germination

Seed germination values were measured with the Germinator package (Joosen *et al.*, 2010). Seeds were sown in trays containing two layers of blue germination paper (Anchor paper Co) with 50 ml of deionized water and depending on the treatment, were cold stratified at 4˚C for 72 hours prior to incubation or incubated immediately. For germination, the trays were placed under constant light at either 22˚C ± 0.5 or 25˚C ± 0.5. Three biological replicates were measured, each consisting of approximately 50 seeds collected from a different mother plant. Plants were all grown simultaneously at 26 ± 1˚C, 50-70% relative humidity, 4 weeks under 8 h:16 h and an additional 6 weeks under 16 h:8 h L:D photoperiod, with a light intensity of 200 µmol $m^{-2}$ $s^{-1}$. At the start of the experiments, seeds had experienced 2.5 weeks of dry storage.

### Statistics and data processing

Data were tested for normality and homogeneity of variances using the Shapiro test and Levene's test. In case of significant departure from normality, data were assessed with the Mann-Whitney U test (2 groups), or the Kruskall Wallis test (> 2 groups) otherwise. Normal distributed data were analysed using a one-way ANOVA. Aphid development was analysed with a Cox proportional hazards model. Honeydew droplet excretion rate was calculated as the slope of the linear model between cumulative number of droplets and time. Climate data were used to test for a haplotype effect using two approaches: (1) a Mann-Whitney U test, and (2) a mixed model correcting for population structure. For the latter, haplotype was defined as fixed effect and the first 8 principal components of the marker-based kinship matrix were included as random effects. Only European Arabidopsis accessions were tested for geographical distribution patterns and climate variables (latitude: 36°00'N to 63°00'N, longitude: 26°00'E to 40°00'W), due to the dense sampling compared to other regions. Statistical tests and graphs were constructed with R, using the packages "survival", "sciplot", "lawstat" and "asreml" (Butler *et al.*,

2009; Morales, 2011; R-Core-Team, 2013; Therneau, 2015).

**Acknowledgements**

This work was supported by The Netherlands Organization for Scientific Research (NWO) through the Technology Foundation Perspective Programme 'Learning from Nature' [STW10989]. We would like to thank Freddy Tjallingii for helpful instructions for the honeydew clock experiment, Diaan Jamar for performing the HPLC measurements, and Joost van Heerwaarden for compiling the climate data set of the Arabidopsis population.

## Supplementary information

**Table S1.** The total time *M. persicae* aphids spent on short probes (< 3 min) on 350 natural Arabidopsis accessions. Aphid behaviour was measured with automated video tracking during a 1.5-h observation. For each accession the mean value and standard error are shown (n=3-6), as well as the allele for the most significant SNP from the GWA mapping (*At3g10680*, pos. 3338114). The Col-0 allele (59%) was associated with an increased total time spent on short probes (Acc= accession, Lat= latitude, Long= longitude, Mean= mean total minutes, Se= standard error, H= haplotype (Y= Col-0 haplotype)).

| Acc | Line | Lat | Long | Mean | Se | H | Acc | Line | Lat | Long | Mean | Se | H |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CS76100 | Bor-4 | 49.4 | 16.23 | 0.32 | 0.1 | N | CS76101 | Br-0 | 49.2 | 16.62 | 1.99 | 0.43 | N |
| CS76195 | Na-1 | 47.5 | 1.5 | 0.58 | 0.24 | Y | CS28550 | NFC-20 | 51.41 | -0.64 | 2.01 | 0.29 | N |
| CS28210 | Do-0 | 50.72 | 8.24 | 0.83 | 0.23 | N | CS76167 | Lillö-1 | 56.15 | 15.78 | 2.08 | 0.63 | N |
| CS76133 | Ga-0 | 50.3 | 8 | 0.98 | 0.38 | Y | CS28564 | No-0 | 51.06 | 13.3 | 2.11 | 0.35 | Y |
| CS28758 | Tha-1 | 52.08 | 4.3 | 1.03 | 0.38 | N | CS28800 | Ven-1 | 52.03 | 5.55 | 2.13 | 0.46 | N |
| CS76165 | LI-OF-095 | 40.78 | -72.91 | 1.05 | 0.21 | N | CS76221 | ROM-1 | 45.53 | 4.85 | 2.18 | 0.27 | Y |
| CS28395 | Kn-0 | 54.9 | 23.89 | 1.07 | 0.28 | Y | CS76280 | UKSE06-062 | 51.3 | 0.5 | 2.19 | 0.28 | N |
| CS22689 | RRS-10 | 41.56 | -86.43 | 1.22 | 0.2 | N | CS28809 | Wag-4 | 51.97 | 5.67 | 2.19 | 0.48 | N |
| CS28651 | Pr-0 | 50.14 | 8.61 | 1.28 | 0.26 | Y | CS28018 | Ang-0 | 50.3 | 5.3 | 2.22 | 0.51 | N |
| CS76210 | Per-1 | 58 | 56.32 | 1.45 | 0.29 | N | CS76160 | LDV-14 | 48.52 | -4.07 | 2.26 | 0.5 | N |
| CS28282 | Go-0 | 51.53 | 9.94 | 1.46 | 0.36 | N | CS76241 | T690 | 55.84 | 13.31 | 2.27 | 0.66 | Y |
| CS76094 | Bay-0 | 49 | 11 | 1.5 | 0.52 | Y | CS76230 | Sq-8 | 51.41 | -0.64 | 2.29 | 0.34 | Y |
| CS76267 | TOU-K-3 | 46.67 | 4.12 | 1.51 | 0.18 | N | CS28364 | Je-0 | 50.93 | 11.59 | 2.33 | 0.55 | Y |
| CS76142 | Hov4-1 | 56.1 | 13.74 | 1.74 | 0.57 | N | CS76147 | In-0 | 47.5 | 11.5 | 2.34 | 0.33 | Y |
| CS76102 | Brö1-6 | 56.3 | 16 | 1.75 | 0.39 | N | CS28097 | Bs-2 | 47.5 | 7.5 | 2.34 | 0.49 | Y |
| CS76132 | Fjä1-5 | 56.06 | 14.29 | 1.79 | 0.26 | N | CS28729 | Sei-0 | 46.54 | 11.56 | 2.37 | 0.53 | Y |
| CS76104 | BUI | 48.37 | 0.93 | 1.79 | 0.57 | Y | CS28743 | Sp-0 | 52.53 | 13.18 | 2.37 | 0.59 | Y |
| CS28822 | WI-0 | 47.93 | 10.81 | 1.79 | 0.37 | Y | CS28007 | Aa-0 | 50.92 | 9.57 | 2.38 | 0.4 | N |
| CS28622 | PHW-22 | 51.42 | -1.72 | 1.8 | 0.3 | N | CS76201 | Ör-1 | 56.45 | 16.11 | 2.39 | 0.31 | N |
| CS76153 | Kin-0 | 44.46 | -85.37 | 1.81 | 0.29 | Y | CS76245 | TDr-1 | 55.77 | 14.14 | 2.4 | 0.42 | N |
| CS76189 | MOG-37 | 48.67 | -4.07 | 1.83 | 0.36 | N | CS28133 | Cha-0 | 46.03 | 7.12 | 2.45 | 0.48 | N |
| CS28193 | Com-1 | 49.42 | 2.82 | 1.83 | 0.34 | N | CS28202 | Db-0 | 50.31 | 8.32 | 2.48 | 0.38 | Y |
| CS76117 | Dra3-1 | 55.76 | 14.12 | 1.83 | 0.29 | N | CS76170 | Lis-2 | 56 | 14.7 | 2.48 | 0.35 | Y |
| CS76125 | Eden-2 | 62.88 | 18.18 | 1.84 | 0.22 | Y | CS76183 | MIB-28 | 47.38 | 5.32 | 2.48 | 0.33 | Y |
| CS28236 | Ep-0 | 50.17 | 8.39 | 1.89 | 0.47 | Y | CS76137 | Gr-1 | 47 | 15.5 | 2.5 | 0.3 | Y |
| CS76203 | Oy-0 | 60.23 | 6.13 | 1.9 | 0.27 | N | CS76089 | ALL1-2 | 45.27 | 1.48 | 2.51 | 0.46 | N |
| CS76239 | T540 | 55.8 | 13.1 | 1.97 | 0.34 | Y | CS28713 | RRS-7 | 41.56 | -86.43 | 2.51 | 0.45 | N |

**Table S1.** (continued)

| Acc | Line | Lat | Long | Mean | Se | H |
|---|---|---|---|---|---|---|
| CS76278 | UKNW-06436 | 54.7 | -3.4 | 2.52 | 0.42 | N |
| CS28063 | Be-1 | 49.68 | 8.62 | 2.54 | 0.38 | Y |
| CS76095 | Belmonte -4-94 | 42.12 | 12.48 | 2.56 | 0.51 | N |
| CS76244 | Tamm-2 | 60 | 23.5 | 2.57 | 0.44 | N |
| CS28786 | Ty-0 | 56.43 | -5.23 | 2.58 | 0.29 | Y |
| CS76300 | VOU-2 | 46.65 | 0.17 | 2.61 | 0.42 | N |
| CS76247 | TDr-18 | 55.77 | 14.12 | 2.63 | 0.55 | N |
| CS76275 | UKNW-06059 | 54.4 | -3 | 2.64 | 0.66 | N |
| CS76161 | LDV-25 | 48.52 | -4.07 | 2.65 | 0.54 | N |
| CS76289 | UKSE06-482 | 51.2 | 0.6 | 2.68 | 0.61 | N |
| CS28663 | Pu2-24 | 49.42 | 16.36 | 2.69 | 0.37 | Y |
| CS28787 | Uk-1 | 48.03 | 7.77 | 2.7 | 0.35 | Y |
| CS76150 | Kas-1 | 35 | 77 | 2.72 | 0.34 | Y |
| CS28724 | Sapporo-0 | 43.06 | 141.35 | 2.72 | 0.36 | Y |
| CS76291 | UKSE06-628 | 51.1 | 0.4 | 2.74 | 0.35 | N |
| CS28760 | Tiv-1 | 41.96 | 12.8 | 2.79 | 0.43 | N |
| CS76124 | Duk | 49.1 | 16.2 | 2.8 | 0.47 | Y |
| CS76231 | St-0 | 59 | 18 | 2.8 | 0.27 | Y |
| CS76253 | TOU-A1-116 | 46.67 | 4.12 | 2.85 | 0.61 | Y |
| CS28158 | Cit-0 | 43.38 | 2.54 | 2.87 | 0.29 | Y |
| CS76226 | Se-0 | 38.33 | -3.53 | 2.95 | 0.75 | N |
| CS76285 | UKSE06-351 | 51.3 | 0.4 | 2.96 | 0.49 | N |
| CS76266 | TOU-J-3 | 46.67 | 4.12 | 3 | 0.55 | N |
| CS76129 | Fei-0 | 40.5 | -8.32 | 3.01 | 0.64 | N |
| CS28461 | Li-7 | 50.38 | 8.07 | 3.02 | 0.26 | Y |
| CS76242 | Ta-0 | 49.5 | 14.5 | 3.02 | 0.61 | Y |
| CS76257 | TOU-A1-67 | 46.67 | 4.12 | 3.02 | 0.5 | Y |
| CS28573 | Nw-0 | 50.5 | 8.5 | 3.02 | 0.83 | Y |
| CS76298 | Vår2-1 | 55.58 | 14.33 | 3.02 | 0.49 | Y |
| CS76136 | Got-7 | 51.53 | 9.94 | 3.05 | 0.75 | N |
| CS76086 | 627ME-4Y1 | 42.09 | -86.36 | 3.1 | 0.72 | Y |
| CS76283 | UKSE06-278 | 51.3 | 0.4 | 3.11 | 0.39 | N |
| CS76085 | 328PNA054 | 42.09 | -86.33 | 3.11 | 0.54 | Y |
| CS76135 | Ge-0 | 46.5 | 6.08 | 3.12 | 0.76 | N |
| CS28181 | CSHL-5 | 40.86 | -73.47 | 3.14 | 0.7 | Y |
| CS76144 | HR-5 | 51.41 | -0.64 | 3.2 | 0.53 | N |
| CS28280 | Gie-0 | 50.58 | 8.68 | 3.22 | 0.63 | Y |
| CS28274 | Ga-2 | 50.3 | 8 | 3.23 | 0.5 | Y |
| CS28142 | CIBC-5 | 51.41 | -0.64 | 3.23 | 0.37 | Y |
| CS76110 | Cen-0 | 49 | 0.5 | 3.24 | 0.64 | Y |
| CS76206 | PAR-4 | 46.65 | -0.25 | 3.25 | 0.58 | N |
| CS76123 | DraIV | 49.41 | 16.28 | 3.26 | 0.46 | Y |
| CS28810 | Wag-5 | 51.97 | 5.67 | 3.26 | 0.38 | Y |
| CS76292 | UKSW-06202 | 50.4 | -4.9 | 3.26 | 0.63 | N |
| CS76114 | Ct-1 | 37.3 | 15 | 3.28 | 0.79 | N |
| CS76090 | ALL1-3 | 45.27 | 1.48 | 3.3 | 0.46 | N |
| CS76287 | UKSE06-429 | 51.3 | 0.4 | 3.3 | 0.61 | N |
| CS28739 | Si-0 | 50.87 | 8.02 | 3.33 | 0.51 | N |
| CS76279 | UKNW-06460 | 54.7 | -3.4 | 3.33 | 0.3 | N |
| CS28344 | Hey-1 | 51.25 | 5.9 | 3.34 | 0.6 | Y |
| CS76172 | LL-0 | 41.59 | 2.49 | 3.34 | 0.52 | N |
| CS76139 | Gy-0 | 49 | 2 | 3.36 | 0.83 | Y |
| CS28252 | Fi-1 | 50.5 | 8.02 | 3.37 | 0.52 | Y |
| CS28017 | An-2 | 51.22 | 4.4 | 3.37 | 0.36 | Y |
| CS76299 | VOU-1 | 46.65 | 0.17 | 3.39 | 0.58 | N |
| CS76209 | Pent-1 | 43.76 | -86.39 | 3.42 | 0.55 | Y |
| CS76217 | Rak-2 | 49 | 16 | 3.42 | 0.97 | Y |
| CS76151 | KBS-Mac-8 | 42.41 | -85.4 | 3.42 | 0.51 | N |
| CS76084 | 11PNA4.101 | 42.09 | -86.33 | 3.43 | 0.91 | N |

**Table S1.** (continued)

| Acc | Line | Lat | Long | Mean | Se | H |
|---|---|---|---|---|---|---|
| CS28628 | PHW-28 | 50.35 | -3.58 | 3.46 | 0.33 | Y |
| CS76252 | TOU-A1-115 | 46.67 | 4.12 | 3.49 | 0.61 | N |
| CS28633 | PHW-33 | 52.25 | 4.57 | 3.5 | 0.69 | N |
| CS76187 | MNF-Pot-48 | 43.59 | -86.27 | 3.5 | 0.48 | N |
| CS76207 | PAR-5 | 46.65 | -0.25 | 3.53 | 0.82 | N |
| CS28575 | Nw-2 | 50.5 | 8.5 | 3.53 | 1.14 | Y |
| CS76191 | Mrk-0 | 49 | 9.3 | 3.56 | 0.64 | Y |
| CS76276 | UKNW-06060 | 54.4 | -3 | 3.59 | 0.46 | N |
| CS76105 | Bur-0 | 54.1 | -6.2 | 3.59 | 0.67 | Y |
| CS76288 | UKSE06-466 | 51.2 | 0.4 | 3.6 | 0.66 | Y |
| CS76162 | LDV-34 | 48.52 | -4.07 | 3.62 | 0.88 | N |
| CS76097 | Bla-1 | 41.68 | 2.8 | 3.62 | 0.79 | N |
| CS76254 | TOU-A1-12 | 46.67 | 4.12 | 3.63 | 0.62 | Y |
| CS76220 | Rmx-A180 | 42.04 | -86.51 | 3.64 | 0.62 | N |
| CS76107 | CAM-16 | 48.27 | -4.58 | 3.67 | 0.81 | N |
| CS76281 | UKSE06-192 | 51.3 | 0.5 | 3.68 | 0.5 | N |
| CS76113 | Col-0 | 38.3 | -92.3 | 3.68 | 0.74 | Y |
| CS28631 | PHW-31 | 51.47 | -3.2 | 3.68 | 0.48 | Y |
| CS76274 | UKID80 | 54.7 | -2.9 | 3.71 | 0.88 | N |
| CS28640 | Pla-0 | 41.5 | 2.25 | 3.71 | 0.8 | N |
| CS28140 | CIBC-2 | 51.41 | -0.64 | 3.72 | 0.64 | Y |
| CS76141 | Hod | 48.8 | 17.1 | 3.73 | 0.69 | N |
| CS76098 | Blh-1 | 48 | 19 | 3.73 | 0.51 | N |
| CS76146 | HSm | 49.33 | 15.76 | 3.74 | 0.49 | Y |
| CS76121 | DraIV | 49.41 | 16.28 | 3.74 | 0.71 | Y |
| CS76213 | Pna-17 | 42.09 | -86.33 | 3.74 | 0.53 | Y |
| CS76176 | Lp2-2 | 49.38 | 16.81 | 3.77 | 0.47 | Y |
| CS76308 | ZdrI | 49.39 | 16.25 | 3.77 | 1.09 | Y |
| CS28051 | Arby-1 | 59.43 | 16.8 | 3.79 | 0.52 | N |
| CS76199 | NFA-8 | 51.41 | -0.64 | 3.81 | 0.65 | Y |
| CS28407 | KNO-11 | 41.28 | -86.62 | 3.82 | 0.47 | N |
| CS28804 | Wa-1 | 52.3 | 21 | 3.85 | 0.67 | Y |
| CS28650 | Pog-0 | 49.27 | -123.21 | 3.86 | 0.32 | N |
| CS76193 | Mz-0 | 50.3 | 8.3 | 3.9 | 0.8 | N |
| CS76156 | Nord-borg | 55.7 | 13.2 | 3.9 | 0.61 | N |
| CS28163 | Co-2 | 40.12 | -8.25 | 3.91 | 0.46 | Y |
| CS28636 | PHW-36 | 48.61 | 2.31 | 3.92 | 0.57 | Y |
| CS76240 | T620 | 55.7 | 13.2 | 3.94 | 0.74 | Y |
| CS76218 | Ren-1 | 48.5 | -1.41 | 3.95 | 0.49 | Y |
| CS76200 | Ömö2-1 | 56.14 | 15.78 | 3.96 | 0.73 | Y |
| CS76295 | Ull3-4 | 56.06 | 13.97 | 4.02 | 0.79 | N |
| CS28614 | PHW-14 | 51.29 | 0.06 | 4.02 | 0.79 | N |
| CS76148 | JEA | 43.68 | 7.33 | 4.03 | 0.57 | N |
| CS76099 | Bor-1 | 49.4 | 16.23 | 4.09 | 0.49 | Y |
| CS76234 | T1060 | 55.65 | 13.22 | 4.09 | 0.33 | Y |
| CS76251 | Tottarp-2 | 55.95 | 13.85 | 4.12 | 0.64 | Y |
| CS76255 | TOU-A1-43 | 46.67 | 4.12 | 4.13 | 0.65 | Y |
| CS28459 | Li-6 | 50.38 | 8.07 | 4.18 | 0.57 | N |
| CS76235 | T1080 | 55.66 | 13.22 | 4.2 | 0.47 | Y |
| CS76215 | Pu2-23 | 49.42 | 16.36 | 4.22 | 0.83 | N |
| CS76261 | TOU-H-12 | 46.67 | 4.12 | 4.24 | 0.63 | Y |
| CS28090 | Blh-2 | 48 | 19 | 4.26 | 0.58 | N |
| CS76302 | Wil-1 | 54.68 | 25.32 | 4.27 | 1.15 | N |
| CS76166 | Liarum | 55.95 | 13.85 | 4.3 | 0.67 | N |
| CS76111 | CIBC-17 | 51.41 | -0.64 | 4.34 | 0.6 | Y |
| CS28725 | Sav-0 | 49.18 | 15.88 | 4.36 | 0.67 | Y |
| CS76202 | Ost-0 | 60.25 | 18.37 | 4.37 | 0.99 | N |
| CS28595 | Pa-2 | 38.07 | 13.22 | 4.37 | 0.74 | N |
| CS28635 | PHW-35 | 48.61 | 2.31 | 4.38 | 0.36 | N |
| CS76256 | TOU-A1-62 | 46.67 | 4.12 | 4.4 | 0.81 | N |
| CS76249 | TDr-8 | 55.77 | 14.13 | 4.4 | 0.37 | Y |

**Table S1.** (continued)

| Acc | Line | Lat | Long | Mean | Se | H |
|---|---|---|---|---|---|---|
| CS28823 | Ws | 52.3 | 30 | 4.41 | 0.65 | Y |
| CS28492 | Mh-0 | 50.95 | 7.5 | 4.41 | 0.56 | Y |
| CS28848 | Ors-1 | 44.72 | 22.4 | 4.43 | 0.54 | N |
| CS28510 | N4 | 61.36 | 34.15 | 4.44 | 1.05 | Y |
| CS76173 | Lm-2 | 48 | 0.5 | 4.47 | 0.65 | N |
| CS76265 | TOU-I-6 | 46.67 | 4.12 | 4.48 | 0.7 | Y |
| CS28165 | Co-4 | 40.12 | -8.25 | 4.48 | 0.6 | N |
| CS28626 | PHW-26 | 50.67 | -3.84 | 4.55 | 0.73 | Y |
| CS76096 | Bg-2 | 47.65 | -122.31 | 4.55 | 0.68 | N |
| CS28135 | Chat-1 | 48.07 | 1.34 | 4.57 | 0.57 | Y |
| CS76194 | N13 | 61.36 | 34.15 | 4.57 | 0.79 | Y |
| CS76305 | Yo-0 | 37.45 | -119.35 | 4.6 | 0.24 | Y |
| CS76119 | DraIV | 49.41 | 16.28 | 4.6 | 0.81 | Y |
| CS28692 | Rou-0 | 49.44 | 1.1 | 4.63 | 0.94 | Y |
| CS76169 | Lis-1 | 56 | 14.7 | 4.63 | 0.64 | N |
| CS76197 | Nd-1 | 50 | 10 | 4.64 | 0.97 | Y |
| CS76205 | PAR-3 | 46.65 | -0.25 | 4.66 | 0.89 | N |
| CS76143 | Hovda-la-2 | 56.1 | 13.74 | 4.67 | 0.84 | Y |
| CS76155 | Köln | 51 | 7 | 4.71 | 0.65 | Y |
| CS28685 | Rhen-1 | 51.97 | 5.57 | 4.75 | 0.8 | Y |
| CS28457 | Li-5:2 | 50.38 | 8.07 | 4.76 | 0.97 | Y |
| CS76236 | T1110 | 55.6 | 13.2 | 4.76 | 0.88 | Y |
| CS28578 | Nz1 | -37.79 | 175.28 | 4.78 | 0.45 | N |
| CS76158 | LAC-5 | 47.7 | 6.82 | 4.78 | 0.47 | Y |
| CS28849 | Ors-2 | 44.72 | 22.4 | 4.8 | 0.6 | Y |
| CS28610 | PHW-10 | 51.29 | 0.06 | 4.81 | 0.76 | Y |
| CS28201 | Da(1)-12 | NA | NA | 4.82 | 0.64 | N |
| CS28814 | Wc-2 | 52.6 | 10.07 | 4.83 | 0.74 | Y |
| CS76168 | Lip-0 | 50 | 19.3 | 4.84 | 0.62 | Y |
| CS76290 | UKSE06-520 | 51.3 | 1.1 | 4.84 | 1.14 | N |
| CS28812 | WAR | 41.73 | -71.28 | 4.86 | 1.05 | Y |
| CS76174 | Lom1-1 | 56.09 | 13.9 | 4.86 | 0.63 | Y |
| CS28053 | Ba-1 | 56.55 | -4.8 | 4.87 | 0.95 | N |
| CS76127 | Est-1 | 58.3 | 25.3 | 4.87 | 0.72 | Y |
| CS76108 | CAM-61 | 48.27 | -4.58 | 4.87 | 0.56 | N |
| CS76179 | Lz-0 | 46 | 3.3 | 4.9 | 0.97 | N |
| CS76185 | MNF-Che-2 | 43.53 | -86.18 | 4.92 | 0.52 | Y |
| CS76269 | Udul | 49.28 | 16.63 | 4.95 | 0.9 | Y |
| CS76286 | UKSE06-414 | 51.3 | 0.4 | 4.95 | 0.71 | N |
| CS76284 | UKSE06-349 | 51.3 | 0.4 | 4.97 | 0.72 | N |
| CS76260 | TOU-E-11 | 46.67 | 4.12 | 4.97 | 0.48 | Y |
| CS76115 | CUR-3 | 45 | 1.75 | 4.98 | 0.32 | N |
| CS76112 | CLE-6 | 48.92 | -0.48 | 4.99 | 0.69 | Y |
| CS28141 | CIBC-4 | 51.41 | -0.64 | 5.01 | 0.46 | Y |
| CS76164 | Ler-1 | 47.98 | 10.87 | 5.02 | 0.96 | Y |
| CS76216 | Ra-0 | 46 | 3.3 | 5.03 | 0.71 | N |
| CS76128 | Fäb-4 | 63.02 | 18.32 | 5.06 | 0.81 | Y |
| CS76219 | Rev-2 | 55.7 | 13.4 | 5.06 | 1.23 | Y |
| CS28423 | Krot-2 | 49.63 | 11.57 | 5.06 | 0.92 | N |
| CS76149 | Ka-0 | 47 | 14 | 5.06 | 0.56 | Y |
| CS76263 | TOU-I-17 | 46.67 | 4.12 | 5.08 | 0.62 | N |
| CS28779 | Tscha-1 | 47.07 | 9.9 | 5.1 | 0.41 | N |
| CS76188 | MNF-Pot-68 | 43.59 | -86.27 | 5.15 | 0.45 | Y |
| CS28160 | Cnt-1 | 51.3 | 1.1 | 5.15 | 0.67 | Y |
| CS28513 | N7 | 61.36 | 34.15 | 5.16 | 0.95 | N |
| CS28200 | Da-0 | 49.87 | 8.65 | 5.16 | 0.69 | Y |
| CS76154 | Kno-18 | 41.28 | -86.62 | 5.18 | 1.1 | Y |
| CS76228 | SLSP-30 | 43.66 | -86.5 | 5.18 | 0.57 | Y |
| CS28613 | PHW-13 | 51.29 | 0.06 | 5.22 | 0.75 | N |
| CS76116 | Cvi-0 | 15.11 | -23.62 | 5.23 | 0.86 | Y |
| CS28382 | Kelster-bach-2 | 50.07 | 8.53 | 5.24 | 0.9 | Y |
| CS76273 | UKID48 | 54.7 | -2.7 | 5.27 | 0.51 | Y |

**Table S1.** (continued)

| Acc | Line | Lat | Long | Mean | Se | H |
|---|---|---|---|---|---|---|
| CS28780 | Tsu-0 | 34.43 | 136.31 | 5.27 | 0.56 | Y |
| CS76222 | Rsch-4 | 56.3 | 34 | 5.27 | 0.55 | Y |
| CS76109 | Can-0 | 29.21 | -13.48 | 5.3 | 0.66 | Y |
| CS76214 | Pro-0 | 43.25 | -6 | 5.31 | 0.69 | N |
| CS28369 | Jl-3 | 49.2 | 16.62 | 5.33 | 0.77 | Y |
| CS76083 | 11ME-1.32 | 42.09 | -86.36 | 5.34 | 0.41 | N |
| CS28108 | Bu-8 | 50.5 | 9.5 | 5.36 | 0.85 | Y |
| CS28373 | Jm-1 | 49 | 15 | 5.38 | 1.27 | Y |
| CS76227 | Shah-dara | 38.35 | 68.48 | 5.4 | 0.87 | Y |
| CS28419 | Kr-0 | 51.33 | 6.56 | 5.4 | 0.8 | Y |
| CS76192 | Mt-0 | 32.34 | 22.46 | 5.41 | 0.97 | Y |
| CS76196 | NC-6 | 35 | -79.18 | 5.42 | 0.47 | Y |
| CS76270 | UKID101 | 53.2 | -1.4 | 5.43 | 1.06 | Y |
| CS28750 | Ste-0 | 52.61 | 11.86 | 5.44 | 0.71 | Y |
| CS76106 | C24 | 41.25 | -8.45 | 5.44 | 0.9 | N |
| CS28128 | Ca-0 | 50.3 | 8.27 | 5.46 | 0.74 | Y |
| CS76225 | Sav-0 | 49.18 | 15.88 | 5.47 | 0.84 | Y |
| CS28336 | Ha-0 | 52.37 | 9.74 | 5.49 | 0.77 | Y |
| CS76103 | Bu-0 | 50.5 | 9.5 | 5.49 | 0.96 | N |
| CS76175 | Löv-5 | 62.8 | 18.08 | 5.58 | 0.46 | Y |
| CS76152 | Kelster-bach-4 | 50.07 | 8.53 | 5.59 | 1.37 | Y |
| CS28645 | Pn-0 | 48.07 | -2.97 | 5.62 | 0.83 | N |
| CS28720 | S96 | NA | NA | 5.64 | 0.83 | Y |
| CS76293 | Ull2-3 | 56.06 | 13.97 | 5.69 | 0.93 | Y |
| CS76182 | MIB-22 | 47.38 | 5.32 | 5.7 | 0.88 | N |
| CS28583 | Old-1 | 53.17 | 8.2 | 5.71 | 0.89 | Y |
| CS28277 | Ge-1 | 46.5 | 6.08 | 5.72 | 0.64 | Y |
| CS76198 | NFA-10 | 51.41 | -0.64 | 5.73 | 0.94 | Y |
| CS28587 | Or-0 | 50.38 | 8.01 | 5.77 | 0.83 | Y |
| CS28241 | Es-0 | 60.2 | 24.57 | 5.78 | 1.86 | Y |
| CS76208 | Paw-3 | 42.15 | -86.43 | 5.78 | 0.63 | N |
| CS76211 | Petergof | 59 | 29 | 5.8 | 1.09 | Y |
| CS76120 | DraIV | 49.41 | 16.28 | 5.81 | 0.67 | Y |
| CS76157 | LAC-3 | 47.7 | 6.82 | 5.82 | 1.29 | Y |
| CS28099 | Bsch-0 | 40.02 | 8.67 | 5.82 | 1.18 | Y |
| CS28420 | Kro-0 | 50.07 | 8.97 | 5.83 | 0.78 | Y |
| CS76180 | Map-42 | 42.17 | -86.41 | 5.87 | 0.88 | N |
| CS76145 | Hs-0 | 52.24 | 9.44 | 5.89 | 0.89 | N |
| CS76223 | Sanna-2 | 62.69 | 18 | 5.9 | 0.89 | N |
| CS76296 | Uod-7 | 48.3 | 14.45 | 5.91 | 1.53 | N |
| CS28350 | Hn-0 | 51.35 | 8.29 | 5.91 | 0.58 | Y |
| CS28064 | Benk-1 | 52 | 5.67 | 5.92 | 0.76 | Y |
| CS76301 | Wei-0 | 47.25 | 8.26 | 5.92 | 0.98 | Y |
| CS76093 | Bå1-2 | 56.4 | 12.9 | 5.94 | 0.65 | Y |
| CS28637 | PHW-37 | 48.61 | 2.31 | 5.94 | 0.86 | N |
| CS76248 | TDr-3 | 55.77 | 14.14 | 5.94 | 0.94 | N |
| CS76122 | DraIV | 49.41 | 16.28 | 5.95 | 0.89 | Y |
| CS76092 | App1-16 | 56.33 | 15.97 | 5.96 | 1.1 | Y |
| CS28049 | Ann-1 | 45.9 | 6.13 | 5.99 | 0.46 | Y |
| CS28833 | Wt-3 | 52.3 | 9.3 | 5.99 | 0.53 | Y |
| CS28527 | Nc-1 | 48.62 | 6.25 | 6.05 | 0.76 | Y |
| CS28243 | Est-0 | 58.3 | 25.3 | 6.05 | 0.84 | N |
| CS76140 | Hi-0 | 52 | 5 | 6.1 | 0.92 | Y |
| CS76232 | Ste-3 | 42.03 | -86.51 | 6.13 | 0.69 | Y |
| CS76088 | Alc-0 | 40.31 | -3.22 | 6.15 | 0.68 | Y |
| CS76238 | T510 | 55.79 | 13.12 | 6.16 | 0.83 | Y |
| CS76304 | Wt-5 | 52.3 | 9.3 | 6.17 | 0.84 | N |
| CS28208 | Di-1 | 47 | 5 | 6.17 | 0.75 | Y |
| CS76181 | MIB-15 | 47.38 | 5.32 | 6.18 | 0.77 | N |
| CS28847 | Zu-1 | 47.37 | 8.55 | 6.19 | 0.62 | Y |
| CS76277 | UKNW-06386 | 54.6 | -3.1 | 6.19 | 1 | N |
| CS28268 | Fr-4 | 50.11 | 8.68 | 6.21 | 0.89 | Y |
| CS76212 | PHW-34 | 48.61 | 2.31 | 6.28 | 0.78 | Y |
| CS76138 | Gul1-2 | 56.3 | 16 | 6.28 | 0.45 | Y |

**Table S1.** (continued)

| Acc | Line | Lat | Long | Mean | Se | H |
|---|---|---|---|---|---|---|
| CS76262 | TOU-H-13 | 46.67 | 4.12 | 6.34 | 1.52 | Y |
| CS28214 | Dra-2 | 49.42 | 16.27 | 6.35 | 1.33 | Y |
| CS76268 | Ts-1 | 41.72 | 2.93 | 6.36 | 1.08 | N |
| CS28732 | Sg-1 | 47.67 | 9.5 | 6.36 | 1.1 | Y |
| CS28394 | Kl-5 | 50.95 | 6.97 | 6.38 | 0.58 | Y |
| CS28332 | Gu-1 | 50.3 | 8 | 6.42 | 1.21 | Y |
| CS76233 | T1040 | 55.65 | 13.21 | 6.51 | 0.88 | Y |
| CS76190 | Mr-0 | 44.15 | 9.65 | 6.51 | 1.58 | Y |
| CS76126 | Edi-0 | 56 | -3 | 6.52 | 0.85 | N |
| CS28568 | Nok-1 | 52.24 | 4.45 | 6.52 | 0.84 | N |
| CS28620 | PHW-20 | 51.29 | 0.06 | 6.64 | 0.81 | Y |
| CS76258 | TOU-A1-96 | 46.67 | 4.12 | 6.71 | 1.2 | N |
| CS76131 | Fjä1-2 | 56.06 | 14.29 | 6.72 | 1.4 | Y |
| CS28795 | Utrecht | 52.09 | 5.11 | 6.76 | 0.91 | Y |
| CS28054 | Baa-1 | 51.33 | 6.1 | 6.85 | 0.75 | Y |
| CS76250 | To-megap-2 | 55.7 | 13.2 | 6.94 | 1 | Y |
| CS28734 | Sh-0 | 51.68 | 10.21 | 6.96 | 0.92 | Y |
| CS76134 | Gd-1 | 53.5 | 10.5 | 6.96 | 1.51 | Y |
| CS76272 | UKID37 | 51.3 | 1.1 | 6.97 | 0.72 | Y |
| CS28788 | Uk-2 | 48.03 | 7.77 | 6.98 | 0.8 | Y |
| CS76163 | LDV-58 | 48.52 | -4.07 | 7.09 | 0.75 | Y |
| CS28014 | Amel-1 | 53.45 | 5.73 | 7.15 | 1.11 | Y |
| CS76297 | Van-0 | 49.3 | -123 | 7.17 | 1.54 | N |
| CS28454 | Li-3 | 50.38 | 8.07 | 7.22 | 0.61 | Y |
| CS28808 | Wag-3 | 51.97 | 5.67 | 7.23 | 1.07 | Y |
| CS76259 | TOU-C-3 | 46.67 | 4.12 | 7.29 | 0.77 | N |
| CS28013 | Alst-1 | 54.8 | -2.43 | 7.33 | 1.79 | N |
| CS76177 | Lp2-6 | 49.38 | 16.81 | 7.36 | 0.47 | Y |
| CS76303 | Ws-0 | 52.3 | 30 | 7.47 | 1.03 | Y |
| CS76186 | MNF-Jac-32 | 43.52 | -86.17 | 7.5 | 0.85 | Y |
| CS28490 | Mc-0 | 54.62 | -2.3 | 7.62 | 1.02 | N |
| CS28326 | Gr-5 | 47 | 15.5 | 7.67 | 0.9 | Y |
| CS76229 | Sparta-1 | 55.71 | 13.05 | 7.69 | 1.02 | Y |
| CS76224 | Sap-0 | 49.49 | 14.24 | 7.7 | 0.95 | N |
| CS28759 | Ting-1 | 56.5 | 14.9 | 7.87 | 2.17 | N |
| CS28495 | Mnz-0 | 50 | 8.27 | 7.97 | 1.08 | N |
| CS76087 | Ag-0 | 45 | 1.3 | 8.04 | 1.27 | N |
| CS76159 | Lc-0 | 57 | -4 | 8.16 | 0.32 | Y |
| CS76306 | Zdr-6 | 49.39 | 16.25 | 8.18 | 2.24 | Y |
| CS76282 | UKSE06-272 | 51.3 | 0.4 | 8.21 | 1.17 | Y |
| CS76264 | TOU-I-2 | 46.67 | 4.12 | 8.27 | 0.54 | Y |
| CS28279 | Gel-1 | 51.02 | 5.87 | 8.3 | 0.76 | Y |
| CS28345 | Hh-0 | 54.42 | 9.89 | 8.3 | 0.82 | Y |
| CS76243 | TÅD | 62.87 | 18.34 | 8.42 | 1.2 | Y |
| CS28217 | Ede-1 | 52.03 | 5.67 | 8.54 | 0.98 | Y |
| CS76307 | Zdrl | 49.39 | 16.25 | 9.49 | 1.01 | Y |
| CS76091 | An-1 | 51.22 | 4.4 | 9.85 | 1.62 | Y |
| CS28091 | Boot-1 | 54.4 | -3.27 | 10.23 | 1.06 | N |
| CS28580 | Ob-1 | 50.2 | 8.58 | 10.34 | 1.29 | Y |
| CS76184 | MIB-84 | 47.38 | 5.32 | 10.38 | 1.58 | Y |
| CS76171 | Lisse | 52.25 | 4.57 | 12.92 | 1.18 | Y |

**Table S2.** Genes in linkage with SNP chr.3, position 3338114. Linkage Disequilibrium (LD) was taken into account for polymorphisms in a 150 kb window around the SNP, taking an LD of 0.3 as the lower threshold.

| AGI code | Name | Full gene name and description | Max. LD | Minor allele frequency | SNPs in exons |
|---|---|---|---|---|---|
| *At3G10650* | *NUP1* | Nucleoporin involved in mRNA export from the nucleus | 0.55 | 49% | non-synonymous |
| *At3G10670* | *ABCI6, NAP7* | Plastidic SufC-like ATP-binding cassette/ ATPase essential for Arabidopsis embryo-genesis | 1.00 | 41% | non-synonymous |
| *At3G10680** | - | HSP20-like chaperones superfamily protein | * | 41% | non-synonymous |
| *At3G10690* | *GYRA* | DNA GYRASE A, involved in DNA topoisom-erase activity | 1.00 | 41% | silent |
| *At3G10700* | *GALAK* | GHMP kinase family protein that acts as a galacturonic acid-1-phosphate kinase | 0.55 | 50% | non-synonymous |

* Gene with most significant SNP (chr.3 pos.3338114). Gene names and descriptions are according to The Arabidopsis Information Resource (Lamesch et al., 2011).

**Table S3.** *PHLO* DNA alignment of Col-0 and C24 (www.1001genomes.org) (Cao *et al.*, 2011; Sievers *et al.*, 2011). Exons are annotated in capital letters (*= conserved base pair). Non-synonymous polymorphisms are annotated in grey, the polymorphism in black may create an extra splice site in the Col-0 intron in the alpha-crystallin domain without disrupting the open reading frame.

| Accession | | Position |
|---|---|---|
| Col-0 | ATGGCGGCGATTTTCAAGAGACCAAGGCCGGGAGGTGGCCGTCATCCGCCACCGCTGGCT | 3337773- |
| C24 | ATGGCGGCGATTTTCAAGAGGCCAAGGCCGGGAGGTGGCCGTCATCCGCCACCGCTGGCT | 3337832 |
| | ********** ********** ********** ********** ********** | |
| Col-0 | CCGACAGTGAGTAGTTTCAAGCCGAGAGCTCAATGGACTAACTCCGGATCTTCCATTTTC | 3337833- |
| C24 | CCGACAGTGAGTAGTTTCAAGCCGAGAGCCCAATGGACTAACTCCGGATCTTCCATTTTC | 3337892 |
| | ********** ********** ********** ********** ********** | |
| Col-0 | CTCTACGTTAATCTTCCTGgtctttttttcttctctctttctttctttttgtttcttgagga | 3337893- |
| C24 | CTCTACGTTAATCTTCCTGgtctttttctcttctctctttctttctttttgtttcttgatta | 3337952 |
| | ********** ********** ********** ********** ********** * | |
| Col-0 | tcaaactttaattgataacaaactttatgaatttctcgtttattgtgtggaaccggcagG | 3337953- |
| C24 | tcgaactttaattgataacatactttatgaatttctcgtttattgtgtggaaccgacagG | 3338012 |
| | ** ********** ********** ********** ********** **** | |
| Col-0 | GTTTTATAGGGATCAAATAGAAATAAAAAAGGATGAGAGGACAAGGACTGTCCAGATTCA | 3338013- |
| C24 | GTTTTATAGGGATCAAATAGAAATAAAAAAGGATGAGAGGACAAGGACTGTCCAGATTCA | 3338072 |
| | ********** ********** ********** ********** ********** | |
| Col-0 | AGGCCAGCGACCGCTCTCTGCTCAGACTAAGGCTCGTTTCAGTGAAGCTTACCGTGTCCC | 3338073- |
| C24 | AGGCCAGCGACCGCTCTCTGCTCAGACTAAGGCTCGTTTCAATGAAGCTTACCGTGTCCC | 3338132 |
| | ********** ********** ********** ********** ********** | |

**Table S3.** (continued)

```
Col-0    TGATACCTGTGACATGACCAAATTGAGCACCTCCTTCTCACATGGACTCTTGACCATCGA  3338133-
C24      TGATACCTGTGACATGACCAAATTGAGCACCTCCTTCTCACATGGACTCTTGACCATCGA  3338192
         ************************************************************

Col-0    GTTTCCAGCAATCGTGGAAGCTAACAAACAGGAGAAAGCAGTCCAAGATCAGGAGAAGAT  3338193-
C24      GTTTCCAGCAATCGTGGAAGATAAAAAACAGGAGAAAGCAGTCCAAGATCAGGGGAAGAT  3338252
         ******************** *** ************************** ******

Col-0    TGGACAGAGGTCTAATCAGGAGAAAAGTGGAGGGCCTGGTCCTAATGGGAGTACCTTGGG  3338253-
C24      TGGACAGAGGTCTAATCAGGAGAAAAGTGGAGGGTCTGTTCCTAATGGGAGTACCTTGGG  3338312
         ********************************** *** **********************

Col-0    ACGAAAGAAAGCTTTGGAGGAGGAGAAACAAGTGGGAACAAGTCAAGAGAAAACCACTCC  3338313-
C24      ACGAAAGAAAGCTTTGGAGGAGGAGAAACAAGTGGGAACAAGTCAAGAGAAAACCACTCC  3338372
         ************************************************************

Col-0    AACGTTGAATGAAGAAGCACCCAAAACGTATAAATCAGTGGTTGAGGGTAAAAGAGCAGT  3338373-
C24      AACGTTGAATGAAGAAGCACCCAAAACGTATAAATCAGTGGTTGAGGGTAAAAGAGCTGT  3338432
         ************************************************************ **

Col-0    GCCAACGGGCAGTCAAGAGAAGTCTGAAGCGAAAGTTAAAGCGCGAGAGGCAATCCCTAG  3338433-
C24      GCCAACTGGCATTCAAGAGAAGTCTGAACCGAAAGTAAAAGCGCGAGAAGCGATCCCGAG  3338492
         ****** **** ********** ******* ********** ** ***** **

Col-0    CTTAGGGGGAAGAGAACGTGCGAAAGAAGAGAAGGTGGTTGAGAGAAAGGAGGCTGCTCA  3338493-
C24      CTTAGGGGGAAGAGAACGTGCGAAAGAAGAGAAGGTAGTTGAGAGAAAGGAGGCTGCTCA  3338552
         ************************************ ************************

Col-0    CATAGTTCAGCAAAAGATAGGAGATAAATTGAAGGAGGAAGAGGCTAAAAGTACACCAAC  3338553-
C24      AATAGGTCAGCAAAAGATAGGACAAAAATTGAAGGAGGAAGAGGCTAAAAGTACATCAAC  3338612
          **** ********** * ***************************** ****

Col-0    CCTTGGTGGTAGCTTAAAACCTAAGGTACAGGCAAAAGAAGAGAAAGTAATTGAGAGAAA  3338613-
C24      CCTTGGTGGTAGCTTAAAACCTAAGGTACAGGCAAAAGAAGAGAAAGTAATTGAGAGAAA  3338672
         ************************************************************

Col-0    GAAAGATGATGATATAAGTCAGCTTAAGACTGGACAGAAGGTGAAGGAAAAAGAGATTAG  3338673-
C24      GAAAGATGATGATATAAGTCAGCTTAAGACTGGACAGAAGGTGAAGGAAAAGGAGATTAG  3338732
         ************************************************** ********
```

**Table S3.** (continued)

```
Col-0      CAGAACGCCGACCATAGATGCTAGAGTGGAATCCAAGGAACATGCCAAAGTTGTTGAGAA   3338733-
C24        CAGAACGCCGACCATAGATGCTAGAGTGGAATCCAAGGAACATGCCAAAGTTGTTGAGAA   3338792
           ************************************************************

Col-0      AAAGGAAGATGGTGAAATAGGTCAGAAGCTGGAGGGAGGAAGGATTAACTTGGGACCAAA   3338793-
C24        AAAGGAAGATGGTGAAATAGGTCAGAAGCTGGAGGGAGGAAGGATTAACTTGGGACCAAA   3338852
           ************************************************************

Col-0      GAAAGAAGATAAATTTACAAAACCAGTTGTTGGTGATGAGGCAAGAAGAATAGAGAAGAA   3338853-
C24        GAAAGAAGATAAATTTACAAAACCAGTTGTTGGTGATGAGGCAAGAAGAATAGAGAAGAA   3338912
           ************************************************************

Col-0      TGTTGCTCAGAGAAACCAAGCAGAGCCTAAGCTAAAAACGAAAGAGGGAGATGAAAGAAT   3338913-
C24        TGTTGCTCAGAGAAACCAAGCAGAGCCTAAGCTAAAAACGAAAGAGGGAGATGAAAGAAT   3338972
           ************************************************************

Col-0      CAAACTGGATGTTGATGATGGCGTGAGAAACAAGCAGGATGAGAATAATGAGATGGTTGG   3338973-
C24        CAAACTGGATGTTGATGATGGCGTGAGAAACAAGCAGGATGAAAATAATGAGATGGTTGG   3339032
           ****************************************** ******************

Col-0      AGACATGGTTAGTGAAGGAGAGATTCAAGACAGAGTGGAGCAGAAGAAGATAGACGAAGC   3339033-
C24        AGACATGGTTAGTGAAGGAGAGATTCAAGACAGAGTGGAGCAGAAGAAGATAGACGAAGC   3339092
           ************************************************************

Col-0      TGGTTTAGCCAAAGATAGAGGCGATCTAAAGGATAATGCTGAAGTCGTGGAACCAGAAAC   3339093-
C24        TGGTTTAGCCAAAGATAGAGGCGATCTAAAGGATAATGCTGAAGTCGTGGAACCAGAAAC   3339152
           ************************************************************

Col-0      CGGTCCATTACTAGTCAAAGAAGGACAGAAGAAGAGCGACATGGATCCTCCTCTTGCAGT   3339153-
C24        CGGTCCATTACTAGTCAAAGAAGGACAGAAGAAGAGCGACATGGATCCTCCTCTTGCAGT   3339212
           ************************************************************

Col-0      AGGAGGAAGAGGAATGGGTGAAGAAGAGAGTCGGACATATGACATACCTTTGGTGAATGT   3339213-
C24        AGGAGGAAGAGGAATGGGTGAAGAAGAGAGTCGGACATATGACATACCTTTGGTGAATGT   3339272
           ************************************************************

Col-0      TGGTGTAGCTGCTCTTGTGATTATGGGGTTTGGTGCTTACGTCTTTGTACCACTAGTAAA   3339273-
C24        TGGTGTAGCTGCTCTTGTGATTATGGGGTTTGGTGCTTACGTCTTTGTACCACTTGTAAA   3339332
           ********************************************************* *****

Col-0      ATATTTCTCCTGA                                                  3339333-
C24        ATATTTCTCCTGA                                                  3339345
           *************
```

**Table S4.** Protein alignment of PHLO (At3g10680) and RTM2 (At5g04890) in Col-0. The alignment was performed with Clustal Omega (Sievers *et al.*, 2011), protein domain predictions were retrieved from Interpro (Mitchell *et al.*, 2015).

```
At3g10680    MAAIFKRPRPGGGRHPPPLAPTVSSFKPRAQWTNSGSSIFLYVNLPGFYRDQIEIKKDER
At5g04890    MAARQQ-------QKGTGFGVQYEDFVPKSEWKDQPEATILNIDLTGFAKEQMKVTYVHS
             ***  :       ::   :.   ..* *:::*.:. .: :* ::* ** ::*:::.  .
At3g10680    TRTVQIQGQRPLSAQTKARFSEAYRVPDTCDMTKLSTSFSHGLLTIEFPAIVEANKQEKA
At5g04890    SKMIRVTGERPLANRKWSRFNEVFTVPQNCLVDKIHGSFKNNVLTITMPKETITKVAYLP
             :: ::: *:***: :. :**.*.: **:.* : *:  **.. :*** :*  . ::
At3g10680    VQDQEKIGQRSNQEKSGGPGPNGSTLGRKKALEEEKQVGTSQEKTTPTLNEEAPKTYKSV
At5g04890    --------ETSRT--------EAAALEKAAKLEEKRLLEESRRKE--KEEEEAKQMKKQL
                     : *.         :.::* :   ***:: :  *:.*   . :*** :  *.:
At3g10680    VEGKRAVPTGSQEKSEAKVKAREAIPSLGGRERAKEEKVVERKEAAHIVQQKIGDKLKEE
At5g04890    LEEKEALIRKLQEEAKA--KEEAEMRKLQEEAKAKEEA------AAKKLQEEI-------
             :* *.*:    **:::*  * .  : .*  . :****       **: :*::*
At3g10680    EAKSTPTLGGSLKPKVQAKEEKVIERKK-DDDISQLKTGQKVKEKEISRTPTIDARVESK
At5g04890    ------------EAKEKLEERKLEERRLEERKLEDMKLAEEAKLKKIQ------------
                        : * : :*.*: **:  : .:.::* .::.* *:*.
At3g10680    EHAKVVEKKEDGEIGQKLEGGRINLGPKKEDKFTKPVVGDEARRIEKNVAQRNQAEPKLK
At5g04890    ------ERKSVDESGE-------------KEKILKPEVVYT--------KSGHVATPKP-
                   *:*.  * *:             ::*: ** *            . . * **
At3g10680    TKEGDERIKLDVDDGVRNKQDENNEMVGDMVSEGEIQDRVEQKKIDEAGLAKDRGDLKDN
At5g04890    ---------------------------------------------ESGSGLKSGFGGVGEV
                                                          . :** .  * : :
At3g10680    AEVVEPETGPLLVKEGQKKSDMDPPLAVGGRGMG-----EEESRTYDIPLVNVGVAALVI
At5g04890    VKSAEEKLGNLVEKEKK-----------MGKGIMEKIRRKEITSEEKKLMMNVGVAALVI
             .: .* : * *: ** :            *:*:      :* :   .  ::*********
At3g10680    MGFGAYVFVPLVKYFS-----------------
At5g04890    FALGAYVSYTFCSSSSSSSSSPSSSSSTKPE
             :.:****   : .  *
```

alpha-crystallin domain

coiled coil domain (At5g04890)

Trans-membrane domain

*= conserved amino acid (AA), := AA with highly similar properties (>0.5 Gonnet PAM 250 matrix), .= AA with weak similarity (<0.5 Gonnet PAM 250 matrix), underlined= AA belonging to predicted domain.

**Table S5.** Climate variables tested for the *At3g10680* haplotype distribution.

| Description | Number of variables | Source | Resolution | Link |
|---|---|---|---|---|
| Monthly vegetation index (average MOD13C2, 2000-2010, month 3-6) | 4 | MODIS | 0.5 km | http://modis.gsfc.nasa.gov/data/dataprod/mod13.php |
| Evapotranspiration month 3-6, annual aridity | 5 | CGIAR[1] | 1 km | http://www.cgiar-csi.org |
| Cloud cover, diurnal temp. range, mean seasonal temp., solar radiation, frost days, wet days, precipitation, humidity, soil pH and carbon proportion | 19 | CRU[2] | 50 km | http://www.cgiar-csi.org |
| Elevation | 1 | NOAA[3] | 10 km | http://nelson.wisc.edu/sage/ |
| Soil pH, carbon proportion | 2 | IGBP-DIS[4] | 50 km | http://nelson.wisc.edu/sage/ |
| Bioclim BIO1-BIO19 temperature and precipitation | 19 | World-clim[5] | 1 km | http://www.worldclim.org |
| Net primary production biomass, growing period | 2 | FAO | 10 km | http://www.fao.org/geonetwork/srv/en/main.home |

[1]Trabucco, A., and Zomer, R.J. 2009. Global Aridity Index (Global-Aridity) and Global Potential Evapo-Transpiration (Global-PET) Geospatial Database. CGIAR Consortium for Spatial Information. Published online, available from the CGIAR-CSI GeoPortal at: http://www.csi.cgiar.org.
[2]Mitchell, T.D., Jones, P.D. (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology. 25: 693-712.
[3]National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colo.
[4]IGBP-DIS (1998) SoilData(V.0) A program for creating global soil-property databases, IGBP Global Soils Data Task, France.
[5]Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis (2005) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology. 25: 1965-1978.

**Table S6.** Climate variables with a significant haplotype effect for the SNP on chromosome 3, pos. 3338114 in the alpha-crystallin domain of At3g10680.

| Climate Variable | Source | Col-0 allele (59%) | Rare allele (41%) |
|---|---|---|---|
| Temp. seasonality (standard deviation) (°C) | worldclim.org | 6.1 ± 0.1 | 5.5 ± 0.2*** |
| Temp. annual range (°C) | worldclim.org | 24.5 ± 0.4 | 22.6 ± 0.6*** |
| Isothermality (diurnal/annual range) | worldclim.org | 3.0 ± 0 | 3.2 ± 0.1*** |
| Mean temp. wettest quarter (°C) | worldclim.org | 13.5 ± 0.4 | 10.8 ± 0.5*** |
| Mean temp. coldest quarter (°C) | worldclim.org | 0.6 ± 0.2 | 2.3 ± 0.3***/* |
| Mean temp. driest quarter (°C) | worldclim.org | 3.6 ± 0.4 | 7.3 ± 0.6***/** |
| Min temp. coldest month (°C) | worldclim.org | -2.7 ± 0.3 | -1 ± 0.3*** |
| Mean monthly temp. Nov-Feb (°C) | nelson.wisc.edu/sage | 1.4 ± 0.2 | 3 ± 0.3***/* |
| Number of frost days 1980-2010 | cgiar-csi.org | 2409 ± 46 | 2113 ± 72*** |
| Precipitation warmest quarter (mm) | worldclim.org | 209.9 ± 5 | 183.6 ± 6***/* |
| Relative humidity in May (%) | cru.uea.ac.uk/cru/data | 72.6 ± 0.3 | 74.6 ± 0.4*** |
| Vegetation index MOD13C2 in March | modis.gsfc.nasa.gov | $7.0e^{7}$ ± $2e^{6}$ | $8.0e^{7}$ ± $2e^{6}$*** |

Means ± standard error, asterisks before slash: ***P<0.001 (Mann-Whitney U test, Bonferroni threshold=0.001), asterisks after slash: /* P<0.05, /** P<0.01 (mixed model with correction for kinship).

**Table S7.** Primers used for DNA and cDNA amplification.

| Gene | AGI code | T-DNA line | Forward primer PCR | Reverse primer PCR | |
|---|---|---|---|---|---|
| | | | **PCR** | **PCR** | **T-DNA primer** |
| *PHLO* | *At3g10680* | *phlo-1* | TTCAACGTTGGAGTGGTTTTC | TTGGAACGAATTGAGCTATGG | ATTTTGCCGATTTCGGAAC |
| *PHLO* | *At3g10680* | *phlo-2* | GCAAAATCCAATCTGCATCTC | TACACCAACCCTTGGTGGTAG | TAGCATCTGAATTTCATAACCAATCTCGATACAC |
| | | | **RT-qPCR** | **RT-qPCR** | **Comment** |
| *PHLO* | *At3g10680* | *phlo-1* | GGCTCCGACAGTGAGTAGTT | TGGCCTTGAATCTGGACAGT | 1st-2nd exon |
| *PHLO* | *At3g10680* | *phlo-2* | AATGCTGAAGTCGTGGAACC | TCACAAGAGCAGCTACACCA | 2nd exon |
| *ABCI6* | *At3g10670* | *phlo-1* | ATCCCGTTGCTTGAAGTTAGG | TGAACCGTTCTTCCCCATCACT | 1st-2nd exon |
| *PEX4* | *At5g25760* | - | TGCAACCTCCTCAAGTTCG | CACAGACTGAAGCGTCCAAG | reference gene |
| *ACT2* | *At3g18780* | - | CCCGATGGGCAAGTCATCACGAT | GTCTCGTGGATTCCAG-CAGCTTCC | reference gene |

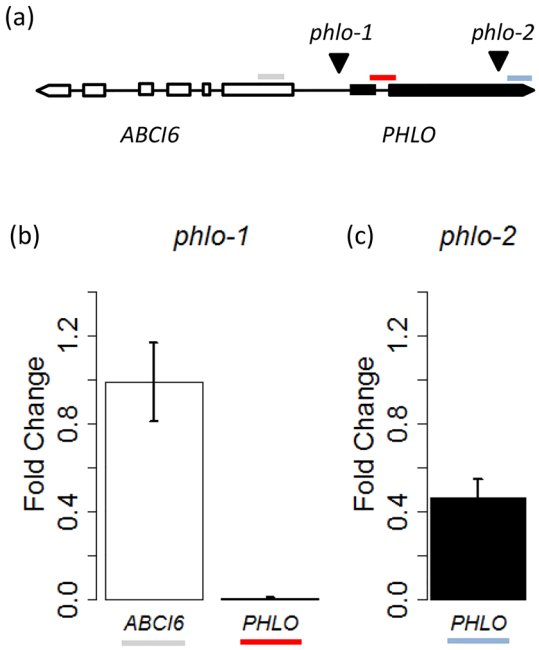



**Figure S1. Gene expression of *phlo-1* and *phlo-2*.** **a.** Position of the T-DNA insertions. Flanking regions of the inserts were sequenced to confirm the position. Coloured bars indicate DNA fragments amplified by RT-qPCR. **b.** Gene expression of *phlo-1* (SALK_027475) compared to Col-0. The expression of both *PHLO* and *ATP-BINDING CASSETTE I6* (*ABCI6*) were validated since the insert is located in the promoter of both genes. **c.** Gene expression of *PHLO* in *phlo-2* (SAIL_1269_C01) compared to Col-0 (bars represent mean, error bars represent standard error).

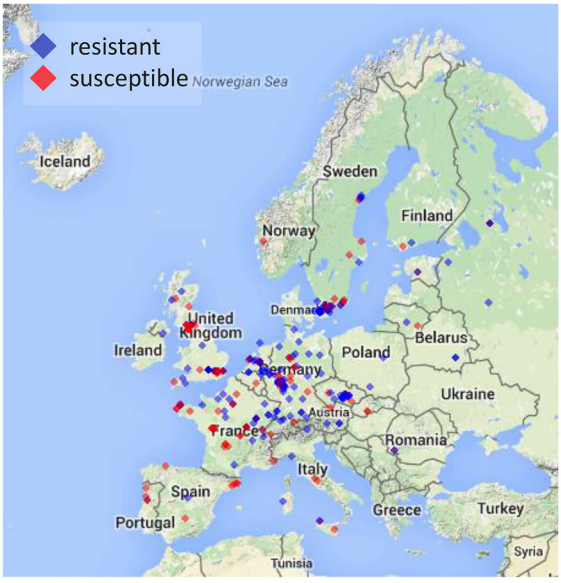



**Figure S2.** ***PHLO*** **haplotype distribution within Europe.** Haplotypes were defined by the SNP with the most significant association with aphid behaviour (chr.3 pos. 3338114, $P=3.2*10^{-5}$, Mann-Whitney U test on longitude, P=0.13, Mann-Whitney U test on latitude). Aphids spent more time on short probes (< 3 min) on accessions with the resistant haplotype compared to accessions with the susceptible haplotype (see Table S1).

# Chapter 6

## AtWRKY22 promotes susceptibility to aphids and integrates jasmonic acid and salicylic acid signalling

Karen J. Kloth, Gerrie L. Wiegers, Jacqueline Busscher-Lange, Jan C. van Haarst, Willem Kruijer, Harro J. Bouwmeester, Marcel Dicke and Maarten A. Jongsma

*under review*

## Abstract

Aphids induce many transcriptional perturbations in their host plants, but the responsible signalling cascades and effects on plant resistance are largely unknown. Through a genome-wide association (GWA) mapping study in *Arabidopsis thaliana*, we identified *WRKY22* as a candidate gene associated with feeding behaviour of the green peach aphid, *Myzus persicae*. The transcription factor WRKY22 is known to be involved in pathogen-triggered immunity and *WRKY22* gene expression has been shown to be induced by aphids. Assessment of aphid population development and feeding behaviour on knockout mutants and overexpression lines, showed that WRKY22 increases susceptibility to *M. persicae* via a mesophyll-located mechanism. mRNA sequencing analysis of aphid-infested *wrky22* knockout plants revealed the upregulation of genes involved in salicylic acid (SA) signalling and downregulation of genes involved in plant growth and cell-wall loosening. In addition, mechanostimulation of knockout plants by clip cages upregulated jasmonic-acid (JA)-responsive genes, resulting in substantial negative JA-SA crosstalk. Together with previous studies, our data indicate that WRKY22 modulates the interplay between SA and JA in response to a wide range of biotic and abiotic stimuli. Its induction by aphids, and its role in suppressing SA and JA signalling, make *WRKY22* a potential target for aphids to manipulate host plant defences.

## Background

As plants are sessile organisms in often dynamically changing environments, plasticity is fundamental to survival. Via transcriptional regulation, plants increase the possibilities to cope with environmental stimuli. In Arabidopsis approximately 50 transcription factor families have been identified, accounting for approximately 2000 genes (Guo *et al.*, 2005; Mitsuda and Ohme-Takagi, 2009). Together with signal perception and transduction elements, these transcription factors participate in complex and dynamic networks that regulate developmental processes and responses to (a)biotic stress. Insect infestations are typical situations that require quick transcriptional reprogramming in order to mount an effective defence response. Aphids are phloem-feeding insects, that manoeuvre their piercing-sucking mouthparts between cells and reach the vascular bundle without inflicting major physical damage (Minks and Harrewijn, 1989). They are vectors of many plant viruses and deprive the plant of photoassimilates. Aphids cause strong transcriptional perturbations in plants, inducing or repressing up to several thousand genes, whereas other insects such as caterpillars and cell-content feeders alter

the expression of only up to several hundreds of genes (Appel *et al.*, 2014; Barah *et al.*, 2013; De Vos *et al.*, 2005; Dubey *et al.*, 2013; Foyer *et al.*, 2015; Kerchev *et al.*, 2013; Kusnierczyk *et al.*, 2007). An open question is, however, whether these transcriptional changes lead to enhanced resistance to aphids or whether they are unsuccessful or even counter-effective modulations. Aphids are known to secrete effectors via their saliva into the apoplast and the vascular bundle (Rodriguez and Bos, 2012), and might be able to manipulate the host plant physiology for their own benefit. In this study, GWA mapping revealed *WRKY22* (*At4g01250*) as one of the candidate genes for affecting feeding behaviour of the generalist aphid *Myzus persicae* (Sulzer) on *Arabidopsis thaliana*. WRKY22 is member of the WRKY transcription factor family, which was discovered in the 1990's and named after its binding affinity to the W-box promoter motif (Eulgem *et al.*, 2000). WRKY22 and its homologue WRKY29 are part of group IIe WRKYs and are both established markers for pathogen-triggered immunity (PTI). Pathogen-associated molecular patterns (PAMPs) such as flagellin, chitin and cellulysin are recognised elicitors of the MAPK cascade that induce *WRKY22* and *WRKY29* within 30 min post inoculation (Asai *et al.*, 2002; Dong *et al.*, 2003; González-Lamothe *et al.*, 2012; Mészáros *et al.*, 2006; Navarro *et al.*, 2004; Schikora *et al.*, 2011; Shi *et al.*, 2015; Thilmony *et al.*, 2006). In general, PTI results in the accumulation of reactive oxygen species and callose deposition and involves SA, JA and ethylene (ET) signalling (Yi *et al.*, 2014). Although the exact role of WRKY22 and WRKY29 in PTI is unknown, WRKY22 has been shown to be required for resistance to the hemibiotrophic pathogen *Pseudomonas syringae* (Hsu *et al.*, 2013) and WRKY29 has been described to confer resistance to *P. syringae*, as well as the necrotrophic pathogen *Botrytis cinerea* (Asai *et al.*, 2002). In this study, we assessed the involvement of WRKY22 in plant resistance to *M. persicae* and its downstream transcriptional effects.

## Results

### GWA mapping

To identify genes involved in resistance to the green peach aphid, *M. persicae*, GWA mapping was performed on 344 natural accessions of Arabidopsis, using a selection of approximately 214k single nucleotide polymorphisms (SNPs) (Atwell *et al.*, 2010; Li *et al.*, 2010). The behaviour of aphids was screened on these accessions with an automated video-tracking platform (Kloth *et al.*, 2015). The number and duration of plant penetrations was estimated by analysing the location and movement of aphids on single leaf discs. It is known that aphids need on average 25 minutes to penetrate

through the epidermis and mesophyll before they reach the vascular bundle (Prado and Tjallingii, 2007; Tjallingii, 1994; van Helden and Tjallingii, 1993). Therefore, we used the proportion of aphids making long probes (> 25 min) as a proxy for the success rate of phloem ingestion. The majority of the Arabidopsis accessions did not show indications of resistance to aphids, but on 10% of the accessions at least half of the aphids were unsuccessful in feeding after 4.5 hours of infestation (**Figure 1a**, **Table S1**). GWA mapping of aphid feeding behaviour revealed seven genomic regions with a -10log(P) value above 4 and a heritability of 10% (**Figure 1b**, **Table 1**). *WRKY22* (*At4g01250*) was identified as a candidate gene in a 40 kb region around a polymorphism with a

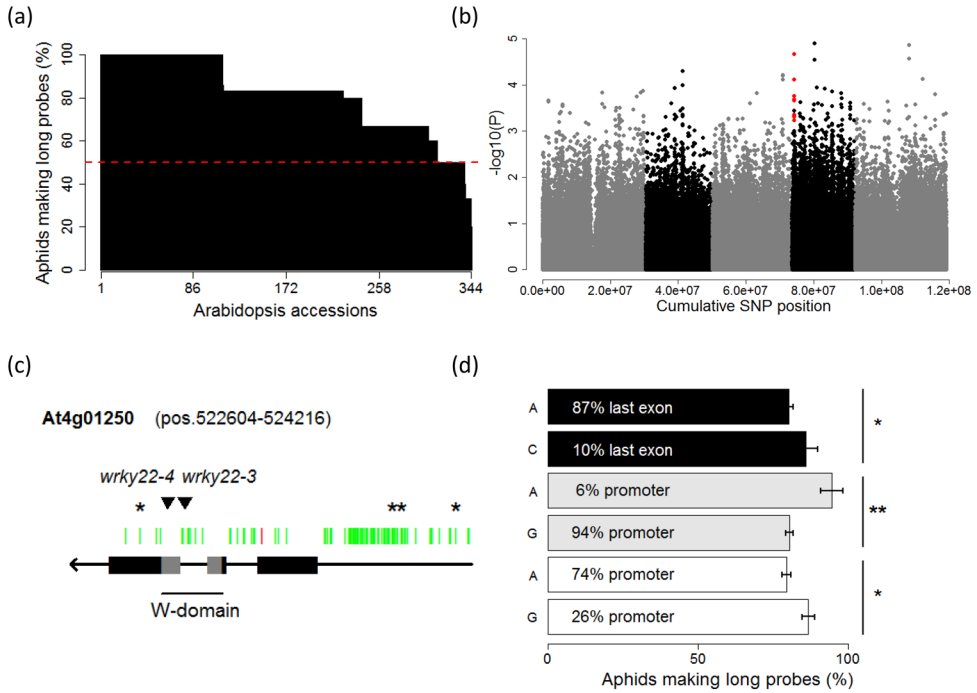



**Figure 1. Genome-wide association mapping of aphid feeding behaviour. a.** Phenotypic distribution of the proportion of aphids making long probes (> 25 min) during a 1.5 hour recording on plants from 344 natural Arabidopsis accessions 4.5 hours post inoculation. On accessions below the dashed line, at least half of the aphids was unsuccessful in feeding. **b.** Genome-wide associations with approximately 214k SNPs. SNPs in red are positioned in a 40 kb region around *WRKY22* (highest -10log(P)= 4.7). **c.** All SNPs in *WRKY22* and its 1000 kb promoter region according to 173 resequenced Arabidopsis accessions (green= silent, red= non-synonymous). Predicted gene domains are shown in gray, unknown domains in black. Triangles represent T-DNA insertions. **d.** One synonymous SNP in the last exon and two SNPs in the promoter had an effect on aphid feeding behaviour (*P<0.05, **P<0.01, Student's t-test, chr.4, pos. 523037, 524726 and 525079).

-10log(P) value of 4.7 (chr.4 pos.543516). Other candidates in the region included a gene with unknown function (*At4g01290*), a methyltransferase and a gene (*At4g01240*) with a MYB-like domain (*At4g01280*). Re-sequenced data of 173 accessions (Cao *et al.*, 2011) showed that *WRKY22* contained one non-synonymous SNP in its coding region, and that most of the polymorphisms were confined to the introns and the promoter region (**Figure 1c**). A silent SNP in the last exon and two SNPs in the promoter were correlated with aphid feeding behaviour (**Figure 1d**). Both polymorphisms in the promoter coincided with an AT-hook DNA-binding motif of AHL20, a transcription factor involved in plant defence to bacteria (Lu *et al.*, 2010). Because WRKY22 is involved in PAMP-triggered immune responses (Asai *et al.*, 2002; Navarro *et al.*, 2004) and its expression is induced by *M. persicae* and *Brevicoryne brassicae* aphids (Barah *et al.*, 2013; De Vos *et al.*, 2005), we conducted further experiments to assess whether WRKY22 is involved in resistance to aphids.

## Mesophyll-located susceptibility to aphids

To validate the previously reported induction of *WRKY22* by aphid infestation (Barah *et al.*, 2013; De Vos *et al.*, 2005), RT-qPCR was performed on wild-type plants with aphids and without aphids. *WRKY22* expression was unaffected at 6 hours post infestation (hpi), and showed a non-significant increase at 48 hpi (**Figure 2a**). Two *wrky22* transfer

**Table 1.** SNPs and corresponding genes associated with the proportion of aphids making long probes (> 25 min, -10log(P) value> 4). Only the highest scoring SNP is shown per gene. Genes were grouped in one linkage disequilibrium (LD) region, if they were located within 20 kb from each other.

| LD region | Chrom. | Position | -10log(P) | AGI code | Description |
|---|---|---|---|---|---|
| 1 | 1 | 28995670 | 6.6 | *At1g77160* | Protein of unknown function (DUF506) |
| 2 | 2 | 10866313 | 4.3 | *At2g25530* | AFG1-like ATPase family protein |
| 3 | 3 | 20709836 | 4.2 | *At3g55800* | Chloroplast enzyme sedoheptulose-1,7-bisphosphatase (SBPase) |
| 4 | 4 | 519513 | 4.1 | *At4g01240* | S-adenosyl-L-methionine-dependent methyltransferase super-family protein |
| 4 | 4 | 536493 | 4.1 | *At4g01280* | Homeodomain-like superfamily protein, SANT DNA-binding MYB-like domain |
| 4 | 4 | 543516 | 4.7 | *At4g01290* | unknown protein |
| 5 | 4 | 6641192 | 4.5 | *At4g10790* | UBX domain-containing protein |
| 5 | 4 | 6644022 | 4.9 | *At4g10800* | BTB/POZ domain-containing protein |
| 6 | 5 | 15927540 | 4.9 | *At5g39770* | Pseudogene homologous to AtMSU81, restriction endonuclease |
| 7 | 5 | 19854700 | 4.1 | *At5g48965* | Mutator-like transposase family |
| 7 | 5 | 19858466 | 4.1 | *At5g48970* | Mitochondrial substrate carrier family protein |

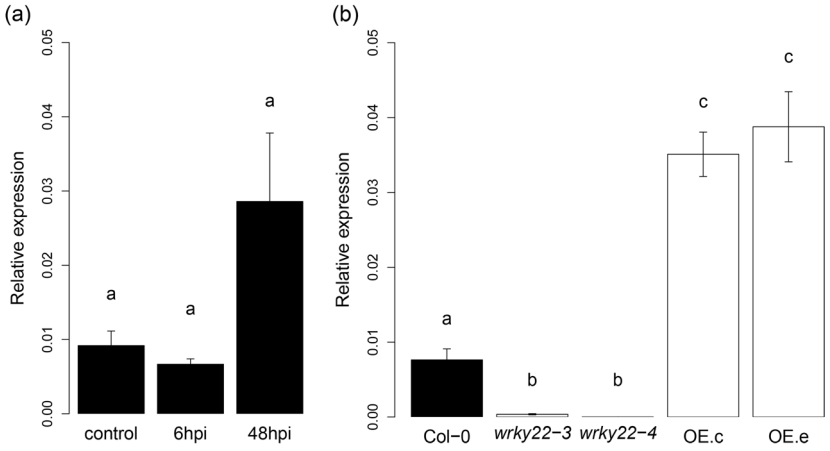



**Figure 2.** *WRKY22* **expression. a.** *WRKY22* expression in the wild type without aphids (control) and after 6 and 48 hours of aphid infestation. **b.** Expression in the wild type (Col-0), *wrky22-3* and *wrky22-4* knockout lines, and *WRKY22* inducible overexpression lines OE.c and OE.e. Overexpression lines were induced with estradiol 24 h before sampling (one-way ANOVA and Student's t-tests, different letters refer to significant differences).

(T)-DNA insertion lines and two *WRKY22* inducible overexpression lines (Coego *et al.*, 2014) were selected for further experiments. RT-qPCR confirmed that both T-DNA lines were true knockouts, and that the overexpression lines showed a 3- to 5-fold upregulation of *WRKY22* at 24 hours after induction with estradiol (**Figure 2b**). For a detailed insight in aphid feeding behaviour on knockout and overexpression lines, we used Electrical Penetration Graph (EPG) recordings (McLean and Kinsey, 1964; Tjallingii, 1988). Aphid feeding behaviour was affected on both *wrky22* knockout lines; on *wrky22-3* aphids spent almost 20% more time on penetrating the epidermis and mesophyll, and on *wrky22-4* aphids showed an hour delay in reaching the vascular bundle compared to the wild type (Col-0) (**Figure 3a-c**, **Table S2**). One of the *WRKY22* inducible overexpression lines showed the opposite trend, with aphids arriving almost an hour earlier at the vascular bundle compared to the wild type (**Figure 3d-f**, **Table S3**). The other *WRKY22* overexpression line did not show any differences compared to the wild type. The total time of phloem ingestion was not affected in any of the (mutant) lines, suggesting that in the first 8 hours of infestation, the overall effects are small and confined to activities in the epidermis and/or mesophyll. An aphid population development assay on *wrky22-3* and *wrky22-4* showed that after two weeks of infestation, aphid populations were approximately 20% smaller on the knockouts compared to the wild type (**Figure 4**). Both behavioural experiments and population assays indicate that WRKY22 increases susceptibility to *M. persicae* aphids.

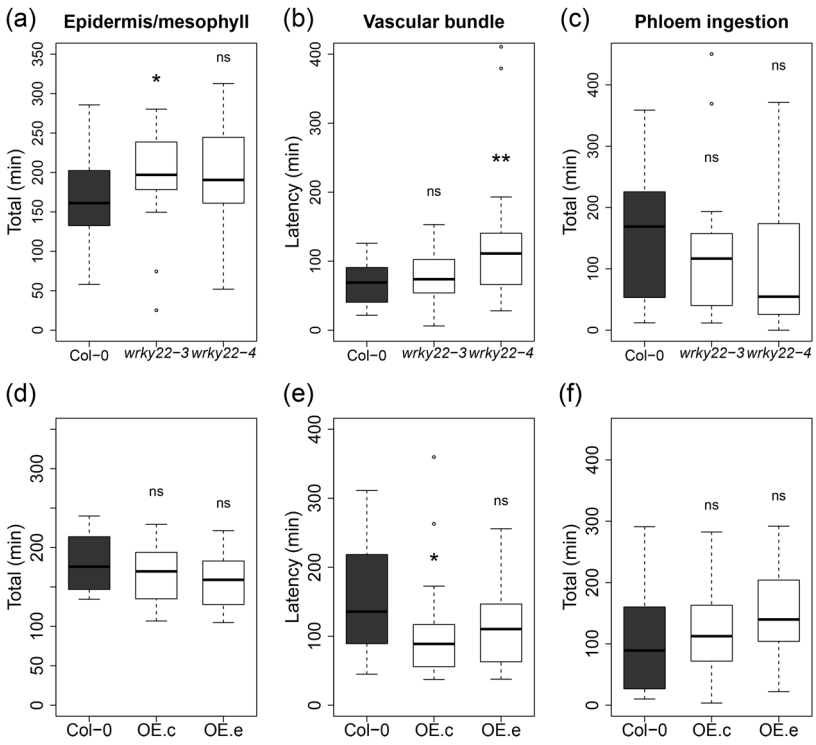



**Figure 3. Aphid behaviour on *wrky22* knockout lines (upper panels) and *WRKY22* overexpression lines (lower panels). a.** The total time aphids were penetrating the epidermis and mesophyll during 8-hour recordings on knockout lines *wrky22-3* and *wrky22-4*, **d.** and overexpression lines OE.c and OE.e. **b.** Time between the start of the recording and the first contact with either a phloem or xylem bundle measured on knockout, and **e.** overexpression lines. **c.** The total time aphids were ingesting phloem on knockout, and **f.** overexpression lines (knockout and overexpression lines were compared with the wild type with Mann Whitney U tests, *P<0.05; **P<0.01). To test the effect of overexpression, all plants were induced with estradiol 24 h before the assay.

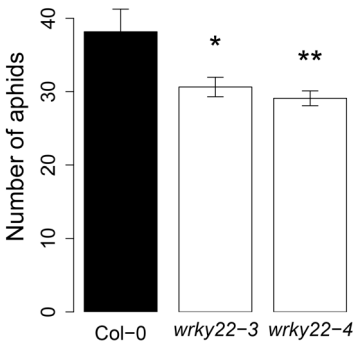



**Figure 4. Aphid population size on wild-type and knockout plants.** The total number of aphids per plant was counted 2 weeks after infestation with 1 neonate aphid. Mutant lines were compared with the wild type with Student's t-tests (*P<0.05; **P<0.01).

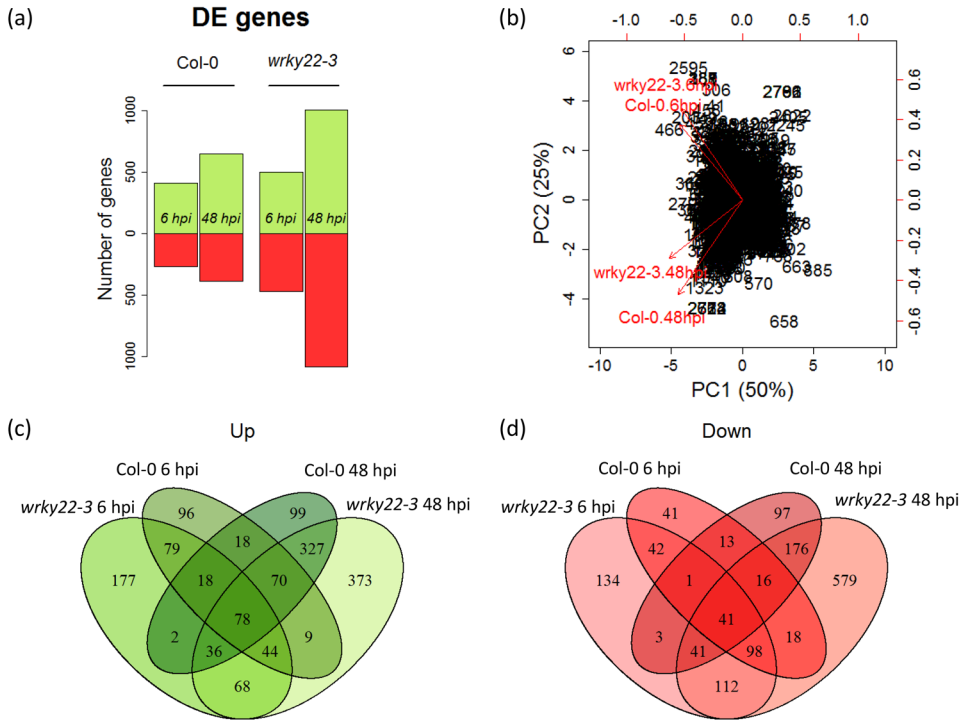



**Figure 5. Differentially expressed (DE) genes between treatments with and without aphids in the wild type and** *wrky22-3*. **a.** The number of DE genes between control and infested treatments (green bars: upregulated, red bars: downregulated). **b.** Biplot of the two first principal components of differentially expressed genes between control and infested treatments (DE genes >= 2 fold). **c.** Overlap in upregulated genes, and **d.** downregulated genes.

## Transcriptomic signature of *wrky22-3*

To study the role of *WRKY22* in resistance to aphids, mRNA sequencing (RNA-seq) analysis was performed on *wrky22-3* and wild-type plants that had received one of three treatments: (1) an empty clip cage for 48 hours, (2) a clip cage for 48 hours with addition of aphids in the last 6 hours, (3) a clip cage with aphids for 48 hours. For each treatment three biological replicates were sampled, each consisting of a pool of five leaves from different plants. Samples were sequenced for single end 50-bp reads and for each sample at least 9.5 million reads mapped to unique loci on the Arabidopsis reference genome. The total number of differentially expressed (DE) genes increased with the duration of infestation, from 700 DE genes at 6 hpi to 1000 DE genes at 48 hpi in the wild type. In both treatments, *wrky22-3* contained twice as many up- and downregulated genes as the wild type (**Figure 5a**). Principal component analysis showed that the duration of infestation was the major factor explaining differential expression (**Figure 5b**), and the

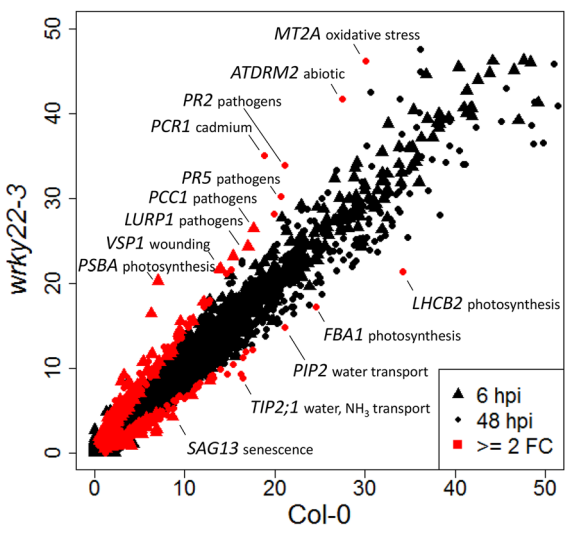



**Figure 6. Gene transcripts of aphid-infested wild-type and *wrky22-3* plants.** Differentially expressed genes between wild-type and knockout plants (>= 2-fold change) are shown in red. Axes depict the square-root transformation of the normalized number of transcripts (the number of fragments per kilobase of transcript per million reads mapped (FPKM)), genes <= 2500 FPKM are shown (including all DE genes in the dataset), annotations include gene name and biological process (wounding= wound responsive, pathogens= pathogen responsive, cadmium= responsive to cadmium, abiotic= responsive to several abiotic stresses).

majority of the DE genes were unique to either the early or late stage of infestation (**Figure 5c-d**). Highly abundant transcripts that were upregulated in *wrky22-3* compared to the wild type included the JA reporter *VSP1* (6 hpi) and pathogenesis-related genes such as *PR2* and *PR5* (48 hpi) (**Figure 6**). Photosynthesis- and water-transport-related genes were abundant in the wild type, but downregulated in *wrky22-3* at 48 hpi (**Figure 6**). Gene ontology (GO) enrichment analysis of the total set of DE genes revealed an overrepresentation of upregulated JA-, SA- and abscisic acid (ABA)-responsive genes in *wrky22-3* both at 6 and 48 hpi (**Figure 7a**). The JA pathway was mainly characterised by upregulation of genes of the ethylene response factor (ERF)-branch (Vos *et al.*, 2015), e.g. the AP2/ERF transcription factor *RAP2.6*, and *PDF1.2* (**Table 2**). The majority of the DE genes associated with JA and ABA showed a peak at 6 hpi in *wrky22-3*, whereas most SA-responsive genes reached their highest level at 48 hpi (**Figure 7b**). The induction of SA-responsive genes in the T-DNA line at 48 hpi coincided with a suppression of genes associated with auxin (AUX) responsiveness, plant growth and cell wall loosening (**Figure 7a-b**).

## Enhanced negative JA-SA crosstalk in *wrky22-3*

Upon aphid infestation, the *wrky22-3* transcriptome showed evidence of initial suppression, but eventual upregulation of SA signalling. The expression of *PR1*, a robust SA-reporter gene (Pieterse *et al.*, 2012), was 2-fold downregulated at 6 hpi, but 3.5-fold upregulated at 48 hpi in *wrky22-3* compared to the wild type. Transcript levels of JA-reporter genes *PDF1.2* and *VSP1* were consistently more abundant in *wrky22-3*

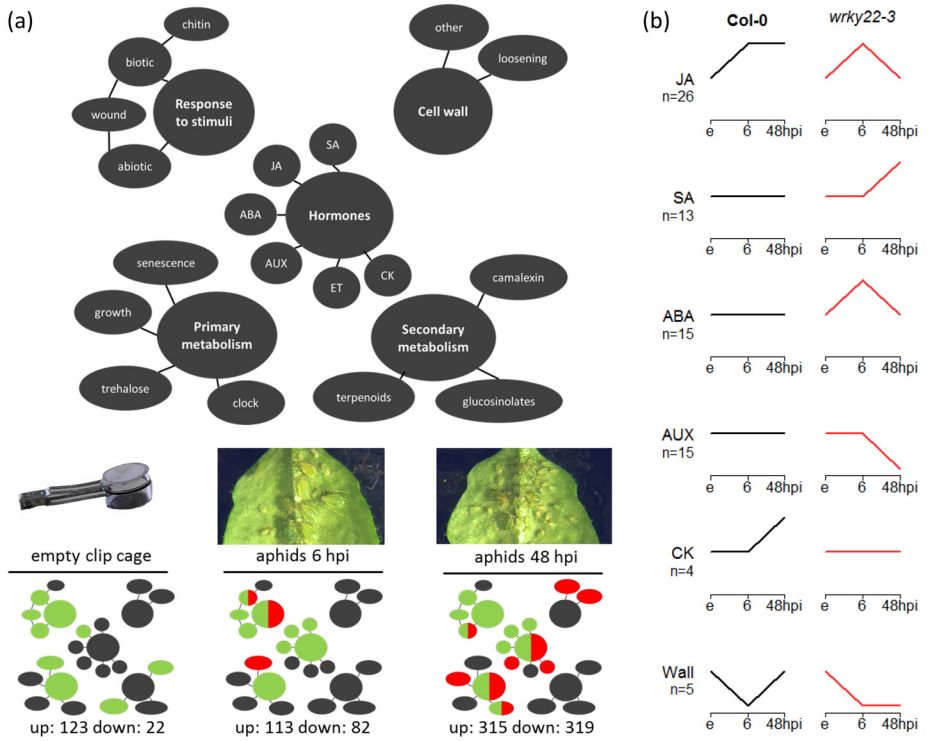



**Figure 7. Enriched biological processes in** *wrky22-3*. **a.** Overrepresentation of biological processes in the knockout relative to the wild type. Balloons refer to a process, or to the biosynthesis of, or responsiveness to the respective compound (SA=salicylic acid, JA=jasmonic acid, ABA=abscisic acid, AUX=auxin, ET=ethylene, CK=cytokinin, clock=circadian clock). Balloon colour indicates enrichment in the knockout (green=upregulated, red=downregulated, green/red=both up- and downregulated, the total number of DE genes is depicted below the charts). **b.** Relative expression patterns between treatments within each plant line. Only the dominant pattern (>=50% of the genes) of significant perturbations (>= 2 fold, q-value < 0.05) between treatments with and without aphids is shown (Wall= cell wall loosening, n=number of genes associated with the biological process, e= empty clip cage, 6= 6 hpi).

(**Table 3**), suggesting a possible role of negative JA-SA crosstalk (Pieterse *et al.*, 2012; Spoel and Dong, 2008). A potential antagonising candidate is *NIMIN-2*, encoding an SA-suppressing protein (Weigel *et al.*, 2005), which was upregulated in *wrky22-3* at 6 hpi (**Table 3**). Apart from SA antagonism, there were also signs of JA antagonism. Several upregulated genes in *wrky22-3*, i.e. *GRX480*, *WRKY51*, and *WRKY62* (**Table 3**), have previously been implicated as potential suppressors of JA signalling (Gao *et al.*, 2011; Mao *et al.*, 2007; Ndamukong *et al.*, 2007). Altogether, the aphid-induced *wrky22-3* transcriptome showed evidence of more elaborate negative crosstalk between the JA and SA pathways than the aphid-induced wild-type transcriptome.

**Table 2.** Differentially expressed genes (>= 2 fold) of overrepresented biological processes in *wrky22-3* relative to the wild type. GO enrichment and gene classification are according to the BiNGO Cytoscape app (SA = Salicylic Acid, JA = Jasmonic Acid, ABA = Abscisic acid, AUX = Auxin, ET = Ethylene, CK = Cytokinin) (Cline *et al.*, 2007; Maere *et al.*, 2005).

| Process | Treatment | Direction | Name | AGI codes | Description |
|---|---|---|---|---|---|
| ABA | empty cage, 6 hpi | up | *ANNAT4* | *At2g38750* | Annexin, Golgi-mediated secretion |
| | 6 hpi | up | *HAI1* | *At5g59220* | *HIGHLY ABA-INDUCED PP2C gene 1* |
| | 6 hpi | up | *HD-Zip-I* | *At3g61890* | Homeodomain leucine zipper class I |
| | 48 hpi | up | *ACR8* | *At1g12420* | *ACT DOMAIN REPEAT 8* |
| | 48 hpi | up | *AMY1* | *At4g25000* | *ALPHA-AMYLASE-LIKE 1*, starch mobilisation |
| | 48 hpi | up | Dehydrins | *At3g50970, At1g20440* | Membrane located, freeze tolerance |
| | 48 hpi | up | *ERF48* | *At2g40340* | ABA responsive AP2/ERF transcription factor |
| | 48 hpi | up | *LTI78* | *At5g52310* | *LOW-TEMPERATURE-INDUCED 78* |
| | 48 hpi | up | *WRKY63* | *At1g66600* | ABA responsive WRKY transcription factor |
| AUX | 6, 48 hpi | down | *CCA1* | *At2g46830* | Negative regulator of circadian rhythm |
| | 48 hpi | down | *AXR3* | *At1g04250* | *AUXIN RESISTANT 3* |
| | 48 hpi | down | *GH3*s | *At2g47750, At5g13360* | *GH3* auxin responsive gene family |
| | 48 hpi | down | *SAUR*s | *At1g20470, At1g29500, At1g29510, At3g03820, At4g22620, At4g38840, At4g38850, At4g38860, At5g18020, At5g18030, At5g18050* | SAUR(-like) auxin-responsive proteins |
| CK | 48 hpi | down | *ARR*s | *At1g19050, At1g74890, At3g57040, At5g62920* | Arabidopsis response regulator (ARR) family |
| JA | empty cage, 6, 48 hpi | up | *JAZ*s | *At2g34600, At5g13220, At1g17380, At1g19180* | JAZ7, JAZ10, JAZ5, JAZ1, Jasmonate-Zim-domain proteins |
| | empty cage, 6, 48 hpi | up | *MDHAR4* | *At3g09940* | Monodehydroascorbate reductase |
| | empty cage, 6, 48 hpi | up | *MYB47* | *At1g18710* | JA-responsive MYB transcription factor |
| | empty cage, 6, 48 hpi | up | *TAT3* | *At2g24850* | Tyrosine aminotransferase, JA responsive |
| | empty cage | up | *VSP2* | *At5g24770* | *VEGETATIVE STORAGE PROTEIN 2* |
| | empty cage, 6, 48 hpi | up | *VSP1* | *At5g24780* | *VEGETATIVE STORAGE PROTEIN 1* |
| | empty cage, 6 hpi | up | *AOC*s | *At3g25760, At3g25780* | Allene Oxide Cyclase family, JA biosynthesis |

**Table 2.** (continued)

| Process | Treatment | Direction | Name | AGI codes | Description |
|---|---|---|---|---|---|
| **JA** | empty cage, 6 hpi | up | *OPR3* | *At2g06050* | *OXOPHYTODIENOATE-REDUCTASE 3*, JA biosynthesis |
| | empty cage | up | *EXT4* | *At1g76930* | Extensin |
| | empty cage | up | *JR1* | *At3g16470* | *JASMONAtE RESPONSIVE 1* |
| | 6, 48 hpi | up | *PDF1.2* | *At5g44420* | *PLANT DEFENSIN 1.2* |
| | 6 hpi | up | *DAD1* | *At2g44810* | *DEFECTIVE ANTHER DEHISCENCE 1*, JA biosynthesis |
| | 6 hpi | up | *JAR1* | *At2g46370* | Jasmonate-amido synthetase |
| | 6 hpi | up | *LOX3* | *At1g17420* | *LIPOXYGENASE 3* |
| | 6 hpi | up | *RAP2.6* | *At1g43160* | AP2/ERF transcription factor |
| **SA** | 6, 48 hpi | up | *GRX480* | *At1g28480* | Glutaredoxin family, suppresses PDF1.2 |
| | 6, 48 hpi | up | *LURP1* | *At2g14560* | Resistance to *Hyaloperonospora parasitica* |
| | 6, 48 hpi | up | *WRKY18* | *At4g31800* | *WRKY18* |
| | 48 hpi | up | *WRKY*s | *At5g01900, At5g22570* | *WRKY38*, *WRKY62* |
| | 48 hpi | up | *MYB77* | *At3g50060* | *MYB77* |
| | 48 hpi | up | *WAK1* | *At1g21250* | *CELL WALL-ASSOCIAtED KINASE 1* |
| **JA, SA, ABA** | 6 hpi | up | *CIR1* | *At5g37260* | MYB transcription factor |
| | 48 hpi | up | *MYB*s | *At1g06180, At1g57560, At5g67300, At2g16720* | *MYB13*, *MYB50*, *MYB44*, *MYB7* |
| | 48 hpi | up | *MPK11* | *At1g01560* | *MAP KINASE 11* |
| | 48 hpi | up | *PDR12* | *At1g15520* | ABC transporter family, MAPK cascade |
| **Camalexin** | empty cage, 48 hpi | up | *PAD3* | *At3g26830* | *PHYTOALEXIN DEFICIENT 3*, camalexin biosynthesis |
| | 48 hpi | up | *P450* | *At4g39950* | Cytochrome P450, indo-3-acetaldoxime (IAOx) biosynthesis |
| **Terpenoids** | empty cage, 6, 48 hpi | up | *TSP4* | *At1g61120* | *TERPENE SYNTHASE 4* |
| | empty cage | up | *TPS10* | *At2g24210* | *TERPENE SYNTHASE 10* |
| **Cell wall** | 48 hpi | down | Expansins | *At1g20190, At1g26770, At1g69530, At2g20750, At2g40610* | Expansin family, cell wall loosening and multidimensional cell growth |

**Table 3.** JA- and SA-signalling-related gene expression in *wrky22-3* plants compared to wild-type plants with and without aphids. Differential expressed genes with at least 2-fold absolute change are shown (emp.= empty clip cage, signa= signalling, suppr= suppression of signalling, ns= not significant).

| | | | Fold change | | | |
|---|---|---|---|---|---|---|
| Gene | Name | Role | Emp. | 6 hpi | 48 hpi | **Reference** |
| *VSP1* | *VEGETATIVE STORAGE PROTEIN 1* | JA signa | 13.0 | 2.5 | 2.8 | (Anderson *et al.*, 2004; Lorenzo *et al.*, 2004) |
| *VSP2* | *VEGETATIVE STORAGE PROTEIN 2* | JA signa | 7.5 | ns | ns | (Anderson *et al.*, 2004; Lorenzo *et al.*, 2004) |
| *PDF1.2, 1.2C* | *PLANT DEFENSIN 1.2A, 1.2C* | JA signa | ns | 2.3 | 3.5 | (Lorenzo *et al.*, 2003; Penninckx *et al.*, 1998) |
| *PR1* | *PATHOGENESIS-RELATED GENE 1* | SA signa | ns | 0.4 | 3.5 | (Loon *et al.*, 2006) |
| *NIMIN-1* | NIM1-INTERACTING 1 | SA suppr | ns | ns | 2.5 | (Weigel *et al.*, 2001; Weigel *et al.*, 2005) |
| *NIMIN-2* | *NIM1-INTERACTING 2* | SA suppr | ns | 2.0 | 2.5 | (Weigel *et al.*, 2001; Weigel *et al.*, 2005) |
| *GRX480* | Glutaredoxin | JA suppr | ns | 2.8 | 2.8 | (Ndamukong *et al.*, 2007) |
| *WRKY51* | *WRKY DNA-BINDING PROTEIN 51* | JA suppr | ns | ns | 4.3 | (Gao *et al.*, 2011)2011 |
| *WRKY62* | *WRKY DNA-BINDING PROTEIN 62* | JA suppr | ns | ns | 2.8 | (Mao *et al.*, 2007) |

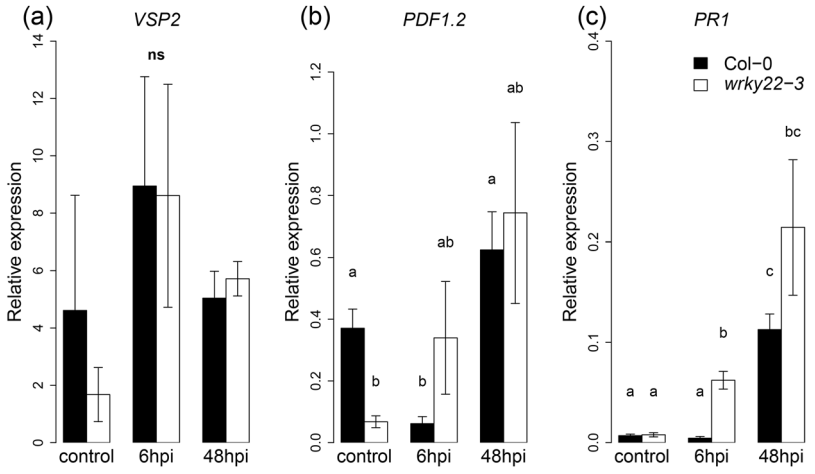



**Figure 8. The effect of aphid infestation without clip cage on the expression of JA- and SA-reporter genes.** RT-qPCR measurements of expression of the JA reporters **a.** *VSP2*, and **b.** *PDF1.2*, and **c.** the SA reporter *PR1* in wild-type and *wrky22-3* plants (Student's t-tests and Mann-Whitney U tests, different letters annotate significant differences). Aphids were contained on the leaves without inflicting major mechanical stimulation (see Materials and methods).

### JA induction by mechanostimulation in *wrky22-3*

Remarkably, the treatment with empty clip cages changed the expression of almost 150 genes in the knockout relative to the wild type. Most of them were upregulated in *wrky22-3* and showed a significant overrepresentation of JA-responsive genes, including *VSP2* of the wound-responsive MYC-branch of the JA signalling pathway (Vos *et al.*, 2015) (**Figure 7**, **Table 2**), even though we did not see obvious signs of wounding. In order to see if mechanostimulation had affected the plant's response to aphids, RT-qPCR was conducted on aphid-infested leaves without clip cages (see Materials and methods). We found that in clean *wrky22-3* plants *PDF1.2* expression was lower compared to wild-type plants (**Figure 8**). After six hours of aphid infestation, *PR1* was upregulated in *wrky22*-3 while *PDF1.2* only showed a non-significant increase (**Figure 8**). These results suggest that, in the RNA-seq analysis, the overrepresented JA response may have been clip cage-induced and was likely involved in the suppression of the SA response to aphids at 6 hpi.

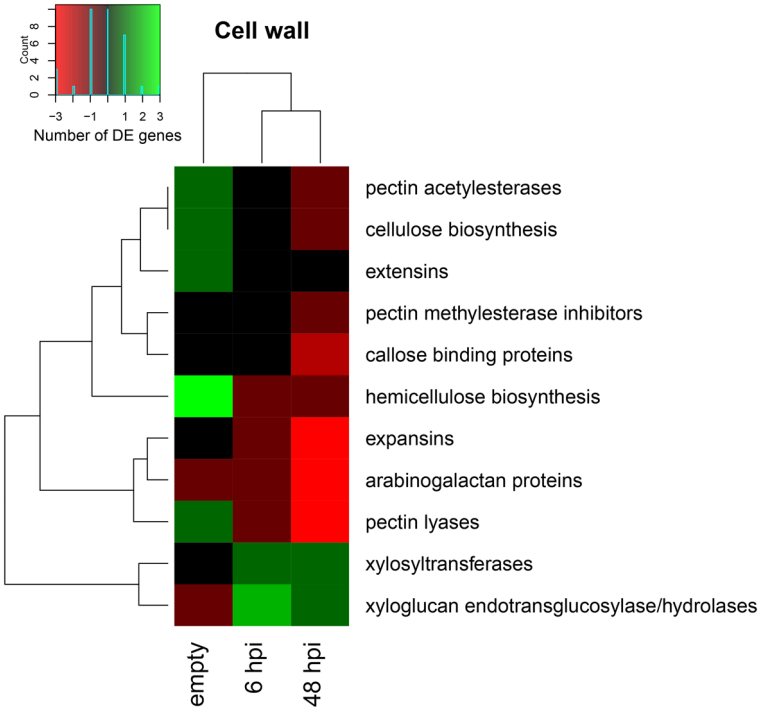



**Figure 9. Empty clip cage- and aphid-induced changes in expression of cell-wall related genes in *wrky22-3*.** Genes and treatments are clustered according to the number of differentially expressed genes (>= 2 fold change), using Ward's minimum variance method (red = downregulated, green = upregulated in *wrky22-3* compared to the wild type).

### Differential expression of cell wall-related genes

After 48 hours of aphid infestation, the *wrky22-3* transcriptome was characterised by downregulation of genes associated with cell wall loosening (**Figure 7**, **Table 2**). To assess all cell wall-related processes, DE genes with cell wall annotation were selected and grouped into categories based on their name and function (**Figure 9**). Although JA was upregulated in all treatments and is required for responses to touch and mechanostimulation (Benikhlef *et al.*, 2013; Lange and Lange, 2015; Lee *et al.*, 2004), we did not observe an upregulation of touch-responsive xyloglucan endotransglucosylases/hydrolases (XTHs) by the empty clip cage. The empty clip cage did, however, cause a 3- to 5-fold upregulation of the cellulose synthase-like genes *CSLA1*, *CSLA10*, and *CSLA15*, involved in hemicellulose biosynthesis (Liepman *et al.*, 2005). Aphids upregulated XTHs and downregulated e.g. expansins, involved in cell wall loosening, and pectin lyases, involved in pectin breakdown. While cell wall loosening is a prerequisite for cell elongation, a process mainly regulated by CK and AUX (Albersheim *et al.*, 2011; Taiz, 1984; Yadav *et al.*, 2009), the transcriptomic patterns indicate an aphid-induced arrestment of symplastic cell growth in *wrky22-3*.

## Discussion

### The effect of WRKY22 on *M. persicae* aphids

Plant responses to aphids are known to involve many transcriptional perturbations including multiple phytohormonal pathways (De Vos *et al.*, 2007; De Vos *et al.*, 2005; Foyer *et al.*, 2015; Smith and Boyko, 2007). It is, therefore, a challenge to unravel the genetic basis of effective defence mechanisms against aphids. In this study, we explored natural variation in Arabidopsis to find genes related to impaired feeding behaviour of *M. persicae*. Natural variation in the occurrence of long probes, a proxy for the success rate of phloem ingestion, was associated to several genomic regions, including the *WRKY22* locus. Polymorphisms in the *WRKY22* promoter and in the last exon most strongly correlated with variation in aphid feeding behaviour. Even though the associations had low statistical power and heritability, knockout lines confirmed an effect of *WRKY22* on aphid performance. Without a functional WRKY22 protein, it was more difficult for aphids to penetrate the epidermis and mesophyll and they arrived later at the vascular bundle. One *WRKY22* overexpression line showed the opposite trend, although the impact was smaller than in the *wrky22* knockouts, most likely due to the moderate extent and short time frame of the overexpression (3- to 5-fold change, induced 24

hours before the experiments). The effects of WRKY22 on aphid performance were marginal in the first eight hours, but more substantial after an infestation period of two weeks; *wrky22* knockout lines eventually showed a reduction in aphid population size of 20% compared to the wild type. Overall, these assays indicate that WRKY22 promotes susceptibility to *M. persicae* aphids. Our RNA-seq and RT-qPCR analyses showed that aphid infestation of *wrky22-3* plants resulted in faster and potentially stronger upregulation of the SA pathway than in wild-type plants. Even though it has been described that JA-induced defences are most effective against aphids (De Vos *et al.*, 2007; Walling, 2008), SA-induced mechanisms have been shown to impose detrimental impacts on aphids as well (Li *et al.*, 2006; Moloi and Westhuizen, 2006).

### Involvement of WRKY22 in biotic and abiotic stress responses

Apart from *WRKY22*'s responsiveness to aphids, we observed a strong activation of the JA pathway in *wrky22-3* as a result of the use of clip cages. JA accumulation is known to be induced by mechanical stimuli (Chehab *et al.*, 2012; Ichimura *et al.*, 2000), wounding, and damage-associated molecular patterns (DAMPs) (Denoux *et al.*, 2008; Doares *et al.*, 1995; Vidhyasekaran, 2014). Since there were no obvious signs of plant damage, the upregulation of the JA pathway was most likely triggered by touch-induced surface stimulation of our samples. *WRKY22* has been shown to be induced in response to touch and wounding (Kilian *et al.*, 2007; Lee *et al.*, 2004). Our data suggest that WRKY22 acts as a suppressor of JA signalling in response to these stimuli. Many other abiotic stimuli have been described to induce *WRKY22* as well, such as prolonged darkness, submergence, cold acclimation, light perception, salinity, potassium starvation, and exposure to ozone (Chawade *et al.*, 2007; Folta *et al.*, 2003; Göhre *et al.*, 2012; Hampton *et al.*, 2004; Hsu *et al.*, 2013; Kilian *et al.*, 2007; Kim *et al.*, 2013; Lee *et al.*, 2005; Monte *et al.*, 2004; Sugimoto *et al.*, 2014; Tosti *et al.*, 2006; Zhou *et al.*, 2011). Hsu *et al.* (2013) identified several potential downstream targets of WRKY22, including genes involved in drought resistance and phosphate starvation. The accumulating evidence for its involvement in abiotic stress responses, warrants a change of view, i.e. that WRKY22 is not solely involved in PTI. A parallel can be drawn to WRKY40, previously known as a repressor of PTI, but recently also recognised as a central player in ABA inhibition during abiotic stress (Chen *et al.*, 2010a; Friedel *et al.*, 2012). It is tempting to look for common patterns among the stimuli that induce *WRKY22*. A shared feature of plant responses to touch, submergence, darkness and temperature, is that they involve adjustments of plant growth and development (Albersheim *et al.*, 2011; Chehab *et al.*, 2009; Jaffe, 1973; Lee *et al.*, 2004). Apart from being defence hormones, SA and JA also contribute to the hormonal modulation of plant growth and development. Specifically, SA is known to be involved in the regulation of photosynthesis, stomatal closure,

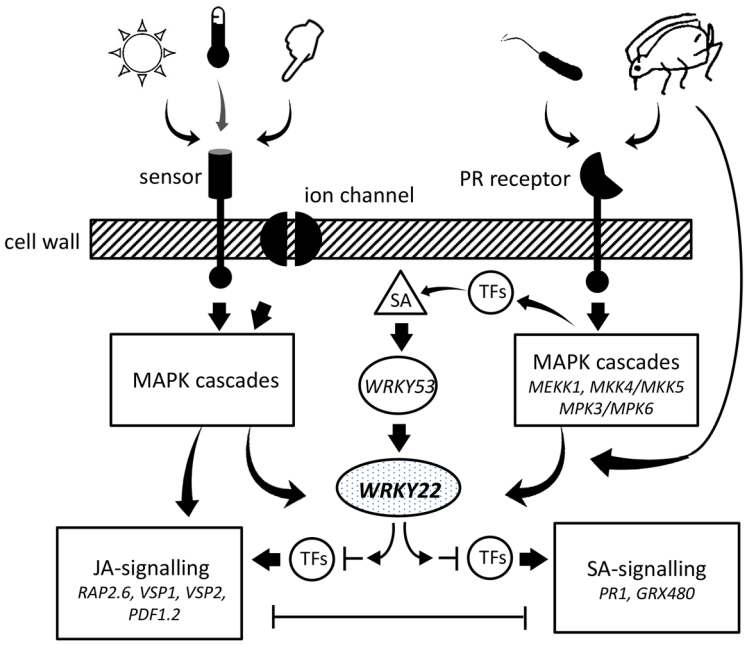



**Figure 10. Hypothetical model of WRKY22's role in plant response to abiotic (left) and biotic (right) stresses.** Changes in, for example, light, temperature and touch, are perceived via sensors and ion channels; plant invasion by organisms, such as bacteria and aphids is mainly perceived via pattern-recognition (PR) receptors. These stimuli induce *WRKY22* directly via MAPK cascades (Asai *et al.*, 2002; Ichimura *et al.*, 2000), or indirectly via SA accumulation (Miao *et al.*, 2004; Miao and Zentgraf, 2007). Alternatively, aphid effectors secreted via the saliva could directly induce *WRKY22* via PR-receptor-independent routes. WRKY22 subsequently integrates signalling of the JA and SA pathway, by inhibiting or activating specific transcription factors (TFs) and other regulatory genes.

respiration, vegetative growth and flowering (Rivas-San Vicente and Plasencia, 2011), and JA particularly in root and inflorescence development (Wasternack and Kombrink, 2010). Although abiotic and biotic stimuli are most likely perceived via stress-specific mechanisms and require differential plant responses, signal-transduction pathways might converge via common regulators, such as WRKY22, in order to fine-tune the interplay between phytohormones.



## SA-JA signal integration

One of the major questions is if WRKY22 is an activator or repressor of SA and JA signalling. Our transcriptome analysis of *wrky22-3* revealed upregulation of JA signalling upon mechanostimulation and upregulation of SA signalling upon aphid

infestation. This would suggest that in wild-type plants, WRKY22 is a suppressor of JA and SA signalling. From previous studies we know, however, that WRKY22 and WRKY29 confer resistance to (hemi)biotrophic and necrotrophic pathogens (Asai *et al.*, 2002; Hsu *et al.*, 2013), and that they are induced by PAMP-triggered MAPK cascades which result in the activation of SA, JA and ET signalling (Zipfel *et al.*, 2004). There is no direct evidence that WRKY22 and its homologue WRKY29 induce SA, JA and ET signalling, and the possibility exists that they are involved in MAPK-triggered processes independent of SA and JA signalling. Nevertheless, their requirement for PTI makes them unlikely candidates for consistent suppression of plant defence hormones. Rather, WRKY22 could be an integrator of SA and JA signals, inhibiting or enforcing both pathways, depending on their interaction with other transcription factors and signalling pathways (**Figure 10**). Since most SA and JA interactions in Arabidopsis are characterised by antagonism (Koornneef *et al.*, 2008; Pieterse *et al.*, 2012), it is remarkable that both pathways can be repressed and induced by the same transcription factor, but not uncommon. WRKY70 has been proposed to be capable of inducing and inhibiting both SA and JA, depending on the strength of the induction (Li *et al.*, 2004). *wrky70* T-DNA lines also showed upregulation of both the SA and the JA pathway (Ülker *et al.*, 2007), comparable to our observations in the *wrky22* T-DNA line. Although many questions remain with regard to the underlying mechanism, our study shows that WRKY22 plays a role in both SA and JA signalling and is involved in transcriptional reprogramming in response to mechanostimulation and aphid infestation. To understand the function of *WRKY22*, its transcriptional network needs to be further unravelled under multiple biotic and abiotic stress conditions. With respect to aphids, WRKY22 plays a controversial role. Its responsiveness to aphid infestation and its potential to suppress JA and SA signalling, would make *WRKY22* an excellent target for aphids to manipulate JA- and SA-dependent host plant defences for their own benefit.

## Materials and methods

### Plants and insects

A collection of 344 natural accessions of *Arabidopsis thaliana* was obtained from the ABRC Stock Center (Baxter *et al.*, 2010). This set was selected in a previous study to represent most intraspecific genetic variation and minimal redundancy (Platt *et al.*, 2010), and was genotyped for 214k SNPs with AtSNPtile1 arrays (Atwell *et al.*, 2010; Horton *et al.*, 2012; Li *et al.*, 2010). T-DNA lines SALK_094892 (*wrky22-3*) and SALK_098205

(*wrky22-4*), and TRANSPLANTA inducible overexpression lines, TPT_4.01250.1C and TPT_4.01250.1E, were obtained from NASC (Coego *et al.*, 2014). Seeds were cold stratified for 72 h at 4˚C before they were sown into pots (5 cm diameter) with pasteurized (4 h at 80˚C) Arabidopsis potting soil (Lentse potgrond, Lent, The Netherlands) in a climate room at 24 ± 1˚C, 50-70% relative humidity, 8h:16 h L:D photoperiod, and a light intensity of 200 µmol $m^{-2}$ $s^{-1}$. Homozygous T-DNA plants were selected based on PCR and harvested for seeds for subsequent experiments. The location of the T-DNA insertion was confirmed via sequencing, and abolishment of *WRKY22* expression was confirmed with RT-qPCR (**Figure 5**, **Table S4**). Expression of *WRKY22* in the TRANSPLANTA inducible overexpression lines (Coego *et al.*, 2014) was measured with RT-qPCR 24 h after application of 10 µM estradiol in water to the plant trays (**Figure 5, Table S4**). Green peach aphids, *Myzus persicae*, were reared on radish, *Raphanus sativus* (L.), at 19˚C, 50-70% relative humidity and an 16h:8h L:D photoperiod.

### Automated video tracking

Aphid behaviour was tracked on 344 natural accessions of Arabidopsis (n=5-6 per accession) according to the methodology of Kloth *et al.* (2015). One aphid was introduced into a well of a 96-well plate containing a leaf disc of 6 mm diameter, abaxial side up, on 1% agar substrate. Wells were covered with cling film, to avoid aphid escape, and 20 aphids were recorded on 20 different accessions simultaneously with a camera mounted above the plate, at 22˚C. EthoVision® XT 8.5 video tracking and analysis software (Noldus Information Technology bv, Wageningen, The Netherlands) was used for automated acquisition of aphid position and velocity. The number and duration of probes were subsequently calculated with the statistical computing program R (R-Core-Team, 2013). Leaf discs were made of intermediately aged leaves of 4- to 5-week-old Arabidopsis plants, one disc per plant. Aphid behaviour was recorded for 85 min, starting at 4.5 hours after inoculation of the aphids. The video-tracking assay was performed in an incomplete block design with each complete replicate consisting of 18 blocks of 20 accessions. Sixty plants were screened each day across 3 blocks, and one replicate of the complete Hapmap collection was acquired in 6 days. An alpha design was generated with Gendex (http://designcomputing.net/gendex/) to assign accessions to each block. Accessions with less than 5 replicates were excluded from the analysis.

### GWA mapping and haplotype analysis

GWA mapping was performed on the proportion of aphids making long probes (> 25 min) with scan_GLS (Kruijer *et al.*, 2015). Within scan_GLS, the statistical program

ASReml (Butler *et al.*, 2009) was used to apply a mixed model to correct for population structure, which was estimated with a kinship matrix including all SNPs. SNPs with a minor allele frequency <0.05 were excluded from analysis. Block and replicate were included in the model as covariates. Generalized heritability was calculated by scan_GLS according to Oakey *et al.* (2006). For haplotype analysis, SNPs with a minor allele frequency above 5% were retrieved from the Arabidopsis 1,001 genomes browser for 173 accessions (Cao *et al.*, 2011). For each domain, haplotypes were defined as unique SNP combinations with a frequency above 5%. For exons, only non-synonymous SNPs were included. A promoter region of 1000 kb was used and gene domains were obtained from Interpro (Mitchell *et al.*, 2015). Promoter motifs were retrieved from Athamap (Hehl and Bülow, 2014).

### RT-qPCR

For each sample, two intermediately aged leaves per 4- to 5-week-old Arabidopsis plant were harvested between 12 and 3 pm. Samples were immediately frozen in liquid nitrogen, and stored at -80˚C until processing. RNA was isolated from homogenised leaf material with an InviTrap® Spin Plant RNA kit, and treated with Ambion® TURBO DNA-free™ according to the manufacturer's instructions. RNA was quantified with a NanoDrop® ND-1000 spectrophotometer, and integrity was assessed with gel electrophoresis. DNA-free RNA was converted into cDNA using the Bio-Rad iScript™ cDNA synthesis kit. Quantitative reverse transcription PCR was carried out on a Bio-Rad IQ™5 system using SYBR Green. For each primer combination (**Table S4**), RT-qPCR products were sequenced to validate the region of amplification. To test aphid induction of *WRKY22*, *PR1*, *VSP2*, and *PDF1.2*, plants were treated with and without aphids (n=4). For infested samples, a Petri dish with indentation for the petiole was used to contain 15 adult *M. persicae* aphids on the leaf, to inflict as little mechanostimulation as possible. Four biological replicates were collected for three treatments: (1) an empty Petri dish for 48 h, (2) a Petri dish for 48 h with addition of aphids in the last 6 h, (3) a Petri dish with aphids for 48 h.

### Electrical Penetration Graph (EPG) recording

Feeding behaviour of *M. persicae* aphids was investigated with EPG recording on 4- to 5-week-old Arabidopsis plants, using direct current (DC) according to the methodology of Ten Broeke et al. (2013a). To adjust the radish-reared aphids to Arabidopsis, aphids were transferred to Col-0 Arabidopsis plants 24 h before the experiments. EPG recording was performed at 22 ± 2˚C and light intensity of 120 μmol $m^{-2}$ $s^{-1}$, using clean plants and one aphid per plant. An electrode was inserted in the potting soil and

a thin gold wire of 1.5 cm was gently attached to the dorsum of an adult, wingless aphid with silver glue. The electrical circuit was completed when the aphid's piercing-sucking stylet mouthparts penetrated the plant cuticle. Electrical signals associated with stylet activities were recorded and annotated with EPG Stylet+ software and further processed in R (R-Core-Team, 2013; Tjallingii, 1988) (http://www.epgsystems.eu). Between 20 and 24 biological replicates were measured on T-DNA lines (Col-0: n=24, *wrky22-3*: n=22, *wrky22-4*: n=20) and between 15 and 19 on overexpression lines (Col-0: n=15, OE.c: n=19, OE.e: n=19). *WRKY22* overexpression was induced by supplying estradiol solution to the plants 24 hours before the experiment. To correct for potential side-effects of estradiol, the wild-type plants received the same treatment as the overexpression lines.

### Aphid population development

To assess aphid developmental rate and population size, 2.5-week-old Arabidopsis plants were infested with one *M. persicae* neonate of 0 to 24 hour old and placed in a climate room at 24 ± 1˚C, 50-70% relative humidity, an 8h:16 h L:D photoperiod, 200 μmol $m^{-2}$ $s^{-1}$ light intensity. A soap-diluted water barrier prevented aphids from moving between plants. None of the aphids developed wings. From day 7 onwards, occurrence of the first offspring was checked twice per day using 5x magnification glasses (Col-0: n=18, *wrky22-3*: n=19, *wrky22-4*: n=22). The number of aphids per plant was counted at 14 days after infestation. Plants without an adult aphid eight days after introduction, and plants without any adults or neonates 14 days after introduction were excluded from the analysis.

### Statistics

Data were tested for a normal distribution and homogeneity of variances using the Shapiro test and Levene's test. Non-parametric data sets were assessed with Mann-Whitney U tests (2 groups), or Kruskall Wallis tests (> 2 groups). Data sets with a normal distribution were tested with a Student's t-test (2 groups) or a one-way ANOVA (>2 groups).

### RNA-seq analysis

RNA-seq analysis was conducted on leaves of Col-0 and *wrky22-3* with three biological replicates per treatment. Five leaves of five different 4- to 5-week-old plants were pooled per sample. Plants had been exposed to one of three treatments: (1) an empty clip cage for 48 h, (2) a clip cage for 48 h with addition of 15 aphids in the last 6 h, (3) a clip cage with 15 aphids for 48 h. Only 4$^{th}$ instar nymphs and adult *M. persicae* aphids were used.

Experiments were conducted simultaneously in a climate chamber (24 ± 1°C, 50-70% relative humidity, 8h:16h L:D photoperiod, and a light intensity of 120 µmol $m^{-2}$ $s^{-1}$), but in a separate cages with an air circulation system that prevented contamination of plant volatiles between treatments (Menzel *et al.*, 2014). Samples were harvested in two batches between 1 and 4 pm, immediately frozen in liquid nitrogen, and stored at -80°C until processing. RNA was isolated and checked according to the description above (260/280 OD range: 2.0-2.2, 260/230 OD range: 1.9-2.3). Library preparation was performed with a TruSeq™ RNA Sample Prep Kit (Illumina®) and between 11 to 24 million single-end 50-bp reads were sequenced per sample with Illumina® HiSeq™ 2000 in 3 lanes, multiplexed with 12 samples per lane. Reads were cleaned from adaptors and trimmed to 51 bp using the program Trimmomatic version 0.32 (Bolger *et al.*, 2014). Quality control was performed with FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc). Reads were mapped to the TAIR10 Arabidopsis reference genome (https://www.arabidopsis.org/) with Tophat version 2.0.13, intron length 20-2000 (Trapnell *et al.*, 2013; Trapnell *et al.*, 2012). An index file was built with Bowtie 2 (Langmead *et al.*, 2009). Transcript assembly, quantification, normalisation and differential expression analysis were performed with Cufflinks, using the bias detection and correction algorithm, multi-read correction for reads mapping to multiple locations, and a minimum alignment count of 10. Treatments were both compared between plant lines (Col-0 versus mutant) and within plant line (empty clip cage versus 6 hpi, empty clip cage versus 48 hpi, and 6 hpi versus 48 hpi). Only differentially expressed genes (False Discovery Rate Q-value<0.05) with an absolute fold change ≥ 2 ($\log_2 \geq 1$) were taken into account. Differentially expressed genes were tested for overrepresentation of biological processes against a reference set including all transcripts in the complete data set with at least 1 count, using the application BiNGO in Cytoscape (Cline *et al.*, 2007; Maere *et al.*, 2005). Genes associated with cell-wall processes were selected and classified based upon their TAIR description, the heatmap was constructed with the R package 'gplots' (Warnes *et al.*, 2009).

## Acknowledgements

This work was supported by The Netherlands Organization for Scientific Research (NWO) through the Technology Foundation Perspective Programme 'Learning from Nature' [STW10989] and the ZonMw programme 'Enabling Technologies' [435000017]. We would like to thank Johan Bucher of the Department of Plant Breeding, Wageningen University, for providing the clip cages.

## Supplementary information

**Table S1.** Percentage of aphids making long probes (> 25 min) 4.5 hour after inoculation on 344 natural Arabidopsis accessions.

| Accession | Line | % | Accession | Line | % | Accession | Line | % |
|---|---|---|---|---|---|---|---|---|
| CS76214 | Pro-0 | 20 | CS76208 | Paw-3 | 50 | CS28734 | Sh-0 | 67 |
| CS28163 | Co-2 | 33 | CS76221 | ROM-1 | 50 | CS28739 | Si-0 | 67 |
| CS28808 | Wag-3 | 33 | CS76276 | UKNW06-060 | 50 | CS28787 | Uk-1 | 67 |
| CS76107 | CAM-16 | 33 | CS76281 | UKSE06-192 | 50 | CS28795 | Utrecht | 67 |
| CS76127 | Est-1 | 33 | CS76108 | CAM-61 | 60 | CS28800 | Ven-1 | 67 |
| CS76287 | UKSE06-429 | 33 | CS76109 | Can-0 | 60 | CS28804 | Wa-1 | 67 |
| CS28214 | Dra-2 | 40 | CS76129 | Fei-0 | 60 | CS28812 | WAR | 67 |
| CS28091 | Boot-1 | 50 | CS76131 | Fjä1-2 | 60 | CS76088 | Alc-0 | 67 |
| CS28097 | Bs-2 | 50 | CS76137 | Gr-1 | 60 | CS76090 | ALL1-3 | 67 |
| CS28160 | Cnt-1 | 50 | CS76177 | Lp2-6 | 60 | CS76093 | Bå1-2 | 67 |
| CS28217 | Ede-1 | 50 | CS76252 | TOU-A1-115 | 60 | CS76094 | Bay-0 | 67 |
| CS28345 | Hh-0 | 50 | CS76291 | UKSE06-628 | 60 | CS76097 | Bla-1 | 67 |
| CS28419 | Kr-0 | 50 | CS28049 | Ann-1 | 67 | CS76101 | Br-0 | 67 |
| CS28492 | Mh-0 | 50 | CS28053 | Ba-1 | 67 | CS76104 | BUI | 67 |
| CS28633 | PHW-33 | 50 | CS28135 | Chat-1 | 67 | CS76105 | Bur-0 | 67 |
| CS28635 | PHW-35 | 50 | CS28140 | CIBC-2 | 67 | CS76124 | Duk | 67 |
| CS28645 | Pn-0 | 50 | CS28141 | CIBC-4 | 67 | CS76128 | Fäb-4 | 67 |
| CS28725 | Sav-0 | 50 | CS28252 | Fi-1 | 67 | CS76149 | Ka-0 | 67 |
| CS28758 | Tha-1 | 50 | CS28344 | Hey-1 | 67 | CS76159 | Lc-0 | 67 |
| CS28823 | Ws | 50 | CS28350 | Hn-0 | 67 | CS76164 | Ler-1 | 67 |
| CS76098 | Blh-1 | 50 | CS28373 | Jm-1 | 67 | CS76169 | Lis-1 | 67 |
| CS76106 | C24 | 50 | CS28423 | Krot-2 | 67 | CS76171 | Lisse | 67 |
| CS76150 | Kas-1 | 50 | CS28459 | Li-6 | 67 | CS76174 | Lom1-1 | 67 |
| CS76152 | Kelsterbach-4 | 50 | CS28614 | PHW-14 | 67 | CS76181 | MIB-15 | 67 |
| CS76165 | LI-OF-095 | 50 | CS28628 | PHW-28 | 67 | CS76195 | Na-1 | 67 |
| CS76172 | LL-0 | 50 | CS28631 | PHW-31 | 67 | CS76200 | Ömö2-1 | 67 |
| CS76190 | Mr-0 | 50 | CS28685 | Rhen-1 | 67 | CS76202 | Ost-0 | 67 |
| CS76196 | NC-6 | 50 | CS28692 | Rou-0 | 67 | CS76203 | Oy-0 | 67 |

**Table S1.** (continued)

| Accession | Line | % | Accession | Line | % | Accession | Line | % |
|---|---|---|---|---|---|---|---|---|
| CS76207 | PAR-5 | 67 | CS76293 | Ull2-3 | 80 | CS28750 | Ste-0 | 83 |
| CS76226 | Se-0 | 67 | CS76298 | Vår2-1 | 80 | CS28779 | Tscha-1 | 83 |
| CS76227 | Shahdara | 67 | CS28014 | Amel-1 | 83 | CS28788 | Uk-2 | 83 |
| CS76228 | SLSP-30 | 67 | CS28017 | An-2 | 83 | CS28810 | Wag-5 | 83 |
| CS76231 | St-0 | 67 | CS28051 | Arby-1 | 83 | CS28814 | Wc-2 | 83 |
| CS76232 | Ste-3 | 67 | CS28064 | Benk-1 | 83 | CS28822 | Wl-0 | 83 |
| CS76233 | T1040 | 67 | CS28090 | Blh-2 | 83 | CS28833 | Wt-3 | 83 |
| CS76234 | T1060 | 67 | CS28099 | Bsch-0 | 83 | CS28848 | Ors-1 | 83 |
| CS76236 | T1110 | 67 | CS28165 | Co-4 | 83 | CS28849 | Ors-2 | 83 |
| CS76244 | Tamm-2 | 67 | CS28193 | Com-1 | 83 | CS76083 | 11ME1.32 | 83 |
| CS76254 | TOU-A1-12 | 67 | CS28200 | Da-0 | 83 | CS76084 | 11PNA4.101 | 83 |
| CS76257 | TOU-A1-67 | 67 | CS28243 | Est-0 | 83 | CS76091 | An-1 | 83 |
| CS76262 | TOU-H-13 | 67 | CS28268 | Fr-4 | 83 | CS76095 | Belmonte-4-94 | 83 |
| CS76263 | TOU-I-17 | 67 | CS28277 | Ge-1 | 83 | CS76096 | Bg-2 | 83 |
| CS76280 | UKSE06-062 | 67 | CS28279 | Gel-1 | 83 | CS76099 | Bor-1 | 83 |
| CS76282 | UKSE06-272 | 67 | CS28364 | Je-0 | 83 | CS76103 | Bu-0 | 83 |
| CS76288 | UKSE06-466 | 67 | CS28369 | Jl-3 | 83 | CS76110 | Cen-0 | 83 |
| CS76306 | Zdr-6 | 67 | CS28394 | Kl-5 | 83 | CS76113 | Col-0 | 83 |
| CS28108 | Bu-8 | 80 | CS28457 | Li-5:2 | 83 | CS76122 | DraIV | 83 |
| CS28241 | Es-0 | 80 | CS28490 | Mc-0 | 83 | CS76125 | Eden-2 | 83 |
| CS28420 | Kro-0 | 80 | CS28495 | Mnz-0 | 83 | CS76126 | Edi-0 | 83 |
| CS28583 | Old-1 | 80 | CS28510 | N4 | 83 | CS76132 | Fjä1-5 | 83 |
| CS28729 | Sei-0 | 80 | CS28513 | N7 | 83 | CS76134 | Gd-1 | 83 |
| CS76086 | 627ME-4Y1 | 80 | CS28580 | Ob-1 | 83 | CS76138 | Gul1-2 | 83 |
| CS76114 | Ct-1 | 80 | CS28610 | PHW-10 | 83 | CS76139 | Gy-0 | 83 |
| CS76115 | CUR-3 | 80 | CS28613 | PHW-13 | 83 | CS76140 | Hi-0 | 83 |
| CS76157 | LAC-3 | 80 | CS28620 | PHW-20 | 83 | CS76141 | Hod | 83 |
| CS76184 | MIB-84 | 80 | CS28622 | PHW-22 | 83 | CS76144 | HR-5 | 83 |
| CS76191 | Mrk-0 | 80 | CS28636 | PHW-36 | 83 | CS76145 | Hs-0 | 83 |
| CS76224 | Sap-0 | 80 | CS28637 | PHW-37 | 83 | CS76148 | JEA | 83 |
| CS76230 | Sq-8 | 80 | CS28640 | Pla-0 | 83 | CS76151 | KBS-Mac-8 | 83 |
| CS76251 | Tottarp-2 | 80 | CS28650 | Pog-0 | 83 | CS76155 | Köln | 83 |
| CS76259 | TOU-C-3 | 80 | CS28743 | Sp-0 | 83 | CS76161 | LDV-25 | 83 |

**Table S1.** (continued)

| Accession | Line | % | Accession | Line | % | Accession | Line | % |
|---|---|---|---|---|---|---|---|---|
| CS76163 | LDV-58 | 83 | CS76278 | UKNW06-436 | 83 | CS28326 | Gr-5 | 100 |
| CS76168 | Lip-0 | 83 | CS76279 | UKNW06-460 | 83 | CS28332 | Gu-1 | 100 |
| CS76170 | Lis-2 | 83 | CS76283 | UKSE06-278 | 83 | CS28336 | Ha-0 | 100 |
| CS76176 | Lp2-2 | 83 | CS76284 | UKSE06-349 | 83 | CS28382 | Kelsterbach-2 | 100 |
| CS76180 | Map-42 | 83 | CS76286 | UKSE06-414 | 83 | CS28395 | Kn-0 | 100 |
| CS76185 | MNF-Che-2 | 83 | CS76289 | UKSE06-482 | 83 | CS28407 | KNO-11 | 100 |
| CS76186 | MNF-Jac-32 | 83 | CS76292 | UKSW06-202 | 83 | CS28454 | Li-3 | 100 |
| CS76187 | MNF-Pot-48 | 83 | CS76297 | Van-0 | 83 | CS28461 | Li-7 | 100 |
| CS76188 | MNF-Pot-68 | 83 | CS76299 | VOU-1 | 83 | CS28527 | Nc-1 | 100 |
| CS76192 | Mt-0 | 83 | CS76300 | VOU-2 | 83 | CS28550 | NFC-20 | 100 |
| CS76193 | Mz-0 | 83 | CS76302 | Wil-1 | 83 | CS28564 | No-0 | 100 |
| CS76197 | Nd-1 | 83 | CS76304 | Wt-5 | 83 | CS28568 | Nok-1 | 100 |
| CS76205 | PAR-3 | 83 | CS76305 | Yo-0 | 83 | CS28573 | Nw-0 | 100 |
| CS76209 | Pent-1 | 83 | CS76308 | Zdrl | 83 | CS28575 | Nw-2 | 100 |
| CS76210 | Per-1 | 83 | CS76213 | Pna-17 | 86 | CS28578 | Nz1 | 100 |
| CS76211 | Petergof | 83 | CS22689 | RRS-10 | 100 | CS28587 | Or-0 | 100 |
| CS76215 | Pu2-23 | 83 | CS28007 | Aa-0 | 100 | CS28626 | PHW-26 | 100 |
| CS76216 | Ra-0 | 83 | CS28013 | Alst-1 | 100 | CS28651 | Pr-0 | 100 |
| CS76217 | Rak-2 | 83 | CS28018 | Ang-0 | 100 | CS28663 | Pu2-24 | 100 |
| CS76220 | Rmx-A180 | 83 | CS28054 | Baa-1 | 100 | CS28713 | RRS-7 | 100 |
| CS76222 | Rsch-4 | 83 | CS28063 | Be-1 | 100 | CS28720 | S96 | 100 |
| CS76225 | Sav-0 | 83 | CS28128 | Ca-0 | 100 | CS28724 | Sapporo-0 | 100 |
| CS76229 | Sparta-1 | 83 | CS28133 | Cha-0 | 100 | CS28732 | Sg-1 | 100 |
| CS76235 | T1080 | 83 | CS28142 | CIBC-5 | 100 | CS28760 | Tiv-1 | 100 |
| CS76239 | T540 | 83 | CS28158 | Cit-0 | 100 | CS28780 | Tsu-0 | 100 |
| CS76247 | TDr-18 | 83 | CS28181 | CSHL-5 | 100 | CS28786 | Ty-0 | 100 |
| CS76248 | TDr-3 | 83 | CS28201 | Da(1)-12 | 100 | CS28809 | Wag-4 | 100 |
| CS76253 | TOU-A1-116 | 83 | CS28202 | Db-0 | 100 | CS28847 | Zu-1 | 100 |
| CS76256 | TOU-A1-62 | 83 | CS28208 | Di-1 | 100 | CS76085 | 328PNA054 | 100 |
| CS76264 | TOU-I-2 | 83 | CS28210 | Do-0 | 100 | CS76087 | Ag-0 | 100 |
| CS76265 | TOU-I-6 | 83 | CS28236 | Ep-0 | 100 | CS76089 | ALL1-2 | 100 |
| CS76272 | UKID37 | 83 | CS28280 | Gie-0 | 100 | CS76092 | App1-16 | 100 |
| CS76273 | UKID48 | 83 | CS28282 | Go-0 | 100 | CS76100 | Bor-4 | 100 |

**Table S1.** (continued)

| Accession | Line | % | Accession | Line | % |
|---|---|---|---|---|---|
| CS76102 | Brö1-6 | 100 | CS76218 | Ren-1 | 100 |
| CS76111 | CIBC-17 | 100 | CS76219 | Rev-2 | 100 |
| CS76116 | Cvi-0 | 100 | CS76223 | Sanna-2 | 100 |
| CS76119 | DraIV | 100 | CS76238 | T510 | 100 |
| CS76120 | DraIV | 100 | CS76240 | T620 | 100 |
| CS76121 | DraIV | 100 | CS76242 | Ta-0 | 100 |
| CS76123 | DraIV | 100 | CS76243 | TÅD | 100 |
| CS76135 | Ge-0 | 100 | CS76245 | TDr-1 | 100 |
| CS76136 | Got-7 | 100 | CS76249 | TDr-8 | 100 |
| CS76142 | Hov4-1 | 100 | CS76250 | Tomegap-2 | 100 |
| CS76143 | Hovdala-2 | 100 | CS76255 | TOU-A1-43 | 100 |
| CS76146 | HSm | 100 | CS76258 | TOU-A1-96 | 100 |
| CS76147 | In-0 | 100 | CS76260 | TOU-E-11 | 100 |
| CS76153 | Kin-0 | 100 | CS76261 | TOU-H-12 | 100 |
| CS76154 | Kno-18 | 100 | CS76266 | TOU-J-3 | 100 |
| CS76156 | Nordborg | 100 | CS76267 | TOU-K-3 | 100 |
| CS76158 | LAC-5 | 100 | CS76268 | Ts-1 | 100 |
| CS76160 | LDV-14 | 100 | CS76269 | Udul | 100 |
| CS76162 | LDV-34 | 100 | CS76270 | UKID101 | 100 |
| CS76166 | Liarum | 100 | CS76274 | UKID80 | 100 |
| CS76167 | Lillö-1 | 100 | CS76275 | UKNW06-059 | 100 |
| CS76173 | Lm-2 | 100 | CS76277 | UKNW06-386 | 100 |
| CS76175 | Löv-5 | 100 | CS76285 | UKSE06-351 | 100 |
| CS76179 | Lz-0 | 100 | CS76290 | UKSE06-520 | 100 |
| CS76182 | MIB-22 | 100 | CS76295 | Ull3-4 | 100 |
| CS76183 | MIB-28 | 100 | CS76296 | Uod-7 | 100 |
| CS76189 | MOG-37 | 100 | CS76301 | Wei-0 | 100 |
| CS76194 | N13 | 100 | CS76303 | Ws-0 | 100 |
| CS76198 | NFA-10 | 100 | CS76307 | Zdrl | 100 |
| CS76199 | NFA-8 | 100 | | | |
| CS76201 | Ör-1 | 100 | | | |
| CS76206 | PAR-4 | 100 | | | |
| CS76212 | PHW-34 | 100 | | | |

**Table S2.** Aphid feeding behaviour, measured by 8-hour EPG recordings on wild type (Col-0) and *wrky22* T-DNA lines (*wrky22-3* and *wrky22-4*).

| Tissue | Variable | Col-0 | *wrky22-3* | *wrky22-4* |
|---|---|---|---|---|
| Epidermis/mesophyll | Total duration NP (min) | 111.6 ± 11.8 | 105.5 ± 13.6ns | 153.1 ± 17.2ns |
| | Latency to 1st C (min) | 6 ± 1.2 | 7.5 ± 1.8ns | 12.4 ± 2.5ns |
| | Total duration C (min) | 167.5 ± 11.2 | **197.6 ± 12.8*** | 192 ± 14.1ns |
| | Number of C | 43.1 ± 4.4 | 41.4 ± 4.9ns | 51.6 ± 5.5ns |
| | Number of C < 3 min | 29.3 ± 4.1 | 25.5 ± 4.2ns | 38 ± 4.9ns |
| | Number of C < 20 sec | 12.1 ± 2.5 | 8 ± 2.1ns | 17.6 ± 2.9ns |
| | Mean duration C (min) | 4.6 ± 0.4 | **6.7 ± 1*** | 4.4 ± 0.4ns |
| | Potential drops in C ($min^{-1}$) | 1.2 ± 0 | 1.2 ± 0ns | 1.2 ± 0.1ns |
| | Total duration F (min) | 2.8 ± 2.8 | 0 ± 0ns | 6.2 ± 4.4ns |
| | Number of F | 0.1 ± 0.1 | 0.1 ± 0.1ns | 0.4 ± 0.2ns |
| Vascular bundle | Latency to 1st E1 after C (min) | 67.8 ± 7.3 | 74.9 ± 8.1ns | 119.8 ± 23.7ns |
| | Mean latency E1 after C (min) | 10.5 ± 0.7 | 12.7 ± 1.3ns | 14.2 ± 1.6ns |
| | Total duration E1 (min) | 10.4 ± 1.3 | 13.3 ± 1.7ns | 7.1 ± 0.6ns |
| | Number of E1 | 11.3 ± 1.4 | 12.5 ± 1.3ns | 8.1 ± 0.9ns |
| | Mean duration E1 (min) | 1.1 ± 0.2 | 1.2 ± 0.2ns | 1 ± 0.1ns |
| | Number of single E1 | 1.2 ± 0.3 | 1.4 ± 0.3ns | 1.4 ± 0.3ns |
| | E1 in E (%) | 8.4 ± 1.5 | 10.8 ± 1.7ns | 12.7 ± 3.6ns |
| | Latency to 1st E2 (min) | 96.5 ± 12.5 | 107.3 ± 17.8ns | 144.1 ± 24.6ns |
| | Total duration E2 (min) | 168.3 ± 19.9 | 154 ± 21.3ns | 117.9 ± 24.1ns |
| | Number of E2 | 8.2 ± 1 | 9 ± 0.9ns | 5.8 ± 0.8ns |
| | Mean duration E2 (min) | 32.5 ± 7 | 41.8 ± 21ns | 24.1 ± 6.9ns |
| | Total duration E2s (min) | 148.8 ± 20.9 | 129.1 ± 22.9ns | 103.5 ± 24.1ns |
| | Aphids with E2 (%) | 100 ± 0 | 100 ± 0ns | 100 ± 0ns |
| | E2/C ratio | 1.3 ± 0.3 | 1.6 ± 0.8ns | **1 ± 0.4*** |
| | Latency to vascular bundle (min) | 69.3 ± 6.5 | 78.7 ± 8.6ns | **130.2 ± 22.5**** |
| | Total duration G (min) | 19.4 ± 6.3 | 9.5 ± 3.8ns | **3.6 ± 1.8*** |
| | Number of G | 0.6 ± 0.2 | 0.4 ± 0.1ns | 0.2 ± 0.1ns |

Means ± standard error, *P<0.05; **P<0.01; ***P<0.001 (Mann-Whitney U pairwise comparisons: wrky22-3 versus Col-0 and wrky22-4 versus Col-0). NP = non-penetration, C = pathway, , F = penetration difficulties, E1 = phloem salivation, E2 = phloem ingestion, E2s = sustained phloem ingestion (> 10 min), G = xylem ingestion. Single E1s were phloem salivations that were not directly followed or preceded by phloem uptake. Latency to the vascular bundle was calculated as the time from the start of the recording to the first contact with either a phloem or xylem vessel.

**Table S3.** Aphid feeding behaviour, measured by 8-hour EPG recordings on wild type (Col-0) and *wrky22* inducible overexpression lines (OE.c and OE.e). *WRKY22* overexpression was induced by supplying estradiol solution to all plants 24 hours before the experiment.

| Tissue | Variable | Col-0 | OE.c | OE.e |
|---|---|---|---|---|
| Epidermis/mesophyll | Total duration NP (min) | 171.6 ± 20.5 | 178.5 ± 20ns | 160.6 ± 16.4ns |
| | Latency to 1st C (min) | 16.3 ± 3.4 | 12.6 ± 2.6ns | 17 ± 4.7ns |
| | Total duration C (min) | 179.9 ± 9.6 | 164.6 ± 8.2ns | 158.1 ± 8.4ns |
| | Number of C | 43.2 ± 4.7 | 39.5 ± 3.8ns | 39.4 ± 3.3ns |
| | Number of C < 3 min | 30.3 ± 4.5 | 27.5 ± 3.9ns | 27.7 ± 3ns |
| | Number of C < 20 sec | 13.1 ± 2.5 | 11.7 ± 2.5ns | 12.1 ± 1.7ns |
| | Mean duration C (min) | 5 ± 0.6 | 4.7 ± 0.4ns | 4.3 ± 0.3ns |
| | Total duration F (min) | 14.2 ± 8.8 | 8.3 ± 4.8ns | 0.7 ± 0.6ns |
| | Number of F | 0.5 ± 0.3 | 0.6 ± 0.2ns | 0.2 ± 0.1ns |
| Vascular bundle | Latency to 1st E1 after C (min) | 139.5 ± 21 | 94.9 ± 19ns | 104.6 ± 13.2ns |
| | Mean latency E1 after C (min) | 13.2 ± 1.2 | 11.5 ± 0.7ns | 13.6 ± 1.4ns |
| | Total duration E1 (min) | 6.6 ± 0.7 | 8.3 ± 1.2ns | 8.5 ± 1.2ns |
| | Number of E1 | 8.9 ± 1.3 | 11.6 ± 1.2ns | 9.6 ± 1ns |
| | Mean duration E1 (min) | 0.8 ± 0.1 | 0.7 ± 0ns | 0.9 ± 0.1ns |
| | Number of single E1 | 1 ± 0.2 | 0.8 ± 0.3ns | 1 ± 0.3ns |
| | E1 in E (%) | 11.3 ± 2.9 | 9.3 ± 1.6ns | 7.4 ± 1.4ns |
| | Latency to 1st E2 (min) | 170.3 ± 24 | **108.8 ± 18.7*** | 130.9 ± 16.9ns |
| | Total duration E2 (min) | 107.7 ± 23.4 | 120.3 ± 17.1ns | 150.1 ± 17.8ns |
| | Number of E2 | 6.8 ± 1.1 | 9.4 ± 1ns | 7.6 ± 1ns |
| | Mean duration E2 (min) | 19.2 ± 5.2 | 12.5 ± 1.7ns | 25.4 ± 4.5ns |
| | Total duration E2s (min) | 91.7 ± 24.3 | 89.6 ± 15.5ns | 131 ± 17.9ns |
| | Aphids with E2 (%) | 100 ± 0 | 100 ± 0ns | 100 ± 0ns |
| | E2/C ratio | 0.7 ± 0.2 | 0.8 ± 0.1ns | 1 ± 0.2ns |
| | Latency to vascular bundle (min) | 155.7 ± 21.2 | **107.5 ± 18.8*** | 118.6 ± 14.7ns |
| | Total duration G (min) | 0 ± 0 | 0 ± 0ns | 2.1 ± 2ns |
| | Number of G | 0 ± 0 | 0 ± 0ns | 0.1 ± 0.1ns |

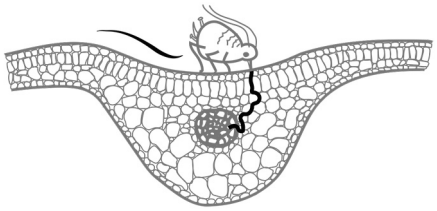
Means ± standard error, *P<0.05; **P<0.01; ***P<0.001 (Mann-Whitney U pairwise comparisons: OE.c versus Col-0 and OE.e versus Col-0). NP = non-penetration, C = pathway, , F = penetration difficulties, E1 = phloem salivation, E2 = phloem ingestion, E2s = sustained phloem ingestion (> 10 min), G = xylem ingestion. Single E1s were phloem salivations that were not directly followed or preceded by phloem uptake. Latency to the vascular bundle was calculated as the time from the start of the recording to the first contact with either a phloem or xylem vessel.

**Table S4.** Primers used for PCR and RT-qPCR.

| **PCR** | | | |
|---|---|---|---|
| **Mutant/insertion** | **SALK code** | **Direction** | **Primer** |
| T-DNA insertion | LBb1.3 | | ATTTTGCCGATTTCGGAAC |
| *wrky22-3* | SALK_094892 | Fw | AAACCCTATATTCACCGTCGG |
| | | Rev | ATATTCCTCCGGTGGTAGTGG |
| *wrky22-4* | SALK_098205 | Fw | CACAGAACCAGAAACGTCCTC |
| | | Rev | ATATTCCTCCGGTGGTAGTGG |
| **RT-qPCR** | | | |
| **Gene name** | **AGI code** | **Direction** | **Primer** |
| *PEX4* | *At5g25760* | Fw | TGCAACCTCCTCAAGTTCG |
| | *At5g25760* | Rev | CACAGACTGAAGCGTCCAAG |
| *ACT2* | *At3g18780* | Fw | CCCGATGGGCAAGTCATCACGAT |
| | *At3g18780* | Rev | GTCTCGTGGAtTCCAGCAGCTTCC |
| *WRKY22* | *At4g01250* | Fw | CATGGCGAAAGTACGGACAG |
| | *At4g01250* | Rev | GAATTACGGTGTGTCGGAGC |
| *PDF1.2* | *At5G44420* | Fw | TTGCTGCTTTCGACGCA |
| | *At5G44420* | Rev | TGTCCCACTTGGCTTCTCG |
| *VSP2* | *At5G24770* | Fw | CATAGACTTCGACACGGTGC |
| | *At5G24770* | Rev | TTTGCACACGGTTTTGGAGT |
| *PR1* | *At2G14610* | Fw | GTTGCAGCCTATGCTCGGAG |
| | *At2G14610* | Rev | CCGCTACCCCAGGCTAAGTT |

# Chapter 7

## General discussion



Karen J. Kloth

## Introduction

To unravel plant-aphid interactions, one needs to understand the biology of both. Particularly, when exploring links between the plant genome and aphid behaviour, small details in insect and plant performance can be of major importance. Plant resistance is often a complex trait with variation in individual genes which play a minor role in a polygenic framework. The step from one plant gene, to the behaviour of a herbivorous insect is huge. Transcriptional and translational regulation, tissue- and developmental-specificity, protein-protein interactions, and metabolic effects are just a few processes connecting the DNA code with a plant phenotype. In addition, the aphid's behavioural and reproductive response is affected by intrinsic traits of the aphid, such as its developmental stage, and plant characteristics including smell, taste, accessibility of the phloem, and the digestibility, toxicity and nutritional value of the phloem sap. This complexity makes the identification of resistance genes challenging, particularly when studying natural genetic variation with often only subtle effects on gene efficacy. Taking a 'blind' approach, and exploring whole genomes consisting of several thousand polymorphisms to identify a single base pair change that affects aphid performance, is like finding a needle in a haystack. Gene discovery and functional characterisation, therefore, require both specificity and pragmatism. Here, I will discuss how to reduce missing heritability and exploit phenotypes. First, I will elaborate on the biology of plant-aphid interactions and address some exciting aspects which deserve more attention in future studies.

## The parasite-like nature of aphids

Aphids, and phloem-feeding insects in general, are distinctly different from many herbivorous insects from other feeding guilds by their intimate relation with the host plant. With their delicate feeding apparatus they feed from one single plant cell, i.e. a sieve element, for hours or even days (Tjallingii, 1995). This is beyond comparison to other piercing-sucking insects, such as thrips, which damage and feed on several hundred or thousand epidermal and mesophyll cells throughout their live (de Kogel *et al.*, 1998; Lewis, 1973). For aphids, maintaining a sieve element functional is essential for long-term harvesting of nutrients. The interest of aphids in the vitality of the host extends even beyond its own lifetime, since colonies stay on a suitable host plant for several generations. Overall, the resemblance between plant-aphid interactions and parasitism can hardly escape. Parasitism has been defined in several ways, but is generally regarded as a non-mutualistic symbiotic relationship, where one organism benefits at the expense

of another organism, whilst living in or on its host for extended time without directly killing it (Crofton, 1970; Loker and Hofkin, 2015). However, in most plant-insect studies, aphids are not referred to as parasites. Herbivorous insects are most commonly categorised by their feeding guild (biting-chewing, leaf-mining, piercing-sucking) or host specificity (specialist, oligophagous or generalist insects) (Ali and Agrawal, 2012; Panda and Khush, 1995; Schoonhoven *et al.*, 2005), and not by the degree of intimacy with a host plant in terms of duration of the infestation and dependency on the physiological vitality of the host. The latter point of view would, however, help in acquiring new insights in plant-insect interactions. Having a closer look, there are more parasitic aspects to aphids (**Table 1**). High reproduction rates and sexual and asexual morphs are typical characteristics of parasites, since they often have a low survival rate due to their dependency on many biotic and abiotic factors (Loker and Hofkin, 2015). In addition, parasites are often specialised in reaching and feeding from a specific tissue within their host. Parasitic weeds, such as *Striga* spp., for example, invade the vascular bundle of the host from where they retrieve photoassimilates, water and nutrients (Bouwmeester *et al.*, 2007). *Striga* spp. are even obligate parasites, since they have lost photosynthetic capacity and completely rely on the retrieval of photoassimilates from their host. In comparison, aphids, have developed a highly specialised feeding apparatus to reach the phloem vessels, and some aphids, such as *Myzus persicae*, have lost the ability to convert glucuronic acid to ascorbic acid, (Mittler *et al.*, 1970), which is highly abundant in the phloem of many plant species (Franceschi and Tarlyn, 2002).

## Host plant manipulation and aphid saliva

Many parasites manipulate the physiology of their host plant for their own benefit. An impressive example is the formation of pseudoflowers in *Boechera stricta*, induced by the plant parasitic rust fungus *Puccinia monoica* (Cano *et al.*, 2013). These pseudoflowers have no obvious function, other than attracting pollinating insects to facilitate dispersal and sexual reproduction of the fungus. Another type of host plant manipulation by parasites, is the secretion of effectors by bacteria, which often engage in a long-term relationship with their host and suppress plant defence mechanisms (Dean, 2011). In return, plants have evolved effector-triggered immunity, resulting in an evolutionary arms race between plants and pathogenic parasites (Jones and Dangl, 2006). In plant-insect studies, host plant manipulation is often referred to as induced plant susceptibility (Karban and Baldwin, 1997). Rhoades (1985) proposed two strategies for host manipulation by herbivores: stealth and opportunism. Stealthy host manipulation involves the suppression or avoidance of plant defences. Larvae of the Alder sawfly,

**Table 1.** Characteristics of a parasitic lifestyle (Loker and Hofkin, 2015), and the occurrence of these aspects in aphids (V= occurring, X= not occuring, ~= depending).

| Parasitic aspects | Aphids | Comments | References |
|---|---|---|---|
| High host species specificity | ~ | Depending on the aphid species | (Blackman and Eastop, 2006) |
| High specificity for a host tissue | V | Phloem feeders | (Dixon, 1998) |
| Establishment of long-term relationship with a host | V | Often several generations on one host plant | (Dixon, 1998) |
| Utilising different hosts for complete life cycle | V | Host-alternating life cycle | (Moran, 1992) |
| Transition through several development stages in/on the host | V | Reproduction, moulting | (Moran, 1992) |
| High phenotypic plasticity and complex life cycles | V | (A)sexual, wing(less) morphs | (Moran, 1992) |
| Tolerance to ambient physiological conditions of the host | V | Adjusted to sugar-rich and nitrogen-poor diet | (Dixon, 1998) |
| Dependency on certain host resources (obligate parasites) | ~ | Loss of conversion to ascorbic acid in *M. persicae* | (Mittler *et al.*, 1970) |
| High reproduction rates | V | Telescoping generations | (Moran, 1992) |
| Both sexual and asexual reproduction | V | Sexual morphs in autumn, asexual in other seasons | (Moran, 1992) |
| Simplified genome | X | Genome size comparable to many other insects | (Hanrahan and Johnston, 2011) |
| Imposing non-lethal effects on the host | ~ | Depending on aphid density | (Dixon, 1998) |
| Detrimental for growth and development of the host | V | Reduction of meristem growth and wood biomass | (Dixon, 1971a; Girousse *et al.*, 2003) |
| Suppression or manipulation of host immune responses | V | Suppression of JA via SA, induction of *WRKY22* | (De Vos *et al.*, 2007), Chapter 6 |
| Reliance on symbionts to subvert host immune responses | V | Endosymbiont presence and activity correlates with host plant susceptibility | (Pinheiro *et al.*, 2014; Wulff and White, 2015) |
| Induction of physiological and behavioural changes in the host | V | Gall-forming aphids, changes in source-sink relationships | (Aoki *et al.*, 1977; Girousse *et al.*, 2005; Larson and Whitham, 1991) |

*Eriocampa ovata*, for example, cut all the main veins in a leaf before feeding (Mackay and Wellington, 1977), most likely to prevent systemic defence signalling and the allocation of defence compounds to the feeding site. Opportunistic host manipulation is the induction and utilisation of processes in the host plant for the benefit of the herbivore itself. This strategy could be considered as a more evolved form of manipulation because it changes host-plant physiology at a local or even systemic level (**Figure 1**). An example is the wood wasp, *Sirex noctilio*, which injects effectors in its host plant during oviposition, which suppress the carbohydrate efflux from the leaf and change local secondary metabolism to the advantage of the wasp's offspring (Madden, 1977). But what is known about host plant manipulation by aphids? The most outstanding example of opportunistic host exploitation by Aphididae is found in gall-forming aphids. They induce the formation of abnormal outgrowths on leaf veins that provide shelter for aphid colonies and function as a nutrient sink (Aoki *et al.*, 1977; Larson and Whitham, 1991). Also free-living aphids, which represent the vast majority of the Aphididae family, engage in host manipulation. In recent years, proteomic studies have revealed that aphid saliva, secreted in the cell wall matrix and in the phloem, contains proteins, some of which elicit plant defence responses (Bos *et al.*, 2010; De Vos and Jander, 2009), and some of which increase aphid performance via yet unknown mechanisms (Bos *et al.*, 2010; Hogenhout and Bos, 2011; Mutti *et al.*, 2008; Pitino and Hogenhout, 2012; Rodriguez and Bos, 2012). The salivary effector C002, for example, increased the duration of phloem ingestion, and is most likely involved in a phloem-located suppression of host-plant defence (Mutti *et al.*, 2008). Pectinases in aphid saliva have been shown to impact host plants both on a local and a systemic scale. Detailed microscopy showed that during intercellular probing, the stylet bundle pushes the cell walls aside and thereby slightly compresses the surrounding cells (Tjallingii and Hogen Esch, 1993). Pectin depolymerisation most likely increases the elasticity of the middle lamellae and cell wall and thereby facilitate stylet penetration while keeping the cells intact (Dreyer and Campbell, 1987). It can therefore be considered as local suppression of host plant resistance. Liu *et al.* (2009) showed that aphid pectinase also is an elicitor of systemic resistance. It induces the release of plant volatiles that attract aphid parasitoids and eventually decrease performance of *Sitobion avenae* aphids.

## The vascular network: the highway between local and systemic plant responses

It is not surprising that aphids induce effects at the whole plant level, considering the fact that they feed on, and secrete watery saliva in the vascular signalling network of the

host plant. Saliva components secreted in one sieve element, and small enough to pass through the sieve plate pores, could in theory be transported by the sap to (other) sink organs, and elicit a response in uninfested plant parts. As far as known, the transport capacity of aphid saliva through the phloem has never been studied. Nevertheless, the access to the vascular signalling network, provides aphids with the possibility to manipulate the physiology at a systemic scale, in virtually all organs both below and above ground. The other side of the coin is, that the host can take advantage of a phloem-located infection. The vascular infrastructure facilitates the plant in building a quick systemic defence response. Local defence signals, such as messenger RNAs, jasmonic acid (JA), methyljasmonate and electrical signals and free calcium, have been described to disperse via the phloem and to activate defence responses in distal plant parts (Gaupels and Vlot, 2013; Salvador-Recatalà *et al.*, 2014; van Bel *et al.*, 2014). In return, defence compounds can be produced systemically and allocated to the infested sieve element. Although the phloem-mobility of several secondary metabolites is still debatable, glucosinolates, terpenoids, and alkaloids have been described to travel via the phloem sap (Gowan *et al.*, 1995; Hagel *et al.*, 2013; Jørgensen *et al.*, 2015; Ober and Kaltenegger, 2009). The fast systemic signalling and allocation of compounds to the feeding site must have consequences for phloem-feeding insects. An important asset of phloem feeders is, however, that they ingest pure sap containing stable compounds of secondary metabolites. Glucosinolates, for example, only become toxic after hydrolysis by myrosinases, which are stored in separate cells in the plant (Fahey *et al.*, 2001). Since aphids do not disrupt cells, unlike chewing insects, they ingest glucosinolates in their intact, non-toxic state (Bones and Rossiter, 1996). Yet, some degree of hydrolysis still takes place during digestion and excretion, as has been illustrated with indole-glucosinolate breakdown products in the honeydew of *M. persicae* (Kim *et al.*, 2008b). It would be interesting to assess whether other phloem-mobile secondary metabolites, such as terpenoids and alkaloids, also have reduced impact on phloem feeders when they are ingested in their intact form.

## Systemic acquired resistance or susceptibility?

Suppose that aphids secrete effectors in the vascular network which could modulate virtually any physiological process in the host plant, what would then be the ultimate target? Particularly for generalist aphids, such as *M. persicae*, it might be most advantageous to hijack a fundamental regulator of plant defence responses occurring in a wide array of plant families. Indeed, there is a candidate target that meets this requirement: SA. While most herbivorous insects elicit mainly the defence-related hormones JA and

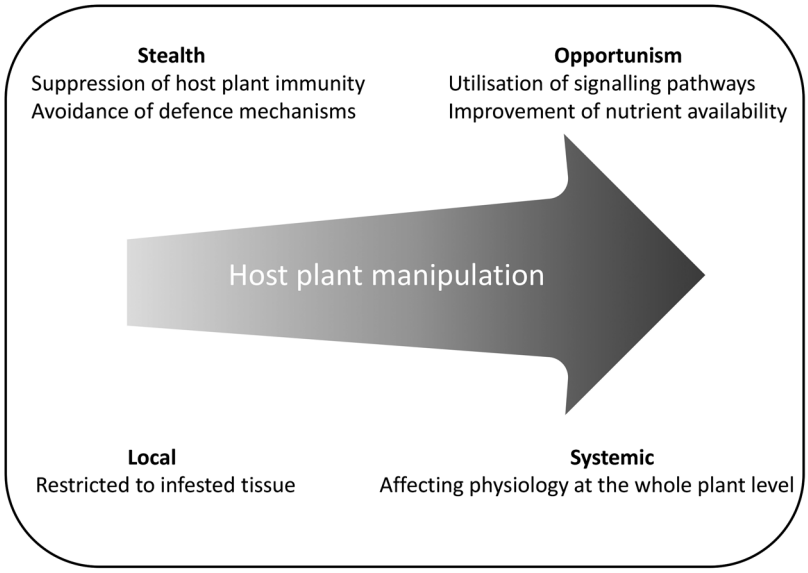



**Figure 1.** Degree of host-pant manipulation by herbivorous insects, determined by stealth or opportunism (Rhoades, 1985), and whether the insects manipulate host-plant physiology at a local or systemic scale.

ethylene (ET), aphids induce apart from JA and ET also the SA pathway (De Vos *et al.*, 2005; Moran *et al.*, 2002; Thompson and Goggin, 2006). SA influences a wide array of processes in the plant, including cell growth and respiration, but also activates systemic acquired resistance (SAR), involving a hypersensitive response and the accumulation of reactive oxygen species (Pieterse *et al.*, 2012; Vlot *et al.*, 2009). It is not surprising that aphids elicit SA, since SAR is triggered by pathogen-associated molecular patterns and effectors. Due to the secretion of saliva and the penetration by the stylet bundle, the plant is exposed to numerous non-self molecules which could potentially elicit SAR. Interestingly, however, it remains unclear whether activation of the SA pathway is beneficial or detrimental to aphids. Studies with Arabidopsis SA- and JA- biosynthesis and signalling mutants show that JA-mediated defences, involving the accumulation of glucosinolates and terpenoids, are generally more effective against aphids than SA-induced responses (Ellis *et al.*, 2002; Mewis *et al.*, 2005; Moran and Thompson, 2001; Pegadaraju *et al.*, 2005; Thompson and Goggin, 2006). Since SA and JA act antagonistically in many plant species (Bostock, 2005), it has been postulated that aphids induce SA to interfere with JA-mediated defences (De Vos *et al.*, 2007; Walling, 2008). This would be a major form of opportunistic host-plant manipulation. Unfortunately, this theory is not unambiguous. First of all, because SA and JA exert pleiotropic effects on many stress responses. The beneficial impacts of SA-JA antagonism on aphids could merely be an

evolutionary by-product of an overall effective defence strategy against other biotic stresses. Several studies indicate that negative SA-JA crosstalk improves plant fitness by prioritising costly defence responses when plants are attacked by multiple herbivores and pathogens (Pieterse *et al.*, 2012; Vos *et al.*, 2015). For many plant species, aphids might simply not have imposed a substantial selection pressure to evolve an alternative defence mechanism. A second reason why SA induction is not a straightforward example of host manipulation is, that SA also confers detrimental effects on aphids. The Mi-1-mediated resistance to the potato aphid, *Macrosiphum euphorbiae*, is an example of a successful SA-mediated defence (Kaloshian, 2004; Li *et al.*, 2006; Rossi *et al.*, 1998). Furthermore, the WRKY22 transcription factor has been shown to suppress effective SA-mediated defences against *M. persicae* in Arabidopsis (Chapter 6). These observations show that SA signalling plays a dualistic role in plant resistance to aphids. It possibly reflects the ongoing arms race between plants and aphids, resulting in a continuous development of new or improved mechanisms in both plants and aphids. Nevertheless, the induction of *WRKY22* by aphids (Barah *et al.*, 2013; De Vos *et al.*, 2005) with subsequent beneficial effects on aphid population development (Chapter 6), might in itself be an example of opportunistic host plant manipulation. Accumulating evidence indicates that WRKY22 is an integrator of SA-JA crosstalk in response to a plethora of abiotic and biotic stimuli (Chapter 6). WRKY22's function in suppressing two major plant defence hormones and the occurrence of WRKY orthologues throughout the plant kingdom, makes it an interesting target for a generalist aphid to manipulate. If it comes to influencing phytohormones, it would be naive to think that only one gene is the ultimate target. Phytohormonal modulation involves numerous regulatory pathways for the biosynthesis, metabolism, transport, perception and signal transduction, each regulated by a complex genetic network (Seyfferth and Tsuda, 2014; Tran and Pal, 2014). Instead of hijacking a single gene, it would be more realistic to expect host plant manipulation at the level of one or more signalling cascades. WRKY22 is possibly a small hub in a larger signalling network, which could be manipulated by aphids. To draw further conclusions about host manipulation, it would be necessary to study if and which aphid effectors induce *WRKY22* and its upstream regulators.

## Reprogramming of the host-plant transcriptome

In comparison to other herbivorous insects, aphids induce a tremendous amount of transcriptional perturbations in host plants. They cause, depending on the duration of infestation, differential expression of thousands of genes, whereas chewing insects and cell-content feeders alter expression of only several hundreds of genes (Appel *et*

*al.*, 2014; Barah *et al.*, 2013; De Vos *et al.*, 2005; Dubey *et al.*, 2013; Foyer *et al.*, 2015; Kerchev *et al.*, 2013; Kusnierczyk *et al.*, 2007) (Chapter 6). It is, however, arguable whether this strong aphid-induced transcriptional reprogramming of the host plant concerns a large-scale induction of plant defences or rather modulations that benefit the aphid. There is, unfortunately, still a lack of understanding of the meaning and consequences of the plethora of aphid-induced transcriptomic perturbations. Maybe partially, because most transcriptomic studies search for plant defence mechanisms and not for host-manipulation processes. As a consequence, the focus is on a small subset of characterised defence genes while the majority of the data are ignored. Nevertheless, some studies show that, depending on the aphid-plant combination, the transcriptional changes involve several defence mechanisms with quantitative detrimental effects on aphids, including cell-wall modifications (Divol *et al.*, 2007; Dreyer and Campbell, 1984), and the accumulation of glucosinolates (Levy *et al.*, 2005; Mewis *et al.*, 2012), camalexin (Kettles *et al.*, 2013; Mewis *et al.*, 2012), and proteinase inhibitors (Casaretto and Corcuera, 1998). Although there are still many unknown resistance mechanisms to be discovered, the overall impression remains that the scale of aphid-induced transcriptional changes in host plants generally outnumbers the amount and effectiveness of plant defence responses.

## Reprogramming of host-plant primary metabolism

Interestingly, the most impressive aphid-induced changes in host-plant physiology are not related to secondary, but primary metabolism. This was initially discovered due to the large biomass losses in plants during infestation. In woody plants, such as lime and sycamore trees, at least 50% reduction in wood formation was observed in aphid-infested trees compared to uninfested trees (Dixon, 1971a, b). In a small experiment with Arabidopsis, I also found a 3-fold reduction in rosette fresh weight when plants were infested with 5 neonate *M. persicae* for 5 days (**Figure 2**). Based on calculations of dry matter weight, nitrogen content and aphid metabolism, Dixon (1971a) concluded that the extent of biomass loss could not be explained by the removal of photoassimilates via aphid feeding alone. Subsequent studies revealed that aphid-infested plants displayed a different phloem translocation than uninfested plants (Hawkins *et al.*, 1987; Veen, 1985). Recent studies show that aphids induce sink strength and redirect nutrients to infested leaves at the expense of their meristems (Girousse *et al.*, 2003; Girousse *et al.*, 2005; Wilson *et al.*, 2011). Aphid-infested leaves indeed contained increased levels of sucrose and starch, indicative of increased carbohydrate allocation (Singh *et al.*, 2011).

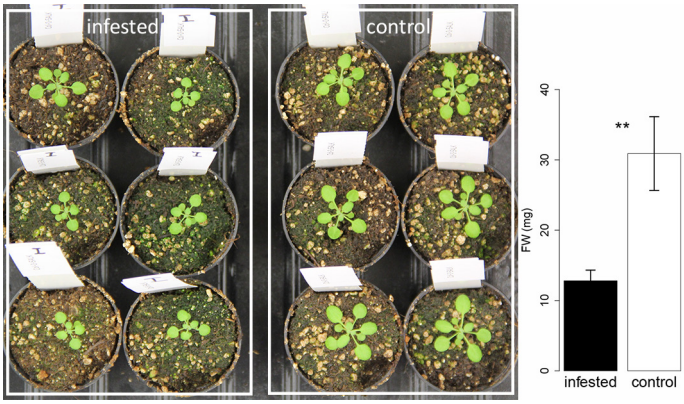



**Figure 2.** Rosette fresh weight of Col-0 Arabidopsis plants after 5 days of infestation with 5 *M. persicae* neonates (left) or without aphids (right). Data represent mean ± standard error (Mann-Whitney U test, **P<0.01, n=6).

Sink strength is also known to be induced by mites, pathogens and parasitic plants (Lemoine *et al.*, 2013; Storms, 1971). It has been postulated that the increased supply of phloem-mobile compounds to infested tissues, supports the host plant in mounting a costly defence response (Roitsch, 1999). In addition, some sugars act as a priming agent for plant defences (Herbers *et al.*, 1996; Morkunas and Ratajczak, 2014). For aphids, little is known about the role of sugars in inducing plant defence, although there are indications that the sugars trehalose and trehalose-6-phosphate enforce plant resistance to aphids (Singh *et al.*, 2011). The function of aphid-induced sink strength remains, however, controversial. By changing infested tissue into sinks, aphids are guaranteed of a large food supply for a prolonged time. Further studies could shed more light on the possible effect of aphid-induced sink strength on plant resistance, for example, by studying aphid performance on plants with an impaired source-sink regulation.

## Phloem-located obstructions: elusive success stories

Two main antibiosis strategies against aphid infestation can be distinguished: the production or release of detrimental compounds, such as secondary metabolites, or, the reduction of nutrient availability and accessibility. The latter strategy might be a very successful one, in view of the fact that aphids need to process large quantities of phloem to acquire enough nitrogen. Most likely, the Nr-mediated near-complete resistance in lettuce (ten Broeke, 2013) and the PHLOEM HEAT-RESPONSE LOCUS (PHLO)-mechanism in Arabidopsis (Chapter 5) are both strategies to reduce food ingestion by aphids. The Nr-based resistance is probably the most effective plant

resistance mechanism that has been reported so far. Ten Broeke *et al.* (2013a) showed that *N. ribisnigri* biotype Nr:0 aphids were virtually unable to ingest phloem on resistant lettuce cultivars. Within three days all aphids died on the resistant cultivar. The aphids could reach the phloem vessels, but were either unable to ingest phloem at all, or could ingest phloem only for several minutes. The possibility that toxic compounds are involved cannot be excluded, but in view of the fast and irreversible interruption of phloem feeding, a physical obstruction of the aphid food channel or the respective sieve element would be the most likely explanation. Although the effects are definitely not that strong, the PHLO-mediated resistance to *M. persicae* in Arabidopsis shows remarkable similarities to the Nr-mechanism. On plants with a functional *PHLO* gene, aphids could reach the phloem vessels, but interrupted feeding more than three times more often and showed a 2- to 3-fold reduction in the duration of phloem ingestion compared to aphids on knockout plants (Chapter 5). Further characterisation revealed that not only the total time spent on phloem feeding was lower on wild type plants, but also the rate of ingestion. Based on its co-expression with phloem proteins and its absence in the Poaceae family, which is devoid of phloem proteins, the phenotype strongly supports the idea of a physical, protein-related occlusion of the aphid's food channel. Interestingly, the closest homologue of *PHLO*, *RESTRICTED TOBACCO ETCH POTYVIRUS (TEV) MOVEMENT 2* (*RTM2*), restricts long-distance virus transport through the vascular bundle (Bondino *et al.*, 2012; Cayla *et al.*, 2015; Chisholm *et al.*, 2001; Whitham *et al.*, 2000). Assuming a similar functionality as RTM2, PHLO may not only reduce aphid infestation and potentially transmission of persistent viruses, but possibly also the systemic transport of viruses post transmission. This multifaceted character would make it a useful target for plant breeding programmes. *PHLO* and *RTM2* homologues are present in other dicots (Bondino *et al.*, 2012) and selective breeding for effective haplotypes could potentially improve resistance to aphids and viruses. The *Nr* gene in lettuce and *PHLO* in Arabidopsis reveal most likely the weakest spot of the aphid: occlusion of its food channel. But even for this aspect, the aphid has a weapon which can change a threat into an opportunity. The secretion of watery saliva in the phloem is postulated to counteract proteinaceous aggregations in phloem vessels (Furch *et al.*, 2010; Tjallingii, 2006; Will *et al.*, 2007). Interestingly, when sieve-element-occluding proteins are not successful in obstructing aphid feeding, they increase the nutritional value of the phloem and boost aphid performance. Anstead *et al.* (2012) showed that aphid populations developed faster on Arabidopsis plants with the phloem filament proteins AtSEOR1 and AtSEOR2 compared to *atseor1* and *atseor2* mutants. This last example nicely illustrates the intimate, dualistic interactions between a host and a parasite.

## What shapes natural variation in plant resistance?

Overall, aphids have a negative impact on plant fitness due to the removal of photoassimilates, modification of source-sink relationships, and transmission of plant viruses (Dixon, 1998; Girousse *et al.*, 2003; Minks and Harrewijn, 1989). Plants can evolve effective resistance strategies, as long as the costs of these strategies do not approach the losses due to the herbivory. To what extent aphids constitute a selection pressure on the development of plant resistance mechanisms, depends on the local abundance of aphids, their host-plant specificity, and the synchrony of plant and aphid phenology. In an evolutionary ecological context, it is tempting to consider these factors as the most important in shaping natural variation in plant resistance to aphids. However, from a plant-physiological point of view, this is only the tip of the iceberg. Although plants can mount a stress-specific response, the underlying genetic and physiological basis can have pleiotropic effects on other abiotic and biotic stress responses. Many studies have revealed that plant defences to aphids, pathogens and other biotic stresses share common or sometimes antagonistic phytohormonal pathways (Appel *et al.*, 2014; De Vos *et al.*, 2005; Moran *et al.*, 2002; Thompson and Goggin, 2006). The genetic implications of these shared pathways are nicely illustrated by the positive correlation between the genetic architecture of Arabidopsis responses to aphids and parasitic plants, both SA-inducing organisms (Chapter 4). Interestingly, both aphids and parasitic plants displayed negative genetic correlation with plant responses to caterpillars, thrips and a necrotrophic fungus, which mainly induce the antagonistic JA pathway. Adaptation to one biotic factor will, therefore, most likely also affect plant resistance to other organisms. The same accounts for abiotic factors. Arabidopsis response to moderate heat stress, for example, involves *PHLO*, which is required for fast development of the inflorescence during moderate heat stress and for resistance to aphids (Chapter 5). *PHLO* expression is temperature inducible and the geographic haplotype distribution is correlated with temperature fluctuations. Heat stress, and not aphid abundance, may possibly have been the major selection pressure in shaping natural variation in *PHLO*. Apart from pleiotropy, neutrality is another underestimated source of natural variation. Resistance traits that have lost their relevance for fitness by a change in e.g. climate or plant or insect phenology, will accumulate substantial genetic variation over time due to the fixation of random mutations with (nearly) neutral effects (Drake *et al.*, 1998; Kimura, 1983). Eventually this will lead to the decay of a trait, unless there is still some degree of positive selection. From that perspective, plant resistance to low abundant stresses is expected to display higher levels of natural variation, than resistance traits which are conserved through strong directional selection. Considering these different sources of natural variation, it would thus be a misunderstanding to think that a species

with limited natural exposure to aphids, such as Arabidopsis, does not display natural genetic variation in relevant resistance mechanisms to aphids.

## Finding new leads for plant resistance with genome-wide association (GWA) mapping

Natural genetic variation is a major resource for the identification of gene functionalities. Quantitative trait locus (QTL) mapping on biparental populations is so far the most commonly used approach to find genomic regions that are associated with plant resistance to biotic stresses (Young, 1996). The emergence of Next Generation Sequencing (NGS) techniques, gave access to more plant genomes with more dense genotype maps (Lister *et al.*, 2009). This technological development was essential for GWA mapping. The concept of GWA mapping is a huge step forward compared to QTL mapping: instead of a biparental gene pool and a maximum of several hundred recombinations, now several hundred plant lines with thousands of markers can be explored. For plant resistance to aphids, there are many more mechanisms to be discovered in a large plant population which has been subjected to millions of years of natural selection (Bergelson and Roux, 2010; Mitchell-Olds, 2010). In practice, however, GWA mapping often suffers from "missing" or "hidden" heritability. In those cases, the observed phenotypic variation cannot be explained by genetic polymorphisms. Most studies claim missing heritability either due to a low statistical power of genotype-phenotype associations, or, due to a low proportion of phenotypic variation that can be explained by genomic regions with a significant association (Ingvarsson and Street, 2011; Myles *et al.*, 2009). With regard to the missing heritability in resistance to (a)biotic stress in Arabidopsis, four main causes can be considered: (1) biased genotypic maps, (2) statistical overcorrection, (3) a polygenic character of the trait, and (4) environmental noise. The first two are related to the genotype and mapping method, while the two latter are more related to the phenotype. In the next two subparagraphs their implications are discussed.

### *The genotype and the model*

Genotypic maps are often biased due to their dependency on the reference accession and lack of allelic information. The genomes of the selected 350 Arabidopsis accessions of the Hapmap collection (Platt *et al.*, 2010), for example, have been mapped to the reference genome of Col-0 (Atwell *et al.*, 2010; Horton *et al.*, 2012). They will consequently miss regions that have been lost in the reference line (Gan *et al.*, 2011). In addition, multi-allelic information is missing due to the use of a binary SNP-tiling array. Instead of

base-pair mutations for each locus in the GWA model (A, T, G, C, deletion or insertion), only the presence or absence of a mutation is known from the atSNPtile1 array (Atwell *et al.*, 2010). Re-sequencing of Arabidopsis accessions has in the meantime resolved these issues by using multiple reference accessions and more in depth sequencing (Cao *et al.*, 2011; Gan *et al.*, 2011). It is only a matter of time before these sequences will find their way into GWA mapping studies. A second impediment of GWA mapping studies is the necessity of statistical corrections, with sometimes disputable thresholds. The most renowned one, is the correction for population structure. This procedure is required to exclude false positive genotype-phenotype associations caused by the relatedness between lines (Hirschhorn and Daly, 2005). Plant adaptation to local biotic or abiotic stress, however, does often show a geographical distribution pattern and may thus be closely related with population structure. Correcting for kinship, will in those cases, also remove relevant associations (Kloth *et al.*, 2012; Myles *et al.*, 2009; Yu and Buckler, 2006). Even though the amount of false positive candidate genes will increase, it would be justifiable to perform GWA mapping without correcting for kinship if traits show a strong geographical cline. Another impediment is the multiple testing correction. Because GWA mapping involves several thousand statistical tests, it is obvious that there is a high chance of false positive outcomes when using an alpha threshold of 0.05. The question is, however, what to consider a suitable correction. The Bonferroni threshold is commonly used as a multiple-testing correction, and divides alpha by the total number of independent hypotheses (Hochberg, 1988). In most GWA mapping studies, alpha is divided by the total number of SNPs (Atwell *et al.*, 2010; Chao *et al.*, 2012), which in the case of the Hapmap population with 214k SNPs, results in a conservative alpha of $2.5*10^{-7}$ (-10log(P) value= 6.6). None of the candidate genes that are discussed in Chapter 4, 5 and 6 of this thesis, contained associations with a -10log(P)value above this threshold. It is, however, disputable whether every single locus should be considered as an independent hypothesis. In many cases, polymorphisms have not evolved via point mutations, but via recombination of larger genomic regions (Webster and Hurst, 2012). From that point of view, a less conservative threshold could be determined, using the decay of linkage disequilibrium to estimate the number of independent polymorphisms.

### *The specificity of the phenotype*

Last but not least, GWA mapping suffers from the complexity of traits. In contrast to qualitative traits, where a phenotypic effect is either present or absent and in general caused by a single gene, quantitative traits show a continuous scale of phenotypic variation, in most cases determined by many genes and epistatic interactions (Keurentjes *et al.*, 2008). Plant resistance to insects comes both in qualitative and quantitative

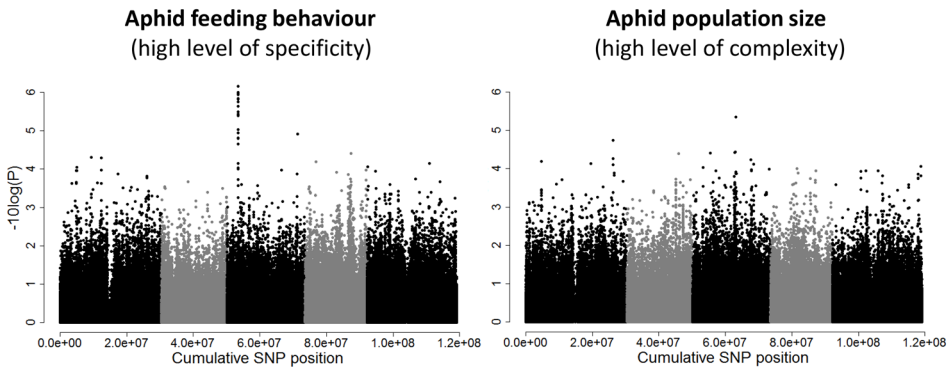



**Figure 3.** Manhattan plots of two traits: (1) aphid feeding behaviour, measured as the total time that aphids spent on making short probes (<3 min) in the first 1.5 h after infestation (n=3-6, highest -10log(P)=6.2) (Chapter 5), and (2) aphid population size on Arabidopsis plants after two weeks of infestation with one neonate *M. persicae* per plant (n=3, highest -10log(P)=5.3). Feeding behaviour is considered as a trait with high specificity, due to the short time frame and its restriction to a specific aspect of plant resistance, i.e. short probes which are correlated to penetrations of the epidermis and mesophyll. In contrast, aphid population size is the outcome of multiple aspects of plant resistance during a longer time frame (GWA mapping was performed on 350 Arabidopsis accessions with a mixed model, correcting for population structure, experimental blocks and replicates were included as covariates (minor allele frequency > 5%)).

mechanisms, but the latter is most common (Kliebenstein, 2014; Rhoades, 1985). GWA mapping is originally designed to identify genes based on their individual effect on the phenotype, and thus often fails in finding the underlying QTLs of highly complex polygenic traits with numerous small effect genes and epistatic interactions (Gibson, 2010; Ingvarsson and Street, 2011; Kooke, 2014; Myles *et al.*, 2009). Multi-locus GWA models have emerged, but it is computational challenging to assess the vast amount of potential genetic interactions within a mixed model (Segura *et al.*, 2012). An alternative solution is to disentangle complex traits into smaller and specific components (Kloth *et al.*, 2012). For example, GWA mapping of a highly specific trait such as aphid feeding behaviour in the mesophyll in the first 1.5 hour after infestation, resulted in associations with higher statistical power than when aphid population size after a two-week infestation was mapped (**Figure 3**). Even though the most significant SNP (-10log(P) value=6.2) of aphid probing behaviour explained only 7% of the phenotype, it was located in a gene, *PHLO*, that negatively affected phloem ingestion and aphid population development (Chapter 5). This is just one case study, but it illustrates, that (1) phenotypic specificity can improve the statistical power of genotype-phenotype associations, and that (2) low trait heritability is not necessarily problematic for traits that are associated with one major QTL. If, however, a large set of genes would have

been causal for only 7% of the phenotypic variation, the individual contribution of each individual SNP would be marginal and most likely of low statistical significance. For complex traits it remains, therefore, important to improve heritability. Rosette dry or fresh weight, for example, could be considered as a complex trait, since it involves several processes such as germination rate, root-shoot ratio, growth rate during a time window of several weeks, resistance to potential stress conditions, water content of leaves, and anatomical features such as petiole length and leaf blade surface area. The heritability of rosette fresh weight varied from low to high (24-85%), depending on the environmental conditions (Bac-Molenaar, 2015). In this case, a multi-environment GWA model was used as a solution to retrieve robust QTLs, shared across different environmental conditions. In most cases, when it is not feasible to screen large plant populations under multiple conditions, it is essential to reduce unintended fluctuations in the environment and measurement errors. When exploring whole genomes, highly controlled experimental conditions with many replicates are, therefore, favourable over more natural, but dynamic field conditions with a low number of replicates. To minimize missing heritability, GWA mapping thus requires highly specific and accurate phenotypic data for several hundred plant lines. These high demands are the reason that GWA mapping studies face a phenotyping bottleneck. Particularly, for plant-insect studies it is challenging in terms of time, space and labour to screen 350 plant lines in multiple replicates. It would, for example, involve manual counting of thousands of insects or observing insects one by one for host plant preference in choice tests. Therefore, more pragmatic and high-throughput solutions are required, such as the automated video-tracking platform to screen aphid feeding behaviour as described in Chapter 3 (Kloth *et al.*, 2015). Another asset of such a technique is the dense data acquisition in terms of time points and variables, which increases the number of parameters for different phenotypic aspects.

## Future perspectives

Although GWA mapping is still under development, it is a promising tool for unravelling the genetic basis of natural plant resistance mechanisms to herbivorous insects. In the next years, improved multi-locus, -trait and -environment GWA models will most probably arise that will facilitate in unravelling the genetic architecture of these complex traits. Nevertheless, on the short term I expect the highest impact from the phenomics field. The current bottleneck is not the GWA algorithm, but rather the input of the model. This thesis illustrates that the specificity of the phenotype is crucial; not the eventual aphid population size, but small aspects of aphid behaviour were required

to pinpoint novel resistance genes (Chapter 5 & 6). By screening traits with a close association to only one or a few plant physiological processes, the model space can be reduced from several thousand to a handful of genes. This is an achievement that is not easily paralleled by a multi-locus model. The roadmap to GWA-proof phenotypes relies on the pragmatic utilisation of existing biological knowledge in combination with novel high-throughput techniques. Currently, many high-throughput and high-specificity phenotyping methods are under development in the plant science community, including imaging techniques for the evaluation of root architecture and chlorophyll content (Goggin *et al.*, 2015; Hartmann *et al.*, 2011; Julkowska *et al.*, 2014; Montes *et al.*, 2007). Although computerised phenotyping of plant-insect interactions is still lagging behind, the technical advances in the field of digital image acquisition and analysis will open up many opportunities. Furthermore, it would be interesting to take a step above the level of individual plant stresses and perform GWA mapping on fundamental and cellular processes underlying plant resistance, such as the occurrence and magnitude of oxidative bursts and hypersensitive responses (Olukolu *et al.*, 2014). GWA mapping of these specific physiological processes has a high chance of finding causal genes and would have a broad relevance for many different plant stresses and plant species. Together with the growing availability of plant genomes, these approaches will be instrumental in the discovery of novel resistance genes. Eventually, this information can be implemented in marker-assisted plant breeding programmes to improve crop resistance and reduce the excessive application of environment-malign insecticides.

## Acknowledgements

I would like to thank Maarten Jongsma, Harro Bouwmeester and Marcel Dicke for constructive comments that helped to improve this chapter.

# References

**Abbott RJ, Gomes MF**. 1989. Population genetic structure and outcrossing rate of *Arabidopsis thaliana* (L.) Heynh. *Heredity* **62**, 411-418.

**Abe H, Ohnishi J, Narusaka M, Seo S, Narusaka Y, Tsuda S, Kobayashi M**. 2008. Function of jasmonate in response and tolerance of *Arabidopsis* to thrip feeding. *Plant and Cell Physiology* **49**, 68-80.

**Agrawal AA, Kurashige NS**. 2003. A role for isothiocyanates in plant resistance against the specialist herbivore *Pieris rapae*. *Journal of Chemical Ecology* **29**, 1403-1415.

**Albersheim P, Darvill A, Roberts K, Sederoff R, Staehelin A**. 2011. *Plant cell walls*. New York: Garland Science, Taylor & Francis Group, LLC.

**Ali JG, Agrawal AA**. 2012. Specialist versus generalist insect herbivores and plant defense. *Trends in Plant Science* **17**, 293-302.

**Alimi NA, Bink MCAM, Dieleman JA, Magán JJ, Wubs AM, Palloix A, Van Eeuwijk FA**. 2013. Multi-trait and multi-environment QTL analyses of yield and a set of physiological traits in pepper. *Theoretical and Applied Genetics* **126**, 2597-2625.

**Allemand R, Pompanon F, Fleury F, Fouillet P, Bouletreau M**. 1994. Behavioral circadian-rhythms measured in real-time by automatic image-analysis - applications in parasitoid insects. *Physiological Entomology* **19**, 1-8.

**Alonso-Blanco C, Aarts MGM, Bentsink L, Keurentjes JJB, Reymond M, Vreugdenhil D, Koornneef M**. 2009. What has natural variation taught us about plant development, physiology, and adaptation? *The Plant Cell* **21**, 1877-1896.

**Alonso-Blanco C, Koornneef M**. 2000. Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. *Trends in Plant Science* **5**, 22-29.

**Alvarez AE, Tjallingii WF, Garzo E, Vleeshouwers V, Dicke M, Vosman B**. 2006. Location of resistance factors in the leaves of potato and wild tuber-bearing *Solanum* species to the aphid *Myzus persicae*. *Entomologia Experimentalis et Applicata* **121**, 145-157.

**Anderson JP, Badruzsaufari E, Schenk PM, Manners JM, Desmond OJ, Ehlert C, Maclean DJ, Ebert PR, Kazan K**. 2004. Antagonistic interaction between abscisic acid and jasmonate-ethylene signaling pathways modulates defense gene expression and disease resistance in Arabidopsis. *The Plant Cell* **16**, 3460-3479.

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

**Anstead JA, Froelich DR, Knoblauch M, Thompson GA**. 2012. Arabidopsis P-protein filament formation requires both AtSEOR1 and AtSEOR2. *Plant & Cell Physiology* **53**, 1033-1042.

**Aoki S, Yamane S, Kiuchi M**. 1977. On the biters of *Astegopteryx styracicola* (Homoptera Aphidoidea). *The Entomological Society of Japan, Kontyu* **45**, 563-570.

**Appel HM, Fescemyer H, Ehlting J, Weston D, Rehrig E, Joshi T, Xu D, Bohlmann J, Schultz J**. 2014. Transcriptional responses of *Arabidopsis thaliana* to chewing and sucking insect herbivores. *Frontiers in Plant Science* **5**, 1-20.

**Arabidopsis Genome Initiative T**. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.

**Arany AM**. 2006. Ecology of *Arabidopsis thaliana*: Local adaptation and interaction with herbivores, Leiden University, Wageningen, the Netherlands.

**Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C, Toomajian C, Traw B, Zheng H, Bergelson J, Dean C, Marjoram P, Nordborg M**. 2005. Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. *PLoS Genetics* **1**, e60.

**Armengaud P, Zambaux K, Hills A, Sulpice R, Pattison RJ, Blatt MR, Amtmann A**. 2009. EZ-Rhizo: integrated software for the fast and accurate measurement of root system architecture. *Plant Journal* **57**, 945-956.

**Asai T, Tena G, Plotnikova J, Willmann MR, Chiu W, Gomez-Gomez L, Boller T, Ausubel FM, Sheen J**. 2002. MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature* **415**, 977-983.

**Asjes CJ**. 2000. Control of aphid-borne Lily symptomless virus and Lily mottle virus in Lilium in the Netherlands. *Virus Research* **71**, 23-32.

**Asjes CJ, Blom-Barnhoorn GJ**. 2001. Control of aphid-vectored and thrips-borne virus spread in lily, tulip, iris and dahlia by sprays of mineral oil, polydimethylsiloxane and pyrethroid insecticide in the field. *Annals of Applied Biology* **139**, 11-19.

**Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JDG, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M**. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627-631.

**Awmack CS, Leather SR**. 2002. Host plant quality and fecundity in herbivorous insects. *Annual Review of Entomology* **47**, 817-844.

**Bac-Molenaar JA, Fradin EF, Becker FFM, Rienstra JA, van der Schoot J, Vreugdenhil D, Keurentjes JJB**. 2015a. Genome-wide association mapping of fertility reduction upon heat stress reveals developmental stage-specific QTLs in *Arabidopsis thaliana*. *The Plant Cell* **27**, 1857-1874.

**Bac-Molenaar JA, Granier C, Keurentjes JJB, Vreugdenhil D**. 2015b. Genome-wide association mapping of time-dependent growth responses to moderate drought stress in *Arabidopsis*. *Plant Cell and Environment*, 10.1111/pce.12595.

**Backus EA, Bennett WH**. 2009. The AC-DC Correlation Monitor: New EPG design with flexible input resistors to detect both R and emf components for any piercing-sucking hemipteran. *Journal of Insect Physiology* **55**, 869-884.

**Bakchine E, Pham-Delegue MH, Kaiser L, Masson C**. 1990. Brief communication. Computer analysis of the exploratory behavior of insects and mites in an olfactometer. *Physiology & Behavior* **48**, 183-187.

**Balasubramanian S, Schwartz C, Singh A, Warthmann N, Kim MC, Maloof JN, Loudet O, Trainer GT, Dabi T, Borevitz JO, Chory J, Weigel D**. 2009. QTL mapping in new

*Arabidopsis thaliana* advanced intercross-recombinant inbred lines. *PLoS ONE* **4**.

**Baldacci-Cresp F, Maucourt M, Deborde C, Pierre O, Moing A, Brouquisse R, Favery B, Frendo P**. 2015. Maturation of nematode-induced galls in *Medicago truncatula* is related to water status and primary metabolism modifications. *Plant Science* **232**, 77-85.

**Ballaré CL**. 2009. Illuminated behavior. Phytochrome as a key regulator of light foraging and plant herbivore defense. *Plant Cell and Environment* **32**, 713-725.

**Ballesteros DC, Mason RE, Addison CK, Acuna MA, Arguello MN, Subramanian N, Miller RG, Sater H, Gbur EE, Miller D, Griffey CA, Barnett MD, Tucker D**. 2015. Tolerance of wheat to vegetative stage soil waterlogging is conditioned by both constitutive and adaptive QTL. *Euphytica* **201**, 329-343.

**Barah P, Winge P, Kusnierczyk A, Tran DH, Bones AM**. 2013. Molecular signatures in *Arabidopsis thaliana* in response to insect attack and bacterial infection. *PLoS ONE* **8**, e58987.

**Barratt DHP, Derbyshire P, Findlay K, Pike M, Wellner N, Lunn J, Fell R, Simpson C, Maule AJ, Smith AM**. 2009. Normal growth of Arabidopsis requires cytosolic invertase but not sucrose synthase. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13124–13129.

**Barratt DHP, Kolling K, Graf A, Pike M, Calder G, Findlay K, Zeeman SC, Smith AM**. 2011. Callose Synthase GSL7 Is Necessary for Normal Phloem Transport and Inflorescence Growth in Arabidopsis. *Plant Physiology* **155**, 328–341.

**Batelli G, Massarelli I, Van Oosten M, Nurcato R, Vannini C, Raimondi G, Leone A, Zhu JK, Maggio A, Grillo S**. 2012. Asg1 is a stress-inducible gene which increases stomatal resistance in salt stressed potato. *Journal of Plant Physiology* **169**, 1849-1857.

**Baxter I, Brazelton JN, Yu D, Huang YS, Lahner B, Yakubova E, Li Y, Bergelson J, Borevitz JO, Nordborg M, Vitek O, Salt DE**. 2010. A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS Genetics* **6**, e1001193.

**Becerra JX**. 2007. The impact of herbivore-plant coevolution on plant community structure. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 7483-7488.

**Beeuwkes J, Spitzen J, Spoor CW, van Leeuwen JL, Takken W**. 2008. 3-D flight behaviour of the malaria mosquito *Anopheles gambiae* s.s. inside an odour plume. *Proceedings of the Netherlands Entomological Society Meeting* **19**, 137-146.

**Beketov MA, Kefford BJ, Schäfer RB, Liess M**. 2013. Pesticides reduce regional biodiversity of stream invertebrates. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11039-11043.

**Bel AJEv, Ehlers K, Knoblauch M**. 2002. Sieve elements caught in the act. *Trends in Plant Science* **7**, 126-132.

**Benfey PN, Mitchell-Olds T**. 2008. Perspective - From genotype to phenotype: Systems biology meets natural variation. *Science* **320**, 495-497.

**Benikhlef L, L'Haridon F, Abou-Mansour E, Serrano M, Binda M, Costa A, Lehmann S, Métraux J**. 2013. Perception of soft mechanical stress in Arabidopsis leaves activates

disease resistance. *BMC Plant Biology* **13**.

**Bentsink L, Koornneef M**. 2008. Seed dormancy and germination. *The Arabidopsis Book*, Vol. 6: The American Society of Plant Biologists.

**Bergelson J, Roux F**. 2010. Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nature Reviews Genetics* **11**, 867-879.

**Bidart-Bouzat M, Kliebenstein D**. 2011. An ecological genomic approach challenging the paradigm of differential plant responses to specialist versus generalist insect herbivores. *Oecologia* **167**, 677-689.

**Blackman RL, Eastop VF**. 2006. *Aphids on the World's Herbaceous Plants and Schrubs*: John Wiley & Sons, Ltd. & Natural History Museum, London.

**Boer MP, Wright D, Feng L, Podlich DW, Luo L, Cooper M, Van Eeuwijk FA**. 2007. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by- environment interactions, with an example in maize. *Genetics* **177**, 1801-1813.

**Bolger AM, Lohse M, Usadel B**. 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*.

**Bond DM, Wilson IW, Dennis ES, Pogson BJ, Jean Finnegan E**. 2009. VERNALIZATION INSENSITIVE 3 (VIN3) is required for the response of *Arabidopsis thaliana* seedlings exposed to low oxygen conditions. *The Plant Journal* **59**, 576-587.

**Bondino HG, Valle EM, Have At**. 2012. Evolution and functional diversification of the small heat shock protein/α-crystallin family in higher plants. *Planta* **235**, 1299-1313.

**Bones AM, Rossiter JT**. 1996. The myrosinase-glucosinolate system, its organisation and biochemistry. *Physiologia Plantarum* **97**, 194-208.

**Boquel S, Giordanengo P, Ameline A**. 2011. Divergent effects of PVY-infected potato plant on aphids. *European Journal of Plant Pathology* **129**, 507-510.

**Bos JIB, Prince D, Pitino M, Maffei ME, Win J, Hogenhout SA**. 2010. A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (Green Peach Aphid). *PLoS Genet* **6**, e1001216.

**Bostock RM**. 2005. Signal crosstalk and induced resistance: Straddling the line between cost and benefit. *Annual Review of Phytopathology* **43**, 545-580.

**Bouwmeester HJ, Roux C, Lopez-Raez JA, Bécard G**. 2007. Rhizosphere communication of plants, parasitic plants and AM fungi. *Trends in Plant Science* **12**, 224-230.

**Braam J, Davis RW**. 1990. Rain-induced, wind-induced, and touch-induced expression of calmodulin and calmodulin-related genes in *Arabidopsis*. *Cell* **60**, 357-364.

**Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F**. 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genetics* **6**.

**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 of the United States of America* **112**, 4032-4037.

**Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES**. 2007. Tassel: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633-2635.

**Brady KU, Kruckeberg AR, Bradshaw HD**. 2005. Evolutionary ecology of plant adaptation to serpentine soils. *Annual Review of Ecology Evolution and Systematics* **36**, 243-266.

**Broekgaarden C, Poelman EH, Steenhuis G, Voorrips RE, Dicke M, Vosman B**. 2008. Responses of *Brassica oleracea* cultivars to infestation by the aphid *Brevicoryne brassicae*: an ecological and molecular approach. *Plant Cell and Environment* **31**, 1592-1605.

**Brotman Y, Riewe D, Lisec J, Meyer RC, Willmitzer L, Altmann T**. 2011. Identification of enzymatic and regulatory genes of plant metabolism through QTL analysis in *Arabidopsis*. *Journal of Plant Physiology* **168**, 1387-1394.

**Bruessow F, Gouhier-Darimont C, Buchala A, Metraux JP, Reymond P**. 2010. Insect eggs suppress plant defence against chewing herbivores. *The Plant Journal* **62**, 876-885.

**Burgueno J, de los Campos G, Weigel K, Crossa J**. 2012. Genomic prediction of breeding values when modeling genotype x environment interaction using pedigree and dense molecular markers. *Crop Science* **52**, 707-719.

**Buschmann C, Lichtenthaler HK**. 1998. Principles and characteristics of multi-colour fluorescence imaging of plants. *Journal of Plant Physiology* **152**, 297-314.

**Butler DG, Cullis BR, Gilmour AR, Gogel BJ**. 2009. *ASReml-R Reference Manual*. Brisbane, Australia: Department of Primary Industries and Fisheries.

**Cano LM, Raffaele S, Haugen RH, Saunders DGO, Leonelli L, MacLean D, Hogenhout SA, Kamoun S**. 2013. Major transcriptome reprogramming underlies floral mimicry induced by the rust fungus *Puccinia monoica* in *Boechera stricta*. *PLoS ONE* **8**, e75293.

**Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Müller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D**. 2011. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nature Genetics* **43**, 956-965.

**Cartieaux F, Contesto C, Gallou A, Desbrosses G, Kopka J, Taconnat L, Renoum J, Touraine B**. 2008. Simultaneous interaction of *Arabidopsis thaliana* with *Bradyrhizobium* Sp. strain ORS278 and *Pseudomonas syringae* pv. tomato DC3000 leads to complex transcriptome changes. *Molecular Plant-Microbe Interactions* **21**, 244-259.

**Casaretto JA, Corcuera LJ**. 1998. Proteinase inhibitor accumulation in aphid-infested barley leaves. *Phytochemistry* **49**, 2279-2286.

**Cayla T, Batailler B, Le Hir R, Revers F, Anstead JA, Thompson GA, Grandjean O, Dinant S**. 2015. Live imaging of companion cells and sieve elements in Arabidopsis leaves. *PLoS ONE* **10**, e0118122.

**Ceci LR, Volpicella M, Rahbé Y, Gallerani R, Beekwilder J, Jongsma MA**. 2003. Selection by phage display of a variant mustard trypsin inhibitor toxic against aphids. *Plant Journal* **33**, 557-566.

**Centraal Bureau voor de Statistiek H**. 2015. Gebruik gewasbeschermingsmiddelen in de landbouw; werkzame stof, toepassing. http://statline.cbs.nl/StatWeb/publicat

ion/?VW=T&DM=SLNL&PA=37606&D1=0,3&D2=a&D3=a&HD=100713-1430&HDR=T,G2&STB=G1. Heerlen, the Netherlands.

**Chaerle L, Van der Straeten D**. 2000. Imaging techniques and the early detection of plant stress. *Trends in Plant Science* **5**, 495-501.

**Chan EKF, Rowe HC, Corwin JA, Joseph B, Kliebenstein DJ**. 2011. Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in *Arabidopsis thaliana*. *PLoS Biology* **9**, e1001125.

**Chan EKF, Rowe HC, Kliebenstein DJ**. 2010. Understanding the evolution of defense metabolites in *Arabidopsis thaliana* using genome-wide association mapping. *Genetics* **185**, 991-1007.

**Chao DY, Silva A, Baxter I, Huang YS, Nordborg M, Danku J, Lahner B, Yakubova E, Salt DE**. 2012. Genome-wide association studies identify heavy metal ATPase3 as the primary determinant of natural variation in leaf cadmium in *Arabidopsis thaliana*. *PLoS Genetics* **8**, e1002923.

**Chawade A, Brautigam M, Lindlof A, Olsson O, Olsson B**. 2007. Putative cold acclimation pathways in *Arabidopsis thaliana* identified by a combined analysis of mRNA co-expression patterns, promoter motifs and transcription factors. *BMC Genomics* **8**, 304.

**Chehab EW, Eich E, Braam J**. 2009. Thigmomorphogenesis: a complex plant response to mechano-stimulation. *Journal of Experimental Botany* **60**, 43-56.

**Chehab EW, Yao C, Henderson Z, Kim S, Braam J**. 2012. Arabidopsis touch-induced morphogenesis is jasmonate mediated and protects against pests. *Current Biology* **22**, 701-706.

**Chen H, Lai Z, Shi J, Chen Z, Xu X**. 2010a. Roles of Arabidopsis WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC Plant Biology* **10**, 281.

**Chen H, Zhang Z, Teng K, Lai J, Zhang Y, Huang Y, Li Y, Liang L, Wang Y, Chu C, Guo H, Xie Q**. 2010b. Up-regulation of LSB1/GDU3 affects geminivirus infection by activating the salicylic acid pathway. *The Plant Journal* **62**, 12-23.

**Chen X, Vosman B, Visser RGF, van der Vlugt RAA, Broekgaarden C**. 2012. High throughput phenotyping for aphid resistance in large plant collections. *Plant Methods* **8**, 33.

**Chin K, DeFalco TA, Moeder W, Yoshioka K**. 2013. The Arabidopsis cyclic nucleotide-gated ion channels AtCNGC2 and AtCNGC4 work in the same signaling pathway to regulate pathogen defense and floral transition. *Plant Physiology* **163**, 611-624.

**Chisholm ST, Parra MA, Anderberg RJ, Carrington JC**. 2001. Arabidopsis *RTM1* and *RTM2* genes function in phloem to restrict long-distance movement of Tobacco Etch Virus. *Plant Physiology* **127**, 1667-1675.

**Christ B, Schelbert S, Aubry S, Sussenbacher I, Muller T, Krautler B, Hortensteiner S**. 2012. MES16, a member of the methyl esterase protein family, specifically demethylates fluorescent chlorophyll catabolites during chlorophyll breakdown in Arabidopsis *Plant Physiology* **158**, 628-641.

**Clark SE, Jacobsen SE, Levin JZ, Meyerowitz EM**. 1996. The CLAVATA and SHOOT

MERISTEMLESS loci competitively regulate meristem activity in Arabidopsis. *Development* **122**, 1567-1575.

**Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vailaya A, Wang P-L, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, Ideker T, Bader GD**. 2007. Integration of biological networks and gene expression data using Cytoscape. **2**, 2366-2382.

**Coego A, Brizuela E, Castillejo P, Ruíz S, Koncz C, Pozo JCd, Piñeiro M, Jarillo JA, Paz-Ares J, León J**. 2014. The TRANSPLANTA collection of Arabidopsis lines: a resource for functional analysis of transcription factors based on their conditional overexpression. *The Plant Journal* **77**, 944-953.

**Cosson P, Schurdi-Levraud V, Le QH, Sicard O, Caballero M, Roux F, Le Gall O, Candresse T, Revers F**. 2012. The RTM resistance to potyviruses in *Arabidopsis thaliana*: Natural variation of the *RTM* genes and evidence for the implication of additional genes. *PLoS ONE* **7**, e39169.

**Cosson P, Sofer L, Le QH, Léger V, Schurdi-Levraud V, Whitham SA, Yamamoto ML, Gopalan S, Le Gall O, Candresse T, Carrington JC, Revers F**. 2010. *RTM3*, which controls long-distance movement of potyviruses, is a member of a new plant gene family encoding a meprin and TRAF homology domain-containing protein. *Plant Physiology* **154**, 222-232.

**Coste AT, Ramsdale M, Ischer F, Sanglard D**. 2008. Divergent functions of three Candida albicans zinc-cluster transcription factors (CTA4, ASG1 and CTF1) complementing pleiotropic drug resistance in Saccharomyces cerevisiae. *Microbiology* **154**, 1491-1501.

**Cozzolino D**. 2009. Near infrared spectroscopy in natural products analysis. *Planta Medica* **75**, 746-756.

**Crofton HD**. 1970. A quantitative approach to parasitism. *Parasitology* **62**, 179-193.

**de Kogel W, van der Hoek M, Dik MA, van Dijken F, Mollema C**. 1998. Variation in performance of western flower thrips populations on a susceptible and a partially resistant chrysanthemum cultivar. *Euphytica* **103**, 181-186.

**De los Campos G, Hickey JM, Pong-Wong R, Daetwyler HD, Calus MPL**. 2013. Whole-genome regression and prediction methods applied to plant and animal breeding. *Genetics* **193**, 327-345.

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

**De Vos M, Kim JH, Jander G**. 2007. Biochemistry and molecular biology of Arabidopsis-aphid interactions. *Bioessays* **29**, 871-883.

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

**De Vos M, Van Zaanen W, Koornneef A, Korzelius JP, Dicke M, Van Loon LC, Pieterse CMJ**. 2006. Herbivore-induced resistance against microbial pathogens in Arabidopsis. *Plant Physiology* **142**, 352-363.

**Dean P**. 2011. Functional domains and motifs of bacterial type III effector proteins and their roles in infection. *FEMS Microbiology Reviews* **35**, 1100-1125.

**Dedryver C, Le Ralec A, Fabre F**. 2010. The conflicting relationships between aphids and men: A review of aphid damage and control strategies. *Comptes rendus Biologies* **333**, 539-553.

**Dell AI, Bender JA, Branson K, Couzin ID, de Polavieja GG, Noldus LPJJ, Pérez-Escudero A, Perona P, Straw AD, Wikelski M, Brose U**. 2014. Automated image-based tracking and its application in ecology. *Trends in Ecology & Evolution* **29**, 417-428.

**DellaPenna D, Last RL**. 2008. Genome-enabled approaches shed new light on plant metabolism. *Science* **320**, 479-481.

**Denoux C, Galletti R, Mammarella N, Gopalan S, Werck D, De Lorenzo G, Ferrari S, Ausubel FM, Dewdney J**. 2008. Activation of defense response pathways by OGs and Flg22 elicitors in *Arabidopsis* seedlings. *Molecular Plant* **1**, 423-445.

**Devlin B, Roeder K**. 1999. Genomic control for association studies. *Biometrics* **55**, 997-1004.

**Devlin B, Roeder K, Wasserman L**. 2001. Genomic control, a new approach to genetic-based association studies. *Theoretical Population Biology* **60**, 155-166.

**Dicke M**. 1999. Evolution of induced indirect defence of plants. In: Tollrian R, Harvell CD, eds. *The Ecology and Evolution of Inducible Defenses*. Princeton, NJ: Princeton University Press, 62-88.

**Dicke M, Baldwin IT**. 2010. The evolutionary context for herbivore-induced plant volatiles: beyond the 'cry for help'. *Trends in Plant Science* **15**, 167-175.

**Dinant S, Clark AM, Zhu Y, Vilaine F, Palauqui J, Kusiak C, Thompson GA**. 2003. Diversity of the superfamily of phloem lectins (Phloem Protein 2) in Angiosperms. *Plant Physiology* **131**, 114-128.

**Divol F, Vilaine F, Thibivilliers S, Kusiak C, Sauge MH, Dinant S**. 2007. Involvement of the xyloglucan endotransglycosylase/hydrolases encoded by celery *XTH1* and Arabidopsis *XTH33* in the phloem response to aphids. *Plant, Cell and Environment* **30**, 187-201.

**Dixon AFG**. 1971a. The role of aphids in wood formation. I. The effect of the Sycamore aphids *Dreopanosiphum platanoides* (Schr.) (Aphididae), on the growth of the sycamore, *Acer pseudoplatanus* (L.). *Journal of Applied Ecology* **8**, 165-179.

**Dixon AFG**. 1971b. The role of aphids in wood formation. II. The effect of the Lime aphid, *Eucallipterus tiliae* L. (Aphididae), on the growth of Lime, *Tilia* x *vulgaris* Hayne. *Journal of Applied Ecology* **8**, 393-399.

**Dixon AFG**. 1998. *Aphid ecology*. London: Chapman and Hall.

**Doares SH, Syrovets T, Weiler EW, Ryan CA**. 1995. Oligogalacturonides and chitosan activate plant defensive genes through the octadecanoid pathway. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 4095-4098.

**Dobón A, Canet JV, Perales L, Tornero P**. 2011. Quantitative genetic analysis of salicylic acid perception in Arabidopsis. *Planta* **234**, 671-684.

**Dogimont C, Chovelon V, Pauquet J, Boualem A, Bendahmane A**. 2014. The *Vat* locus encodes for a CC-NBS-LRR protein that confers resistance to *Aphis gossypii* infestation and *A. gossypii*-mediated virus resistance. *The Plant Journal* **80**, 993-1004.

**Dong J, Chen C, Chen Z**. 2003. Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. *Plant Molecular Biology* **51**, 21-37.

**Dorr I, Kollmann R**. 1995. Symplasmic sieve element continuity between *Orobanche* and its host. *Botanica Acta* **108**, 47-55.

**Drake JW, Charlesworth B, Charlesworth D, Crow JF**. 1998. Rates of spontaneous mutation. *Genetics* **148**, 1667-1686.

**Dreyer DL, Campbell BC**. 1984. Association of the degree of methylation of intercellular pectin with plant resistance to aphids and with induction of aphid biotypes. *Experientia* **40**, 224-226.

**Dreyer DL, Campbell BC**. 1987. Chemical basis of host plant resistance to aphids. *Plant, Cell and Environment* **10**, 553-561.

**Du Y, Poppy GM, Powell W, Pickett JA, Wadhams LJ, Woodcock CM**. 1998. Identification of semiochemicals released during aphid feeding that attract parasitoid *Aphidius ervi*. *Journal of Chemical Ecology* **24**, 1355-1368.

**Dubey NK, Goel R, Ranjan A, Idris A, Singh SK, Bag SK, Chandrashekar K, Pandey KD, Singh PK, Sawant SV**. 2013. Comparative transcriptome analysis of *Gossypium hirsutum* L. in response to sap sucking insects: aphid and whitefly. *BMC Genomics* **14**, 241.

**Eberius M, Lima-Guerra J**. 2009. High-throughput plant phenotyping - Data acquisition, transformation, and analysis. In: Edwards D, Stajich J, Hansen D, eds. *Bioinformatics: Tools and applications*: Springer Science+Business Media, 259 - 278.

**Eenink AH, Dieleman FL**. 1977. Screening *Lactuca* for resistance to *Myzus persicae*. *Netherlands Journal of Plant Pathology* **83**, 139-151.

**Eenink AH, Groenwold R, Dieleman FL**. 1982. Resistance of lettuce *(Lactuca*) to the leaf aphid *Nasonovia ribisnigri*. 1. Transfer of resistance from *L. virosa* to *L.sativa* by interspecific crosses and selection of resistant breeding lines. *Euphytica* **31**, 291-300.

**Ehlers K, Knoblauch M, van Bel AJE**. 2000. Ultrastructural features of well-preserved and injured sieve elements: minute clamps keep the phloem transport conduits free for mass flow. *Protoplasma* **214**, 80-92.

**Ehrenreich IM, Hanzawa Y, Chou L, Roe JL, Kover PX, Purugganan MD**. 2009. Candidate gene association mapping of Arabidopsis flowering time. *Genetics* **183**, 325-335.

**Eigenbrode SD, Espelie KE**. 1995. Effects of plant epicuticular lipids on insect herbivores. *Annual Review of Entomology* **40**, 171-194.

**Eleftheriou EP**. 1990. Monocotyledons. In: Behnke HD, Sjolund RD, eds. *Sieve elements: comparative structure, induction and development*. Berlin: Springer-Verlag, 139-159.

**Ellis C, Karafyllidis L, Turner JG**. 2002. Constitutive activation of jasmonate signaling in an Arabidopsis mutant correlates with enhanced resistance to *Erysiphe cichoracearum*, *Pseudomonas syringae*, and *Myzus persicae*. *Molecular Plant-Microbe Interactions* **15**, 1025-1030.

**Ernst AM, Jekat SB, Zielonka S, Müller B, Neumann U, Rüping B, Twyman RM,**

**Krzyzanek V, Prüfer D, Noll GA**. 2012. Sieve element occlusion (SEO) genes encode structural phloem proteins involved in wound sealing of the phloem. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E1980–E1989.

**Eulgem T, Rushton PJ, Robatzek S, Somssich IE**. 2000. The WRKY superfamily of plant transcription factors. *Trends in Plant Science* **5**, 199-206.

**Evert RF**. 1990. Dicotyledons. In: H.D. B, R.D. S, eds. *Sieve elements: Comparative structure, induction and development*. Berlin: Springer, 103-137.

**Fahey JW, Zalcmann AT, Talalay P**. 2001. The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* **56**, 5-51.

**FAO**. 2015. The status and development of conifer aphid damage in Malawi. http://www.fao.org/3/a-u4778e/u4778e05.htm.

**Fernie AR, Schauer N**. 2009. Metabolomics-assisted breeding: a viable option for crop improvement? *Trends in Genetics* **25**, 39-48.

**Finnegan EJ, Bond DM, Buzas DM, Goodrich J, Helliwell CA, Tamada Y, Yun JY, Amasino RM, Dennis ES**. 2011. Polycomb proteins regulate the quantitative induction of VERNALIZATION INSENSITIVE 3 in response to low temperatures. *The Plant Journal* **65**, 382-391.

**Fischer RA**. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society Edinburgh* **52**, 399-433.

**Fisher DB, Frame JM**. 1984. A guide to the use of the exuding-stylet technique in phloem physiology. *Planta* **161**, 385-393.

**Fletcher JC, Brand U, Running MP, Simon R, Meyerowitz EM**. 1999. Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science* **283**, 1911-1914.

**Folta K, Pontin MA, Karlin-Neumann G, Bottini R, Spalding EP**. 2003. Genomic and physiological studies of early cryptochrome 1 action demonstrate roles for auxin and gibberellin in the control of hypocotyl growth by blue light. *The Plant Journal* **36**, 203-214.

**Foster SP, Denholm I, Thompson R, Poppy GM, Powell W**. 2005. Reduced response of insecticide-resistant aphids and attraction of parasitoids to aphid alarm pheromone; a potential fitness trade-off. *Bulletin of Entomological Research* **95**, 37-46.

**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.

**Franceschi VR, Tarlyn NM**. 2002. l-ascorbic acid is accumulated in source leaf phloem and transported to sink tissues in plants. *Plant Physiology* **130**, 649-656.

**Friedel S, Usadel B, von Wirén N, Sreenivasulu N**. 2012. Reverse engineering: a key component of systems biology to unravel global abiotic stress cross-talk. *Frontiers in Plant Science* **3**, 294.

**Froelich DR, Mullendore DL, Jensen KH, Ross-Elliott TJ, Anstead JA, Thompson GA, Pélissier HC, Knoblauch M**. 2011. Phloem ultrastructure and pressure flow: sieve-element-occlusion-related agglomerations do not affect translocation. *The Plant Cell* **23**,

4428-4445.

**Furch ACU, Hafke JB, Schulz A, van Bel AJE**. 2007. Ca2+-mediated remote control of reversible sieve tube occlusion in *Vicia faba*. *Journal of Experimental Botany* **58**, 2827-2838.

**Furch ACU, Zimmermann MR, Will T, Hafke JB, van Bel AJE**. 2010. Remote-controlled stop of phloem mass flow by biphasic occlusion in *Cucurbita maxima*. *Journal of Experimental Botany* **61**, 3697-3708.

**Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Ratsch G, Mott R**. 2011. Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. *Nature* **477**, 419-423.

**Gao Q-M, Venugopal S, Navarre D, Kachroo A**. 2011. Low oleic acid-derived repression of jasmonic acid-inducible defense responses requires the WRKY50 and WRKY51 proteins. *Plant Physiology* **155**, 464-476.

**Gaupels F, Vlot AC**. 2013. Plant defense and long-distance signaling in the phloem. In: Thompson GA, Bel AJEv, eds. *Phloem: Molecular Cell Biology, Systemic Communication, Biotic Interactions*. Oxford: John Wiley & Sons, Inc., 227-247.

**Gibson G**. 2010. Hints of hidden heritability in GWAS. **42**, 558-560.

**Girousse C, Faucher M, Kleinpeter C, Bonnemain J**. 2003. Dissection of the effects of the aphid *Acyrthosiphon pisum* feeding on assimilate partitioning in *Medicago sativa*. *New Phytologist* **157**, 83-92.

**Girousse C, Moulia B, Silk W, Bonnemain J**. 2005. Aphid infestation causes different changes in carbon and nitrogen allocation in Alfalfa stems as well as different inhibitions of longitudinal and radial expansion. *Plant Physiology* **137**, 1474-1484.

**Goggin FL, Lorence A, Topp CN**. 2015. Applying high-throughput phenotyping to plant–insect interactions: picturing more resistant crops. *Current Opinion in Insect Science* **9**, 69-76.

**Göhre V, Jones AME, Sklenář J, Robatzek S, Weber APM**. 2012. Molecular crosstalk between PAMP-triggered immunity and photosynthesis *Molecular Plant-Microbe Interactions* **25**, 1083-1092.

**Gols R, Bukovinszky T, van Dam NM, Dicke M, Bullock JM, Harvey JA**. 2008. Performance of generalist and specialist herbivores and their endoparasitoids differs on cultivated and wild *Brassica* populations. *Journal of Chemical Ecology* **34**, 132-143.

**Gómez G, Pallás V**. 2004. A long-distance translocatable phloem protein from cucumber forms a ribonucleoprotein complex in vivo with *Hop Stunt Viroid* RNA. *Journal of Virology* **78**, 10104-10110.

**González-Lamothe R, El Oirdi M, Brisson N, Bouarab K**. 2012. The conjugated auxin indole-3-acetic acid–aspartic acid promotes plant disease development. *The Plant Cell* **24**, 762-777.

**Govind G, Mittapalli O, Griebel T, Allmann S, Böcker S, Baldwin IT**. 2010. Unbiased transcriptional comparisons of generalist and specialist herbivores feeding on progressively defenseless *Nicotiana attenuata* plants. *PLoS ONE* **5**, e8735.

**Gowan E, Lewis BA, Turgeon R**. 1995. Phloem transport of antirrhinoside, an iridoid glycoside, in *Asarina scandens* (Scrophulariaceae). *Journal of Chemical Ecology* **21**, 1781-1788.

**Granier C, Aguirrezabal L, Chenu K, Cookson SJ, Dauzat M, Hamard P, Thioux JJ, Rolland G, Bouchier-Combaud S, Lebaudy A, Muller B, Simonneau T, Tardieu F**. 2006. PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytologist* **169**, 623-635.

**Grieco F, Loijens L, Krips OE, Smit G, Spink AJ, Zimmerman P**. 2010. *EthoVision XT Reference manual*. Wageningen: Noldus Information Technology.

**Guo A, He K, Liu D, Bai S, Gu X, Wei L, Luo J**. 2005. DATF: a database of Arabidopsis transcription factors. *Bioinformatics* **21**, 2568-2569.

**Hagel JM, Onoyovwi A, Yeung EC, Facchini PJ**. 2013. Role of phloem metabolites in plant defense. In: Thompson GA, Bel AJEv, eds. *Phloem: Molecular Cell Biology, Systemic Communication, Biotic Interactions*. Oxford: John Wiley & Sons, Inc., 251-270.

**Hagler JR, Jackson CG**. 2001. Methods for marking insects: Current techniques and future prospects. *Annual Review of Entomology* **46**, 511-543.

**Hall D, Tegstrom C, Ingvarsson PK**. 2010. Using association mapping to dissect the genetic basis of complex traits in plants. *Briefings in Functional Genomics* **9**, 157-165.

**Hall SM, Milburn JA**. 1973. Phloem transport in *Ricinus*: Its dependence on the water balance of the tissues. *Planta* **109**, 1-10.

**Hammer G, Cooper M, Tardieu F, Welch S, Walsh B, van Eeuwijk F, Chapman S, Podlich D**. 2006. Models for navigating biological complexity in breeding improved crop plants. *Trends in Plant Science* **11**, 587-593.

**Hampton CR, Bowen HC, Broadley MR, Hammond JP, Mead A, Payne KA, Pritchard J, White PJ**. 2004. Cesium toxicity in Arabidopsis. *Plant Physiology* **136**, 3824-3837.

**Hanrahan S, Johnston JS**. 2011. New genome size estimates of 134 species of arthropods. *Chromosome Research* **19**, 809-823.

**Hardie J, Holyoak M, Taylor NJ, Griffiths DC**. 1992. The combination of electronic monitoring and video-assisted observations of plant penetration by aphids and behavioural effects of polygodial. *Entomologica Experimentalis et Applicata* **62**, 233-239.

**Hardie J, Powell G**. 2000. Close-up video combined with electronic monitoring of plant penetration and behavioral effects of an aphid (Homoptera: Aphididae) antifeedant. In: Walker GP, Backus EA, eds. *Principles and applications of electronic monitoring and other techniques in the study of Homopteran feeding behavior*. US: Entomological Society of America, 201-211.

**Hartmann A, Czauderna T, Hoffmann R, Stein N, Schreiber F**. 2011. HTPheno: An image analysis pipeline for high-throughput plant phenotyping. *BMC Bioinformatics* **12**, 148.

**Hawkins CDB, Whitecross MI, Aston MJ**. 1987. The effect of short-term aphid feeding on the partitioning of 14CO2 photoassimilate in three legume species. *Canadian Journal of Botany* **65**, 666-672.

**He H, de Souza Vidigal D, Snoek LB, Schnabel S, Nijveen H, Hilhorst H, Bentsink L**. 2014. Interaction between parental environment and genotype affects plant and seed

performance in Arabidopsis. *Journal of Experimental Botany* **65**, 6603-6615.

**Hehl R, Bülow L**. 2014. Athamap web tools for the analysis of transcriptional and posttranscriptional regulation of gene expression in *Arabidopsis thaliana*. In: Staiger D, ed. *Plant Circadian networks*, Vol. 1158. New York: Springer Science+Business Media, 139-156.

**Heil M**. 2008. Indirect defence via tritrophic interactions. *New Phytologist* **178**, 41-61.

**Herbers K, Meuwly P, Métraux J-P, Sonnewald U**. 1996. Salicylic acid-independent induction of pathogenesis-related protein transcripts by sugars is dependent on leaf developmental stage. *FEBS Letters* **397**, 239-244.

**Hilfiker O, Groux R, Bruessow F, Kiefer K, Zeier J, Reymond P**. 2014. Insect eggs induce a systemic acquired resistance in Arabidopsis. *The Plant Journal* **80**, 1085-1094.

**Hilker M, Meiners T**. 2006. Early herbivore alert: Insect eggs induce plant defense. *Journal of Chemical Ecology* **32**, 1379-1397.

**Hill DS**. 1987. *Agricultural insect pests of temperate regions and their control*. Cambridge: Cambridge University Press.

**Hipper C, Brault V, Ziegler-Graff V, Revers F**. 2013. Viral and cellular factors involved in phloem transport of plant viruses. *Frontiers in Plant Science* **4**, 1-24.

**Hirschhorn JN, Daly MJ**. 2005. Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics* **6**, 95-108.

**Hochberg Y**. 1988. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75**, 800-802.

**Hogenhout SA, Bos JIB**. 2011. Effector proteins that modulate plant–insect interactions. *Current Opinion in Plant Biology* **14**, 422-428.

**Holopainen JK, Gershenzon J**. 2010. Multiple stress factors and the emission of plant VOCs. *Trends in Plant Science* **15**, 176-184.

**Hölttä T, Mencuccini M, Nikinmaa E**. 2009. Linking phloem function to structure: Analysis with a coupled xylem–phloem transport model. *Journal of Theoretical Biology* **259**, 325-337.

**Horton MW, Hancock AM, Huang YS, Toomajian C, Atwell S, Auton A, Muliyati NW, Platt A, Sperone FG, Vilhjalmsson BJ, Nordborg M, Borevitz JO, Bergelson J**. 2012. Genome-wide patterns of genetic variation in worldwide *Arabidopsis thaliana* accessions from the RegMap panel. *Nature Genetics* **44**, 212-216.

**Howe GA, Jander G**. 2008. Plant immunity to insect herbivores. *annual Review of Plant Biology* **59**, 41-66.

**Hsu F-C, Chou M-Y, Chou S-J, Li Y-R, Peng H-P, Shih M-C**. 2013. Submergence confers immunity mediated by the WRKY22 transcription factor in Arabidopsis. *The Plant Cell* **25**, 2699-2713.

**Huang MS, Abel C, Sohrabi R, Petri J, Haupt I, Cosimano J, Gershenzon J, Tholl D**. 2010. Variation of herbivore-induced volatile terpenes among Arabidopsis ecotypes depends on allelic differences and subcellular targeting of two terpene synthases, TPS02 and TPS03. *Plant Physiology* **153**, 1293-1310.

**Husson SJ, Steuer Costa W, Schmitt C, Gottschalk A**. September 10, 2012. Keeping track of worm trackers. In: Community eTCeR, ed. *Wormbook*. http://www.wormbook.org, 1-17.

**Ichimura K, Mizoguchi T, Yoshida R, Yuasa T, Shinozaki K**. 2000. Various abiotic stresses rapidly activate *Arabidopsis* MAP kinases ATMPK4 and ATMPK6. *The Plant Journal* **24**, 655-665.

**Iliev EA, Xu W, Polisensky DH, Oh MH, Torisky RS, Clouse SD, Braam J**. 2002. Transcriptional and posttranscriptional regulation of Arabidopsis *TCH4* expression by diverse stimuli. Roles of cis regions and brassinosteroids. *Plant Physiology* **130**, 770-783.

**Ingvarsson PK, Street NR**. 2011. Association genetics of complex traits in plants. *New Phytologist* **189**, 909-922.

**Jaffe MJ**. 1973. Thigmomorphogenesis: The respone of plant growth and development to mechanical stimulation. *Planta* **114**, 143-157.

**Jansen RC**. 1993. Interval mapping of multiple quantitative trait loci. *Genetics* **135**, 205-211.

**Jones JDG, Dangl JL**. 2006. The plant immune system. **444**, 323-329.

**Joosen RV, Kodde J, Willems LAJ, Ligterink W, van der Plas LHW, Hilhorst HWM**. 2010. GERMINATOR: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. *The Plant Journal* **62**, 148-159.

**Jørgensen ME, Nour-Eldin HH, Halkier BA**. 2015. Transport of defense compounds from source to sink: lessons learned from glucosinolates. *Trends in Plant Science* **20**, 508-514.

**Joukhadar R, El-Bouhssini M, Jighly A**. 2013. Genome-wide association mapping for five major pest resistances in wheat. *Molecular Breeding* **32**, 943-960.

**Julkowska MM, Hoefsloot HCJ, Mol S, Feron R, de Boer G-J, Haring MA, Testerink C**. 2014. Capturing Arabidopsis root architecture dynamics with root-fit reveals diversity in responses to salinity. *Plant Physiology* **166**, 1387-1402.

**Julkowska MM, Testerink C**. 2015. Tuning plant signaling and growth to survive salt. *Trends in Plant Science* **20**, 586-594.

**Jurkowski GI, Smith RK, Jr., Yu IC, Ham JH, Sharma SB, Klessig DF, Fengler KA, Bent AF**. 2004. Arabidopsis DND2, a second cyclic nucleotide-gated ion channel gene for which mutation causes the "defense, no death" phenotype. *Molecular Plant-Microbe Interactions* **17**, 511-520.

**Kaloshian I**. 2004. Gene-for-gene disease resistance: Bridging insect pest and pathogen defense. *Journal of Chemical Ecology* **30**, 2419-2434.

**Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E**. 2008. Efficient control of population structure in model organism association mapping. *Genetics* **178**, 1709-1723.

**Karban R, Agrawal AA**. 2002. Herbivore offense. *Annual Review of Ecology and Systematics* **33**, 641-664.

**Karban R, Baldwin IT**. 1997. *Induced responses to herbivory*. Chicago: Chicago University Press.

**Keisa A, Kanberga-Silina K, Nakurte I, Kunga L, Rostoks N**. 2011. Differential disease resistance response in the barley necrotic mutant nec1. *BMC Plant Biology* **11**, 66.

**Kennedy JS, Booth CO, Kershaw WJS**. 1961. Host finding by aphids in the field. III. Visual attraction. *Annals of Applied Biology* **49**, 1-21.

**Kennedy JS, Mittler TE**. 1953. A method of obtaining phloem sap via the mouth-parts of

aphids. *Nature* **171**, 528.

**Kerchev PI, Karpinska B, Morris JA, Hussain A, Verrall SR, Hedley PE, Fenton B, Foyer CH, Hancock RD**. 2013. Vitamin C and the abscisic acid-insensitive 4 transcription factor are important determinants of aphid resistance in Arabidopsis. *Antioxidants & redox signaling* **18**, 2091-2105.

**Kerwin R, Feusier J, Corwin J, Rubin M, Lin C, Muok A, Larson B, Li BH, Joseph B, Francisco M, Copeland D, Weinig C, Kliebenstein DJ**. 2015. Natural genetic variation in *Arabidopsis thaliana* defense metabolism genes modulates field fitness. *eLife* **4**, e05604.

**Kettles GJ, Drurey C, Schoonbeek H-j, Maule AJ, Hogenhout SA**. 2013. Resistance of *Arabidopsis thaliana* to the green peach aphid, *Myzus persicae*, involves camalexin and is regulated by microRNAs. *The New Phytologist* **198**, 1178-1190.

**Keurentjes JJB, Angenent GC, Dicke M, Dos Santos V, Molenaar J, van der Putten WH, de Ruiter PC, Struik PC, Thomma B**. 2011. Redefining plant systems biology: from cell to ecosystem. *Trends in Plant Science* **16**, 183-190.

**Keurentjes JJB, Bentsink L, Alonso-Blanco C, Hanhart CJ, Blankestijn-De Vries H, Effgen S, Vreugdenhil D, Koornneef M**. 2007. Development of a near-isogenic line population of *Arabidopsis thaliana* and comparison of mapping power with a recombinant inbred line population. *Genetics* **175**, 891-905.

**Keurentjes JJB, Fu JY, de Vos CHR, Lommen A, Hall RD, Bino RJ, van der Plas LHW, Jansen RC, Vreugdenhil D, Koornneef M**. 2006. The genetics of plant metabolism. *Nature Genetics* **38**, 842-849.

**Keurentjes JJB, Koornneef M, Vreugdenhil D**. 2008. Quantitative genetics in the age of omics. *Current Opinion in Plant Biology* **11**, 123-128.

**Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K**. 2007. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *The Plant Journal* **50**, 347-363.

**Kim CS, Schaible G, Garrett L, Lubowski R, Lee D**. 2008a. Economic impacts of the U.S. soybean aphid infestation: a multi-regional competitive dynamic analysis. *Agricultural and Resource Economics Review* **37**, 227-242.

**Kim JH, Lee BW, Schroeder FC, Jander G**. 2008b. Identification of indole glucosinolate breakdown products with antifeedant effects on *Myzus persicae* (green peach aphid). *The Plant Journal* **54**, 1015-1026.

**Kim KC, Lai ZB, Fan BF, Chen ZX**. 2008c. Arabidopsis WRKY38 and WRKY62 transcription factors interact with histone deacetylase 19 in basal defense. *The Plant Cell* **20**, 2357-2371.

**Kim S, Plagnol V, Hu TT, Toomajian C, Clark RM, Ossowski S, Ecker JR, Weigel D, Nordborg M**. 2007. Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics* **39**, 1151-1155.

**Kim Y-S, Sakuraba Y, Han S-H, Yoo S-C, Paek N-C**. 2013. Mutation of the Arabidopsis NAC016 transcription factor delays leaf senescence. *Plant and Cell Physiology* **54**, 1660-1672.

**Kimura M**. 1983. *The neutral theory of molecular evolution*. Cambridge: Cambridge University Press.

**Kissoudis C, Chowdhury R, van Heusden S, van de Wiel C, Finkers R, Visser RF, Bai Y, van der Linden G**. 2015. Combined biotic and abiotic stress resistance in tomato. *Euphytica* **202**, 317-332.

**Kliebenstein DJ**. 2014. Quantitative genetics and genomics of plant resistance to insects. *Annual Plant Reviews* **47**, 235-262.

**Kloth KJ, ten Broeke CJM, Thoen MPM, Hanhart-van den Brink M, Wiegers GL, Krips OE, Noldus LPJJ, Dicke M, Jongsma MA**. 2015. High-throughput phenotyping of plant resistance to aphids by automated video tracking. *Plant Methods* **11**, 4.

**Kloth KJ, Thoen MPM, Bouwmeester HJ, Jongsma MA, Dicke M**. 2012. Association mapping of plant resistance to insects. *Trends in Plant Science* **17**, 311-319.

**Knoblauch M, Froelich DR, Pickard WF, Peters WS**. 2014. SEORious business: structural proteins in sieve tubes and their involvement in sieve element occlusion. *Journal of Experimental Botany* **65**, 1879-1893.

**Knoblauch M, Mullendore DL**. 2013. Sieve element occlusion. In: Thompson GA, van Bel AJE, eds. *Phloem: Molecular Cell Biology, Systemic Communication, Biotic Interactions*. Oxford: John Wiley & Sons, Inc., 141-153.

**Knoblauch M, Oparka K**. 2012. The structure of the phloem – still more questions than answers. *The Plant Journal* **70**, 147–156.

**Knoblauch M, Van Bel AJE**. 1998. Sieve tubes in action. *The Plant Cell* **10**, 35-50.

**Kokorian J, Polder G, Keurentjes JJB, Vreugdenhil D, Olortegui Guzman M**. 2010. An ImageJ based measurement setup for automated phenotyping of plants. In: Jahnen A, Moll C, eds. *Proc. of the ImageJ user and developer conference 2010*. Luxembourg: Centre de Recherche Public Henri Tudor, 178-182.

**Kollner TG, Held M, Lenk C, Hiltpold I, Turlings TCJ, Gershenzon J, Degenhardt J**. 2008. A maize (E)-beta-caryophyllene synthase implicated in indirect defense responses against herbivores is not expressed in most American maize varieties. *The Plant Cell* **20**, 482-494.

**Kooke R**. 2014. Missing heritability and soft inheritance of morphology and metabolism in Arabidopsis, Wageningen University, Wageningen, the Netherlands.

**Koornneef A, Leon-Reyes A, Ritsema T, Verhage A, Den Otter FC, Van Loon LC, Pieterse CMJ**. 2008. Kinetics of salicylate-mediated suppression of jasmonate signaling reveal a role for redox modulation. *Plant Physiology* **147**, 1358-1368.

**Koornneef M, Meinke D**. 2010. The development of Arabidopsis as a model plant. *The Plant Journal* **61**, 909-921.

**Korte A, Vilhjalmsson BJ, Segura V, Platt A, Long Q, Nordborg M**. 2012. A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature Genetics* **44**, 1066–1071.

**Kramer KJ, Morgan TD, Throne JE, Dowell FE, Bailey M, Howard JA**. 2000. Transgenic avidin maize is resistant to storage insect pests. *Nature Biotechnology* **18**, 670-674.

**Krips OE, A. W, Willems PEL, Dicke M**. 1998. Intrinsic rate of population increase of the spider mite *Tetranychus urticae* on the ornamental crop gerbera: intraspecific variation in host

plant and herbivore. *Entomologia Experimentalis et Applicata* **89**, 159-168.

**Kruijer W, Boer MP, Malosetti M, Flood PJ, Engel B, Kooke R, Keurentjes JJB, van Eeuwijk FA**. 2015. Marker-based estimation of heritability in immortal populations. *Genetics* **199**, 379-398.

**Kusnierczyk A, Winge P, Jorstad TS, Troczynska J, Rossiter JT, Bones AM**. 2008. Towards global understanding of plant defence against aphids - timing and dynamics of early *Arabidopsis* defence responses to cabbage aphid (*Brevicoryne brassicae*) attack. *Plant Cell and Environment* **31**, 1097-1115.

**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.

**Lacey ES, Carde RT**. 2011. Activation, orientation and landing of female *Culex quinquefasciatus* in response to carbon dioxide and odour from human feet: 3-D flight analysis in a wind tunnel. *Medical and Veterinary Entomology* **25**, 94-103.

**Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E**. 2011. The Arabidopsis information resource (TAIR): Improved gene annotation and new tools. *Nucleic Acids Research* **40**, D1202–D1210.

**Lange MJP, Lange T**. 2015. Touch-induced changes in Arabidopsis morphology dependent on gibberellin breakdown. **1**, 14025.

**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.

**Larson K, Whitham T**. 1991. Manipulation of food resources by a gall-forming aphid: the physiology of sink-source interactions. *Oecologia* **88**, 15-21.

**Lauvergeat V, Lacomme C, Lacombe E, Lasserre E, Roby D, Grima-Pettenati J**. 2001. Two cinnamoyl-CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria. *Phytochemistry* **57**, 1187-1195.

**Lee B-H, Henderson DA, Zhu J-K**. 2005. The Arabidopsis cold-responsive transcriptome and its regulation by ICE1. *The Plant Cell* **17**, 3155-3175.

**Lee D, Polisensky DH, Braam J**. 2004. Genome-wide identification of touch- and darkness-regulated Arabidopsis genes: a focus on calmodulin-like and *XTH* genes. *New Phytologist* **165**, 429-444.

**Lee SM, Hoang MH, Han HJ, Kim HS, Lee K, Kim KE, Kim DH, Lee SY, Chung WS**. 2009. Pathogen inducible voltage-dependent anion channel (AtVDAC) isoforms are localized to mitochondria membrane in Arabidopsis. *Molecules and Cells* **27**, 321-327.

**Lemoine R, Camera SL, Atanassova R, Dédaldéchamp F, Allario T, Pourtau N, Bonnemain J, Laloi M, Coutos-Thévenot P, Maurousset L, Faucher M, Girousse G, Lemonnier P, Parrilla J, Durand M**. 2013. Source-to-sink transport of sugar and regulation by environmental factors. *Frontiers in Plant Science* **4**, 1-21.

**Leon J, Rojo E, Sanchez-Serrano JJ**. 2001. Wound signalling in plants. *Journal of Experimental Botany* **52**, 1-9.

**Levy M, Wang Q, Kaspi R, Parrella MP, Abel S**. 2005. Arabidopsis IQD1, a novel calmodulin-binding nuclear protein, stimulates glucosinolate accumulation and plant defense. *The Plant Journal* **43**, 79-96.

**Lewis T**. 1973. *Thrips their biology, ecology and economic importance*. New York: Academic Press.

**Li J, Brader G, Palva ET**. 2004. The WRKY70 transcription factor: A node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *The Plant Cell* **16**, 319-331.

**Li Q, Xie Q, Smith-Becker J, Navarre D, Kaloshian I**. 2006. *Mi-1*-mediated aphid resistance involves salicylic acid and mitogen-activated protein kinase signaling cascades. *Molecular Plant-Microbe Interactions* **19**, 655-664.

**Li X, Yan W, Agrama H, Jia L, Shen X, Jackson A, Moldenhauer K, Yeater K, McClung A, Wu D**. 2011. Mapping QTLs for improving grain yield using the USDA rice mini-core collection. *Planta* **234**, 347-361.

**Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO**. 2010. Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21199-21204.

**Liepman AH, Wilkerson CG, Keegstra K**. 2005. Expression of cellulose synthase-like (*Csl*) genes in insect cells reveals that *CslA* family members encode mannan synthases. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 2221-2226.

**Lister R, Gregory BD, Ecker JR**. 2009. Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Current Opinion in Plant Biology* **12**, 107-118.

**Liu Y, McCreight JD**. 2006. Responses of *Nasonovia ribisnigri* (Homoptera: Aphididae) to susceptible and resistant lettuce. *Journal of Economic Entomology* **99**, 972-978.

**Liu Y, Wang W-L, Guo G-X, Ji X-L**. 2009. Volatile emission in wheat and parasitism by Aphidius avenae after exogenous application of salivary enzymes of Sitobion avenae. *Entomologia Experimentalis et Applicata* **130**, 215-221.

**Liu ZQ, Yan L, Wu Z, Mei C, Lu K, Yu YT, Liang S, Zhang XF, Wang XF, Zhang DP**. 2012. Cooperation of three WRKY-domain transcription factors WRKY18, WRKY40, and WRKY60 in repressing two ABA-responsive genes ABI4 and ABI5 in Arabidopsis. *Journal of Experimental Botany* **63**, 6371-6392.

**Loker ES, Hofkin BV**. 2015. *Parasitology. A conceptual framework*. Abingdon: Garland Science, Taylor & Francis Group, LLC.

**Loon LCv, Rep M, Pieterse CMJ**. 2006. Significance of inducible defense-related proteins in infected plants. *Annual Review of Phytopathology* **44**, 135-162.

**Lorenzo O, Chico JM, Sánchez-Serrano JJ, Solano R**. 2004. *JASMONATE-INSENSITIVE1* encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in Arabidopsis. *The Plant Cell* **16**, 1938-1950.

**Lorenzo O, Piqueras R, Sánchez-Serrano JJ, Solano R**. 2003. ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense. *The*

*Plant Cell* **15**, 165-178.

**Loxdale HD, Lushai G, Harvey JA**. 2011. The evolutionary improbability of 'generalism' in nature, with special reference to insects. *Biological Journal of the Linnean Society* **103**, 1-18.

**Lu H, Zou Y, Feng N**. 2010. Overexpression of *AHL20* negatively regulates defenses in Arabidopsis. *Journal of Integrated Plant Biology* **52**, 801-808.

**Lu X, Tintor N, Mentzel T, Kombrink E, Boller T, Robatzek S, Schulze-Lefert P, Saijo Y**. 2009. Uncoupling of sustained MAMP receptor signaling from early outputs in an Arabidopsis endoplasmic reticulum glucosidase II allele. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 22522-22527.

**Luderitz T, Grisebach H**. 1981. Enzymic synthesis of lignin precursors. Comparison of cinnamoyl-CoA reductase and cinnamyl alcohol:NADP+ dehydrogenase from spruce (*Picea abies* L.) and soybean (*Glycine max* L.). *European Journal of Biochemistry* **119**, 115-124.

**Macel M, van Dam NM, Keurentjes JJB**. 2010. Metabolomics: the chemistry between ecology and genetics. *Molecular Ecology Resources* **10**, 583-593.

**Mackay PA, Wellington WG**. 1977. Notes on the life history and habits of the red-backed sawfly, *Eriocampa ovata* (Hymenoptera: Tenthredinidae). *The Canadian Entomologist* **109**, 53-58.

**Madden JL**. 1977. Physiological reactions of *Pinus radiata* to attack by woodwasp, *Sirex noctilio* F. (Hymenoptera: Siricidae). *Bulletin of Entomological Research* **67**, 405-426.

**Maere S, Heymans K, Kuiper M**. 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics* **21**, 3448-3449.

**Mahajan SK, Chisholm ST, Whitham SA, Carrington JC**. 1998. Identification and characterization of a locus (RTM1) that restricts long-distance movement of tobacco etch virus in Arabidopsis thaliana. *The Plant Journal* **14**, 177-186.

**Makumburage GB, Richbourg HL, LaTorre KD, Capps A, Chen CX, Stapleton AE**. 2013. Genotype to phenotype maps: multiple input abiotic signals combine to produce growth effects via attenuating signaling interactions in maize. *G3-Genes Genomes Genetics* **3**, 2195-2204.

**Malosetti M, Ribaut JM, Vargas M, Crossa J, Van Eeuwijk FA**. 2008. A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). *Euphytica* **161**, 241-257.

**Malosetti M, van der Linden CG, Vosman B, van Eeuwijk FA**. 2007. A mixed-model approach to association mapping using pedigree information with an illustration of resistance to *Phytophthora infestans* in potato. *Genetics* **175**, 879-889.

**Malosetti M, van Eeuwijk FA, Boer MP, Casas AM, Elia M, Moralejo M, Bhat PR, Ramsay L, Molina-Cano JL**. 2011. Gene and QTL detection in a three-way barley cross under selection by a mixed model with kinship information using SNPs. *Theoretical and Applied Genetics* **122**, 1605-1616.

**Mao P, Duan M, Wei C, Li Y**. 2007. WRKY62 transcription factor acts downstream of cytosolic NPR1 and negatively regulates jasmonate-responsive gene expression. *Plant and Cell Physiology* **48**, 833-842.

**Mare C, Mazzucotelli E, Crosatti C, Francia E, Stanca AM, Cattivelli L**. 2004. Hv-WRKY38: a new transcription factor involved in cold- and drought-response in barley. *Plant Molecular Biology* **55**, 399-416.

**Martín-López B, Varela I, Marnotes S, Cabaleiro C**. 2006. Use of oils combined with low doses of insecticide for the control of Myzus persicae and PVY epidemics. *Pest Management Science* **62**, 372-378.

**Martin B, Rahbé Y, Fereres A**. 2003. Blockage of stylet tips as the mechanism of resistance to virus transmission by *Aphis gossypii* in melon lines bearing the Vat gene. *Annals of Applied Biology* **142**, 245-250.

**Martiniere A, Li X, Runions J, Lin J, Maurel C, Luu DT**. 2012. Salt stress triggers enhanced cycling of Arabidopsis root plasma-membrane aquaporins. *Plant Signaling & Behavior* **7**, 529-532.

**Mauricio R, Rausher MD**. 1997. Experimental manipulation of putative selective agents provides evidence for the role of natural enemies in the evolution of plant defense. *Evolution* **5**, 1435-1444.

**McLean DL, Kinsey MG**. 1964. A technique for electronically recording aphid feeding and salivation. *Nature* **202**, 1358-1359.

**McNairn RB**. 1972. Phloem translocation and heat-induced callose formation in field-grown *Gossypium hirsutum* L. *Plant Physiology* **50**, 366-370.

**McNairn RB, Currier HB**. 1968. Translocation blockage by sieve plate callose. *Planta* **82**, 369-380.

**Meijer K**. 2009. Factor-analytic models for genotype × environment type problems and structured covariance matrices. *Genetics Selection Evolution* **41**, 21.

**Menzel TR, Weldegergis BT, David A, Boland W, Gols R, van Loon JJA, Dicke M**. 2014. Synergism in the effect of prior jasmonic acid application on herbivore-induced volatile emission by Lima bean plants: transcription of a monoterpene synthase gene and volatile emission. *Journal of Experimental Botany* **65**, 4821-4831.

**Mészáros T, Helfer A, Hatzimasoura E, Magyar Z, Serazetdinova L, Rios G, Bardóczy V, Teige M, Koncz C, Peck S, Bögre L**. 2006. The Arabidopsis MAP kinase kinase MKK1 participates in defence responses to the bacterial elicitor flagellin. *The Plant Journal* **48**, 485-498.

**Mewis I, Appel HM, Hom A, Raina R, Schultz JC**. 2005. Major signaling pathways modulate *Arabidopsis* glucosinolate accumulation and response to both phloem-feeding and chewing insects. *Plant Physiology* **138**, 1149-1162.

**Mewis I, Khan MAM, Glawischnig E, Schreiner M, Ulrichs C**. 2012. Water stress and aphid feeding differentially influence metabolite composition in *Arabidopsis thaliana* (L.). *PLoS ONE* **7**, e48661.

**Mewis I, Tokuhisa JG, Schultz JC, Appel HM, Ulrichs C, Gershenzon J**. 2006. Gene expression and glucosinolate accumulation in *Arabidopsis thaliana* in response to generalist and specialist herbivores of different feeding guilds and the role of defense signaling pathways. *Phytochemistry* **67**, 2450-2462.

**Miao Y, Laun T, Zimmerman P, Zentgraf U**. 2004. Targets of the WRKY53 transcription factor and its role during leaf senescence in Arabidopsis *Plant Molecular Biology* **55**, 853-867.

**Miao Y, Zentgraf U**. 2007. The antagonist function of *Arabidopsis* WRKY53 and ESR/ESP in leaf senescence is modulated by the jasmonic and salicylic acid equilibrium. *The Plant Cell* **19**, 819-830.

**Minks AK, Harrewijn P**. 1989. *World crop pests. Aphids. Their biology, natural enemies and control*. Amsterdam, the Netherlands: Elsevier Science Publishers.

**Mitchell-Olds T**. 2010. Complex-trait analysis in plants. *Genome Biology* **11**, 113.

**Mitchell A, Chang H-Y, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A, Scheremetjew M, Rato C, Yong S-Y, Bateman A, Punta M, Attwood TK, Sigrist CJA, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P, Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo C, Sillitoe I, Mi H, Thomas PD, Finn RD**. 2015. The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Research* **43**, D213-D221.

**Mitsuda N, Ohme-Takagi M**. 2009. Functional analysis of transcription factors in Arabidopsis. *Plant and Cell Physiology* **50**, 1232-1248.

**Mittler TE**. 1988. Applications of artificial feeding techniques for aphids. In: Minks AK, Harrewijn P, eds. *Aphids, their biology, natural enemies ano control*, Vol. 2B. Amsterdam: Elsevier, 145-170.

**Mittler TE, Tsitsipis JA, Kleinjan JE**. 1970. Utilization of dehydroascorbic acid and some related compounds by the aphid *Myzus persicae* feeding on an improved diet. *Journal of Insect Physiology* **16**, 2315-2326.

**Moloi MJ, Westhuizen AJvd**. 2006. The reactive oxygen species are involved in resistance responses of wheat to the Russian wheat aphid. *Journal of Plant Physiology* **163**, 1118-1125.

**Monte E, Tepperman JM, Al-Sady B, Kaczorowski KA, Alonso JM, Ecker JR, Li X, Zhang Y, Quail PH**. 2004. The phytochrome-interacting transcription factor, PIF3, acts early, selectively, and positively in light-induced chloroplast development. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 16091-16098.

**Montes JM, Melchinger AE, Reif JC**. 2007. Novel throughput phenotyping platforms in plant genetic studies. *Trends in Plant Science* **12**, 433-436.

**Morales M**. 2011. sciplot: Scientific graphing functions for factorial designs, R package version 1.1-9. http://CRAN.R-project.org/package=sciplot.

**Moran NA**. 1992. The evolution of aphid life cycles. *Annual Review of Entomology* **37**, 321-348.

**Moran PJ, Cheng YF, 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.

**Moran PJ, Thompson GA**. 2001. Molecular responses to aphid feeding in Arabidopsis in relation to plant defense pathways. *Plant Physiology* **125**, 1074-1085.

**Moreno-Delafuente A, Garzo E, Moreno A, Fereres A**. 2013. A Plant Virus Manipulates the Behavior of Its Whitefly Vector to Enhance Its Transmission Efficiency and Spread. *PLoS*

*ONE* **8**, e61543.

**Morkunas I, Ratajczak L**. 2014. The role of sugar signaling in plant defense responses against fungal pathogens. *Acta Physiologiae Plantarum* **36**, 1607-1619.

**Münch E**. 1926. Dynamik der Saftströmungen. *Berichte der Deutschen Botanischen Gesellschaft* **44**, 68-71.

**Münch E**. 1927. Versuche über den Saftkreislauf. *Berichte der Deutschen Botanischen Gesellschaft* **45**, 340-356.

**Münch E**. 1930. *Die Stoffbewegingen in der Pflanze*. Jena: G. Fischer.

**Mutti NS, Louis J, Pappan LK, Pappan K, Begum K, Chen M-S, Park Y, Dittmer N, Marshall J, Reese JC, Reeck GR**. 2008. A protein from the salivary glands of the pea aphid, *Acyrthosiphon pisum*, is essential in feeding on a host plant. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 9965-9969.

**Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES**. 2009. Association mapping: critical considerations shift from genotyping to experimental design. *The Plant Cell* **21**, 2194-2202.

**Nalam VJ, Keereetaweep J, Shah J**. 2012a. The green peach aphid, *Myzus persicae*, acquires a LIPOXYGENASE5-derived oxylipin from Arabidopsis thaliana, which promotes colonization of the host plant. *Plant Signalling & Behavior* **8**.

**Nalam VJ, Keeretaweep J, Sarowar S, Shah J**. 2012b. Root-derived oxylipins promote green peach aphid performance on arabidopsis foliage. *The Plant Cell* **24**, 1643-1653.

**Nault LR**. 1997. Arthropod transmission of plant viruses: A new synthesis. *Annals of the Entomological Society of America* **90**, 521-541.

**Navarro L, Zipfel C, Rowland O, Keller I, Robatzek S, Boller T, Jones JDG**. 2004. The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiology* **135**, 1113-1128.

**Ndamukong I, Abdallat AA, Thurow C, Fode B, Zander M, Weigel RR, Gatz C**. 2007. SA-inducible Arabidopsis glutaredoxin interacts with TGA factors and suppresses JA-responsive *PDF1.2* transcription. *The Plant Journal* **50**, 128-139.

**Nemri A, Atwell S, Tarone AM, Huang YS, Zhao K, Studholme DJ, Nordborg M, Jones JDG**. 2010. Genome-wide survey of Arabidopsis natural variation in downy mildew resistance using combined association and linkage mapping. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 10302-10307.

**Noldus LPJJ, Spink AJ, Tegelenbosch RAJ**. 2001. EthoVision: A versatile video tracking system for automation of behavioral experiments. *Behavior Research Methods, Instruments, and Computers* **33**, 398-414.

**Noldus LPJJ, Spink AJ, Tegelenbosch RAJ**. 2002. Computerised video tracking, movement analysis and behaviour recognition in insects. *Computers and Electronics in Agriculture* **35**, 201-227.

**Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, Bakker E, Calabrese P, Gladstone J, Goyal R, Jakobsson M, Kim S, Morozov Y, Padhukasahasram B, Plagnol V, Rosenberg NA, Shah C, Wall JD, Wang J, Zhao K, Kalbfleisch T, Schulz**

**V, Kreitman M, Bergelson J**. 2005. The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biology* **3**, e196.

**Nordborg M, Weigel D**. 2008. Next-generation genetics in plants. *Nature* **456**, 720-723.

**Oakey H, Verbyla A, Pitchford W, Cullis B, Kuchel H**. 2006. Joint modeling of additive and non-additive genetic line effects in single field trials. *Theoretical and Applied Genetics* **113**, 809-819.

**Obayashi T, Okamura Y, Ito S, Aoki Y, Shirota M, Kinoshita K**. 2014. ATTED-II in 2014: Evaluation of gene coexpression in agriculturally important plants. *Plant & Cell Physiology* **55**, e6(1-7).

**Ober D, Kaltenegger E**. 2009. Pyrrolizidine alkaloid biosynthesis, evolution of a pathway in plant secondary metabolism. *Phytochemistry* **70**, 1687-1695.

**Obrycki JJ, Harwood JD, Kring TJ, O'Neil RJ**. 2009. Aphidophagy by Coccinellidae: Application of biological control in agroecosystems. *Biological Control* **51**, 244-254.

**Odong TL, van Heerwaarden J, van Hintum TJL, van Eeuwijk FA, Jansen J**. 2013. Improving hierarchical clustering of genotypic data via Principal Component Analysis. *Crop Science* **53**, 1546-1554.

**Ogura T, Busch W**. 2015. From phenotypes to causal sequences: using genome wide association studies to dissect the sequence basis for variation of plant development. *Current Opinion in Plant Biology* **23**, 98-108.

**Olukolu BA, Wang G-F, Vontimitta V, Venkata BP, Marla S, Ji J, Gachomo E, Chu K, Negeri A, Benson J, Nelson R, Bradbury P, Nielsen D, Holland JB, Balint-Kurti PJ, Johal G**. 2014. A genome-wide association study of the Maize hypersensitive defense response identifies genes that cluster in related pathways. *PLoS Genetics* **10**, e1004562.

**Oparka KJ, Turgeon R**. 1999. Sieve elements and companion cells - traffic control centers of the phloem. *The Plant Cell* **11**, 739-750.

**Panda N, Khush GS**. 1995. *Host plant resistance to insects*. Wallingford, UK: CABI Intl.

**Pandey SP, Roccaro M, Schon M, Logemann E, Somssich IE**. 2010. Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of Arabidopsis. *The Plant Journal* **64**, 912-923.

**Payne RW**. 2009. GenStat. *WIREs Computational Statistics* **1**, 255-258.

**Pegadaraju V, Knepper C, Reese J, Shah J**. 2005. Premature leaf senescence modulated by the Arabidopsis *PHYTOALEXIN DEFICIENT4* gene is associated with defense against the phloem-feeding green peach aphid. *Plant Physiology* **139**, 1927-1934.

**Pegadaraju V, Louis J, Singh V, Reese JC, Bautor J, Feys BJ, Cook G, Parker JE, Shah J**. 2007. Phloem-based resistance to green peach aphid is controlled by Arabidopsis *PHYTOALEXIN DEFICIENT4* without its signaling partner *ENHANCED DISEASE SUSCEPTIBILITY1*. *The Plant Journal* **52**, 332-341.

**Penninckx IAMA, Thomma BPHJ, Buchala A, Metraux JP, Broekaert WF**. 1998. Concomitant activation of jasmonate and ethylene response pathways is required for induction of a plant defensin gene in Arabidopsis. *The Plant Cell* **10**, 2103-2113.

**Peret B, Li G, Zhao J, Band LR, Voss U, Postaire O, Luu DT, Da Ines O, Casimiro I,**

**Lucas M, Wells DM, Lazzerini L, Nacry P, King JR, Jensen OE, Schaffner AR, Maurel C, Bennett MJ**. 2012. Auxin regulates aquaporin function to facilitate lateral root emergence. *Nature Cell Biology* **14**, 991-998.

**Pfalz M, Vogel H, Kroymann J**. 2009. The gene controlling the *Indole glucosinolate modifier1* quantitative trait locus alters indole glucosinolate structures and aphid resistance in *Arabidopsis*. *The Plant Cell* **21**, 985-999.

**Pickett JA, Wadhams LJ, Woodcock CM, Hardie J**. 1992. The chemical ecology of aphids. *Annual Review of Entomology* **37**, 67-90.

**Piepho HP**. 1997. Analyzing genotype-environment data by mixed models with multiplicative effects. *Biometrics* **53**, 761-766.

**Pierik R, Testerink C**. 2014. The art of being flexible: how to escape from shade, salt, and drought. *Plant Physiology* **166**, 5-22.

**Pieterse CMJ, Leon-Reyes A, Van der Ent S, Van Wees SCM**. 2009. Networking by small-molecule hormones in plant immunity. *Nature Chemical Biology* **5**, 308-316.

**Pieterse CMJ, Van der Does D, Zamioudis C, Leon-Reyes A, Van Wees SCM**. 2012. Hormonal modulation of plant immunity. *Annual Review of Cell and Developmental Biology* **28**, 489-521.

**Pinheiro P, Bereman MS, Burd J, Pals M, Armstrong S, Howe KJ, Thannhauser TW, MacCoss MJ, Gray SM, Cilia M**. 2014. Evidence of the biochemical basis of host virulence in the greenbug aphid, *Schizaphis graminum* (Homoptera: Aphididae). *Journal of Proteome Research* **13**, 2094-2108.

**Pinkiewicz TH, Purser GJ, Williams RN**. 2011. A computer vision system to analyse the swimming behaviour of farmed fish in commercial aquaculture facilities: A case study using cage-held atlantic salmon. *Aquacultural Engineering* **45**, 20-27.

**Pistori H, Viana Aguiar Odakura VV, Oliveira Monteiro JB, Goncalves WN, Roel AR, Silva JdA, Machado BB**. 2010. Mice and larvae tracking using a particle filter with an auto-adjustable observation model. *Pattern Recognition Letters* **31**, 337-346.

**Pitino M, Hogenhout SA**. 2012. Aphid protein effectors promote aphid colonization in a plant species-specific manner. *Molecular Plant-Microbe Interactions* **26**, 130-139.

**Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Ågren J, Bossdorf O, Byers D, Donohue K, Dunning M, Holub EB, Hudson A, Le Corre V, Loudet O, Roux F, Warthmann N, Weigel D, Rivero L, Scholl R, Nordborg M, Bergelson J, Borevitz JO**. 2010. The scale of population structure in *Arabidopsis thaliana*. *PLoS Genetics* **6**, e1000843.

**Poelman EH, Galiart R, Raaijmakers CE, van Loon JJA, van Dam NM**. 2008a. Performance of specialist and generalist herbivores feeding on cabbage cultivars is not explained by glucosinolate profiles. *Entomologia Experimentalis et Applicata* **127**, 218-228.

**Poelman EH, van Loon JJA, Dicke M**. 2008b. Consequences of variation in plant defense for biodiversity at higher trophic levels. *Trends in Plant Science* **13**, 534-541.

**Poelman EH, Van Loon JJA, Van Dam NM, Vet LEM, Dicke M**. 2010. Herbivore-induced plant responses in *Brassica oleracea* prevail over effects of constitutive resistance and result

in enhanced herbivore attack. *Ecological Entomology* **35**, 240-247.

**Pompon J, Quiring D, Giordanengo P, Pelletier Y**. 2010. Role of xylem consumption on osmoregulation in *Macrosiphum euphorbiae* (Thomas). *Journal of Insect Physiology* **56**, 610-615.

**Powell G, Tosh CR, Hardie J**. 2006. Host plant selection by aphids: behavioral, evolutionary and applied perspectives. *Annual Review of Entomology* **51**, 309-330.

**Prado E, Tjallingii WF**. 1994. Aphid activities during sieve element punctures. *Entomologia Experimentalis et Applicata* **72**, 157-165.

**Prado E, Tjallingii WF**. 1999. Effects of experimental stress factors on probing behaviour by aphids. *Entomologia Experimentalis et Applicata* **90**, 289-300.

**Prado E, Tjallingii WF**. 2007. Behavioral evidence for local reduction of aphid-induced resistance. *Journal of Insect Science* **7**, 1-8.

**Prasifka JR, Hellmich RL, Crespo ALB, Siegfried BD, Onstad DW**. 2010. Video-tracking and on-plant tests show Cry1Ab resistance influences behavior and survival of neonate *Ostrinia nubilalis* following exposure to Bt maize. *Journal of Insect Behavior* **23**, 1-11.

**Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D**. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* **38**, 904-909.

**Price PW, Denno RF, Eubanks MD, Finke DL, Kaplan I**. 2011. Insect ecology: behavior, populations and communities. Cambridge: Cambridge University Press.

**Price RA, Al-Shehbaz IA, Palmer JD**. 1994. Systematic relationships of *Arabidopsis*: a molecular and morphological approach. In: E. M, Somerville C, eds. *Arabidopsis*. Cold Spring Harbor, New York: Cold Spring Harbor Press, 7-19.

**Purugganan MM, Braam J, Fry SC**. 1997. The arabidopsis TCH4 xyloglucan endotransglycosylase - Substrate specificity, pH optimum, and cold tolerance. *Plant Physiology* **115**, 181-190.

**R-Core-Team**. 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

**Raffaele S, Bayer E, Lafarge D, Cluzet S, Retana SG, Boubekeur T, Leborgne-Castel N, Carde JP, Lherminier J, Noirot E, Satiat-Jeunemaitre B, Laroche-Traineau J, Moreau P, Ott T, Maule AJ, Reymond P, Simon-Plas F, Farmer EE, Bessoule JJ, Mongrand S**. 2009. Remorin, a Solanaceae protein resident in membrane rafts and plasmodesmata, impairs Potato Virus X movement. *Plant Cell* **21**, 1541-1555.

**Rasmussen S, Barah P, Suarez-Rodriguez MC, Bressendorff S, Friis P, Costantino P, Bones AM, Nielsen HB, Mundy J**. 2013. Transcriptome responses to combinations of stresses in *Arabidopsis*. *Plant Physiology* **161**, 1783-1794.

**Reinink K, Dieleman FL**. 1989. Comparison of sources of resistance to leaf aphids in lettuce (*lactuca sativa* L.). *Euphytica* **40**, 21-29.

**Reymond P, Bodenhausen N, Van Poecke RMP, Krishnamurthy V, Dicke M, Farmer EE**. 2004. A conserved transcript pattern in response to a specialist and a generalist herbivore. *The Plant Cell* **16**, 3132-3147.

**Reynolds DR, Riley JR**. 2002. Remote-sensing, telemetric and computer-based technologies

for investigating insect movement: a survey of existing and potential techniques. *Computers and Electronics in Agriculture* **35**, 271-307.

**Rhoades DF**. 1985. Offensive-defensive interactions between herbivores and plants: their relevance in herbivore population dynamics and ecological theory. *The American Naturalist* **125**, 205-238.

**Rivas-San Vicente M, Plasencia J**. 2011. Salicylic acid beyond defence: its role in plant growth and development. *Journal of Experimental Botany* **62**, 3321-3338.

**Rodriguez PA, Bos JIB**. 2012. Toward understanding the role of aphid effectors in plant infestation. *Molecular Plant-Microbe Interactions* **26**, 25-30.

**Rohr F, Ulrichs C, Schreiner M, Nguyen CN, Mewis I**. 2011. Impact of hydroxylated and non-hydroxylated aliphatic glucosinolates in *Arabidopsis thaliana* crosses on plant resistance against a generalist and a specialist herbivore. *Chemoecology* **21**, 171-180.

**Roitsch T**. 1999. Source-sink regulation by sugar and stress. *Current Opinion in Plant Biology* **2**, 198-206.

**Rossi M, Goggin FL, Milligan SB, Kaloshian I, Ullman DE, Williamson VM**. 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 9750-9754.

**Runyon JB, Mescher MC, De Moraes CM**. 2008. Parasitism by *Cuscuta pentagona* attenuates host plant defenses against insect herbivores. *Plant Physiology* **146**, 987-995.

**Ruping B, Ernst A, Jekat S, Nordzieke S, Reineke A, Muller B, Bornberg-Bauer E, Prufer D, Noll G**. 2010. Molecular and phylogenetic characterization of the sieve element occlusion gene family in Fabaceae and non-Fabaceae plants. *BMC Plant Biology* **10**, 219.

**Russell GE**. 1978. *Plant breeding for pest and disease resistance*. London: Butterworths.

**Rutherford RS, vanStaden J**. 1996. Towards a rapid near-infrared technique for prediction of resistance to sugarcane borer Eldana saccharina Walker (Lepidoptera: Pyralidae) using stalk surface wax. *Journal of Chemical Ecology* **22**, 681-694.

**Salvador-Recatalà V, Tjallingii WF, Farmer EEC**. 2014. Real-time, in vivo intracellular recordings of caterpillar-induced depolarization waves in sieve elements using aphid electrodes. *New Phytologist* **203**, 674-684.

**Samayoa LF, Malvar RA, Olukolu BA, Holland JB, Butron A**. 2015. Genome-wide association study reveals a set of genes associated with resistance to the Mediterranean corn borer (*Sesamia nonagrioides* L.) in a maize diversity panel. *BMC Plant Biology* **15**, 1:15.

**Samuels AL, Giddings TH, Staehelin LA**. 1995. Cytokinesis in Tobacco BY-2 and root tip cells: A new model of cell plate formation in higher plants. *The Journal of Cell Biology* **130**, 1345-1357.

**Sanchez-Romera B, Ruiz-Lozano JM, Li G, Luu DT, Martinez-Ballesta Mdel C, Carvajal M, Zamarreno AM, Garcia-Mina JM, Maurel C, Aroca R**. 2014. Enhancement of root hydraulic conductivity by methyl jasmonate and the role of calcium and abscisic acid in this process. *Plant, Cell & Environment* **37**, 995-1008.

**Schikora A, Schenk ST, Stein E, Molitor A, Zuccaro A, Kogel K**. 2011. *N*-Acyl-Homoserine Lactone confers resistance toward biotrophic and hemibiotrophic pathogens via altered

activation of AtMPK61 *Plant Physiology* **157**, 1407-1418.

**Schneider CA, Rasband WS, Eliceiri KW**. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675.

**Schon CC, Utz HF, Groh S, Truberg B, Openshaw S, Melchinger AE**. 2004. Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. *Genetics* **167**, 485-498.

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

**Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M**. 2012. An efficient multi-locus mixed model approach for genome-wide association studies in structured populations. *Nature Genetics* **44**, 825-830.

**Seyfferth C, Tsuda K**. 2014. Salicylic acid signal transduction: the initiation of biosynthesis, perception and transcriptional reprogramming. *Frontiers in Plant Science* **5**, 697.

**Shah PA, Pell JK**. 2003. Entomopathogenic fungi as biological control agents. *Applied Microbiology and Biotechnology* **61**, 413-423.

**Shcherbakov D, Schill RO, Brümmer F, Blum M**. 2010. Movement behaviour and video tracking of *Milnesium tardigradum* Doyère, 1840 (Eutardigrada, Apochela). *Contributions to Zoology* **79**, 33-38.

**Shi Q, Febres VJ, Jones JB, Moore GA**. 2015. Responsiveness of different citrus genotypes to the *Xanthomonas citri* ssp. citri-derived pathogen-associated molecular pattern (PAMP) flg22 correlates with resistance to citrus canker. *Molecular Plant Pathology* **16**, 507-520.

**Shinohara H, Matsubayashi Y**. 2010. Arabinosylated glycopeptide hormones: new insights into CLAVATA3 structure. *Current Opinion in Plant Biology* **13**, 515-519.

**Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG**. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**, 539-539.

**Singh V, Louis J, Ayre BG, Reese JC, Shah J**. 2011. TREHALOSE PHOSPHATE SYNTHASE11-dependent trehalose metabolism promotes *Arabidopsis thaliana* defense against the phloem-feeding insect *Myzus persicae*. *The Plant Journal* **67**, 94-104.

**Smith CM, Boyko EV**. 2007. The molecular bases of plant resistance and defense responses to aphid feeding: Current status. *Entomologia Experimentalis et Applicata* **122**, 1-16.

**Smith CM, Khan ZR, Pathak MD**. 1994. *Techniques for evaluating insect resistance in crop plants*: CRC Press LLC, Florida.

**Smyrnioudis IN, Harrington R, Katis NI**. 2002. A simple test for evaluation of transmission efficiency of barley yellow dwarf virus by aphids. *Phytoparasitica* **30**, 535-538.

**Spitzen J, Spoor CW, Grieco F, ter Braak C, Beeuwkes J, van Brugge SP, Kranenbarg S, Noldus LPJJ, van Leeuwen JL, Takken W**. 2013. A 3D analysis of flight behavior of *Anopheles gambiae* sensu stricto malaria mosquitoes in response to human odor and heat. *PLoS ONE* **8**, e62995.

**Spoel SH, Dong X**. 2008. Making sense of hormone crosstalk during plant immune responses.

*Cell Host & Microbe* **3**, 348-351.

**Stam JM, Kroes A, Li YH, Gols R, van Loon JJA, Poelman EH, Dicke M**. 2014. Plant interactions with multiple insect herbivores: from community to genes. *annual Review of Plant Biology* **65**, 689-713.

**Stelinski L, Tiwari S**. 2013. Vertical T-maze choice assay for arthropod response to odorants. *Journal of visualized experiments* **72**, e50229.

**Steppuhn A, Baldwin IT**. 2007. Resistance management in a native plant: nicotine prevents herbivores from compensating for plant protease inhibitors. *Ecology Letters* **10**, 499-511.

**Steppuhn A, Gase K, Krock B, Halitschke R, Baldwin IT**. 2004. Nicotine's defensive function in nature. *PLoS Biology* **2**, e217.

**Storms JJH**. 1971. Some physiological effects of spider mite infestation on bean plants. *Netherlands Journal of Plant Pathology* **77**, 154-167.

**Stracke R, Ishihara H, Huep G, Mehrtens F, Niehaus K, Weisshaar B, Barsch A**. 2007. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant Journal* **50**, 660-677.

**Stracke R, Werber M, Weisshaar B**. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* **4**, 447-456.

**Sugimoto M, Oono Y, Gusev O, Matsumoto T, Yazawa T, Levinskikh M, Sychev V, Bingham G, Wheeler R, Hummerick M**. 2014. Genome-wide expression analysis of reactive oxygen species gene network in *Mizuna* plants grown in long-term spaceflight. *BMC Plant Biology* **14**, 4.

**Sung S, Schmitz RJ, Amasino R**. 2007. The role of *VIN3-LIKE* genes in environmentally induced epigenetic regulation of flowering. *Plant Signaling & Behavior* **2**, 127-128.

**Sutter R, Muller C**. 2011. Mining for treatment-specific and general changes in target compounds and metabolic fingerprints in response to herbivory and phytohormones in *Plantago lanceolata*. *New Phytologist* **191**, 1069-1082.

**Suzuki N, Rivero RM, Shulaev V, Blumwald E, Mittler R**. 2014. Abiotic and biotic stress combinations. *New Phytologist* **203**, 32-43.

**Taiz L**. 1984. Plant cell expansion: Regulation of cell wall mechanical properties. *Annual Review of Plant Physiology* **35**, 585-657.

**Tamura N, Yoshida T, Tanaka A, Sasaki R, Bando A, Toh S, Lepiniec L, Kawakami N**. 2006. Isolation and characterization of high temperature-resistant germination mutants of *Arabidopsis thaliana*. *Plant & Cell Physiology* **47**, 1081-1094.

**Tateda C, Watanabe K, Kusano T, Takahashi Y**. 2011. Molecular and genetic characterization of the gene family encoding the voltage-dependent anion channel in Arabidopsis. *Journal of Experimental Botany* **62**, 4773-4785.

**ten Broeke CJM**. 2013. Unravelling the resistance mechanism of lettuce against *Nasonovia ribisnigri*, Wageningen University, Wageningen, The Netherlands.

**ten Broeke CJM, Dicke M, van Loon JJA**. 2013a. Performance and feeding behaviour of two biotypes of the black currant-lettuce aphid, *Nasonovia ribisnigri*, on resistant and susceptible

*Lactuca sativa* near-isogenic lines. *Bulletin of Entomological Research* **103**, 511-521.

**ten Broeke CJM, Dicke M, van Loon JJA**. 2013b. Resistance in a *Lactuca virosa* accession to a new biotype of *Nasonovia ribisnigri*. *Euphytica* **193**, 265-275.

**Tetyuk O, Benning UF, Hoffmann-Benning S**. 2013. Collection and analysis of Arabidopsis phloem exudates using the EDTA-facilitated method. *Journal of visualized experiments* **80**, e51111.

**Therneau T**. 2015. A package for survival analysis in S version 2.38.

**Thilmony R, Underwood W, He SH**. 2006. Genome-wide transcriptional analysis of the *Arabidopsis thaliana* interaction with the plant pathogen *Pseudomonas syringae* pv. tomato DC3000 and the human pathogen *Escherichia coli*. *The Plant Journal* **46**, 43-53.

**Thompson GA, Goggin FL**. 2006. Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. *Journal of Experimental Botany* **57**, 755-766.

**Thompson GA, Van Bel AJE, eds.** 2013. *Phloem: Molecular Cell Biology, Systemic Communication, Biotic Interactions*. Oxford: Wiley-Blackwell.

**Thompson JN**. 2005. *The Geographic Mosaic of Coevolution*. Chicago: University of Chicago Press.

**Thorpe MR, Furch ACU, Minchin PEH, Foller J, Van Bel AJE, Hafke JB**. 2010. Rapid cooling triggers forisome dispersion just before phloem transport stops. *Plant, Cell and Environment* **33**, 259-271.

**Tjallingii WF**. 1978. Mechanoreceptors of the aphid labium. *Entomol. Exp. Appl.* **24**, 531-537.

**Tjallingii WF**. 1985. Electrical nature of recorded signals during stylet penetration by aphids. *Entomologia Experimentalis et Applicata* **38**, 177-186.

**Tjallingii WF**. 1988. Electrical recording of stylet penetration activities. In: Minks AK, Harrewijn P, eds. *Aphids, their biology, natural enemies and control.*, Vol. 2B. Amsterdam: Elsevier, 95-108.

**Tjallingii WF**. 1994. Sieve element acceptance by aphids. *European Journal of Entomology* **91**, 47-52.

**Tjallingii WF**. 1995. Regulation of phloem sap feeding by aphids. In: Chapman RF, de Boer G, eds. *Regulatory mechanisms in insect feeding*. New York: Chapman & Hall, 190-209.

**Tjallingii WF**. 2006. Salivary secretions by aphids interacting with proteins of phloem wound responses. *Journal of Experimental Botany* **57**, 739-745.

**Tjallingii WF, Hogen Esch T**. 1993. Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. *Physiological Entomology* **18**, 317-328.

**Törjék O, Meyer RC, Zehnsdorf M, Teltow M, Strompen G, Witucka-Wall H, Blacha A, Altmann T**. 2008. Construction and analysis of 2 reciprocal Arabidopsis introgression line populations. *Journal of Heredity* **99**, 396-406.

**Tosti N, Pasqualini S, Borgogni A, Ederli L, Falistocco E, Crispi S, Paolocci F**. 2006. Gene expression profiles of O3-treated Arabidopsis plants. *Plant, Cell and Environment* **29**, 1686-1702.

**Toufighi K, Brady SM, Austin R, Ly E, Provart NJ**. 2005. The botany array resource: e-northerns, expression angling, and promoter analyses. *The Plant Journal* **43**, 153-163.

**Tran LSP, Pal S**. 2014. *Phytohormones: A window to metabolism, signaling and biotechnological applications*. New York: Springer.

**Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L**. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* **31**, 46-54.

**Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L**. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562-578.

**Turgeon R**. 2010. The puzzle of phloem pressure. *Plant Physiology* **154**, 578-581.

**Ülker B, Mukhtar MS, Somssich IE.** 2007. The WRKY70 transcription factor of Arabidopsis influences both the plant senescence and defense signaling pathways. *Planta* **226**, 125-137.

**Van Bel AJE**. 2003. The phloem, a miracle of ingenuity. *Plant, Cell and Environment* **26**, 125-149.

**van Bel AJE, Furch ACU, Will T, Buxa SV, Musetti R, Hafke JB**. 2014. Spread the news: Systemic dissemination and local impact of Ca2+ signals along the phloem pathway. *Journal of Experimental Botany* **65**, 1761-1787.

**Van der Meijden E**. 1996. Plant defence, an evolutionary dilemma: Contrasting effects of (specialist and generalist) herbivores and natural enemies. *Entomologia Experimentalis et Applicata* **80**, 307-310.

**van Emden HF, Harrington R**. 2007. *Aphids as crop pests*. Oxfordshire: CABI.

**van Heerwaarden J, Hufford MB, Ross-Ibarra J**. 2012. Historical genomics of North American maize. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 12420-12425.

**van Heerwaarden J, Odong TL, van Eeuwijk FA**. 2013. Maximizing genetic differentiation in core collections by PCA-based clustering of molecular marker data. *Theoretical and Applied Genetics* **126**, 763-772.

**van Helden M, Tjallingii WF**. 1993. Tissue localisation of lettuce resistance to the aphid *Nasonovia ribisnigri* using electrical penetration graphs. *Entomologia Experimentalis et Applicata* **68**, 269-278.

**Van Poecke RMP**. 2007. *Arabidopsis*-insect interactions. In: Somerville CR, Meyerowitz EM, eds. *The Arabidopsis Book*. Rockville, MD: American Society of Plant Biologists, doi:10.1199/tab.0107, www.aspb.org/publications/arabidopsis/.

**Vandenborre G, Groten K, Smagghe G, Lannoo N, Baldwin IT, Van Damme EJM**. 2010. *Nicotiana tabacum* agglutinin is active against Lepidopteran pest insects. *Journal of Experimental Botany* **61**, 1003-1014.

**Veen BW**. 1985. Photosynthesis and assimilate transport in potato with top-roll disorder caused by the aphid *Macrosiphum euphorbiae*. *Annals of Applied Biology* **107**, 319-323.

**Verhage A, Vlaardingerbroek I, Raaijmakers C, Van Dam NM, Dicke M, Van Wees SCM, Pieterse CMJ**. 2011. Rewiring of the jasmonate signaling pathway in *Arabidopsis* during insect herbivory. *Frontiers in Plant Science* **2**, 1-12.

**Verheggen F, Haubruge E, De Moraes C, Mescher M**. 2013. Aphid responses to volatile cues from turnip plants (*Brassica rapa*) infested with phloem-feeding and chewing herbivores. *Arthropod-Plant Interactions* **7**, 567-577.

**Verma DPS, Hong Z**. 2001. Plant callose synthase complexes. *Plant Molecular Biology* **47**, 693-

701.

**Vermeer KMCA, Dicke M, de Jong PW**. 2011. The potential of a population genomics approach to analyse geographic mosaics of plant-insect coevolution *Evolutionary Ecology* **25**, 977-992.

**Vidhyasekaran P**. 2014. *PAMP Signals in Plant Innate Immunity: Signal Perception and Transduction*. Dordrecht: Springer Science+Business Media.

**Visscher PM**. 2008. Sizing up human height variation. *Nature Genetics* **40**, 489-490.

**Vlot AC, Dempsey DMA, Klessig DF**. 2009. Salicylic acid, a multifaceted hormone to combat disease. *Annual Review of Phytopathology* **47**, 177-206.

**Vos IA, Moritz L, Pieterse CMJ, Van Wees SCM**. 2015. Impact of hormonal crosstalk on resistance and fitness of plants under multi-attacker conditions. *Frontiers in Plant Science* **6**, 639.

**Wahid A, gelani S, Ashraf M, Foolad MR**. 2007. Heat tolerance in plants: An overview. *Environmental and Experimental Botany* **61**, 199-223.

**Walling LL**. 2008. Avoiding effective defenses: Strategies employed by phloem-feeding insects. *Plant Physiology* **146**, 859-866.

**Wang E, Wang R, DeParasis J, Loughrin JH, Gan S, Wagner GJ**. 2001. Suppression of a P450 hydroxylase gene in plant trichome glands enhances natural-product-based aphid resistance. *Nature Biotechnology* **19**, 371-374.

**Wang Y, Zhang Y, Wang Z, Zhang X, Yang S**. 2013. A missense mutation in CHS1, a TIR-NB protein, induces chilling sensitivity in Arabidopsis. *The Plant Journal* **75**, 553-565.

**War AR, Paulraj MG, Ahmad T, Buhroo AA, Hussain B, Ignacimuthu S, Sharma HC**. 2012. Mechanisms of plant defense against insect herbivores. *Plant Signaling & Behavior* **7**, 1306-1320.

**Warnes GR, Bolker B, Bonebakker L, Gentleman R, Liaw WHA, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B**. 2009. gplots: Various R programming tools for plotting data. *R package*.

**Wasternack C, Kombrink E**. 2010. Jasmonates: Structural requirements for lipid-derived signals active in plant stress responses and development. *ACS Chemical Biology* **5**, 63-77.

**Webster MT, Hurst LD**. 2012. Direct and indirect consequences of meiotic recombination: implications for genome evolution. *Trends in Genetics* **28**, 101-109.

**Weigel R, Bäuscher C, Pfitzner AP, Pfitzner U**. 2001. NIMIN-1, NIMIN-2 and NIMIN-3, members of a novel family of proteins from Arabidopsis that interact with NPR1/NIM1, a key regulator of systemic acquired resistance in plants. *Plant Molecular Biology* **46**, 143-160.

**Weigel RR, Pfitzner UM, Gatz C**. 2005. Interaction of NIMIN1 with NPR1 Modulates *PR* Gene Expression in Arabidopsis. *The Plant Cell* **17**, 1279–1291.

**Whiteman NK, Jander G**. 2010. Genome-Enabled Research on the Ecology of Plant-Insect Interactions. *Plant Physiology* **154**, 475-478.

**Whitham SA, Anderberg RJ, Chisholm ST, Carrington JC**. 2000. Arabidopsis *RTM2* gene is necessary for specific restriction of Tobacco Etch Virus and encodes an unusual small heat shock–like protein. *The Plant Cell* **12**, 569-582.

**Whitham SA, Yamamoto ML, Carrington JC**. 1999. Selectable viruses and altered susceptibility mutants in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 772-777.

**Wijnen CL, Keurentjes JJB**. 2014. Genetic resources for quantitative trait analysis: novelty and efficiency in design from an Arabidopsis perspective. *Current Opinion in Plant Biology* **18**, 103-109.

**Will T, Furch ACU, Zimmermann MR**. 2013. How phloem-feeding insects face the challenge of phloem-located defenses. *Frontiers in Plant Science* **4**, 1-12.

**Will T, Hewer A, Van Bel AJE**. 2008. A novel perfusion system shows that aphid feeding behaviour is altered by decrease of sieve-tube pressure. *Entomologia Experimentalis et Applicata* **127**, 237-245.

**Will T, Kornemann SR, Furch ACU, Tjallingii WF, van Bel AJE**. 2009. Aphid watery saliva counteracts sieve-tube occlusion: a universal phenomenon? *Journal of Experimental Biology* **212**, 3305-3312.

**Will T, Tjallingii WF, Thönnessen A, Van Bel AJE**. 2007. Molecular sabotage of plant defense by aphid saliva. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 10536-10541.

**Will T, van Bel AJE**. 2006. Physical and chemical interactions between aphids and plants. *Journal of Experimental Botany* **57**, 729-737.

**Wilson ACC, Sternberg LdSL, Hurley KB**. 2011. Aphids alter host-plant nitrogen isotope fractionation. *Proc. Natl Acad. Sci. USA* **108**, 10220-10224.

**Windt CW, Vergeldt FJ, De Jager PA, Van As H**. 2006. MRI of long-distance water transport: a comparison of the phloem and xylem flow characteristics and dynamics in poplar, castor bean, tomato and tobacco. *Plant, Cell and Environment* **29**, 1715-1729.

**Wink M**. 1988. Plant breeding: importance of plant secondary metabolites for protection against pathogens and herbivores. *Theoretical and Applied Genetics* **75**, 225-233.

**Wulff JA, White JA**. 2015. The endosymbiont *Arsenophonus* provides a general benefit to soybean aphid (Hemiptera: Aphididae) regardless of host plant resistance (*Rag*). *Environmental Entomology* **44**, 574-581.

**Wyatt IJ, White PF**. 1977. Simple estimation of intrinsic increase rates for aphids and tetranychid mites. *Journal of Applied Ecology* **14**, 757-766.

**Xu W, Campbell P, Vargheese AK, Braam J**. 1996. The Arabidopsis XET-related gene family: Environmental and hormonal regulation of expression. *Plant Journal* **9**, 879-889.

**Yadav S, Yadav PK, Yadav D, Yadav KDS**. 2009. Pectin lyase: A review. *Process Biochemistry* **44**, 1-10.

**Yi SY, Shirasu K, Moon JS, Lee S, Kwon S**. 2014. The activated SA and JA signaling pathways have an influence on flg22-triggered oxidative burst and callose deposition. **9**, e88951.

**Young ND**. 1996. QTL mapping and quantitative disease resistance in plants. *Annual Review of Phytopathology* **34**, 479–501.

**Yu A, Lepere G, Jay F, Wang JY, Bapaume L, Wang Y, Abraham AL, Penterman J, Fischer RL, Voinnet O, Navarro L**. 2013. Dynamics and biological relevance of DNA

demethylation in *Arabidopsis* antibacterial defense. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 2389-2394.

**Yu JM, Buckler ES**. 2006. Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology* **17**, 155-160.

**Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES**. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* **38**, 203-208.

**Zbierzak AM, Porfirova S, Griebel T, Melzer M, Parker JE, Dormann P**. 2013. A TIR-NBS protein encoded by Arabidopsis Chilling Sensitive 1 (CHS1) limits chloroplast damage and cell death at low temperature. *Plant Journal* **75**, 539-552.

**Zhang Z, Ersoz E, Lai CQ, Todhunter RT, Tiwari HK, Gore MA, Bradbury PJ, Yu J, Arnett DK, Ordovas JM, Buckler ES**. 2010. Mixed linear model approach adapted for genome-wide association studies. *Nature Genetics* **42**, 355-360.

**Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, Nordborg M**. 2007. An *Arabidopsis* example of association mapping in structured samples. *PLoS Genetics* **3**, e4.

**Zhao X, Wang YL, Qiao XR, Wang J, Wang LD, Xu CS, Zhang X**. 2013. Phototropins Function in High-Intensity-Blue-Light-Induced Hypocotyl Phototropism in Arabidopsis by Altering Cytosolic Calcium *Plant Physiology* **162**, 1539-1551.

**Zhou J, Spallek, T., Faulkner, C., Robatzek, S.** 2012. CalloseMeasurer: a novel software solution to measure callose deposition and recognise spreading callose patterns. *Plant Methods* **8**, 49.

**Zhou R, Jackson L, Shadle G, Nakashima J, Temple S, Chen F, Dixon RA**. 2010. Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 17803-17808.

**Zhou X, Jiang Y, Yu D**. 2011. WRKY22 transcription factor mediates dark-induced leaf senescence in Arabidopsis. *Molecules and Cells* **31**, 303-313.

**Zhou X, Stephens M**. 2014. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods* **11**, 407-409.

**Zhu C, Gore M, Buckler ES, Yu J**. 2008. Status and prospects of association mapping in plants. *The Plant Genome* **1**, 5-20.

**Zhu XF, Feng T, Tayo BO, Liang JJ, Young JH, Franceschini N, Smith JA, Yanek LR, Sun YV, Edwards TL, Chen W, Nalls M, Fox E, Sale M, Bottinger E, Rotimi C, Liu YM, McKnight B, Liu K, Arnett DK, Chakravati A, Cooper RS, Redline S, Consortium CB**. 2015. Meta-analysis of correlated traits via summary statistics from GWASs with an application in hypertension. *American Journal of Human Genetics* **96**, 21-36.

**Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W**. 2004. GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiology* **136**, 2621-2632.

**Zipfel C, Robatzek S, Navarro L, Oakeley EJ, Jones JD, Felix G, Boller T**. 2004. Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature* **428**, 764-767.

# Summary

Aphids are pest insects in a wide variety of crops. They penetrate the host plant with their piercing-sucking mouthparts and feed from the sugar-rich phloem sap. Apart from the removal of nutrients from the host, in many instances they also transmit plant viruses. To control aphid populations and virus transmission in agriculture, insecticides are applied on a large scale with often negative effects on biodiversity. Improvement of crop resistance to aphids would be a more sustainable pest management strategy. Plants harbour many natural compounds, proteins and morphological structures that protect them against herbivorous insects. By exploiting these natural resistance mechanisms in crops, the incidence of virus transmission and pest outbreaks, as well as the application frequency of insecticides, could potentially be reduced. The genetic and physiological mechanisms underlying resistance mechanisms to aphids are, however, largely elusive, and the identification of resistant genotypes is costly in terms of time, space and labour.

In the last years, genome-wide association (GWA) mapping has evolved as a new statistical method to study the genetic architecture of plant traits. The concept of GWA mapping is to use genetically diverse populations to evaluate the association between genetic polymorphisms and a phenotype. A phenotype can, for example, consist of plant resistance to aphids, expressed as aphid population size. A significant association between aphid population size and a single nucleotide polymorphism (SNP) is an indication that the genes with, or in linkage with, the SNP are affecting aphid population development. Despite its potential, GWA mapping has so far hardly been used to find genes involved in plant resistance to herbivorous insects. In a literature review in **Chapter 2**, two major challenges of GWA mapping in plant-insect studies are discussed: (1) phenotyping large plant populations, and (2) reducing trait complexity. Robust assessment of plant resistance to insects requires highly controlled conditions and time-consuming measurements of multiple replicates per plant line. Phenotyping hundreds of plant lines, which is required for GWA mapping, is therefore a bottleneck. Secondly, plant resistance to herbivorous insects is often a quantitative trait, depending on many genes that each have a minor effect on the phenotype. As a consequence, GWA mapping has a low chance of detecting causal loci. A solution for both problems would be to improve the efficiency and the specificity of phenotyping. Instead of assessing the increase in aphid population size - which in principle is the result of the integration of multiple plant resistance mechanisms over time - assessing different parameters of aphid behaviour could result in a higher chance to discover genes, since it disentangles plant resistance into specific aspects, each controlled by a smaller subset of genes.

One of the objectives of this thesis, therefore, was to develop a high-throughput phenotyping method to efficiently measure different aspects of plant resistance to aphids (**Chapter 3**). With a video-tracking platform the behaviour of aphids was recorded on multiple leaf discs simultaneously. Comparison between automated recording and manual observations showed that non-moving aphids were generally probing the plant tissue with their piercing-sucking mouthparts. By utilising this correlation, several elements of aphid feeding behaviour could be assessed with automated video-tracking software, such as the total time spent on short probes (< 3 min) related to penetration of the epidermis and mesophyll, and the time aphids spent on long probes (>15 min), which are putatively related to phloem ingestion. Benchmarking to the established electrical penetration graph (EPG) technique with two plant and aphid species, showed that video tracking was more efficient in identifying quantitative differences in plant resistance, although less accurate than EPG recording.

The video-tracking platform was subsequently used to screen feeding behaviour of *Myzus persicae*, the green peach aphid, on a population of 350 natural Arabidopsis accessions which had been genotyped for 214,000 SNPs. In addition, *M. persicae* population development was assessed on the same Arabidopsis accessions. In **Chapter 4** these data sets were integrated in a multi-trait GWA mapping analysis of 30 different abiotic and biotic stresses, including salinity, drought, heat, the necrotrophic fungus *Botrytis cinerea*, the generalist thrips *Frankliniella occidentalis*, the specialist caterpilar *Pieris rapae*, and some combinations of these stresses. Interestingly, the genome-wide associations with plant resistance to aphids were positively correlated with associations involving resistance to the parasitic plant *Phelipanche ramosa*. A common genetic architecture of these stress responses could be explained by the fact that both species retrieve resources from the host vascular bundle and are known to induce the phytohormone salicylic acid (SA). In addition, aphids and parasitic plants showed a negative genetic correlation with stresses that are known to induce the antagonistic phytohormone jasmonic acid (JA), such as *B. cinerea* and *P. rapae*. By combining all 30 traits into one multivariate GWA mapping analysis, several candidate genes were identified which might play a role in multiple plant stress responses. Furthermore, SNPs were found with common or antagonistic effects between biotic and abiotic, and below- and aboveground stresses.

In **Chapter 5** the identification of the gene *PHLOEM HEAT-RESPONSE LOCUS* (*PHLO*) is described. GWA mapping of the time *M. persicae* aphids spent on short probes (< 3 min), revealed a strong association with a SNP in *PHLO*, a small heat-shock-like gene with unknown function. *PHLO*'s closest homologue is *RESTRICTED TOBACCO ETCH POTYVIRUS (TEV) MOVEMENT 2* (*RTM2*). *RTM2* restricts virus transport through the phloem via an as yet unknown mechanism. EPG recording on *phlo* knockout mutants and near-isogenic lines revealed several effects on aphid feeding behaviour. Most importantly, PHLO reduced the duration of phloem ingestion and increased the duration of salivation in the phloem. The effects were most pronounced when plants were grown under moderate heat stress (26˚C). At this temperature *PHLO* expression was upregulated. PHLO also affected plant performance. When exposed to moderate heat stress, knockout plants had a stunted growth of the inflorescence and produced seeds with reduced seed dormancy compared to the wild type. Natural Arabidopsis accessions with the resistant haplotype occurred significantly more often in geographical regions with high temperature fluctuations. Overall, these findings suggest that PHLO is required for plant resistance to both aphids and temperature stress. Further characterisation of *phlo* knockout mutants indicated that PHLO increased the phloem flow rate during moderate heat stress via a callose-independent mechanism. Based on the predicted protein structure, co-expression network, and its absence in Poaceae plants which lack phloem proteins, it is postulated that PHLO is a protein clamp. With its membrane anchor and chaperone-like domain PHLO could possibly attach phloem proteins to the lateral side of sieve elements. Membrane-attached protein aggregations would encapsulate the stylet tip and obstruct the aphid's food channel. In addition, protein recruitment to the sieve element margins would reduce proteinaceous obstacles in the lumen and increase phloem flow, particularly during heat stress when water availability and hydrostatic pressure is reduced.

A second gene identified with GWA mapping of aphid behaviour was *WRKY22* (**Chapter 6**). WRKY22 is a transcription factor known to be involved in pathogen-triggered immunity. EPG recording and aphid population assays on knockout mutants and overexpression lines showed that WRKY22 increased plant susceptibility to *M. persicae* aphids via a mesophyll-located mechanism. mRNA sequencing analysis of the knockout transcriptome revealed that upon aphid infestation, WRKY22 suppressed genes involved in salicylic acid (SA) signalling and induced genes involved in cell wall loosening. In addition, WRKY22 inhibited JA responses to mechanostimulation by empty clip cages. Based on these and previous findings, it is proposed that WRKY22

integrates SA and JA signalling in response to a wide variety of biotic and abiotic plant stresses. By inducing *WRKY22* expression, aphids could potentially manipulate host plant defences for their own benefit.

In the general discussion of this thesis (**Chapter 7**) the parasite-like nature of aphids is addressed. Unlike many other insects, aphids engage in an intimate relation with their host plants for several generations. Consequently, aphids have a high interest in optimising host plant vitality and nutrient availability. While most aphid-plant studies have focused on plant defence mechanisms, it would be worthwhile to further investigate if and how aphids manipulate host plant physiology. Finally, the findings of this thesis are integrated to show that, for gene discovery, it is indeed valuable to disentangle complex traits into specific phenotypic aspects.

# Acknowledgements

I still remember the phone call from Marcel Dicke in the early spring of 2011. "Let's make four nice years of it." Well, I think we managed that. Now, almost five years later I am wrapping up the last ends of my thesis and can look back on a great period. I loved every single bit of it! Although I vaguely remember some rainy afternoons counting hundreds of aphids (why are they reproducing so fast? See Chapter 1 and 7), failing subject detections in the video-tracking software, non-cooperative honeydew clocks and incidences of electricity breakdown after having gently attached my aphids to a golden wire. Anyhow, apparently it all was necessary to bring me here. The title of my thesis, 'Mapping *moves* on Arabidopsis', is not only about the movements of aphids, or trials of genome-wide association mapping. No, it is also about the whole journey that has moved me. And not the least because of all the great people whom I have met and with whom I have shared highs and lows. Without your help and presence it would not have been complete! It is impossible to capture it all, but I would like to address some people that played a special role in the past years.

First of all, I would like to thank you, Gerrie. Working with you was a joy. We started off with sowing this huge Hapmap collection of 350 Arabidopsis accessions. It was like a big sudoku. Since we had this wonderful idea of sowing the population in a split-plot design, we had to search through all the boxes and labels and open and close each tube with seeds numerous times. I think you cursed these tiny Arabidopsis seeds more than once, but they grew like weed, so apparently it helped! Also our own aphid rearing prospered during all these years. Everything in your lab was well organised and I could always bother you with questions. I still owe you a prize for winning the counting aphids bingo. Jacqueline, you joined the Ethogenomics project during the third year of my PhD. You guided me in the molecular lab and introduced me in the world of RNA isolations, RT-qPCR and Agrobacterium transformations. This was both new and exciting for me and I am grateful for your patience and the confidence you gave me. You always kept an eye on me and were there to answer my ignorant questions, which I soon was not afraid anymore to ask. Because of you I also learned to appreciate 'het foute uurtje' on the radio and I enjoyed our cosy coffee breaks. Manus, you were my partner in crime (read: science). Together we started as PhD candidates in the Ethogenomics project, you with those tiny, fast thrips and me with my fat, slow aphids. We collaborated in the development of the video-tracking platform (sometimes sitting together in the dark, 2-$m^2$-sized video lab), the GWA mapping, and the writing of the literature review and the consortium manuscript. Going through similar challenges, I think you were the colleague standing closest to my trial and errors and decision making. Thank you for sharing insights and thinking along with me. Maarten, you were a great supervisor. You were always stimulating me with critical comments and new questions

during our weekly meetings. Although tracking insect behaviour was a new field for you as a molecular biologist, you gave many original and practical solutions for the problems we encountered. I am proud that our video-tracking platform is now further optimised for industrial application. Your advice played a crucial role in the validation and characterisation of candidate genes. Thank you for teaching me so much! Marcel, first of all, thank you for giving me the opportunity to do a PhD in your lab. You gave me a lot of freedom and in the same time secured that everything was going according to plan. I could always count on you when necessary and I admire your commitment and sharp comments on my manuscripts, no matter how full your agenda was or the time of the day or week. Harro, I must admit that I left Entomology to join your lab with pain in my heart. But it did not take long to feel at home. Thank you for introducing me into the Lab of Plant Physiology, I learned a tremendous lot during the last 2.5 years! In addition, thank you for your moral support, your improvements on my chapters, and our sportive activities during lab outings. The Noldus team: Lucas, Olga, Andrew, Wil and Albert, thank you for our collaboration. It is amazing to have software that can track tiny, semi-transparent creatures, such as aphids. Olga, thank you for taking the time for me as my external supervisor. The Royal van Zanten team: Sjoukje and Bas, thank you for working together on aphid-lily interactions, I learned a lot about the practical problems and solutions for aphid pests and virus transmission. Tupola, thank you for constructing the hardware of the video platform and honeydew clocks. STW, and specifically Henry and Quirine, thank you for making everything possible. It was a prerequisite to be part of the Learning from Nature (LfN) programme and to share experiences with so many other people who were engaged with GWA mapping on the same Arabidopsis population. Willem, thank you for your relentless statistical and computational support, the LfN programme could not have done without you. Douwe, Ingrid and Ria, the EPS graduate school offered an impressive amount of valuable courses and seminars. Lia, I have learned a lot from you about statistical models and R scripts and enjoyed our collaboration a lot. Nelson, you were (and still are!) updating me with the latest publications of interest, and I could always find you for GWA- or RNA-seq-related stuff, or for an aphid invasion of your climate chamber. Thank you! Silvia, Pingping and Joost, thanks for being part of the LfN consortium dream team. Freddy, thank you for advising me on the use of honeydew clocks; they have revealed an exciting part of the puzzle. Paul Piron, thanks for lending me your handy aphid cages. Johan, thank you for providing hundreds of clip cages. Jan, you helped me to find my way through the bash command language and RNA-seq analysis. Anne, thank you for our nice collaboration in the potato cell wall project. Marianne, Marga and Mabel, it was a pleasure to supervise you during your master theses and thank you for bringing new insights. All the people at the Lab of Entomology: thank you for your companionship

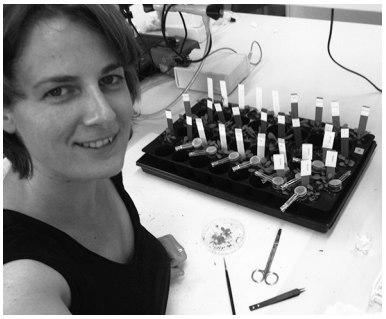
during the coffee breaks, lab outings, student practicals, and Ento borrels and parties! It was an honour to have been a member of the Ento party committee and a lot of fun to organise the multiculti party. Léon, Jeroen, André, Frans and Joop, I could always count on you for the insect rearing and climate chamber facilities. Angelique, thank you for taking care of all the finances and stuff I didn't have to bother about. Patrick, thanks for placing the numerous orders for primers and kits. A special thank you goes to the Ento PhD trip committee: Jeltje, Anneke and Camille. I had a great time visiting the labs in Neuchatel, Lausanne and Basel, sleeping in hostels and forsaken soccer club houses. Hans, Emma and Jitte, thank you for having me at the brains and behaviour lab for my EPG experiments. I hope you will have less electricity breakdowns and escaped aphids in your plants in the future. Jenny, I would not have wanted to miss organising the 'R users meetings' together with you, thank you for your spontaneous and disarming character. Eddie, thank you for taking over the R group organisation in the last years. Bruce, your garden cookery is phenomenal. Cindy and Marjolein, it is wonderful that I got to know you both. Cindy, you were my support for all EPG- and aphid-related things and Marjolein, you were a fantastic co-organiser on several trips and occasions. Thank you for our friendship! I am proud to have you as my paranymphs and I am counting on many more dinners until closing time. Then my 'home' for the last years: the Lab of Plant Physiology. A group with an enormous diversity in research topics and a nice blend of nationalities; some a bit more represented than others (I enjoyed getting more acquainted with Chinese culture and when visiting the Seedlab, I was more than once completely surrounded by Brazilian music and Portuguese-speaking colleagues). I am grateful for the progress meetings with the PPH staff, thank you for your ideas and suggestions that helped me in making the right decisions. Mariëlle, Francel, Diaan, Leo and Juriaan, thank you for having me around in the molecular labs. Gonda, thanks for germinating my seeds and the overload of comments on my manuscript (I will not take it personally). All the people of the Seedlab, strigolactone group, terpenoid group and plant metabolism and transport group and everyone else, thank you for bringing your great atmosphere and personalities to the coffee table and lunch breaks every day. Rina and the activity committee, you have organised superb BBQs, bowling nights, lab outings and christmas dinners. Maria Cecilia, Bing, Esmer, Melissa, and Bea, you organised an unforgettable PhD trip! Playing werewolf each night was a lot of fun (Bing, you are a fantastic storyteller). Wilco and Esmer, you are the best co-drivers and German karaoke singers. Natalia and Elise thanks for organising the PhD-trip quiz together. Bea, Natalia, Maria Cecilia, muchas gracias por todas comidas Españolas muy interessante y divertido! All the members of the 'Glu glu swimming' group, including

the ones that found 7:30 a.m. too early, I enjoyed swimming together! Bea, thank you for your warm personality and always making me smile. Giovanni, let me know when you start your Melandri lab. Alexandre, I will miss your goedemorgen. Bas, thanks for removing seed coats and your presence. Emilie, thank you for your personality, our discussions, your thourough advice about candidate gene validation and your help with the layout of this thesis. You can hire me again for a movie production of the Steam Sisters.

I would not want to end without some special words for my family. Pap en mam, dankjulliewel dat jullie er zijn. Jullie hebben mij altijd gestimuleerd om het beste uit mezelf te halen en hebben mij de ruimte gegeven om dat op mijn eigen manier te doen. Bij jullie logeren in het mooie Zuid-Limburg is heerlijk en doet me alles even vergeten. Judith & John, Esther & Bram en mijn allerliefste neefjes en nichtje, het was misschien niet zo vaak als ik had gewild, maar ik vond het superfijn om bij jullie te zijn en samen met jullie naar Naturalis of de dierentuin te gaan. Dankjulliewel voor jullie interesse en steun. Mama Bekhuis en de families Kuipers, Burema en Bekhuis, bedankt voor jullie enorme gastvrijheid en gezelligheid in de afgelopen jaren. Last but not least, Mariëlle. Het had niet tot stand kunnen komen zonder jou. Je was er altijd en voelde aan als ik druk, gespannen of opgetogen was. Ik mocht uren tegen je aanpraten over luizen en planten. Je kon mij weer met beide benen op de grond zetten of een schop onder mijn kont geven als ik die nodig had. Dankjewel voor het samen terug naar huis fietsen, het wachten op mij als ik laat thuis was, het toelaten van mijn bladluiskweek op de vensterbank in het weekend, de fijne dagen thuis en in de tuin, het samen aangaan van de avonturen, en je liefde.

# About the author

## Curriculum vitae



Karen Kloth was born on 21 July 1980 in Delft, the Netherlands. She obtained her bachelor and master degrees in biology at the Wageningen University. At the Vrije Universiteit, Amsterdam, she did a minor in communicating science and at the North West University, Potchefstroom, South Africa, she studied host-plant preference of maize stem borers. After obtaining her master degree, she worked at a Dutch governmental agency of the Ministry of Economic Affairs and was information analyst at the CITES Scientific Authority, the advisory board for the regulation of trade in endangered plant and animal species, located at the national natural history museum, Naturalis. She, however, remained being drawn to practicing science herself. After two years at Naturalis, she initiated an internship in molecular phylogenetics at the Biosystematics group in Wageningen. On 1 May 2011, she started as a PhD candidate in genetical genomics at the Laboratories of Entomology and Plant Physiology in Wageningen.

## Publications

**Kloth KJ*, Thoen MPM*, Bouwmeester HJ, Jongsma MA, Dicke M**. 2012. Association mapping of plant resistance to insects. *Trends in Plant Science* **17**, 311-319.

**Booij MW, Kloth KJ, Jongsma MA, Dicke M, Hemerik L** 2013. Analysing aphid behaviour with time-to-event techniques to discriminate between susceptible and resistant plants. *Proc. Neth. Entomol. Soc. Meet.* 24, 9-16.

**Kloth KJ, ten Broeke CJM, Thoen MPM, Hanhart-van den Brink M, Wiegers GL, Krips OE, Noldus LPJJ, Dicke M, Jongsma MA**. 2015. High-throughput phenotyping of plant resistance to aphids by automated video tracking. *Plant Methods* **11**, 4.

**Thoen MPM, Kloth KJ, Wiegers GL, Krips OE, Noldus LPJJ, Dicke M, Jongsma MA.** 2016. Automated video tracking of thrips behavior to assess host-plant resistance in multiple parallel two-choice setups. *Plant Methods* **12**, 1.

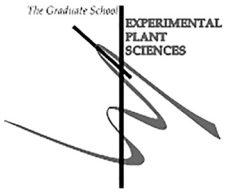
**Thoen MPM*, Davila Olivas NH*, Kloth KJ*, Coolen S*, Huang P-P*, Aarts MGM, Bac-Molenaar JA, Bakker J, Bouwmeester HJ, Broekgaarden C, Bucher J, Busscher-Lange J, Cheng X, Fradin EF, Jongsma MA, Julkowska MM, Keurentjes JJB, Ligterink W, Pieterse CMJ, Ruyter-Spira C, Smant G, Testerink C, Usadel B, van Loon JJA, van Pelt JA, van Schaik CC, van Wees SCM, Visser RGF, Voorrips R, Vosman B, Vreugdenhil D, Warmerdam S, Wiegers GL, van Heerwaarden J, Kruijer W, van Eeuwijk FA, Dicke M.** *under review*. Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping.

**Kloth KJ, Wiegers GL, Busscher-Lange J, van Haarst JC, Kruijer W, Bouwmeester HJ, Dicke M, Jongsma MA.** *under review.* AtWRKY22 promotes susceptibility to aphids and modulates salicylic acid and jasmonic acid signalling.

**Kloth KJ, Busscher-Lange J, Wiegers GL, Kruijer W, Buijs G, Meyer RC, Bouwmeester HJ, Dicke M, Jongsma MA.** *in prep. PHLO*: A gene involved in restricting phloem ingestion by *Myzus persicae* aphids and phloem responses to moderate heat stress.

* authors contributed equally

**Education Statement of the Graduate School**
**Experimental Plant Sciences**




**Issued to:** **Karen Kloth**
**Date:** **11 March 2016**
**Group:** **Laboratories of Entomology & Plant Physiology**
**University:** **Wageningen University & Research Centre**

| **1) Start-up phase** | *date* |
|---|---|
| ► **First presentation of your project** | |
| Presentation kick-off ethogenomics project | Jun 01, 2011 |
| ► **Writing or rewriting a project proposal** | |
| Hotel Grant ZonMw "Gene expression related to plant resistance to aphids" | Oct 01, 2013 |
| Postdoc proposal "Defence at the frontdoor" | Sep 01, 2015 |
| **Writing a review or book chapter** | |
| "Association mapping of plant resistance to insects", Trends in Plant Science, May 2012, Vol.17, no. 5, pp 311-319. | May 01, 2012 |
| ► **MSc courses** | |
| ► **Laboratory use of isotopes** | |
| *Subtotal Start-up Phase* | *10,2 credits** |

| **2) Scientific Exposure** | *date* |
|---|---|
| ► **EPS PhD student days** | |
| EPS PhD student day, Wageningen | May 20, 2011 |
| EPS PhD student day, Wageningen | Nov 30, 2012 |
| EPS PhD student day, Leiden | Nov 29, 2013 |
| ► **EPS theme symposia** | |
| EPS Theme 2 'Interactions between plants and biotic agents', Wageningen | Feb 10, 2012 |
| EPS Theme 2 'Interactions between plants and biotic agents', Utrecht | Feb 20, 2015 |
| ► **NWO Lunteren days and other National Platforms** | |
| NWO-ALW meeting 'Experimental Plant Sciences', Lunteren (NL) | Apr 04, 2011 |
| ALW 'Netherlands Annual Ecology Meeting', Lunteren (NL) | Feb 07-08, 2012 |
| NWO-ALW meeting 'Experimental Plant Sciences', Lunteren (NL) | Apr 02-03, 2012 |
| NWO-ALW meeting 'Experimental Plant Sciences', Lunteren (NL) | Apr 22-23, 2013 |
| NWO-ALW meeting 'Experimental Plant Sciences', Lunteren (NL) | Apr 14-15, 2014 |
| NWO-ALW meeting 'Experimental Plant Sciences', Lunteren (NL) | Apr 13-14, 2015 |

| | | |
|---|---|---|
| ► | **Seminars (series), workshops and symposia** | |
| | Wageningen Evolution and Ecology Seminars (WEES), Dr. Judith Becerra (Univ. of Arizona), Wageningen | May 26, 2011 |
| | Yearly Entomological Laboratory Research Exchange Meeting (YELREM), Renkum | Jun 07, 2011 |
| | National Ecogenomics Day, Amsterdam | Jun 16, 2011 |
| | EPS ExPectationS Day, Wageningen | Nov 18, 2011 |
| | 6th workshop Plant-Insect Interactions, Amsterdam | Nov 23, 2011 |
| | Mini-Symposium Plant Breeding in the Genomics Era, Wageningen | Nov 25, 2011 |
| | EPS Flying seminar, Dr. Jennifer McElwain (Univ. College Dublin), Wageningen | Jan 20, 2012 |
| | SPICY workshop Bioinformatics, statistical genetics and genomics, Wageningen | Mar 07-08, 2012 |
| | Yearly Entomological Laboratory Research Exchange Meeting (YELREM), Renkum | May 30, 2012 |
| | Farewel symposium Maarten Koornneef | Apr 11, 2013 |
| | Yearly Entomological Laboratory Research Exchange Meeting (YELREM) Renkum | May 17, 2013 |
| | 8th workshop Plant-Insect Interactions, Wageningen | Sep 14, 2013 |
| | 6th National Ecogenomics Day, Utrecht | Jun 24, 2014 |
| | Wageningen PhD Symposiums, Wageningen | May 06, 2015 |
| ► | **Seminar plus** | |
| | Phd meeting with Detlef Weigel | Feb 27, 2013 |
| | WEES masterclass with Eric Schranz | Nov 21, 2013 |
| ► | **International symposia and congresses** | |
| | Symposium on Insect-Plant Interactions (SIP14), Wageningen | Aug 13-18, 2011 |
| | Phenodays 2011, Wageningen | Oct 12 &14, 2011 |
| | Measuring Behavior 8th Int. Conference on Methods and techniques in behavioral research, Utrecht | Aug 28-31, 2012 |
| | EUROEEFG Frontiers in Ecological and Evolutionary Genomics, Noordwijkerhout | May 26-29, 2013 |
| | International Symposium Insect-Plant Relationships (SIP15), Neuchatel | Aug 17-22, 2014 |
| ► | **Presentations** | |
| | SIP14 & Phenodays 2011 (Poster) | Oct 12 &14, 2011 |
| | Summerschool Natural Variation of Plants & Measuring Behavior conference 2012 (Poster) | Aug 28-31, 2012 |
| | Noldus EthoVision seminar: “Phenotyping plant resistance to insects” (Talk) | Apr 11, 2013 |
| | ALW meeting ‘Experimental Plant Sciences’: “Arabidopsis genes affecting aphid behaviour” (Talk) | 23-04-2013 |
| | PRI Biointeractions: “Automated video tracking of aphid feeding behaviour” (Talk) | 16-05-2013 |
| | EUROEEFG conference (Poster) | 26/29-05-2013 |
| | Insect-Plant Interaction workshop, Wageningen: “Arabidopsis genes affecting aphid behaviour” (Talk) | Sep 2, 2013 |
| | PRI business unit BioScience: “Genes affecting aphid behaviour - Arabidopsis’ next top candidate” (Talk) | Oct 16, 2013 |
| | Univ. of Lausanne, Dep. Of Plant Molecular Biology: “Arabidopsis genes affecting aphid behaviour” (Talk) | Oct 30, 2013 |

| | | |
|---|---|---|
| | PRI business unit BioScience: "Ethogenomics of plant resistance to *Myzus persicae"* (Talk) | May 07, 2014 |
| | Int. Symposium Insect-Plant Interactions (SIP15), Neuchatel: "High-throughput phenotyping of plant resistance to aphids by automated video tracking" (Talk) | Aug 19, 2014 |
| | MSc course Introduction to Statistics for the Life and Behavioral Sciences (Univ. of Leiden): "Automated video tracking of aphid behaviour" (Talk) | Nov 21, 2014 |
| | PRI business unit BioScience: "Misbehaving aphids: Hormone hijacking and droplet excretion" (Talk) | Mar 06, 2015 |
| | ALW meeting 'Experimental Plant Sciences': "Misbehaving aphids: Hormone hijacking and droplet excretion" (Talk) | Apr 14, 2015 |
| | Rijkszwaan: "Automated video tracking of aphid behaviour" (Talk) | Apr 22, 2015 |
| | Max Planck Institute, Cologne: "Arabidopsis genes affecting aphid behaviour" (Talk) | Apr 23, 2015 |
| | Movie display automated video tracking of aphids in Museum Boerhaave, Leiden (Talk) | May 2015 |
| | Insect-Plant Interaction workshop, Amsterdam: "PHLO: a gene involved in phloem responses to heat stress and restriction of phloem ingestion by aphids" (Talk) | Nov 19, 2015 |
| ► | **IAB interview** | |
| | Interview with a member of the Internional Advisory Board of EPS | Sep 29-09, 2014 |
| ► | **Excursions** | |
| | PhD trip Lab. of Entomology to Univ. of Neuchatel, Univ. of Lausanne, Univ. of Basel | Oct 27-Nov 02, 2013 |
| | PhD trip Lab of Plant Physiology to Max Planck Cologne and Tubingen, Univ. of Tubingen, ETH, Univ. of Zurich, Syngenta Stein, RijksZwaan Fijnaart | Apr 22-May 01, 2015 |
| | *Subtotal Scientific Exposure* | *36.3 credits** |

| **3) In-Depth Studies** | | *date* |
|---|---|---|
| ► | **EPS courses or other PhD courses** | |
| | PhD Summer School Environmental Signaling (6th), Utrecht | Aug 22-24, 2011 |
| | Bioinformatics, a user's approach | Aug 29-Sep 02, 2011 |
| | Introduction to R for Statistical Analysis (Learning from Nature Programme) | Oct 24-25, 2011 |
| | Association Mapping (Learning from Nature programme) | Feb 23, 2012 |
| | Mixed Linear Models (post graduate course) | Jun 21-22, 2012 |
| | Summer School Natural Variation in Plants | Aug 21-24, 2012 |
| | Survival Analysis in Biology | Jan 24-25, 2013 |
| | RNAseq analysis | Dec 16-18, 2013 |
| ► | **Journal club** | |
| | PhD lunch meetings | May 2011-2015 |
| | IPI lunch meetings | May 2011-2015 |
| | Literature discussion Lab. of Entomology, and Lab. of Plant Physiology | May 2011-2015 |
| ► | **Individual research training** | |
| | Training The Observer and Ethovision video tracking software | Jun 20 & 22, 2011 |
| | *Subtotal In-Depth Studies* | *10.2 credits** |

| 4) Personal development | *date* |
|---|---|
| ► **Skill training courses** | |
| Competence Assessment | Mar 20, 2012 |
| Presentation skills | Apr 03, 2013 |
| PhD workshop carousel | Jun 02, 2014 |
| Last stretch PhD | Dec 19, 2014 |
| ► **Organisation of PhD students day, course or conference** | |
| Organiser of 'Writing consortium paper with Learning from Nature programme' | May 01, 2013/2014 |
| Organiser of the monthly R Users Meetings, Wageningen | Aug 01, 2013/2015 |
| ► **Membership of Board, Committee or PhD council** | |
| *Subtotal Personal Development* | *3.6 credits** |

| **TOTAL NUMBER OF CREDIT POINTS*** | **60.3** |
|---|---|

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits

** A credit represents a normative study load of 28 hours of study.*

# Learning from Nature programme

## Learning from Nature to protect crops

This thesis is part of the STW Learning from Nature programme.

### Background of the programme

Plants are under the constant threat of biotic and abiotic stresses. Yet, devastating pests and diseases only rarely occur in nature and plants have managed to sustain for millions of years in this hostile environment. This is due to and has resulted in a tremendous degree of natural variation in mechanisms that plants exploit to defend themselves against pathogens and insects and to deal with abiotic stresses. In agriculture, however, we have exploited only very little of this diversity of defenses and as a consequence environment-malignant pesticides remain a dominant method to control pests and diseases. The current threat of climatic changes and limiting resources for agriculture (water, fertilizer) require improved resistance to abiotic stresses.

### Ambition and goal

With this multidisciplinary and innovative STW programme we want to mine the natural reservoir of plant defence mechanisms. This will be done by using state-of-the-art high-throughput technologies to explore the natural potential and exploit mechanisms, genes and markers to develop novel resistance mechanisms against biotic and abiotic stresses for plant breeding. In nature plants have co-evolved with a large variety of attackers. Therefore, wild species, such as *Arabidopsis thaliana*, harbour a fantastic reservoir of natural adaptive mechanisms to respond to (a)biotic stresses that to date have not been systematically explored. In the past decade, Arabidopsis has been adopted world-wide as the ideal model for plant science and an impressive molecular genetic toolbox has since been developed (e.g. the full genome sequence, the availability of well-characterized Arabidopsis populations, full-genome microarrays and metabolomics protocols). Hence, exploring natural variation in the defense responses of Arabidopsis to a large variety of (a)biotic stresses will yield important new insights into how plants selectively adapt to stresses, and provide novel concepts for sustainable agriculture and resistance breeding.

## Objectives

1. To explore natural variation in resistance to abiotic and biotic stresses in Arabidopsis populations through an integrated multidisciplinary approach

2. To identify mechanisms underlying natural resistance to abiotic and biotic stresses in Arabidopsis

3. To develop methods to analyze complex datasets on different types of resistance

4. To exploit information gained on natural variation in Arabidopsis to identify molecular markers that can assist in breeding for resistance to abiotic and biotic stresses in crop plants.

## Focus and results at the end of the programme

To this end Arabidopsis ecotype and RIL populations can be exploited to analyze the degree of resistance to a diversity of microbial pathogens, herbivorous insects and abiotic stresses and their interaction. Using large-scale bioinformatics this information can be integrated with transcriptomics and metabolomics, to select genotypes and lines that can be used for in-depth analysis of the resistance mechanisms. The information gained from this comprehensive approach will lead to the identification of genes and molecular markers for different resistance mechanisms. These mechanisms will be characterized at the molecular, biochemical and physiological level and can subsequently be used to screen large numbers of lines of various crop species for orthologous genes involved in similar resistance mechanisms.

## Innovation

Never before has the natural variation in plant defenses against different biotic and abiotic stresses and their interaction been investigated in such a comprehensive, multidisciplinary programme. To date, solutions to individual (a)biotic stresses have been sought. However, this has not resulted in a systems approach that results in durable solutions for a range of stresses.

This research was performed at the Laboratories of Entomology and Plant Physiology, Wageningen University, and was supported by the Dutch Technology Foundation STW, which is part of the Netherlands Organisation for Scientific Research (NWO) and partly funded by the Ministry of Economic Affairs, Agriculture and Innovation through the Technology Foundation Perspective Programme 'Learning from Nature' (project number STW10989).

**Cover design by** Karen J. Kloth

**Cover illustration by** Karen J. Kloth

**Thesis layout by** Karen J. Kloth & Emilie F. Fradin

**Printed by** GVO drukkers en vormgevers B.V., Ede