From the cell surface to the field: Receptor-mediated herbivore induced defenses in legume crops.

Natalia Guayazán Palacios

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Adam D. Steinbrenner, Chair

Elizabeth Van Volkenburgh

Takato Imaizumi

Program Authorized to Offer Degree:

Biology

University of Washington

Abstract

From the cell surface to the field: Receptor-mediated herbivore induced defenses in legume crops

Natalia Guayazán Palacios

Chair of the Supervisory Committee:

Adam D. Steinbrenner

Department of Biology

The recognition of herbivore-associated molecular patterns (HAMPs) via pattern recognition receptors (PRRs) initiates signaling cascades that lead to the accumulation of direct and indirect defenses against herbivorous pests. Inceptin 11 (In11) is a peptide HAMP present in the oral secretions of Lepidoptera caterpillars that triggers these defense responses in legume crops. In11 is recognized by the Inceptin Receptor (INR), which induces changes in gene expression preceding the accumulation of antiherbivore defenses. However, the molecular mechanisms regulating In11-induced defenses via INR remain largely unexplored.

This dissertation proposes a molecular mechanism by which In11-induced genes encoding gut-stable direct defenses are transcriptionally regulated by the plant circadian clock. Furthermore, this work demonstrates that INR is required in legumes for the In11-induced attraction of predators, which act as indirect defenses against herbivores in the field. These findings provide the first evidence of a time-dependent response to a specific HAMP and directly link core circadian clock transcription factors to plant immunity against herbivores. Lastly, this study presents the first conclusive molecular and ecological evidence that a specific plant immune receptor directly controls ecologically relevant tritrophic interactions.

Para Ethan y Samuel, my very hungry caterpillars.

Contents

Acknowledgements

I am profoundly thankful to my advisor Dr. Adam Steinbrenner who invested his time and expertise to support my development as an independent scientist and encouraged me to explore the wide range of interests I developed while conducting this research (as evidenced by the three very different chapters in this dissertation). I thank the members of the Steinbrenner, Nemhauser, Imaizumi and Di Stilio labs for conversation and feedback. I am deeply thankful for the support and feedback from the members of my committee and greatly appreciate their commitment to my professional and scientific development through some challenging years. This research was supported by NIH 5R35GM151272 and NSF 2139986 to ADS and NIH R01GM079712 to TI. NGP and ADS were supported by start-up funding from the University of Washington. I was supported by the USDA-AFRI predoctoral fellowship Grant #2023-67011-40362 and was partially supported by the UW Royalty Research Fund grant #A161929 and the Hereensperger and Walter and Margaret Sargent Awards.

Chapter 1. The circadian clock regulates receptor-mediated immune responses to an herbivore-associated molecular pattern.

Natalia Guayazán Palacios, Takato Imaizumi, Adam D. Steinbrenner

1.1 Introduction

Plant survival is dictated by the plant's ability to accurately perceive biotic threats and to activate effective defenses in a timely manner. Plants sense pests and pathogens through recognition of

molecular patterns via cell-surface pattern recognition receptors (Ngou et al., 2022; Zhang et al., 2024). The specific interaction between a molecular pattern and a cognate receptor results in the activation of Pattern Triggered Immunity (PTI) to provide a first line of protection, a hallmark of which is the transcriptional reprogramming that accompanies metabolic and physiological changes associated with immunity (DeFalco & Zipfel, 2021). While gene expression changes and the mechanisms regulating them in response to Pathogen Associated Molecular Patterns (PAMPs) are well documented (Bjornson et al., 2021; B. Li et al., 2016), our understanding of these processes in response to Herbivore Associated Molecular Patterns (HAMPs) is nascent. Although many HAMPs have been identified (Snoeck, Guayazán-Palacios, et al., 2022), detailed transcriptional responses to specific elicitor molecules have only been described for two HAMPs present in lepidopteran oral secretions: the fatty acid-amino acid conjugate (FAC) C18:3-Glu in *Nicotiana* species (W. Zhou et al., 2016), and the peptide Inceptin 11 (In11) active on cowpea (*Vigna unguiculata)* (Steinbrenner et al., 2022). While FAC receptors are not yet fully elucidated (Poretsky et al., 2020), In11 has a defined recognition mechanism through the legume-specific Inceptin Receptor (INR), a leucine-rich repeat receptor in the Receptor Like Protein (RLP) family (Snoeck et al., 2022; Steinbrenner et al., 2020). Since In11-INR is the only HAMP-receptor pair characterized in molecular detail it serves as a model for studying herbivore-specific immune (Steinbrenner et al., 2022).

Consistent with induced defenses in response to herbivory (Erb & Reymond, 2019), In11 elicitation results in a well-characterized set of defensive outputs driven by amplified and accelerated expression of wound-induced genes, as well as In11-specific gene expression in cowpea (Steinbrenner et al., 2022). As a result of rapid transcriptional reprogramming in response to In11, both direct and indirect induced defenses are accumulated to increase

resistance to herbivores, including production of defense-related phytohormones, specialized anti-nutritive proteins and metabolites, and volatile emission (Schmelz et al., 2006). HAMP-induced regulation of these responses is thought to be a mechanism to reduce costs by effectively allocating defenses to times and tissues when and where they are needed (Karban, 2011).

The legume pod borer (*Maruca vitrata*) and the armyworm complex are the most important Lepidoptera pests of cultivated cowpea and other legumes in Africa and South America (Srinivasan et al., 2021) and North America, respectively. Bioactive In11 has been detected in beet (*Spodoptera exigua*) and fall armyworms (*Spodoptera frugiperda*) (Schmelz et al., 2012) oral secretions, therefore In11-induced defenses and the mechanisms that regulate them might be relevant against agriculturally important pests.

Like many other biological processes, plant immunity is regulated by the circadian clock (H. Lu et al., 2017). The core circadian clock components CIRCADIAN CLOCK ASSOCIATED 1 (CCA1), LATE ELONGATED HYPOCOTYL (LHY) and LUX ARRYTHMO (LUX) are Myb-like transcription factors that participate in the rhythmic accumulation of defensive hormones jasmonic (JA) and salicylic acid (SA) (Goodspeed et al., 2012), and resistance genes such as *RECOGNITION OF PERONOSPORA PARASITICA 4* (*RPP4*) (W. Wang et al., 2011) in anticipation to herbivore and pathogen attack, respectively. LUX is also involved in modulating PTI responses in a time-of-day-dependent manner through gated accumulation of ROS and expression of the bacterial marker gene *FLG22-INDUCED RECEPTOR-LIKE KINASE 1* (*FRK1*) in response to flagellin 22 (flg22), a bacterial PAMP, in the early morning (Korneli et al., 2014). Whether HAMP-induced transcriptional changes are also time-of-day dependent and if they are modulated by the plant circadian clock remains unknown.

Modulation of gene expression is directly regulated by clock transcription factors. For example, CCA1 and LHY bind the cis-regulatory elements CCA1 Binding Site (CBS) (Wang et al., 1997) and Evening Element (EE) (Harmer et al., 2000) to regulate target gene expression via repression and/or activation (Nagel et al., 2015; Wang et al., 2011). CBS and EE cis-elements have been found in the promoter of rhythmically expressed bacterial resistance genes (Wang et al., 2011), and the herbivore-induced *Ocimene Synthase* (*PlOS)* in lima bean (*Phaseolus lunatus*), a transcript rhythmically accumulated in response to herbivore feeding and regulated by light and JA (Arimura et al., 2008). However, studies of herbivory and HAMPs have not yet measured whether direct regulation by clock transcription factors extends to genome-wide changes in herbivore-induced gene expression.

Here we present a detailed exploration of the temporal induced response of cowpea to HAMP In11 and provide a molecular link between the In11-induced immune responses and the plant circadian clock. We characterized the global early and late transcriptional changes induced by daytime and nighttime In11 treatment and identified a daytime-induced antiherbivore-related *Kunitz Trypsin Inhibitor* (*KTI*) gene with CBS elements in the promoter region. Using plants with a conditional arrhythmic phenotype under constant light we tested if the daytime *KTI* induction in response to In11 required a functioning clock and found that the misexpression of *VuLHY* homologs under constant light (LL) correlated with lack of repression of *KTI* during nighttime. Furthermore, we show that transient overexpression of cowpea LHY proteins in *Nicotiana benthamiana* modulates the activity of the *Vu*KTI promoter in a CBS-dependent manner. We propose that VuLHY transcriptionally gates In11-induced gene expression at night to ensure that specific defenses are most strongly produced in response to herbivorous threats in daytime.

1.2 Results

1.2.1 In11-induced transcriptional responses are dependent on the time-of-the-day

To examine the contribution of time-of-day on the transcriptional response to a specific HAMP (In11) via a known receptor (INR), we treated cowpea plants by scratch-wounding and adding water (w + H_2O) or In11 (w + In11) at different times of the day (Fig. **1a**)**.** We identified differentially expressed genes (DEGs) by comparing transcriptomes of w + H_2O vs undamaged (i.e. the effect of wounding), and w + In11 vs. w + H_2O treatment (i.e. the additional effect of In11-induced responses) at the corresponding time of the day ($|Log_2$ fold change (FC)$| \geq 1$ and $P_{adj} < 0.05$) (Table S3). A Principal component (PC) analysis across all samples confirmed consistent biological replicates and the effect of the treatments. The largest changes are attributed to damage and time, with clear separation of the early 1 h wound responses, late 6 h wound responses, and undamaged plants. Within those groups, there are also differences in the response to In11 depending on the time-of-day at which the treatment was applied (Fig. **S1a**).

We found a time-of-day dependent response to In11 where daytime treatment resulted in a larger number of transcriptional changes than nighttime treatment (Fig. **1b,c**). While this was true for both 1 h and 6 h responses, time-of-day dependence of the In11 response was particularly striking 6 h after treatment as there were 707 DEGs at ZT10 (light blue, 510 down and 197 up) but only 59 at the corresponding nighttime timepoint ZT22 (dark blue, 16 down and 43 up). Furthermore, most DEGs were unique to ZT10 with only 44 shared with ZT22. A hierarchical clustering analysis of all In11 DEGs further supported the unique In11-induced transcriptional program at ZT10, and revealed that In11-induced nighttime responses were more similar to wounding alone because w+In11 ZT22 samples clustered most closely with day- and nighttime w+H_2O plants, rather than with w+In11 at ZT10 (Fig. **S1b**). Interestingly, these patterns for In11-

regulated genes did not hold for the broader set of 15,842 genes affected by wounding (Fig. **S2a**). In contrast to nearly complete time-of-day dependence of In11-induced downregulation at the 6 h timepoint, wound-induced downregulation of genes was intact at night (ZT22), and affected an even larger number of genes than in daytime (ZT10); nevertheless, nearly 50% of the up and downregulated genes were shared between daytime and nighttime (Fig. **S2b**). Together these results indicate that In11 modulates the wound response in a time-of-day dependent manner.

We compared the types of genes in the daytime and nighttime DEGs to identify shared and unique processes modulated by the recognition of In11 at different times of the day. Gene Ontology (GO) analysis revealed that daytime upregulated genes were significantly enriched for molecular functions related to antiherbivore defense such as lipid biosynthesis and metabolism, acyltransferase activity, protease binding and terpene synthase activity, while the downregulated DEGs were enriched for photosynthesis (Table S4).

1.2.2 Direct and indirect antiherbivore defenses may be directly regulated by the circadian clock

Given the time-of-day dependent response to In11 we hypothesized that the circadian clock could be directly modulating gene expression; specifically, that cowpea homologs of the transcription factors CCA1 or LHY were directly repressing gene expression at nighttime via canonical cis elements CCA1 binding site CBS (Wang et al., 1997) and evening element (EE) (Harmer et al., 2000). To find evidence for direct CCA1/LHY regulation in specific promoter sequences, we calculated the LFC difference (LFC_{diff}) between daytime and nighttime treatments for all In11-induced DEGs, calculated by comparing ZT5 vs ZT17 for 1 h differences, and ZT10 vs ZT22 for 6 h differences (Fig. **2a**). We then annotated 21,768 total CBS and EE sequences in their promoters using FIMO (Fig. **S3a**). We focused on the 6 h comparison because of the strong

effect of time-of-day, and found that 326 out of 722 unique DEGs (45.1%) across ZT10 and ZT22 had at least one of the cis elements (purple dots, Fig. **2a**), and of those 104 had $LFC_{diff} \geq 1$ (Table S5). This was a similar proportion of CBS and EE as in the entire cowpea genome (14,861 out of 31,948, 46.5%), and previously identified Arabidopsis herbivore-responsive DEGs (Reymond et al., 2004) (75 out of 156, 48.1%) (Table S6). We focused on candidate targets with defense-related functions for further analysis (Fig. **2a**), and found that genes with functions in indirect and direct defenses, such as terpene synthases (TPS) and Kunitz Trypsin Inhibitors (KTI), respectively, contain CBS and/or EE sites in promoters, suggesting they are a target of CCA1/LHY (Fig. **2b**). Given that the accumulation of certain induced indirect defenses in response to herbivory is known to be time-of-day dependent (Arimura et al., 2008), we focused on direct defenses and selected *VuKTI* (*Vigun05g143300*) as a marker gene. *VuKTI* encodes a KTI in a gene family known to contribute to direct defenses against chewing herbivores (Eberl et al., 2021; Major & Peter Constabel, 2008; Sultana et al., 2023). We confirmed by qPCR that *VuKTI* was significantly more induced 6 h after daytime application of In11 (ZT10), but not 6 h after nighttime treatment (ZT22) compared to wounding alone (Fig. **2c**).

1.2.3 Expression of VuLHY homologs is disrupted by wounding and constant light in cowpea

To further investigate circadian clock modulation of In11 induced defenses, we first identified LHY in cowpea and profiled its expression pattern under various conditions. We identified two homologs *VuLHY1* (*Vigun10g153300*) and *VuLHY2* (*Vigun09g004100*) (Fig. **3a**) and determined that their transcripts have rhythmic expression that peaks at dawn (ZT4) in LD in undamaged samples (gray lines) according to our transcriptomics (Fig. **3b**) and independent qPCR data (Fig. **3c**); although *VuLHY1* expression was stronger than *VuLHY2*. Furthermore, both genes were significantly downregulated (Table S1) in response to wounding at nighttime (ZT22), without

further effect of In11. Together, these results suggest a reciprocal regulation of the circadian clock by wounding in cowpea.

We also confirm the presence of a free-running clock in cowpea by measuring the expression patterns of several and circadian clock gene homologs under LL: *VuLHY1/2, REVEILLE 8* (*VuRVE8*) (Fig. **S4a,b**), a Myb-like transcriptional activator in the circadian clock (Farinas & Mas, 2011; Hsu et al., 2013; Rawat et al., 2011), and *GIGANTEA* (*VuGI*), involved in protein stability and whose expression is directly regulated by CCA1/LHY1 (Lu et al., 2012). The expression of *VuLHY1/2 and VuRVE8* genes showed clear peaks around subjective dusk of the 2nd day and the expression levels were severely dampened after the 2nd days in constant light conditions (Fig. **4a**). Furthermore, *VuGI* was continuously expressed at medium-high levels throughout the day after 24 h in LL, which is consistent with lack of repression by LHY in free-running conditions. Lastly, the expression pattern of a cowpea chlorophyll A/B-binding protein 2 homolog (*VuLHB1B/CAB2*) (Fig. **S4c, d**), a target of the circadian clock involved in photosynthesis (Anderson & Kay, 1995), became arrhythmic after 48 h in LL (Fig. **4b**), thus confirming the general impaired clock function of cowpea plants under free-running conditions.

1.2.4 The cowpea circadian clock restricts nighttime expression of In11-induced direct defenses

We used the conditional arrhythmic phenotype of cowpea plants under free-running conditions to test if nighttime repression of an In11-induced *VuKTI* was dependent on the circadian clock. We expected the time-of-day differences in In11-induced expression to be lost after 48 h in LL conditions due to reduced expression of the *VuLHY* homologs and therefore lack of repression at nighttime. Briefly, we measured In11-induced *VuKTI* expression 6 h after treatment in plants with 4 to 88 h of LL exposure (Fig. **5a**). Consistent with our prediction and previous observation,

VuKTI was significantly induced by In11 treatment in subjective daytime but not in subjective nighttime after 24 h ~~day~~ in LL conditions. However after 48 h in LL, *VuKTI* induction by In11 was not significantly different in subjective daytime or nighttime conditions (Fig. **5b**). These results suggest that circadian oscillation of *VuLHY* is required for gated expression of In11-induced defenses, likely by suppressing *VuKTI* expression at relative nighttime.

1.2.5 VuLHY homologs regulate pKTI promoter activity in a CBS-dependent manner

To test if transcriptional regulation of *VuKTI* depends on canonical and polymorphic LHY-bound cis element CBS, we performed a transient luciferase reporter assay in *N. benthamiana* leaves. CBS and CBS-L are 8-mer sequences found throughout the cowpea genome and preferentially located in promoters or promoters and coding sequences, respectively (Fig. 6**a**). We generated reporters with the wild type (WT) *pKTI* promoter or mutant versions with changes to the ATCT sequence of the CBS located at position -334 (mutant sequence M1) as well as a CBS-like (CBS-L) sequence identified at position -80 (mutant sequence M2), upstream of the TATA box (Fig. **6b,c**). Compared to the empty vector (EV) control, over-expression of VuLHY1 and VuLHY2 proteins significantly increased the activity of the *pKTI* reporter, but not in the M1, M2 or M1+M2 mutants (Fig. 6**d**). A similar transcriptional activation was also observed when AtLHY was overexpressed (Fig. **S5**). This data indicates that VuLHY1 and VuLHY2 interact with the In11-induced *KTI* promoter in planta via canonical and polymorphic CBSs, and that cowpea and Arabidopsis LHY homologs behave as activators when transiently overexpressed in tobacco in this context.

1.3 Discussion

The time-of-day dependent modulation of transcriptional responses to specific elicitors such as HAMPs underscores the relevance of temporal cues to optimize defensive responses against

herbivores. Our study of the transcriptional response of cowpea plants to HAMP In11 revealed a clear time-of-day dependence of induced gene expression associated with direct and indirect herbivore defenses and provided a mechanistic role for the plant circadian clock in directly modulating such dependence.

We found that 6 h induced gene expression was specific and significantly stronger after daytime than after nighttime elicitation with HAMP In11. The number of DEGs at ZT10 was ~40 times larger than at night, and 50% of the GO term categories were uniquely enriched at ZT10, most of which included genes involved in antiherbivore defense like enzymes involved in volatile biosynthesis and protease inhibitors. This pattern of gene expression was not observed by wounding alone, and thus we propose that timely HAMP-specific expression of antiherbivore-related genes is tightly controlled. These findings expand our knowledge of plant defense against herbivores beyond the anticipation of attack via rhythmic accumulation of defensive hormones (Goodspeed et al., 2012), rhythmic accumulation of green leaf volatile (GLV) biosynthetic enzymes transcripts (Joo et al., 2019), the time-of-day dependent accumulation of GLVs in response to the mix of HAMPs and effectors in caterpillar regurgitant (Joo et al., 2019), and accumulation of plant volatiles in response to nocturnal and diurnal continuous mechanical damage (Arimura et al., 2008).

Owing to the central role of the circadian clock in regulating plant metabolism, we hypothesized that the time-of-day differences in response to In11 were in part due to direct regulation of gene expression by LHY. In support of this hypothesis, our genome-wide promoter analysis in cowpea found canonical CBS and EE located between -250 and -1000 bp upstream the coding region of any given gene (Fig. **S3b**), among which was a previously characterized *VuGI* homolog (Weiss et al., 2018), as well as a subset of In11-responsive genes with strong daytime vs nighttime

differences (Fig. **2a**, Table S5). By leveraging the presence of the cis elements and a strong daytime vs nighttime difference in expression we identified multiple *Terpene synthases* (*TPS*), *chalcone synthases* (*CHS*), *chitinases*, *β-glucosidases* and *Kunitz Trypsin Inhibitors* (*KTIs)* as strong candidate targets of direct regulation by VuLHY. Furthermore, we identified *VuLHY1* and *VuLHY2*, two homologs with a conserved Myb-like DNA binding domain (Fig. **S5**) and an expression pattern that peaked at dawn under LD and LL consistent with other homologs. We also found CBS and EE sites in previously reported herbivore-induced Arabidopsis genes (Reymond et al., 2004), and thus we propose that this regulatory module is conserved in dicots.

Our RNAseq and independent qPCR data also demonstrated that wounding alone was sufficient to cause misexpression of *VuLHY1* and *VuLHY2*, and that the HAMP In11 had no further effect (Fig. **3**). This indicated that mechanical damage might attenuate circadian clock function similar to damping of the circadian oscillation induced by the feedback regulation by hormones, bacterial infection, bacterial PAMPs and toxins (de Leone et al., 2020; Fraser et al., 2024; Gao et al., 2020; Z. Li et al., 2018; Liang et al., 2024; Zhang et al., 2013). Furthermore, this observation suggests that mechanical damage could act as an environmental Zeitgeber and contribute to the entrainment of the cowpea circadian clock by rapidly adjusting the expression of the transcription factors to allow plastic responses under biotic stress (Wang et al., 2022). Previous studies have shown that the expression of the circadian clock genes *LHY* and *TIMING OF CAB EXPRESSION 1* (*TOC1*) in *Nicotiana attenuata* is altered after two days of continuous herbivory, likely because of unidentified HAMPs in herbivore regurgitant (Joo et al., 2019). However, further experimentation is required to determine if pulses of mechanical damage or HAMP recognition, including In11, sets a new pace for the clock to enhance antiherbivore resistance.

Upstream factors such as the inducibility of defense hormones could explain time-of-day dependent responses, which should be apparent from transcriptional signatures of hormone biosynthesis. If hormones control strong daytime responses compared to nighttime we would expect biosynthetic genes with known In11-induced expression (Steinbrenner et al., 2022), such as *Allene oxide synthase* (AOS), *Allene oxide cyclase* (AOC), *Lipoxygenase* (*LOX*) for JA, and *1-aminocyclopropane-1-carboxylic* (ACC) *synthase* (*ACS*) and *ACC oxidase* (*ACO*) for ethylene, to show strong time-of-day dependent changes. Surprisingly, although one ACS and three LOXs are induced by In11, only *VuLOX2* (*Vigun11g163500*) had at least one CBS element and weak daytime vs nighttime differences (ZT10-ZT22 $LFC_{diff} = 0.74$), indicating that In11-induced accumulation of JA in the morning might only be a small factor contributing to the enhanced daytime response to In11. Further studies of JA dynamics in time-of-day dependent In11 responses will clarify this pattern.

Using a classical free-running conditions experiment under constant light; we demonstrated that the time-of-day dependent In11 induced expression of an anti-herbivore *VuKTI* is dependent on the circadian clock. We characterized the expression of *VuLHY1*, *VuLHY2, VuRVE8, VuGI,* and *VuLHB1B/CAB2* under LL and demonstrated that they oscillated only for ~~24 h~~ one day, and became arrhythmic (or severely dampened) after; this timing pattern was similar to that of the clock of petunia leaves under DD (Fenske et al., 2015). Since genetic resources and transformation methods to create cowpea clock mutants are lacking, this conditional arrhythmic phenotype under constant light provided a unique opportunity to study the impact of the circadian regulation in on the clock-influenced output in cowpea.

Therefore, this circadian characteristic in cowpea served our experiments in two ways: 1) the first 24 h after transfer to LL allowed us to address the role of light in the time-of day dependent

response to In11, and 2) the following 24 to 96 h served as the conditional *VuLHY* knockdown (or arrhythmic clock) mutant. We leveraged the conditional arrhythmic plants and demonstrated that In11-induced *VuKTI* expression was higher after daytime treatment than nighttime under constant light conditions in the first 24 h of LL conditions, but that *VuKTI* was equally induced at subjective daytime and nighttime conditions once *VuLHY1* and *VuLHY2* became mis expressed (Fig. **5**). We conclude that light is not a mechanism regulating morning In11-induced *VuKTI* expression, although light does partially regulate herbivore-induced terpene synthesis and emission (Arimura et al., 2008; Joo et al., 2019) and many DEGs with strong daytime and nighttime differences did not have a canonical CBS or EE site in their promoter, indicating that light and indirect regulation by the circadian clock contribute to the overall time-of-day dependent response to In11. While direct mechanisms of regulation are difficult to study in cowpea due to lack of genetic tools, these patterns are consistent with a model where lack of In11-induced expression of *VuKTI* during the night is likely due to transcriptional repression by the VuLHY homologs.

Our transient luciferase reporter assay demonstrated that overexpression of VuLHY1 and VuLHY2 modulated the activity of the *VuKTI* promoter in a CBS dependent manner. By comparing the activity of reporters bearing wild type and mutated variants of the CBS sites, we determined that changes to the ATCT sequence in the ~~5'~~ 3' end of the element are sufficient to alter the interaction between the promoter and the transcription factor. This is similar to the interaction of AtLHY with CBS and EE (Adams et al., 2018; Harmer et al., 2000; Kamioka et al., 2016; Kim et al., 2023; Nagel et al., 2015) via this sequence (Wang et al., 1997), further supporting that VuLHY targets genes via CBS. We also found a polymorphic variant that we have named CBS-like (CBS-L: CAAAATCT) that also requires a conserved ~~5' 3~~' end to interact

with the VuLHY homologs. Based on the distribution and abundance of CBS-L (Table S7), we propose it is likely a novel VuLHY binding site in cowpea. Consistent with our results, Arabidopsis LHY binds other sequences in genome-wide analyses (Adams et al., 2018). The higher background level activity of CBS-L co-expressed with the EV, and the differential effect on interaction with VuLHY1 and VuLHY2 suggests that this site might provide some specificity of binding and an added layer of regulation under certain conditions, although this remains to be explored in detail. In our transient system both VuLHY1 and VuLHY2 functioned as activators, likely due to the regulatory environment in the *N. benthamiana* transient expression system since AtLHY, typically a repressor, also functioned as a weak activator under our experimental conditions (Fig. **S4a**), although a unique activation function has been described for AtLHY in the fatty acid synthesis pathway (Kim et al., 2023). Nevertheless, our free running experiment using the conditional arrhythmic plants clearly demonstrated that *VuGI,* a possible direct target*,* became arrhythmic and highly expressed when *VuLHY* expression was low, thus supporting a repressive function for *VuLHY* against its regulated target genes.

In summary, we describe a molecular link between the plant circadian clock and HAMP-induced gene expression in cowpea. VuLHY gates the expression of In11-induced genes likely fine tuning the herbivore-specific response. At night when *VuLHY* is highly expressed, VuLHY interacts with the promoter of In11-responsive genes involved in antiherbivore defense such as *VuKTI* to repress their expression. When *VuLHY* expression decreases during daytime, the CBS-bearing In11-induced promoters are available for recruitment of the transcriptional machinery required for antiherbivore response. The relevance of this regulation to physiology and metabolism of anti-herbivore defenses is a topic that should be further explored. We expect that gating of HAMP-induced responses by the circadian clock is a mechanism to minimize the effect

of the growth-immunity trade-off by allowing robust and specific response during the day without interfering with nighttime growth, and thus plants could be more resistant to In11-producing herbivores during the morning.

1.4 Materials and Methods

1.4.1 Plant materials and growth conditions

Cowpea (*Vigna unguiculata*) accession IT97K-499-35 was used in all the experiments. For planting, seeds were surface sterilized with 70% ethanol for 2 minutes, followed by two washes with sterilized water. Seeds were sown on sunshine potting mix No.5 and placed in a growth chamber (Conviron PGW-40) at 26 C, 70% relative humidity (RH), 500 µmol/m^2sec light intensity, and 12 h light/dark (LD) cycle for 14 days. Details specific to light/dark (LD) and constant light (LL) experiments are provided in the following sections.

1.4.2 In11 treatment under LD and LL conditions

Inceptin 11 (In11) peptide (ICDINGVCVDA) was synthesized (Genscript Inc.) and dissolved in water. We lightly wounded the middle leaflet of the first fully extended trifoliate on 14-day-old cowpea plants using a new razor blade to remove the cuticle (1 cm^2 per wound). We made four wounds, two on each side of the main vein of the adaxial side of the leaflet, and equally distributed 20 µL of either water or 1 µM In11 with a pipette tip.

For the LD experiment, we applied the In11 and water treatments 4 h after the lights came on (daytime, Zeitgeber time 4: ZT4) or 4 h after the lights went off (nighttime, Zeitgeber time 16: ZT16) in the growth chamber. We collected samples 1 h (ZT5 and ZT17) and 6 h (ZT10 and ZT22) after treatment, along with untreated controls. Treatments and sample collection under dark were performed using a green light to improve visibility.

For the LL experiment, we transferred LD-grown cowpea plants (see above) 10, 11, 12 or 13 days after germination to a separate growth chamber under LL. On day 14, we treated all plants with In11 or water at subjective daytime (time in LL: 4, 28, 52 and 76 h) or subjective nighttime (time in LL: 16, 40, 64 and 88 h). We then collected samples from independent plants 6 h after treatment (time in LL daytime 10, 34, 58, 82 h and nighttime 22, 46, 70, 94 h), along with untreated controls.

1.4.3 RNA extraction, qRT-PCR and transcriptomics

In all experiments we collected samples as follows: two leaf discs were taken from the treated leaflet (one proximal and one distal) using a 0.6 cm^2 leaf punch, placed in a 1.5 mL tube containing a metal bead, frozen in liquid nitrogen, and stored at -80 C. Prior to RNA extraction, the samples were ground using a mixer mill (Retsch MM400).

For qPCR, we extracted total RNA using the Trizol (Invitrogen) method. We performed quality control of the RNA by NanoDrop1000 and gel electrophoresis and 1 µg of RNA was used to synthesize cDNA using the SuperScript IV RT Kit (Thermo). We used the Power SYBR™ Green PCR Master Mix for amplification and quantification in a CFX Connect Real-Time System (Bio-Rad). We calculated relative gene expression by using the $2^{-\Delta\Delta Ct}$ method and *Ubiquitin (UBQ)* (*Vigun07g244400*) as an expression control (Table S1) .

For RNAseq, we extracted total RNA using the NucleoSpin Plant RNA kit (Macherey-Nagel Inc.) and performed quality control as explained above. We further treated the RNA using the TURBO DNA-free kit (Invitrogen) as DNA was still present in the samples. The extracted RNA was used to generate paired-end Illumina 2x150 bp strand-specific libraries with polyA selection that were sequenced in a HiSeq2500 (Azenta). For gene expression analyses, we mapped the

reads to the cowpea genome (Q. Liang et al., 2024) *Vigna unguiculata* v1.2 available in Phytozome13 (Goodstein et al., 2012) and used the –quantMode in STAR to quantify them (Dobin et al., 2013), and then performed differential gene expression analyses using DESeq2 (Love et al., 2014) implemented in R.

1.4.4 Motif search

We searched known CCA1 and LHY binding sites in the promoters from all genes in the cowpea genome (*Vigna unguiculata* v1.2). We retrieved the 1.5 kb region upstream of the start codon from all genes using a custom Python script and the annotation file, and then used the Find Individual Motif Occurrences (FIMO) (Grant et al., 2011) tool to find the 8-mer "AA**MW**ATCT", where M was Adenine (A)/Cytosine (C) and W was Adenine (A)/Thymine (T). We selected this motif because it represented all possible CCA1 binding sites (CBS, CBS-A: AA**AA**ATCT and CBS-B: AA**CA**ATCT) and evening element (EE, AA**AT**ATCT) sequences (Table S2).

1.4.5 Phylogenetic analysis

We used Arabidopsis (*Arabidopsis thaliana*) CCA1 (AT2G46830.1) and LHY (AT1G01060.1) as queries to retrieve sequences from cowpea, common bean (*Phaseolus vulgaris*), lima bean (*Phaseolus lunatus*), soybean (*Glycine max*), and Medicago (*Medicago truncatula*) genomes available in Phytozome13 using tblastn. We retained the top 30 similar sequences and aligned them using MAFFT v7.48 (Katoh et al., 2002) with default parameters. We constructed a phylogenetic tree using RAxML v8 (Stamatakis, 2014), used FigTree (http://tree.bio.ed.ac.uk/software/figtree/) to root and visualize the tree, and manipulated the image in Adobe Illustrator 2024.

1.4.6 Gene expression in cowpea plants with conditional arrhythmic phenotype

We characterized the expression pattern of *VuLHY1*, *VuLHY2*, *VuRVE8, VuGI* and *VuLHB1B/CAB2* in cowpea plants under LL. To this end, cowpea seeds were germinated under LD for 10 days and then transferred to a growth chamber with constant light. Samples were collected from the first fully expanded trifoliate from independent plants every 4 h over 4 days and flash frozen in liquid nitrogen until needed. RNA extraction and gene expression analysis was performed as indicated above.

1.4.7 Molecular Cloning of VuLHY homologs and pKTI promoter

For plant protein expression, the full-length coding sequences of cowpea *VuLHY1* (Vigun10g1533300) and *VuLHY2* (Vigun09g004100) were amplified from a 5' RACE cDNA library using the Q5 High-Fidelity DNA polymerase (NEB) and specific primers (Table S1). The PCR products were cloned into pENTR D-TOPO (Thermo Fisher) and recombined into pB7WG2 for plant expression using the Gateway LR Clonase II (Invitrogen) (Karimi et al., 2002).

For the luciferase reporter assays, the promoter sequence of Kunitz Trypsin Inhibitor (pKTI) Vigun05g143300 (region from the start codon up to 1 kb upstream) was amplified from genomic DNA using DreamTaq DNA polymerase (Thermo Fisher) and specific primers (Table S1) designed against the *V.unguiculata* v1.2 genome. The resulting fragment was cloned into pENTR 5'-TOPO (Thermo Fisher). Mutant versions of the promoter were generated from the wild type (WT) clone using the Q5 Site-Directed Mutagenesis Kit (NEB) and mutagenic primers (Table S1). WT and mutant promoters were re-amplified from the pENTR 5'-TOPO clones using primers with added BpiI restriction enzyme recognition sites and overhangs compatible with Mo-Clo and cloned into a level 0 cloning vector (pICH41295, Addgene plasmid # 47997). The

reporter construct was assembled into a customized pGreenII (Hellens et al., 2000) with BsaI insertion site by combining appropriate ratios of the promoter, luciferase CDS (pICSL80001, Addgene plasmid # 50326) and ocs terminator (pICH41432, Addgene plasmid # 50343) following the recommended XL ligation protocol (Engler et al., 2014; Weber et al., 2011).

1.4.8 Transient luciferase assay in *Nicotiana benthamiana*

To evaluate the in-planta interaction between pKTI and the VuLHY proteins, we individually transformed pGII reporter and pB7WG2 effector constructs into Agrobacterium (*Agrobacterium tumefaciens* GV3101). Cultures were resuspended in infiltration media (10 mM MES pH 5.6, 150 µM Acetosyringone, 10mM MgCl2) and incubated for 3 h. For co-infiltrations, we prepared the appropriate combinations of the reporter and effector. To account for transformation efficiency and to enhance protein expression, we included 35S:Renilla (final OD600=0.1) and the tomato stunt bushy virus silencing-suppressor p19 (final OD600=0.1) plasmid, respectively, in all our assays. The youngest fully expanded leaf on a 6-week-old *Nicotiana benthamiana* plant was infiltrated with the mixture using a needleless syringe at Zeitgeber time 6 (ZT6). All plants were entrained to 12 h light/12 h dark cycles and, after 74 h of incubation at ZT8, we collected leaf punches and immediately froze them in liquid nitrogen. We prepared and analyzed the samples with the Dual-Luciferase Assay System (Promega) according to manufacturer instructions. We measured the activities of firefly (LUC) and Renilla (REN) luciferases using a multi-mode plate reader (Tecan Spark) and calculated the LUC/REN ratio for each reporter - effector combination.

1.4.9 Statistical analysis

All statistical analyses were conducted using R version 4.3.2, and the significance level was set to $\alpha=0.05$. Extreme outliers were identified using the identify_outliers function from the rstatix

(https://CRAN.R-project.org/package=rstatix) package and removed from the qPCR data. A normal distribution of the residuals was confirmed for all data using the Shapiro.test function, and transformations were applied when appropriate to fulfill the assumption of normality for ANOVA and t-test. ANOVA and Tukey post hoc test were used to identify significantly different means of gene expression across treatments, and a two-sided t-test was used to determine significantly different means of promoter activity in the presence of a VuLHY homolog vs an empty vector (EV).

1.5 **Figures**

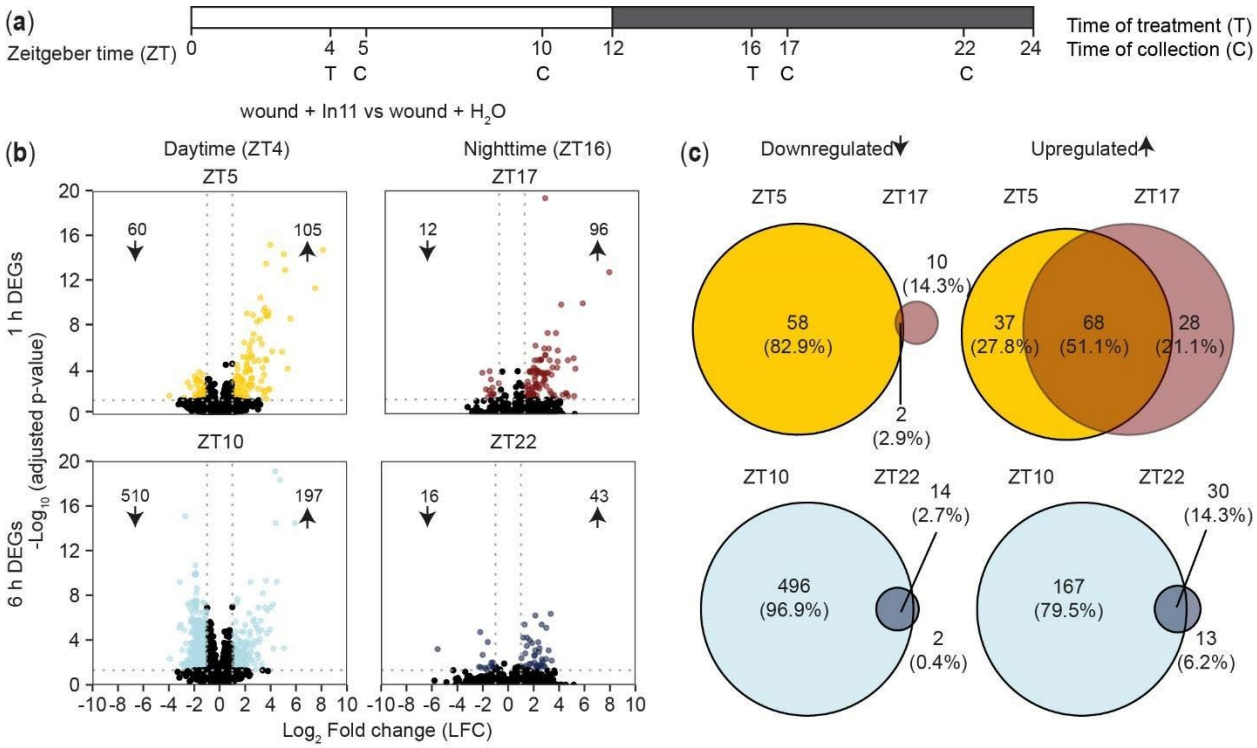

Figure 1. In11 induced responses are time-of-day dependent. (a) Experimental design for RNA-seq. 14-day old cowpea plants grown under light/dark (LD) conditions treated (T) with wound + H_2O or wound + In11 at daytime (ZT4) or nighttime (ZT16), and samples were collected (C) 1h (ZT5 and ZT17) and 6h (ZT10 and ZT22) after treatment (n = 4 individual

plants as biological replicates). **(b)** Volcano plots displaying the number of In11 down (↓) and upregulated (↑) genes (Log_2 Fold Change $|LFC| \geq 1$ relative to wound + H_2O and $p_{adj} < 0.05$) 1h and 6h after treatment. **(c)** Venn diagram indicating the number of shared and unique differentially expressed genes (DEGs) 1h (ZT5 vs ZT17) and 6h (ZT10 vs ZT22) after daytime or nighttime treatment.

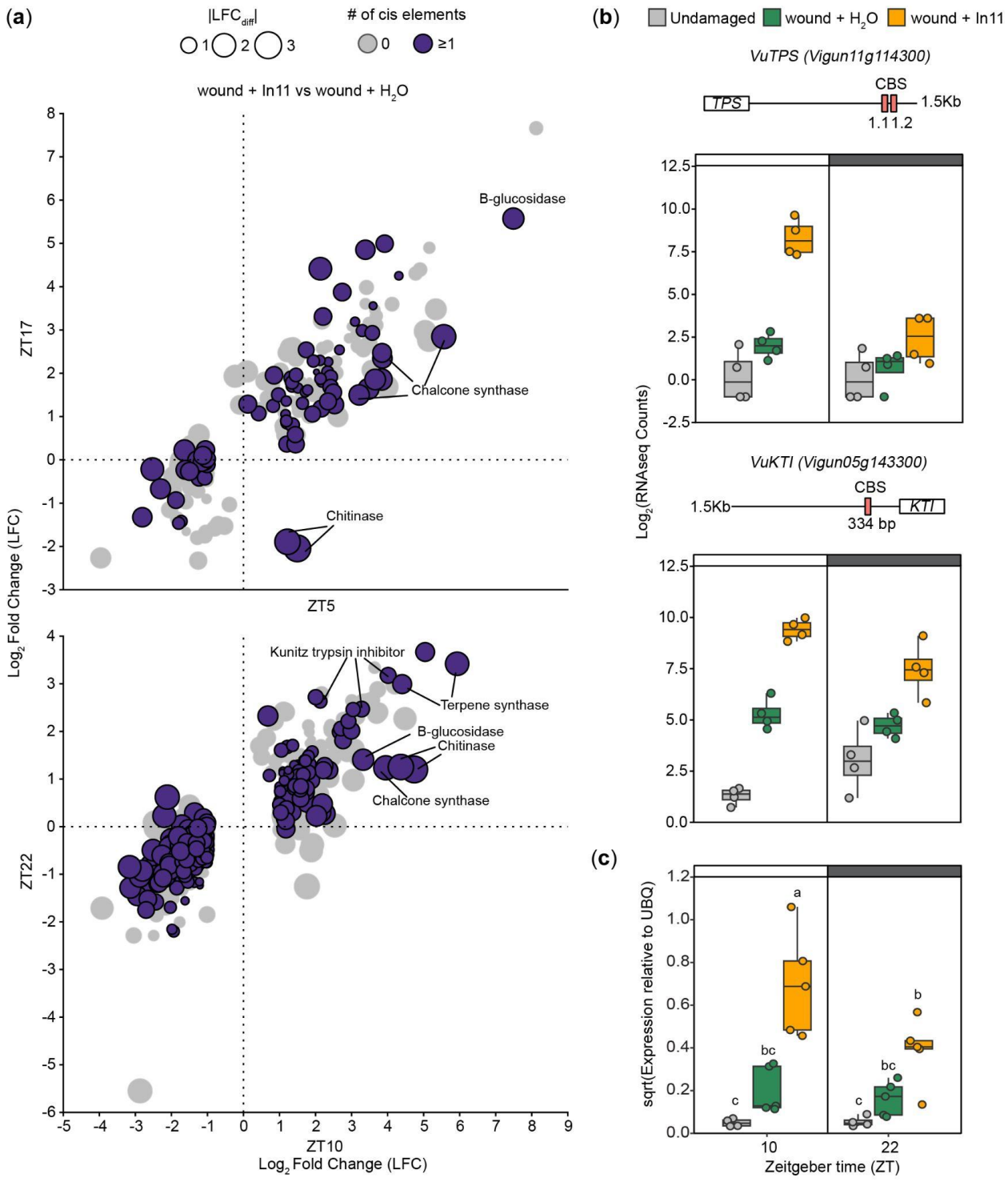

Figure 2. Circadian clock related cis elements CBS and EE are present in the promoters of time-of-day dependent In11-induced defense genes. (a) Scatter plot showing the Log2 Fold Change (LFC) value of In11 DEGs (wound + In11 vs wound + H2O) at 1h (ZT5 and ZT17) and

6h (ZT10 and ZT22) after treatment, and the absence/presence (gray/purple circles) of CBS or EE in their promoter (1.5 kb upstream start codon). The absolute value of the LFC difference ($|LFC_{diff}|$) is represented by the size of the circles, and selected defense-related genes are indicated. **(b-c)** Promoter structure and expression pattern of a Terpene Synthase (*VuTPS*) and Kunitz Trypsin Inhibitor (*VuKTI*) 6h (ZT10 and ZT22) after daytime and nighttime treatment. RNAseq **(b)** and qPCR **(c)** expression data are shown. Different letters indicate significant differences determined by two-way ANOVA followed by Tukey's Honest Significant Difference test (HSD) (n= 4-5 biological replicates, p-value < 0.05). Independent plants were sampled at each treatment - time combination.

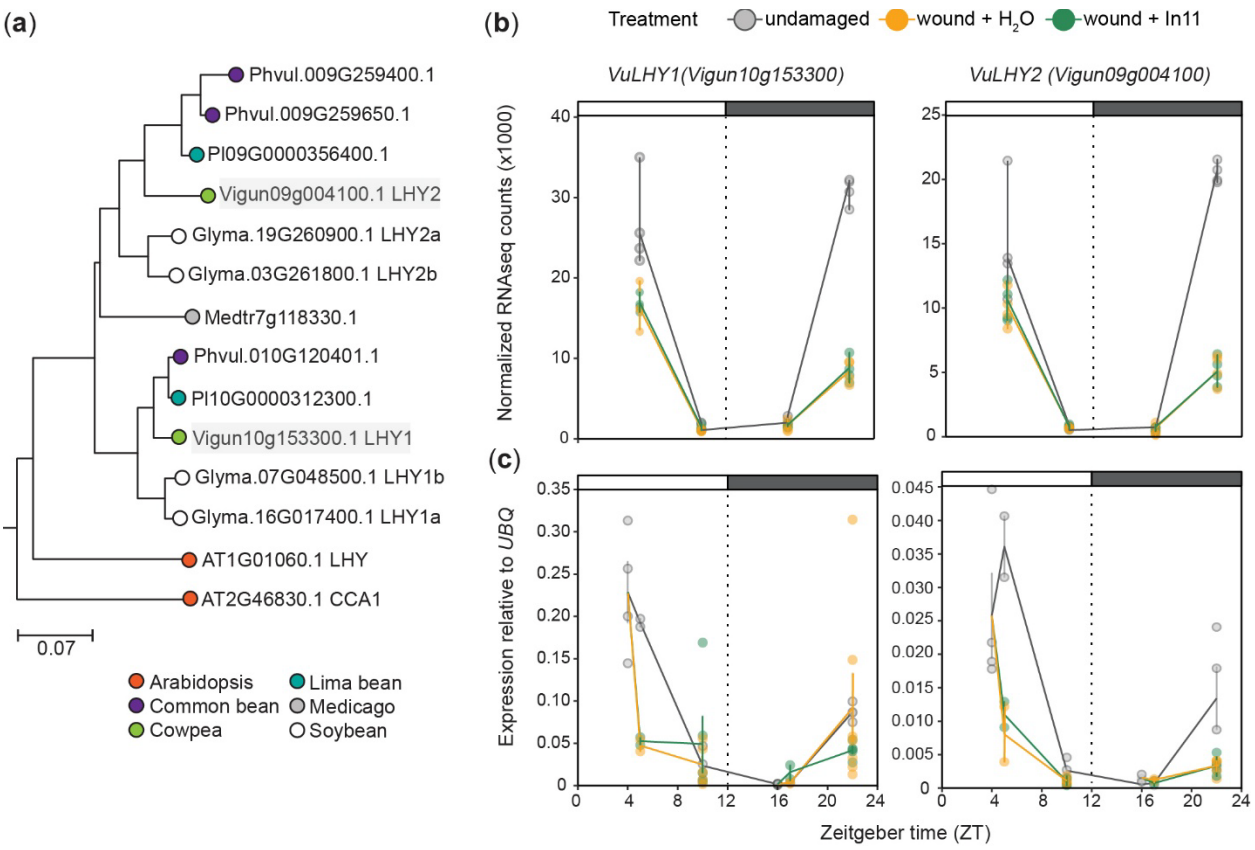

Figure 3. Cowpea Late Elongated Hypocotyl (*VuLHY*) homologs show typical cycling patterns and are downregulated by wounding. (a) Maximum likelihood phylogenetic tree showing 14 *LHY* homologs from five legume species and Arabidopsis. Cowpea homologs *VuLHY1* and *VuLHY2* are highlighted in gray boxes. The scale bar indicates branch length as the mean number substitutions per site. Dielurnal expression pattern of *VuLHY*1 and *VuLHY*2 under LD according to **(b)** RNAseq (n = 4) and **(c)** qPCR data (n = 3-4 biological replicates). Samples were collected at ZT4, ZT5, ZT16, ZT17 and ZT22 from undamaged plants (gray), and 1(ZT5, ZT17) and 6 h (ZT10, ZT22) after daytime (ZT4) or nighttime (ZT16) wound + H_2O (orange) and wound + In11 (green) treatment. Independent plants were sampled at each treatment x time combination. Lines and error bars represent means ± SEM.

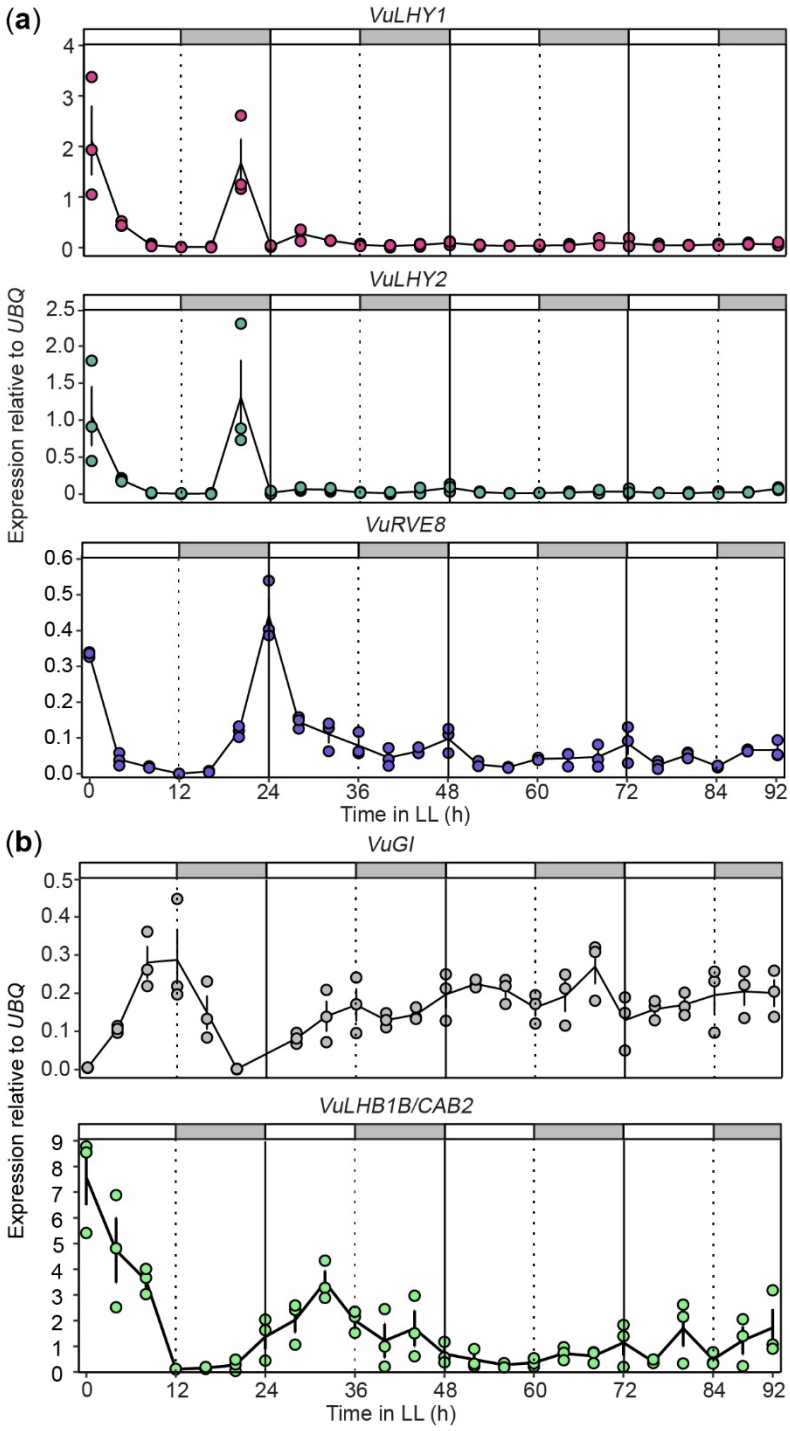

Figure 4. Circadian expression patterns of *VuLHY1*, *VuLHY2, VuRVE8*, *VuGI and VuLHB1B/CAB2* are disrupted under constant light. Expression patterns of **(a)** *VuLHY1*, *VuLHY2, VuRVE8* and **(b)** *VuGI and VuLHB1B/CAB2* under constant light (LL) in cowpea

trifoliates. Cowpea plants were grown under LD for 10 days and then transferred to LL. Leaf samples were taken every 4 hours over the course of four days for gene expression analyses. Lines and error bars represent means ± SEM (n = 3 biological replicates). Independent plants were sampled at each time point.

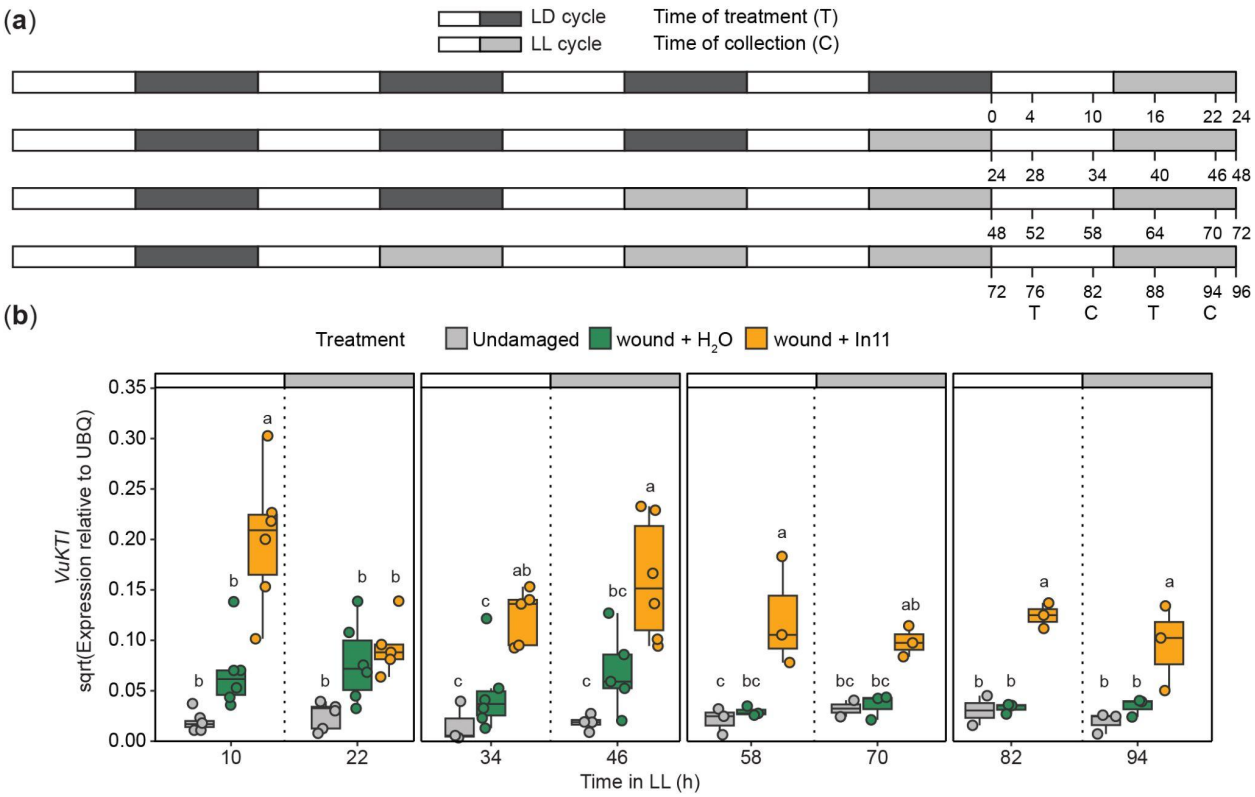

Figure 5. Nighttime repression of In11-induced *VuKTI* is abolished in conditional arrhythmic cowpea plants. (a) Experimental design. Cowpea plants were grown under light/dark (LD) and then transferred to constant light (LL) for one to four days until they were 14 days old. Plants were treated (T) by wound + H_2O (green) or wound + In11 (orange) 4 h after subjective dawn or subjective dusk, and samples were collected (C) 6 h later along with undamaged (gray) controls each day. **(b)** Expression pattern of *VuKTI* according to qPCR data. Different letters indicate significant differences determined by two-way ANOVA followed by Tukey's Honest Significant Difference test (HSD) (n= 3-6 biological replicates, p-value < 0.05) each day. Independent plants were sampled at each treatment - time combination.

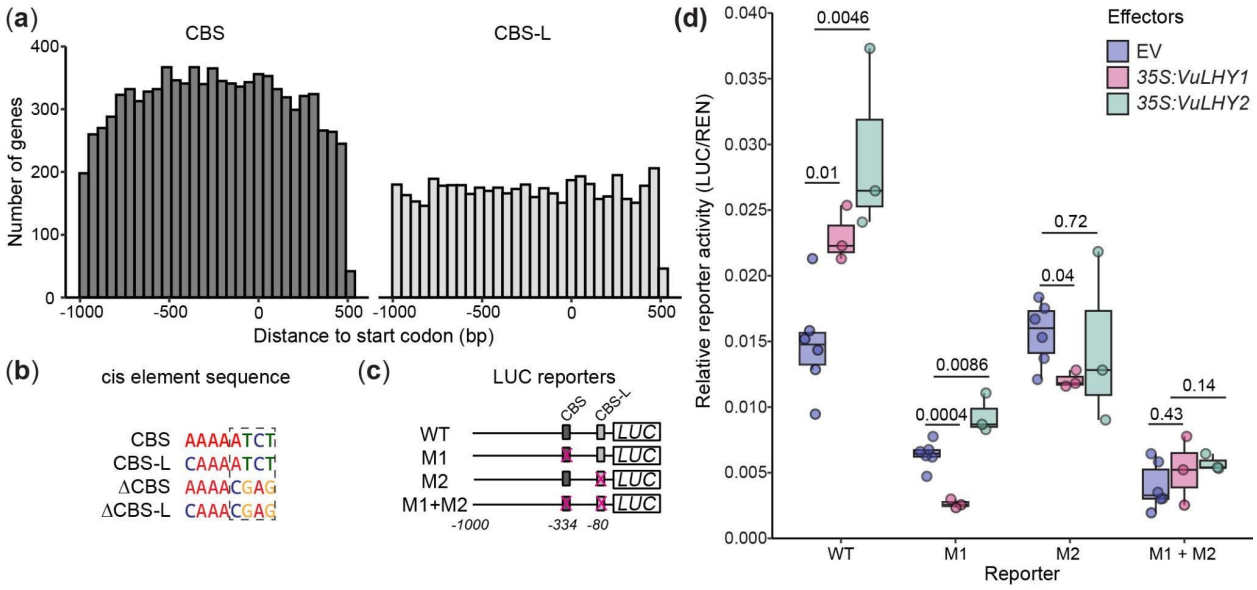

Figure 6. Cowpea LHY homologs modulate the activity of the *VuKTI* promoter in a CBS-dependent manner in tobacco. **(a)** Genome-wide distribution of CBS and CBS-L cis elements in cowpea promoters (-1000 to +500). **(b)** Sequence of the CBS and CBS-like (CBS-L) cis-elements found in the KTI promoter. The CCA1/LHY binding site was mutated on CBS and CBS-L via site directed mutagenesis ~~(underlined)~~ (boxed region) **(c)** LUC reporters used in the assay. WT=CBS, CBS-L, M1 = ΔCBS, CBS-L, M2 = CBS, ΔCBS-L. **(d)** The effect of the VuLHY1 and VuLHY2 proteins on the activity of the LUC reporters. At 72 h LUC activity was measured with 35S:LHY proteins co-expressed in a separate agrobacterium strain. Relative reporter activity was calculated by normalization against *35S:Renilla*. Reporters final OD_{600}=0.3 and effectors final OD_{600}=0.4 Significant differences in the mean (*) were determined by a two-sided t-test of each effector vs. EV (n = 3-6 biological replicates, p-value < 0.05).

1.6 **Supplementary figures**

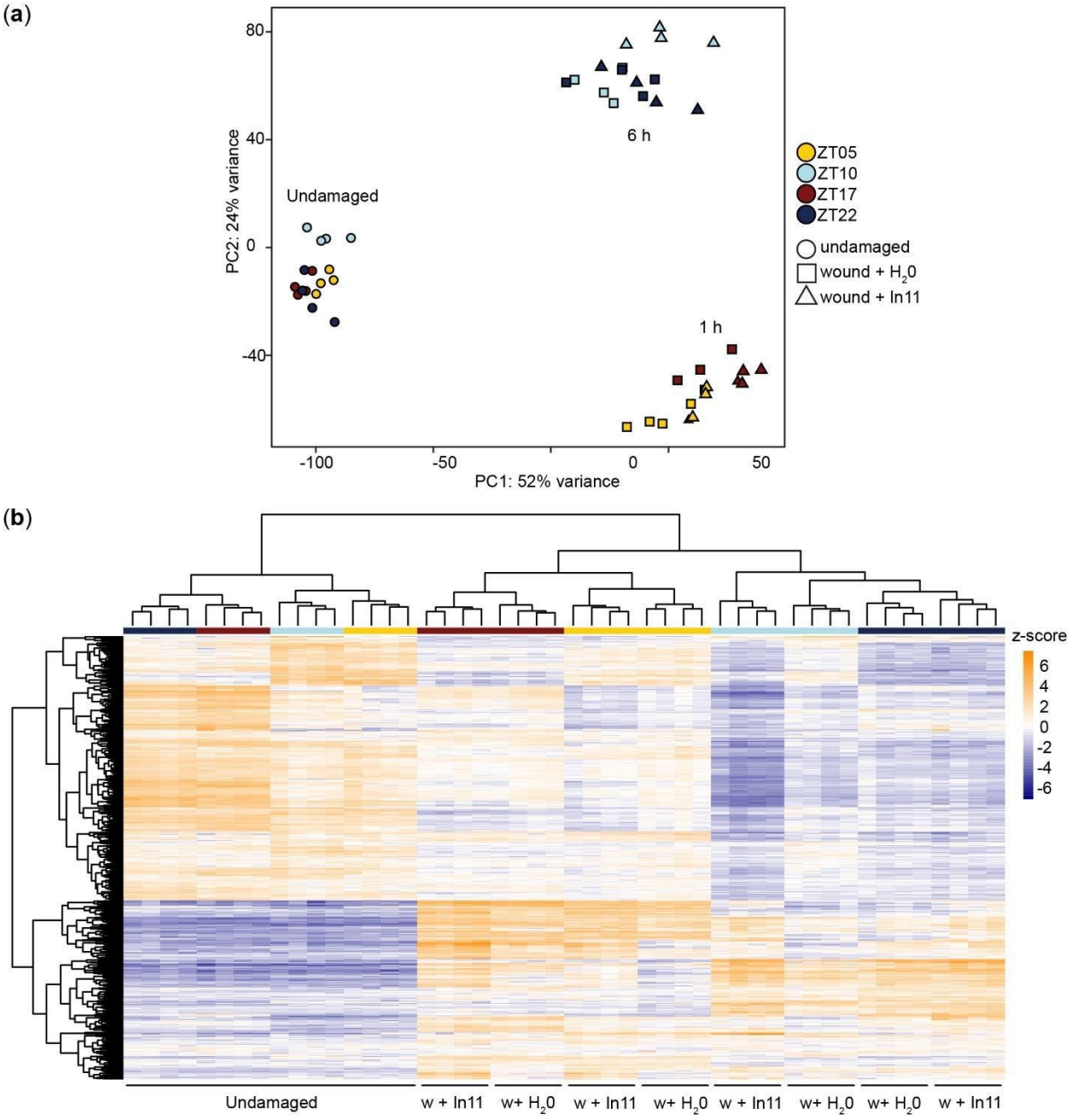

Figure S1. Clustering analysis of DEGs in response to w + H_2O and w + In11. (a) Principal component (PC) analysis of differentially expressed genes (DEGs) across all samples. **(b)** Hierarchical clustering of samples according to the expression pattern of 847 In11-responsive DEGs across all samples.

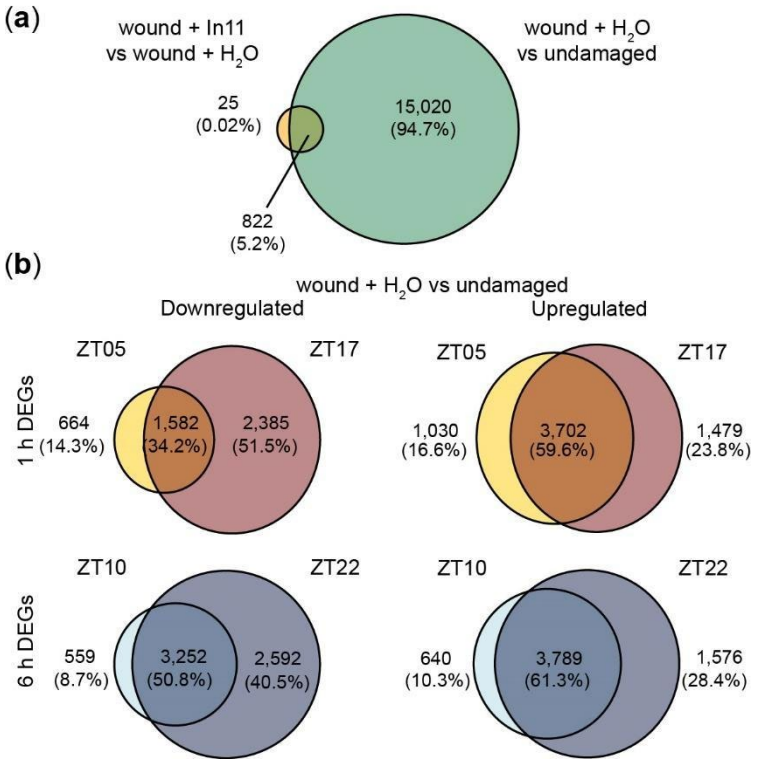

Figure S2. Time -of-day response to wound + H_2O vs undamaged. Venn diagrams indicating the number of shared and unique up and down-regulated genes **(a)** In11 vs wound across all time points, and **(b)** 1h (ZT5 vs ZT17) and 6h (ZT10 vs ZT22) after daytime or nighttime wounding (wound + H_2O vs undamaged).

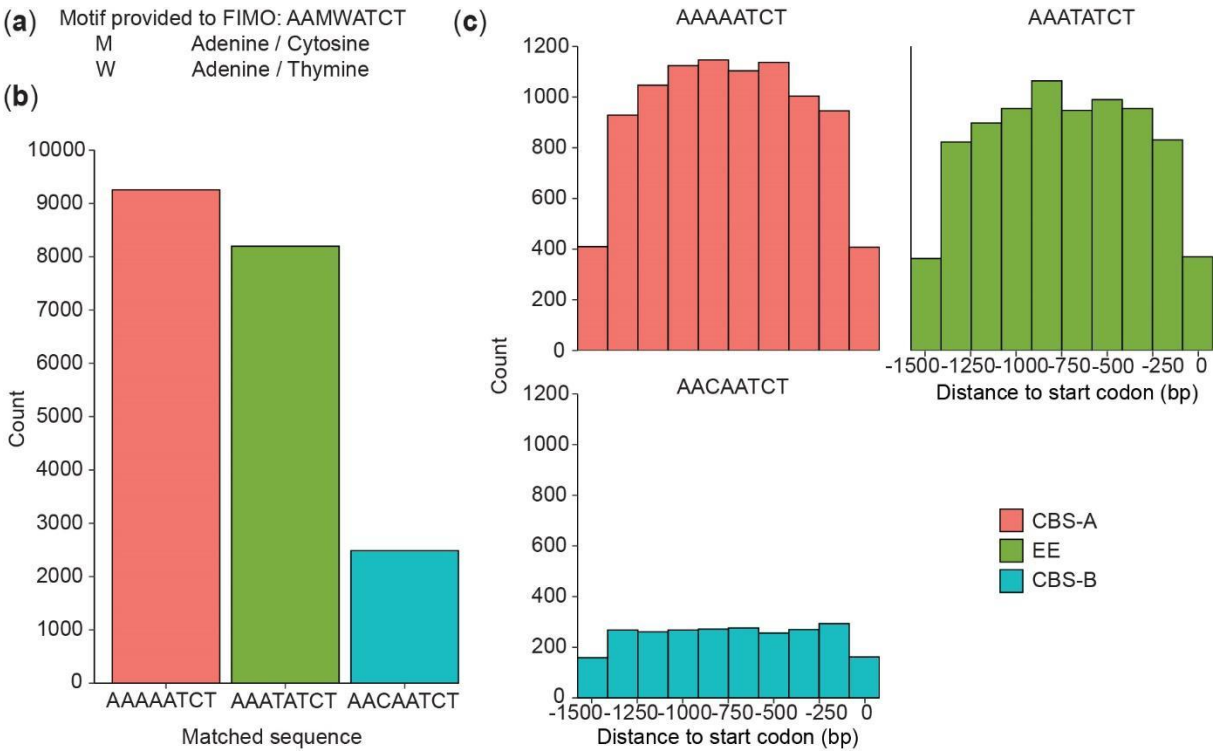

Figure S3. Genome-wide distribution and abundance of CBS and EE motifs in cowpea promoters. Predicted promoter sequences (1.5 kb upstream start codon) were retrieved from the cowpea genome for a circadian clock cis element analysis. **(a)** Motif provided to Find Individual Motif Occurrences (FIMO) software. **(b)** Sequence and count for known and novel motifs found in the promoters. CCA1 Binding site A (CBS-A), Evening element (EE), CBS-B, CBS-like. **(c)** Motif location distribution in the promoters in 250 base pair (bp) bins.

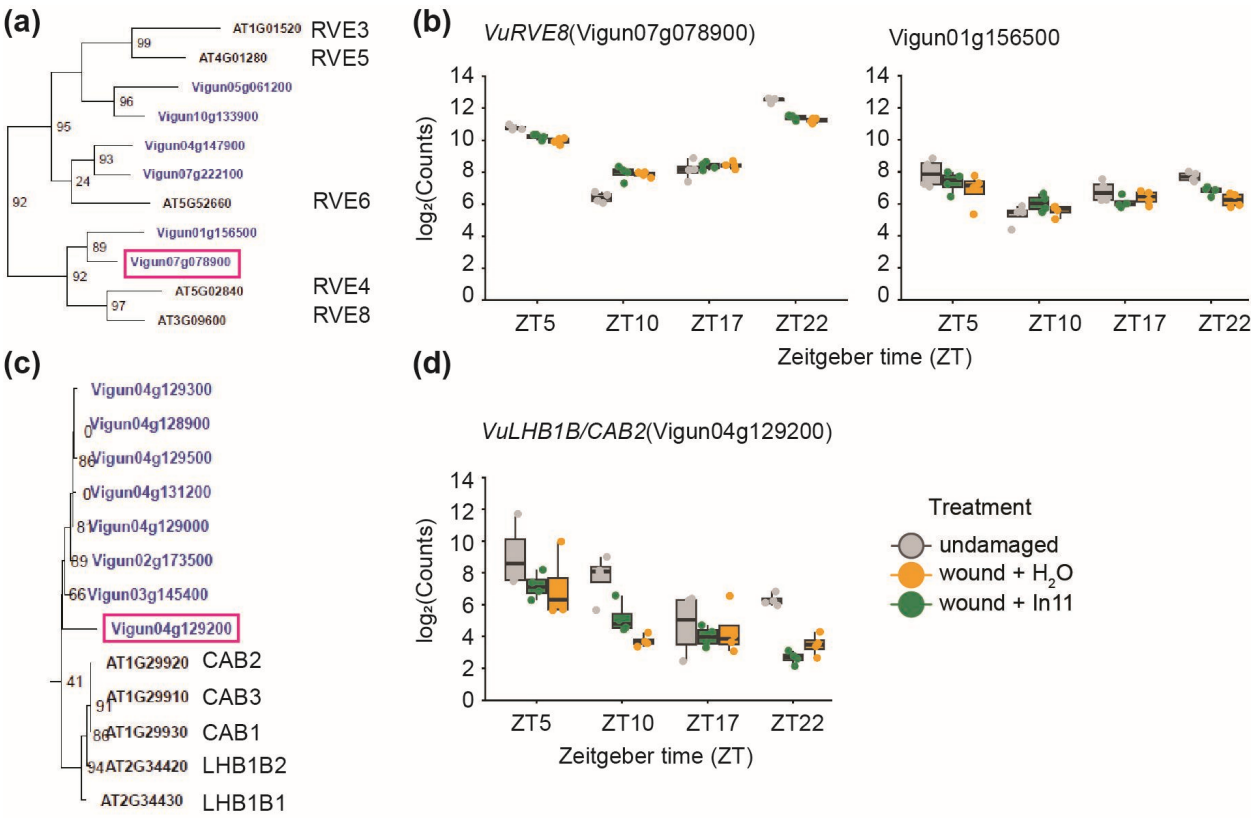

Figure S4. Expression pattern of *VuRVE8* and *VuLHB1B/CAB2* homologs under LD. Trees of the **(a)** REVEILLE and **(c)** light-harvesting complex gene family in cowpea and Arabidopsis. Expression pattern of selected **(b)***VuRVE8* and **(d)** *VuLHB1B/CAB2* according to RNAseq data in Figure 1.

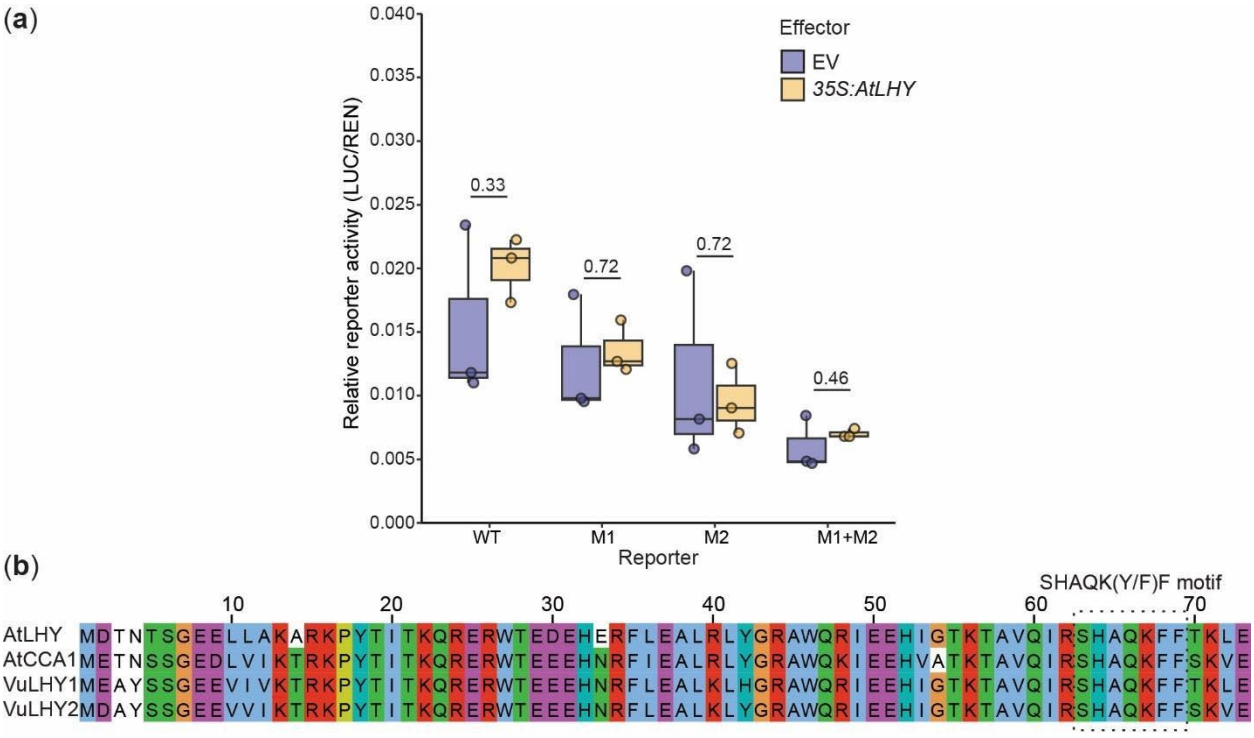

Figure S5. **AtLHY weakly activates the *VuKTI* promoter in a CBS-dependent manner.** The effect of the AtLHY protein on the activity of the LUC reporters. **(a)** At 72 h LUC activity was measured with 35S:AtLHY protein co-expressed in a separate agrobacterium strain. WT=CBS, CBS-L, M1 = ΔCBS, CBS-L, M2 = CBS, ΔCBS-L. Relative reporter activity was calculated by normalization against *35S:Renilla*. Reporters final OD_{600}=0.3 and effectors final OD_{600}=0.4 Significant differences in the mean (*) were determined by a two-sided t-test of each effector vs. EV (n = 3, α= 0.05). **(b)** Alignment of the amino acid sequence of the Myb-like DNA binding domain for AtLHY, AtCCA1, VuLHY1 and VuLHY2. Color scale according to ClustalW.

Chapter 2. The plant immune receptor INR links caterpillar recognition to tritrophic interactions in nature

Natalia Guayazán-Palacios, Patrick Grof-Tisza, Brian Behnken, Carla Marques Arce, Di Wu, Antonio Chaparro, Ted C.J. Turlings, Betty Benrey, and Adam D. Steinbrenner

2.1 Main text

Plants defend themselves against antagonists through complex immune responses that are initiated when cell surface receptors recognize conserved molecular patterns (Boller & Felix, 2009; DeFalco & Zipfel, 2021; Erb & Reymond, 2019; Felton & Tumlinson, 2008; Howe & Jander, 2008; Schuman & Baldwin, 2016; Snoeck, Guayazán-Palacios, et al., 2022). These include herbivore-associated molecular patterns (HAMPs) present in the oral secretions of herbivorous insects (Alborn et al., 1997, 2007; Huffaker et al., 2013; Schmelz et al., 2006; Zeng et al., 2023). The perception of HAMPs can activate direct defenses in the form of chemical and physical changes that directly reduce herbivore performance (Green & Ryan, 1972; Karban & Baldwin, 1997), as well as indirect defenses like the emission of volatiles that attract natural enemies of the herbivores (Kessler & Baldwin, 2001; Turlings et al., 1993). One class of HAMPs, termed inceptins, are peptide fragments of the chloroplast ATP synthase γ-subunit that are generated upon ingestion of leaf tissue by caterpillars and perceived at picomolar concentrations by the leaves of legumes (Schmelz et al., 2006). Recently the Inceptin Receptor (INR), a leucine-rich repeat receptor-like protein (LRR-RLP) found in various Phaseoloid legumes, was identified as the receptor of inceptins, including the 11 amino acid peptide In11 (Snoeck, Abramson, et al., 2022; Steinbrenner et al., 2020). While biochemical and molecular studies have elucidated INR's role in initiating immune signaling and transcriptional reprogramming (Guayazán Palacios et al., 2024; Steinbrenner et al., 2020, 2022), the broader ecological implications of inceptin perception remain unknown. Here, we exploit natural

variation in ln11 perception to demonstrate in laboratory and field experiments that INR mediates direct defense activation as well as the emission of specific volatile blends that recruit natural enemies in a realistic agricultural setting. Our findings establish a direct link between the molecular recognition of herbivores and ecologically relevant tritrophic outcomes, providing insight into the ecological processes that shape the evolution of plant immune systems and offering a mechanistic basis for the effectiveness of (Schmelz et al., 2006) traditional intercropping systems such as the Mesoamerican Milpa (Benrey et al., 2024; Grof-Tisza et al., 2025; Grof-Tisza, Muller, et al., 2024) and push-pull agriculture (Khan et al., 2010).

2.1.1 Natural variation in In11-induced defense is associated with a deletion in the *INR* gene

The lack of efficient gene transformation or silencing methods in common bean has limited reverse genetic study of INR's role in induced defenses. Instead, we aimed to identify naturally occurring In11-insensitive common bean varieties. We screened a panel of Mesoamerican landraces for characteristic induced ethylene gas accumulation in response to wounding (w + H_2O) and application of In11 (w + In11) (Schmelz et al., 2006). While 89 varieties had In11/H_2O induced ethylene ratio of 1.2 or higher, two insensitive landraces, Honduran landrace W6 13807 and Argentinian landrace W6 17491, did not produce induced ethylene in response to In11 (In11/$H_2O \leq 1$) (table S1). We focused on Honduran landrace W6 13807 and, to facilitate crossing and generation of germplasm, we also identified a sexually compatible, closely related landrace PI 311785 with standard In11-induced ethylene response (Fig. 1A, fig. S1A).

The emission of specific blends of herbivore-induced volatiles (HIPVs) to attract natural enemies is a hallmark of plant indirect defenses (Dicke & Baldwin, 2010; Heil, 2014; Schuman & Baldwin, 2016; Turlings & Erb, 2018), as HIPVs induced by live herbivores or artificial

hormone treatment can reduce oviposition and increase predation rates in the field (Gouinguené et al., 2005; Kappers et al., 2005; Kessler & Heil, 2011). The perception of In11 in a related species (cowpea) results in induced expression of biosynthetic terpene synthase (TPS) genes and the accumulation of specific volatile blends that include the homoterpenes (E)-4,8-dimethyl-1,3,7-nonatriene (DMNT) and (3E,7E)-4,8,12-trimethyl-1,3,7,11-tridecatetraene (TMTT) (Schmelz et al., 2006; Steinbrenner et al., 2022), important signaling compounds known to mediate indirect plant defenses (F. Li et al., 2018; W. Li et al., 2021; S. Zhou & Jander, 2022). We hypothesized that HIPV emission is impaired in In11-insensitive common bean plants. Indeed, we found that the responsive variety PI 311785 emitted a distinct volatile profile in response to wounding plus In11 compared to wounding alone, whereas W6 13807 responded with a much-reduced emission that was consistent with lack of In11 recognition (fig. S1B and C). Specifically, methyl salicylate (MeSA) was significantly higher in PI 311785 but not W6 13807 in response to In11, while the less abundant TMTT was uniquely present in In11-treated PI 311785, and DMNT showed a similar trend (Fig. 1B and Fig. 1C). These volatiles are known to serve as foraging cues for natural enemies, contributing to indirect plant defenses (Gouinguené et al., 2005; Rodriguez-Saona et al., 2011; S. Zhou & Jander, 2022).

We additionally measured MAPK3/6 phosphorylation, a canonical marker of plant pattern triggered immunity (PTI) (Ligterink et al., 1997; T. Sun & Zhang, 2022), in response to spray application of In11. We found that MAPK3 and MAPK6 orthologs were phosphorylated in PI 311785 but not in W6 13807 15 minutes after elicitation with In11 (Fig. 1D). Together, these results suggested a lack of activation in the In11-specific signaling cascade upstream of MAPK phosphorylation that might be required for HIPV emission in W6 13807.

To identify the potential genetic basis of In11 insensitivity, we generated short read re-sequencing data for W6 13807. Mapping to the Mesoamerican reference genome *P. vulgaris* cv. Labor Ovalle revealed a candidate 103 base pair (bp) deletion in W6 13807 without read support in the LRR domain of the Inceptin Receptor (INR) (Fig. 1E). The resulting protein product is expected to be truncated at the last LRR before the island domain (Fig. 1F), indicating that attachment to the plasma membrane and ligand-induced interaction with co-receptors is impaired (Snoeck et al., 2023; Y. Sun et al., 2022). Furthermore, F2 populations from the W6 13807 x PI 311785 cross showed segregation of inducible ethylene accumulation, indicating that sensitivity to In11 is conferred by a simply inherited locus. Genotyping of F2 individuals revealed co-segregation of the 103-bp deletion with response to In11 (fig. S1D), and thus we conclude that insensitivity is associated with a naturally occurring null allele, which we termed *inr-1*.

2.1.2 Direct defenses are impaired in *inr-1* deletion lines

To directly compare the effect of the *inr-1* allele on direct and indirect defenses, we generated three sets of near-isogenic lines (NILs) through independent W6 13807 x PI 311685 crosses, followed by six recurrent backcrosses with PI 311785, and final selection of *inr-1/inr-1* and *INR/INR* genotypes (fig. S2A). Consistent with data from parent lines and F2 individuals, the *inr-1/inr-1* NILs did not produce In11-induced ethylene (Fig. 2A) but retained the ability to recognize the bacterial peptide flg22 (fig. S2B), indicating that insensitivity to In11 is not due to a downstream signaling component required for immune responses. Importantly, *Spodoptera exigua* larval growth rate was 72.7% higher on *inr-1/inr-1* relative to *INR/INR* sibling NILs over a 5-day feeding time course (Fig. 2B). Together, these results imply that a functional allele of *INR* is needed for In11 to induce direct defense responses that enhance herbivore resistance in common bean.

To identify downstream components of the INR-dependent immune pathway, we generated RNA sequencing data from *P. vulgaris* cv. Red Hawk 1 hour after treatment with wounding + In11 compared to mock treatment (w + H_2O). Our analysis revealed In11-specific gene expression (fig. S3A and data S1) defined by 527 genes that were up-regulated by In11 treatment relative to wounding alone (fig. S3B). From these, 29.6% belonged to Gene Ontology (GO) categories associated with protein phosphorylation and regulation of transcription (fig. S3C and data S2). We selected the transcription factor *PvMYB* as an induced marker gene and found that it was highly upregulated in *INR/INR* but not in *inr-1/inr-1* NILs 1 hour after In11 treatment (Fig. 2C), suggesting that recognition of In11 by INR results in rapid herbivore-specific transcriptional reprogramming.

2.1.3 Impaired In11-induced volatile emissions in *inr-1* deletion lines negatively impacts predatory wasp attraction in a Mexican field

In their natural environments, plant's emission of specific blends of HIPVs recruit natural enemies (Kessler & Baldwin, 2001) that serve as indirect defenses by predating or parasitizing herbivorous pests (Turlings & Erb, 2018). To determine if *INR* mediates the emission of specific HIPV blends we measured and characterized In11-induced volatiles in the common bean NILs grown under laboratory conditions.

We found that *inr-1/inr-1* NILs did not emit the typical HIPV blend that is normally induced upon In11 perception, but rather emitted volatiles that bean plants release after wounding alone. In contrast, *INR/INR* NILs produced the characteristic blend of volatiles in response to treatment with a physiologically relevant concentration of In11 (Schmelz et al., 2006) or the oral secretion (OS) derived from bean-fed fall armyworm (*Spodoptera frugiperda*) caterpillars (Fig. 3A). Among other compounds, DMNT, MeSA, and TMTT were predominantly detected in the

headspace of In11- or OS-treated *INR/INR* NILs (Fig. 3B) and at significantly higher levels than those with the w + H_2O treatment (Fig. 3C), thus contributing to the distinct volatile blends observed between the NIL genotypes. These results are consistent with previous reports that In11 is the primary elicitor of plant responses found in OS (Schmelz et al., 2006), and indicate that INR is required for specific volatile emission in response to synthetic and natural In11 found in the caterpillar OS.

Lastly, we hypothesized that a functional INR would mediate the recruitment of natural enemies to plants treated with In11 or herbivore-derived OS. To test this in nature, we adapted a sentinel-prey field assay (Grof-Tisza, Turlings, et al., 2024) and recorded wasp visitation to pairs of common bean NILs (Fig. 3D) treated with wounding plus water, OS or a physiologically relevant concentration of In11 in an experimental agricultural field in Oaxaca, Mexico over two field seasons in 2023 and 2024. There was considerable wasp activity across both seasons as indicated by plot visitation rates ranging from 55.1% to 86.9% (Fig. 3E). From the plots visited in 2023, we found that pairs of the responsive parent PI 311785 and *inr-1/inr-1* NIL plants treated with OS derived from bean-fed fall armyworm demonstrated 40% reduction of attack to sentinel caterpillars by predatory *Polybia sp and Mischocyttaurus sp* wasps to *inr-1/inr-1* plants. Similarly in 2024, pairs of NILs treated with a physiologically relevant concentration or excess of In11 showed a 40% reduction of attack to *inr-1/inr-1* across two independent experiments, while there was no difference in visitation for mock treated plants (w + H_2O) (Fig. 3F). These results demonstrate that INR mediates the In11-triggered recruitment of natural enemies in the field.

Taken together, our laboratory and field assays using the response to In11 in common bean landraces and near-isogenic lines (NILs) provide evidence that the INR receptor mediates In11-

induced defenses and ultimately the interaction between plants, herbivorous pests, and predatory wasps in an agricultural ecosystem. In agreement with previous findings, we report that In11 is the primary elicitor of induced defenses found in herbivore OS, and that physiologically relevant concentrations of In11 elicit HIPV emission. The recruitment of predatory wasps to In11- or OS-treated plants in the field requires a functional INR receptor and is associated with the emission of specific volatile blends. Our findings provide a receptor-mediated mechanism for HIPV emission in legumes and suggest that the success of agroecological practices like Milpa and pull-push may be partially attributed to the co-planting of INR-encoding legume species.

2.2 Materials and Methods

2.2.1 Plant growth

Seeds from all common bean lines (*Phaseolus vulgaris*) were sown on Sunshine mix No. 4 and germinated in a reach-in PGC-FLEX growth chamber (Argus/Conviron) at 26ºC, 70% relative humidity, and 500 µMol of light with a 12/12 h light/dark photoperiod. On day 13, plants with a fully expanded first trifoliate were transferred to a walk-in PGW40 chamber (Conviron) at 25°C, 50% relative humidity, and 500 µmol m^{-2} s^{-1} light with a 12/12 h light/dark photoperiod prior to experimentation.

2.2.2 Genotyping and resequencing

Samples were taken from primary leaves of common bean plants and flash frozen in liquid nitrogen. Frozen samples were then ground to a fine powder using a mixer mill (Retsch MM400, 2008, Part No. 20.745.0001) and genomic DNA extraction was performed using the NucleoSpin® Plant II kit (Machery-Nagel, 740770.520) according to the manufacturer's instructions. The concentration and quality of the extracted DNA was assessed by NanoDrop

One (ThermoFisher Scientific ND-ONE-W) and gel electrophoresis. Common bean accessions were genotyped using gene-specific primers (table S2) flanking a ~1000 bp region that contained a 103 bp indel in the PvINR locus. PCR reactions were performed using gene-specific primers with DreamTaq PCR MasterMix (Thermo Fisher Cat. No. K1081).

For resequencing, Illumina libraries were synthesized and 150 bp paired-end reads were sequenced with the Novaseq 6000 instrument by Novogene Inc. Reads were mapped to the *Phaseolus vulgaris* reference genome using BWA-MEM (v 1.2.3) with default settings. Polymorphisms were called with FreeBayes (v 1.1) (41) and relatedness calculated with TASSEL (42).

2.2.3 Generation of near-isogenic lines (NILs)

Mature anthers from opened flowers of *Phaseolus vulgaris* W6 13807 containing the 103-base pair (bp) deletion in the *PvINR* locus (*inr-1*) were excised and used to pollinate the immature stigma from unpollinated, unopened host flowers of Phaseolus vulgaris PI 311785 without a deletion (*INR*). The resulting seeds were planted and genotyped as described. Heterozygous plants were kept and crossed with the recurrent line, PI 311785. This procedure was repeated until backcross six (BC6) lines were generated. For the BC five and six lines, heterozygous plants were allowed to self-pollinate and their offspring genotyped to select homozygous individuals with *INR/INR* or *inr-1/inr-1* genotypes.

2.2.4 In11-induced ethylene gas production and gene expression on common bean

The In11 peptide (ICDINGVCVDA) was synthesized (Genescript Inc) based on the cATP synthase sequence from cowpea (*Vigna unguiculata*) (43) and reconstituted in water.

For ethylene gas measurements, the left and middle leaflet of the first trifoliate or the two primary leaves of a common bean plant were lightly wounded on both sides of the main vein using a fresh razor blade to remove the cuticle, and 10 µL of water alone or 1 µM In11 in 10 µL of water were distributed between the wounds using the tip of a pipette. After 1 hour, leaflets or primary leaves were excised at the petiole using a razor blade and placed in sealed tubes for 1 hour before sampling 1mL of the headspace. Ethylene gas was measured as previously described (44) with a gas chromatograph (HP 5890 series 2, supelco #13018-U, 80/100 Hayesep Q 3FT x 1/8IN x 2.1MM nickel) with flame ionization detection and quantified using a standard curve (Scott, 99.5% ethylene, Cat. No 25881-U).

For gene expression analyses, the middle leaflet of the first trifoliate or the primary leaves of a common bean plant were lightly wounded on both sides of the main vein using a fresh razor blade to remove the cuticle, and 10 µL of water alone or 1 µM In11 in 10 µL of water were distributed between the wounds using the tip of a pipette. After 1 hour, leaf samples were collected by removing the treated area with a 0.6 cm^2 cork borer and flash frozen in liquid nitrogen. All samples were stored at -80ºC until RNA extraction

2.2.5 In11-induced MAPK phosphorylation and immunoblotting

To measure In11-induced MAPK phosphorylation, we sprayed the first fully expanded trifoliate of 17-day old common bean landraces and NILs with approximately 3 mL of a 0.01% Silwet L-77 (PhytoTech) solution with or without 1 µM In11. Leaf punches were collected from the three leaflets 15 minutes after treatment, alongside untreated controls, and combined before freezing in liquid nitrogen.

Total proteins were extracted using a 3X Laemmli (50 mM Tris-Cl pH 6.8, 6% SDS, 30% glycerol, 16% β-mercaptoethanol, and 0.006% bromophenol blue sample buffer) and separated by SDS-PAGE before transfer to a nitrocellulose membrane. To detect phosphorylated MAPK proteins, we used an anti phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) antibody (Cell signaling, #4370) at 1:2000 diluted in 1% Tris-buffered saline with Tween-20 (TBST) buffer, with secondary antibody anti-rabbit IgG-HRP conjugate (Sigma) at 1:10,000, followed by chemiluminescent visualization. Ponceau S solution (Sigma-Aldrich) staining was used to verify equal loading. The molecular weight of MAPK3 (42.66 kDa) and MAPK6 (45.06 kDa) were estimated based on their amino acid sequence.

2.2.6 In11-induced volatile production

The seeds of the parental landrace and NILs were germinated on wet cotton in a Petri dish closed with Parafilm and covered with aluminum foil at 28°C for 48 hours. Subsequently, the seedlings were transplanted in individual cylindrical plastic pots filled with commercial soil (Einheitserdewerke Patzer Gebrüder Patzer GmbH & Co KG, Germany) and kept in a controlled growth chamber (L16:D8 light, 26°C 2, and 60% R.H.) until they were had two fully expanded trifoliates. The plants were watered every other day.

The three leaflets of a trifoliate in 12–16-day old plants were mechanically damaged by scratching the upper leaf surface with a razor blade and applying water (w + H_2O) or 10 µL of 45 nM or 150 nM In11. The plants received the treatment twice: at 3:00 PM in the afternoon before the experiment and at 8:00 AM on the day of the experiment. The volatiles were collected as previously described (45) 1 hour after the second treatment, for 3 hours. Briefly, the plants were individually placed in a glass bottle with clean and humidified air pushed inside through Teflon tubing at a rate of 0.8 L/min and the volatiles were trapped using 25mg of 80-100 mesh of

HayeSep Q absorbent (Sigma, Switzerland) that was connected via Tygon tubing to a vacuum pump that pulled out the air at a rate of 0,5L/min. After 3 hours of volatiles collection each trapping filter was eluted with 100 µL of dichloromethane (Honeywell, Riedel-de Haën, DE) and 200 ng of n-nonyl acetate in 10 µL dichloromethane was added as internal standard. Finally, all the samples were stored in a -80°C freezer before analysis.

The samples were analyzed using a gas chromatograph (Agilent 7890B) coupled to a mass spectrometer (Agilent 5977B GC/MSD) in TIC mode. 2 µL of each sample was injected in pulsed spitless mode onto an Agilent HP-5MS column (30 m length x 0.25 mm diam. and 0.25m thick). After injection, the temperature program was the following: 40°C for 3min, increased to 100°C at a rate of 8°C/min and subsequently at 5°C/min until reaching 200°C and a post run of 3 min at 250°C. Helium was used as a carrier gas with a constant flow of 1.1 mL/min. For the identification of the volatiles, the mass spectrum of each compound was compared with the ones from two reference databases, NIST mass spectral library and PBM Quicksearch library (U.S. Department of commerce). Identifications were further confirmed based on the spectra and retention times of authentic standards. The relative quantities (ng/hours) of the most relevant compounds were calculated considering their response factors and internal standards peak areas.

2.2.7 flg22-induced reactive oxygen species (ROS) accumulation

The flg22 peptide was synthesized (Genescript Inc) according to canonical flagellin sequence (46) and reconstituted in water. For the assay, punches were taken from the primary leaves of common bean NILs with a 4 mm biopsy punch and floated overnight in 150 µL of H_2O using individual cells of a white 96-well white bottom plate (BRANDplates F pureGrade S white). After overnight incubation, ROS production was measured upon addition of a 100 µL assay solution which contained 10 µg/mL luminol-horseradish peroxidase, 17 µg/mL luminol and H_2O

or flg22 to a final concentration of 2 μM. Luminescence was quantified with a TECAN SPARK plate reader every minute for 1 hour using an integration time of 500 ms, and the average ROS production for each plant was the mean of four leaf discs at each time point.

2.2.8 Transcriptomics

To characterize transcriptional changes in response to wound and In11 treatment in common bean plants, we treated the middle leaflet of the first fully expanded trifoliate of *Phaseolus vulgaris* cv. Red Hawk by scratch wounding and adding water (w + H_2O) or 1 μM In11 (w + In11). Samples were collected 1 hour later and flash frozen in liquid nitrogen. Total RNA extraction was performed using the Spectrum Plant RNA kit (Sigma) and mRNA was isolated with the NEBNext High Input Poly(A) mRNA Isolation Module (New England BioLabs). Paired-end Illumina 2x150 bp libraries were prepared with NEBNext Ultra™ RNA Library Prep Kit (New England BioLabs) and multiplexed with NEBNext® Multiplex Oligos for Illumina® (Index Primers Set 1). The sequencing was performed by Novogene Inc. For gene expression analyses, the reads were mapped to the reference *Phaseolus vulgaris* v2.1 genome available in Phytozome13 (47) and quantified using the –quantMode in STAR (48). Differential gene expression analyses were performed using DESeq2 (49) implemented in R.

2.2.9 RNA extraction, cDNA synthesis and qRT-PCR

Total RNA was extracted using the Trizol (Invitrogen) method. Briefly, frozen samples were ground to a fine powder using a mixer mill (Retsch MM400, 2008), and 900 μL of Trizol were added and homogenized using a vortex. The cell debris was separated by centrifugation, and the supernatant recovered in a clean tube. 200 μL of chloroform were added to the supernatant for phase separation. After centrifugation, the aqueous phase (top layer) was recovered, and the RNA was precipitated using an equal amount of isopropanol. RNA pellets were washed using 70%

ethanol and resuspended in water. The concentration and quality of the extracted RNA was assessed by NanoDrop One (ThermoFisher Scientific ND-ONE-W) and gel electrophoresis, respectively. As DNA contamination was evident from the gel electrophoresis, the RNA was treated with TURBO DNA-free kit (Invitrogen) following manufacturer's instructions. cDNA was synthesized using 1µg of RNA using the SuperScript IV Reverse Transcriptase Kit (Thermo Fisher, Cat. No. 18090050). qPCR reactions were performed using gene-specific primers with Power SYBR Green PCR Master Mix (Thermo Fisher, Cat. No. 4367659). Relative expressions were quantified using the ΔΔCq method, with ΔCq values normalized to the common bean UBQ gene and made relative to w + H_2O treatment.

2.2.10 Herbivory assays

First instar beet armyworm (*Spodoptera exigua*) larvae were obtained from Benzon Research Inc (Carlisle, PA). Upon receipt, larvae were incubated in their original packaging (with artificial diet) under dark conditions at 28ºC for 36-48 hours to accelerate molting. 2nd instar larvae were weighed using a digital scale (Mettler Toledo International Inc) and placed in plastic cups without artificial diet. Once all larval masses were recorded, one single larva was placed on the first trifoliate or primary leaf of *INR/INR* and *inr-1/inr-1* plants, and leaves were enclosed in a transparent mesh bag to contain the larva. Infested plants were kept in a growth chamber at 25°C, 50% RH, and 500 µMol light, and larvae weights were recorded again 5 days later.

2.2.11 Sentinel-prey field assay

Field assays were conducted in 2023 and 2024 in a 925 m^2 agricultural field near Bajos de Chila, Oaxaca, Mexico (15,91818° N, 97,16401° W) (Fig. 3).

For the 2023 field season, seeds of the parental PI 311785 landrace and the *inr-1/inr-1* NIL were sown 1.5 L plastic pots containing field-collected soil. To prevent herbivory, plants were grown in large mesh cages (2 x 2 x 4 m; BioQuip, CA, USA) for 13 days and watered every other day. Plants were treated by removing 2 cm^2 of cuticle on the adaxial side of an expanded leaflet with sandpaper and adding water (w + H_2O) or fresh oral secretions (w + OS) derived from bean-fed fall armyworm larvae (*Spodoptera frugiperda*). The treatments were imposed the evening before and the morning of the experiment. After the second damage treatment, one randomly selected freeze-killed caterpillar (defrosted) was pinned to the center of each leaf (two caterpillars per plant). At 9:00 am, pairs of potted PI 311785 and *inr-1/inr-1* treated by either wounding and adding water (w + H_2O) or OS (w + OS) and with pinned caterpillars were arranged in six spatial blocks in plots within the experimental field. Paired pots were approximately 30 cm apart. Blocks were visited continuously over 90 min, and the order in which sentinel caterpillars from the different treatments were subject to committed visits by one or more wasps in each block was recorded. The experiment was repeated a second day using the same plants but arranged in different treatment pairs.

For the 2024 field season, seeds of *INR/INR* and *inr-1/inr-1* sibling NILs were grown as described above. 13-day old plants were treated at 7:30 am by removing 2 cm^2 of cuticle on the adaxial side of two opposing leaflets on the 1st fully expanded trifoliate with a razor blade and adding 45 nM In11 (w + In11) or water (w + H_2O). After imposing the same treatment 24 hours later, freeze-killed fall armyworm caterpillars were pinned to the treated leaflets as described above.

Beginning at approximately 08:00 am, treatment pairs (*w* + In11 $_{INR/INR}$ and *w* + In11 $_{inr\text{-}1/inr\text{-}1}$) were observed for wasp visitation in five spatial blocks of the experimental field. For each pair,

we recorded the time of the first wasp visit, defined as a committed attack on a larva, as well as the species of the first wasp to arrive. Following the initial visit in each block, we recorded the number of wasps on each plant at 5, 15, 30, 45, 60, 75, and 90 minutes. The experiment was repeated on a second day using the same plants, but with treatments reassigned to new pairings. Three days later, the trial was repeated on a new set of In11-treated plants, and seven days after that with plants that were only water treated after the razorblade-inflicted damage

2.2.12 Statistical analysis

All statistical analyses were conducted in R (R version 4.3.2 or 4.3.3, R Core Team 2023). We first confirmed that the residuals fulfilled assumption of normality using a Shapiro Wilk Test and qq plots and then implemented t-test to detect differences in the mean ethylene gas accumulation, gene expression, and mean growth rate of beet armyworm caterpillars feeding on *inr-1/inr-1* or *INR/INR* NILs. Generalized linear models were used to assess the effect of treatments on volatile emissions. A Gamma error distribution was used to account for the non-normality of emission data. Estimated marginal means were calculated using the emmeans package (Lenth, 2023), and pairwise comparisons were conducted with a false discovery rate (FDR) correction for multiple testing.

To visualize overall differences in volatile profiles across treatments, we conducted principal component analysis (PCA). Volatile compound emissions were log-transformed and scaled to unit variance prior to analysis.

Generalized linear mixed models using the glmmTMB package (50) were used to assess the effect of treatment (2023, OS; 2024, In11 and H_2O) on the removal of sentinel caterpillars from bean plant pairs in the field. To account for the hierarchical structure of the experiment, we

included random effects for plots (plant pairs) nested within spatial blocks, which were in turn nested within each trial. Date was included as a fixed effect in the model, as only two trial days were used and random effects with fewer than five levels are generally discouraged due to unstable variance estimates (51).

2.3 Figures

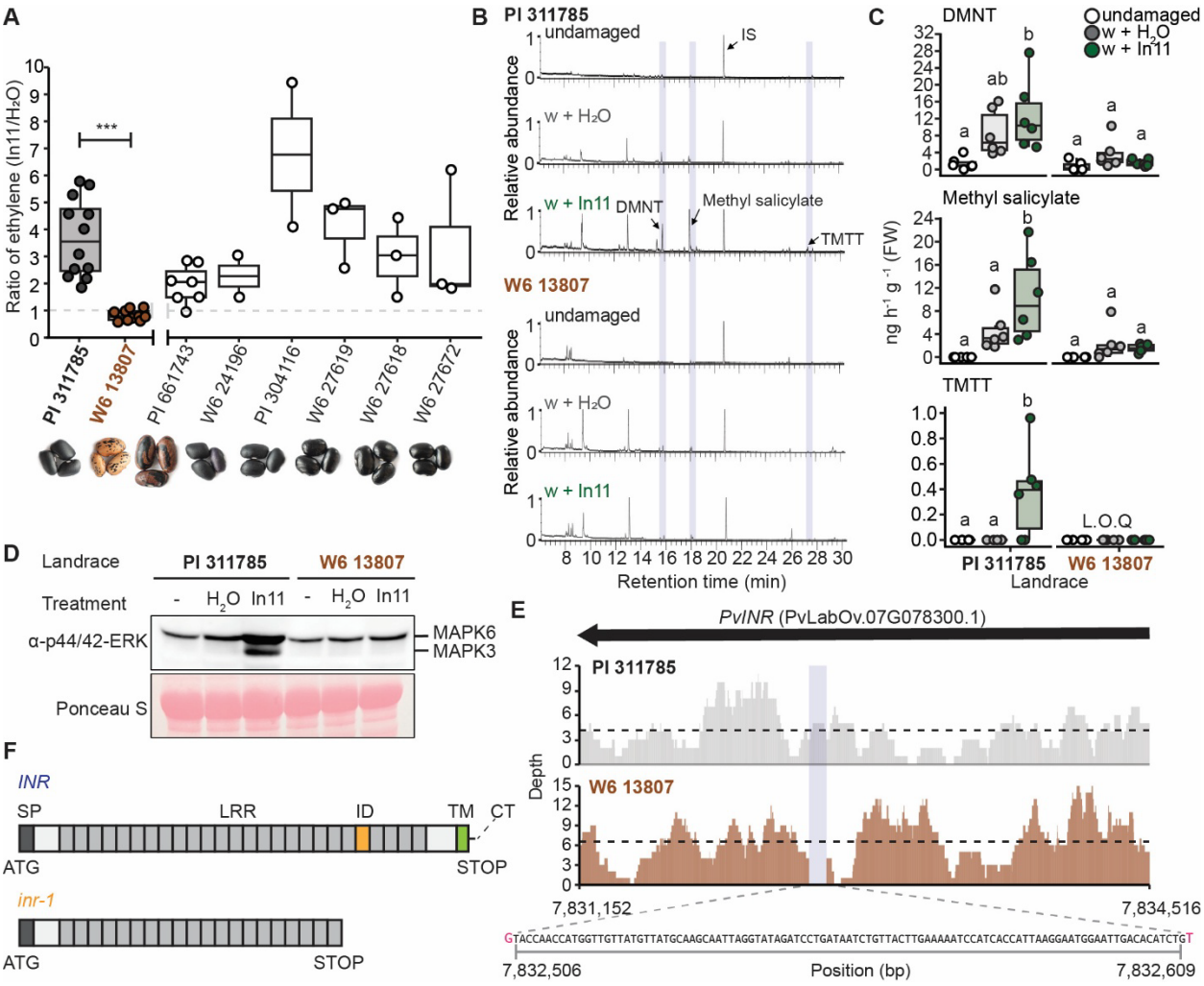

Figure 1. *inr-1* is a naturally occurring deletion allele among Honduran bean landraces. (A) Ratio of ethylene gas induced by wounding (w) + 1 µM In11 treatment relative to mock treatment (w + H_2O) for select common bean (*Phaseolus vulgaris*) landraces. Seed coat diversity among landraces is shown. Boxplots represent the 1st and 3rd quantile and the median. Each dot represents

an independent plant. Comparing crossed varieties PI 311785 and W6 13807, significant differences in the mean were determined by a two-sided t-test (n = 10-12, α= 0.05), ***, $p<0.001$. (**B)** Representative GC-MS chromatograms of volatiles collected from bean varieties PI 311785 and W6 13807 after the plants had either been left unharmed (undamaged), wounded and treated with water (w +H_2O), or wounded and treated with 1 µM ln11 (w +ln11). (**C)** Quantified concentrations of individual volatile organic compounds. Different letters indicate significant differences within treatments of each variety. Linear model [family, Gaussian] followed by pairwise comparisons of least squares means. L.O.Q = Limit of quantification. (**D)** MAPK 3 and 6 phosphorylation measured 15' after spraying H_2O or 1 µM In11 on indicated landraces. Protein loading is shown by Ponceau stain. **(E)** Read depth coverage of re-sequenced PI 311785 and W6 13807 landraces mapped to the *PvINR* locus in the Mesoamerican (Guatemalan) *Phaseolus vulgaris* Labor Ovalle v1.1 genome. The average read depth is shown as dashed lines. The sequence and position of the 103-base pair (bp) deletion is also indicated with a vertical bar. The nucleotides flanking the deletion are shown in magenta. (**F)** Domain architecture of INR receptor and *inr-1* deletion variant with resulting frameshift mutation. SP: signal peptide, ID: Island domain, TM: Transmembrane domain, CT: Cytoplasmic tail.

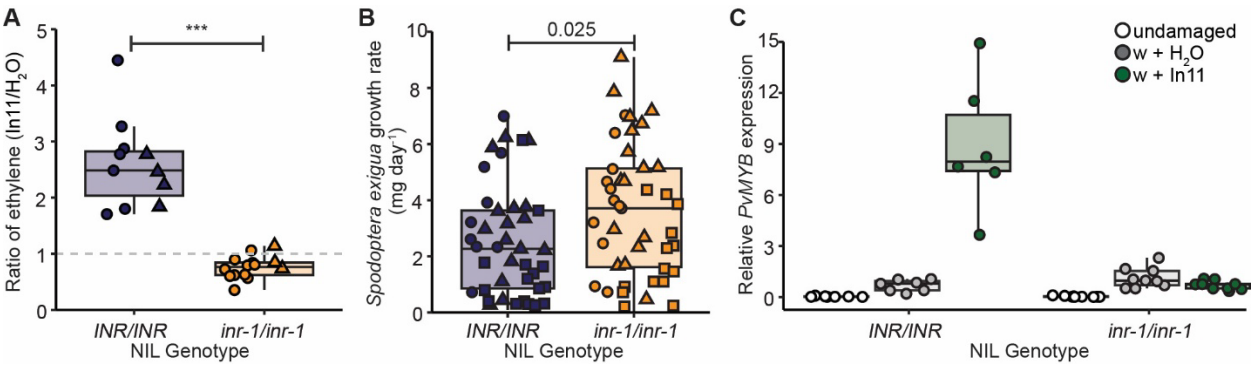

Figure 2. *inr-1* is associated with reduced In11-induced defense responses and increased herbivore growth in near-isogenic lines. (A) Ratio of ethylene gas induced by wounding (w) + 1 µM In11 treatment relative to mock treatment (w + H_2O) for common bean near-isogenic lines (NILs) with *INR/INR* or *inr-1/inr-1* genotype. Each dot represents an independent plant (n=11-15). Boxplots represent the 1st and 3rd quantile and the median. Significant differences in the mean were determined by a two-sided t-test (n = 11-13, α= 0.05), ***, $p<0.001$. (**B)** Growth rate of 2nd instar *Spodoptera exigua* larvae feeding on common bean NILs for 5 days. Each dot represents one larva feeding on an individual plant. Significant differences in the mean were determined by a two-sided t-test (n = 40-43, α= 0.05). Symbols in panels A and B indicate data from independent experiments using a mixture of NILs 3, 4 and 8 (fig. S2C). **(C)** *PvMYB* (*Phvul.001G215100*) marker gene expression 1 hour after wounding and adding water or 1 µM In11. Gene expression was normalized against *PvUBQ*.

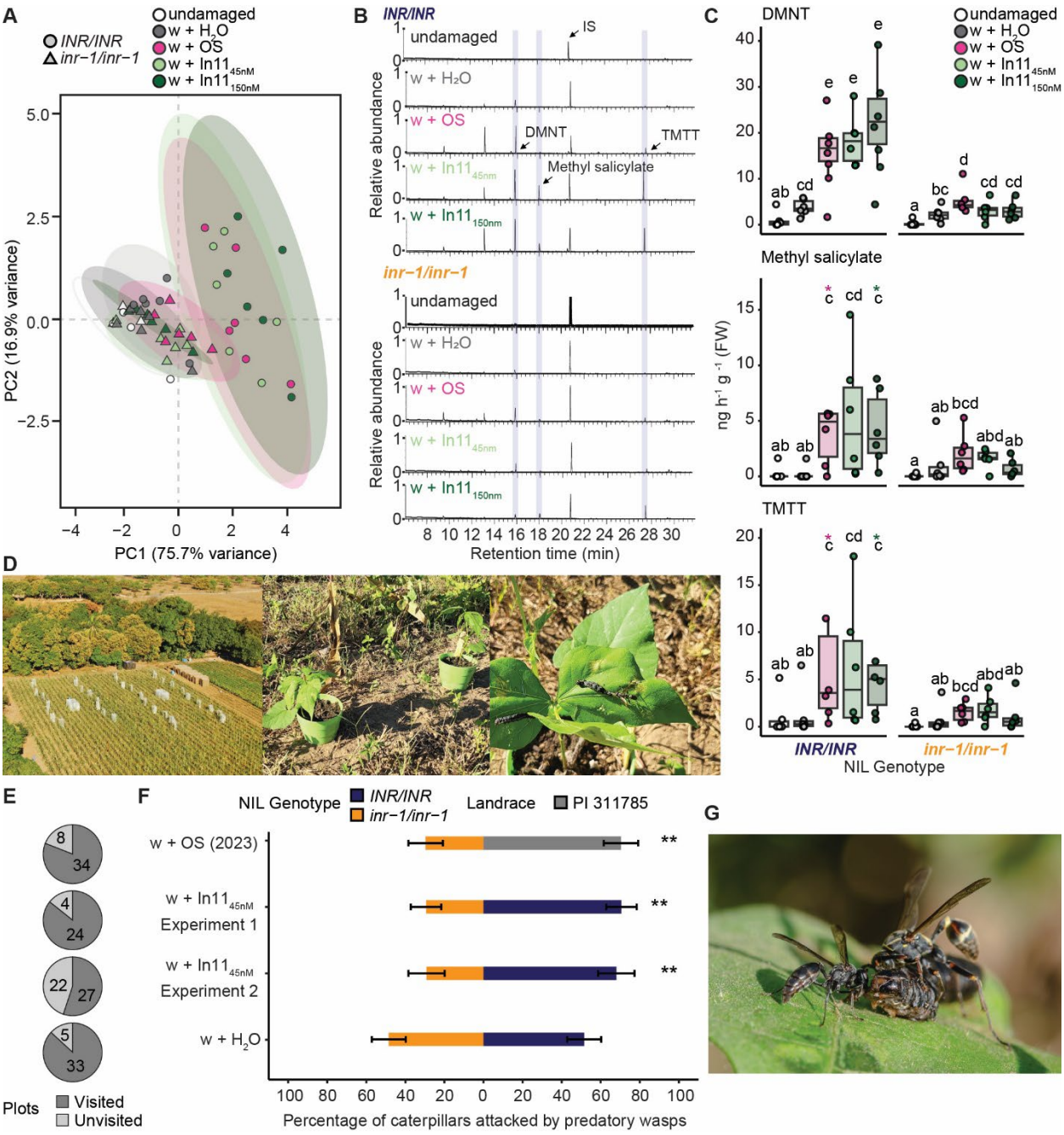

Figure 3. *inr-1* is associated with impaired In11-induced volatile emission and reduced attraction of predatory wasps in near-isogenic lines of common bean plants in Mexico.

Panels **A–C** show results from a shared headspace volatile collection experiment comparing NILs (INR/INR and *inr-1/inr-1*) subjected to five treatments: undamaged, mechanical wounding (w + H_2O), wounding with fall armyworm (FAW, *Spodoptera frugiperda*) oral secretions (w + OS), a physiologically relevant concentration of In11 (45 nM, w + $In11_{45nM}$), or an excess of In11 (150 nM, w + $In11_{150\ nM}$) (n = 6). (**A**) PCA ordination plot of VOC profiles, illustrating treatment- and genotype-driven separation. Ellipses represent 95% confidence intervals. (**B)** Representative GC-MS chromatograms of headspace induced volatile organic compounds (VOCs) (**C)** Boxplots of selected VOCs, known to be important in the indirect defenses of plants, 1 hour after treatment. Letters denote significant differences from a generalized linear model (Gamma error) with pairwise least squares means comparisons ($\alpha = 0.05$). Asterisks indicate extreme values not shown for clarity. (**D**) Field experimental setup in Oaxaca, Mexico, showing block layout of where wasp visits were monitored, treatment-assigned NIL pairs, and pinned sentinel FAW caterpillars. (E) Visitation rates to each NIL/treatment pair. (**F)** Bar plots show the mean percentage of caterpillars attacked by social wasps across treatments and two field seasons (2023, 2024). Statistical analysis was performed using a binomial generalized linear mixed model. **, $p<0.01$. (**G)** Representative image of a committed attack by *Polybia sp* and *Mischocyttaurus sp* on a sentinel FAW caterpillar.

2.4 Supplementary figures

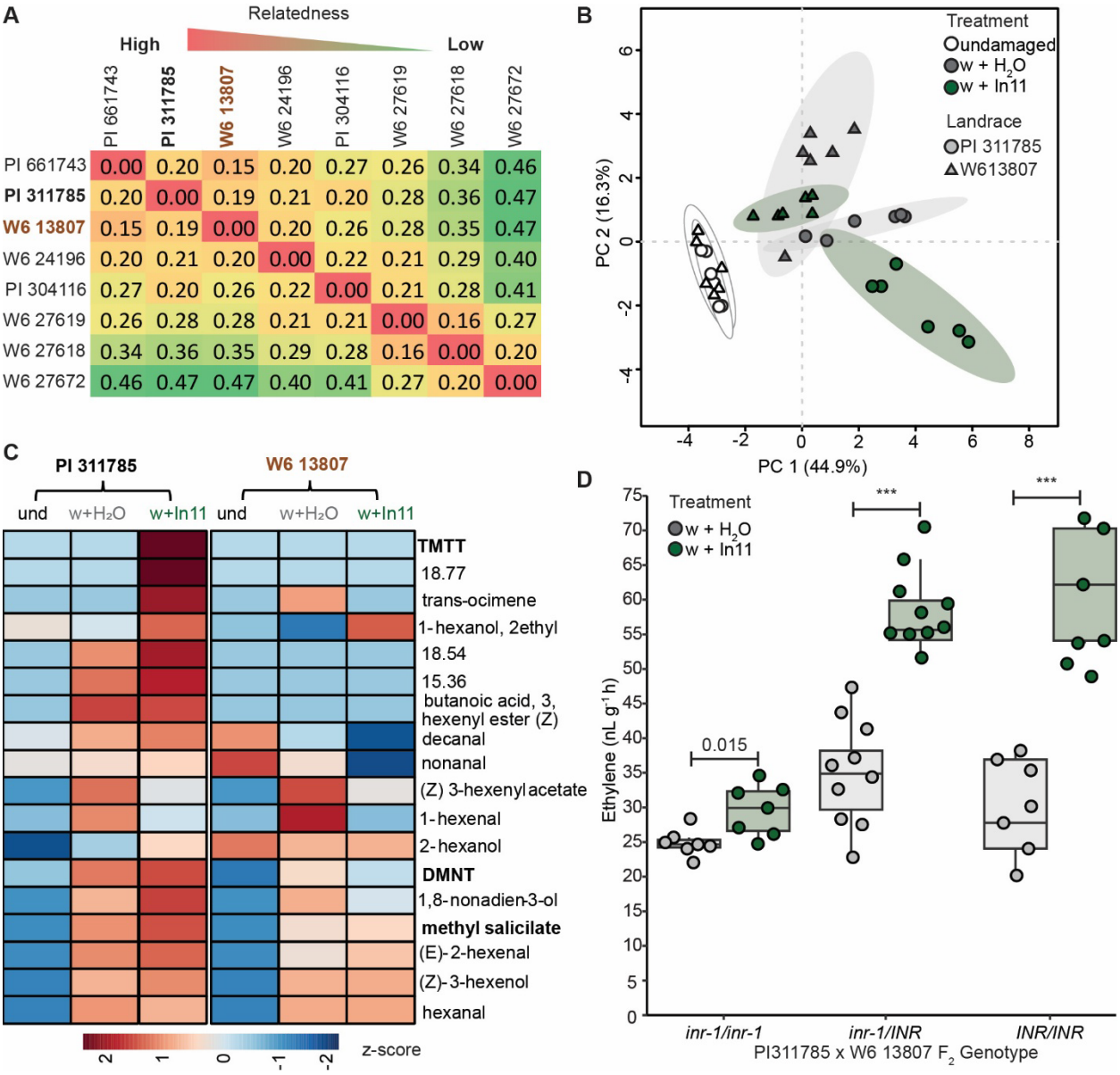

Fig. S1. (**A**) Relatedness matrix of Honduran common bean landraces used in the ethylene screen. Cross-compatible In11-responsive (PI 311785) and non-responsive (W6 13807) landraces are indicated in bold. (**B**) Principal component analysis (PCA) of the volatile profiles from undamaged and 2 hours after wounding (w + H_2O) or In11 (wound + In11) treatment in selected

common bean landraces (n=6). (**C**) Relative abundance of specific compounds in the volatile blends. **D.** Ethylene production of F2 individuals from a PI 311785 x W6 13807 cross.

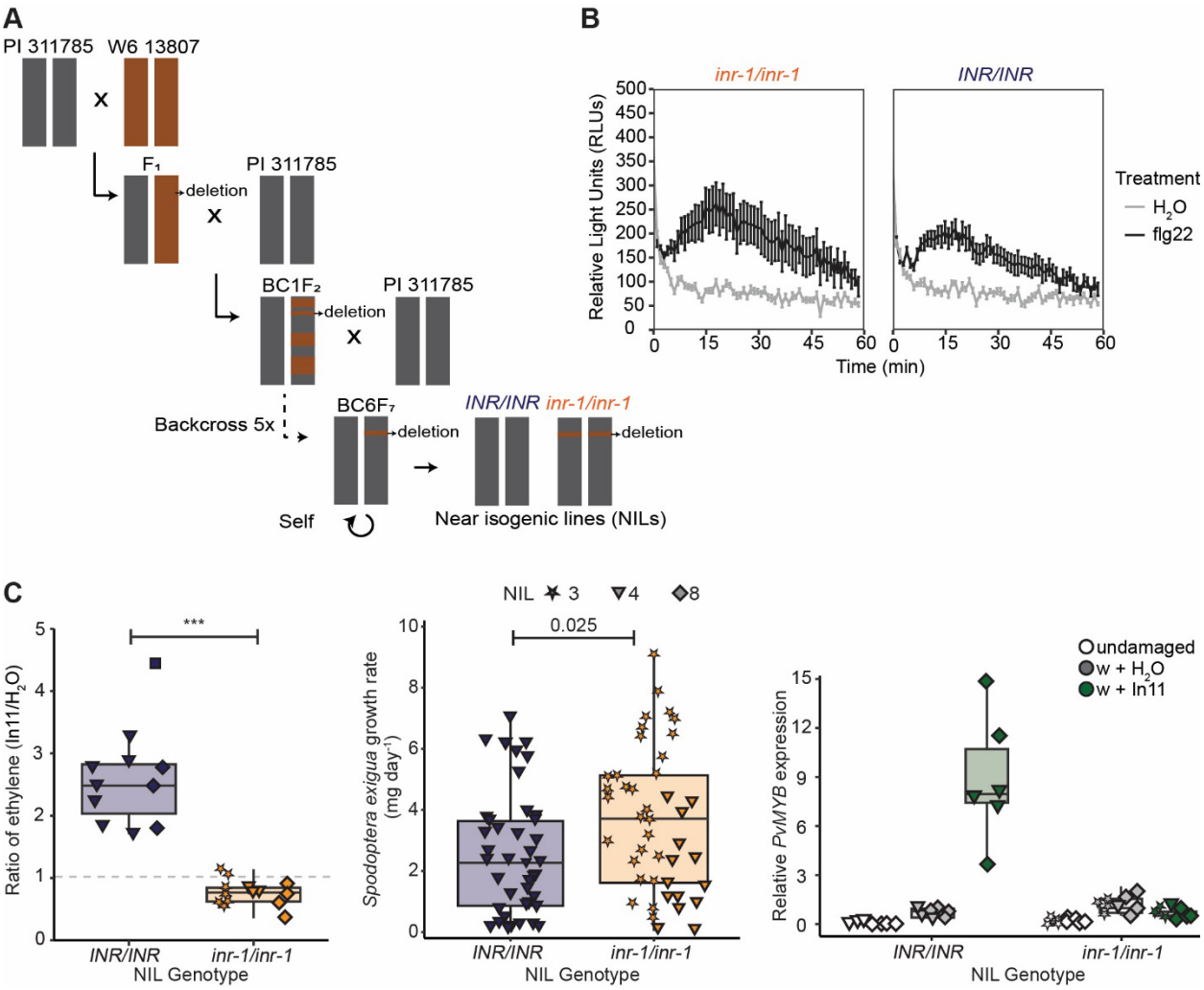

Fig. S2. flg22-induced reactive oxygen species (ROS) burst in common bean NILs. (A) Back-crossing scheme of sexually compatible In11-responsive (PI 311785) and non-responsive (W6 13807) landraces to generate *INR/INR* and *inr-1/inr-1* near-isogenic lines (NILs) in the PI 311785 background. NILs labeled as lines 3, 4, and 7 derived from three independent initial crosses are used in the work, each unresponsive to In11. **(B)** Oxidative burst produced by

primary leaves from common bean NIL 4 treated with 2 µM flg22 peptide and measured in relative luminescence units (RLU). Results shown are means ± s.d. (n = 3 plants).

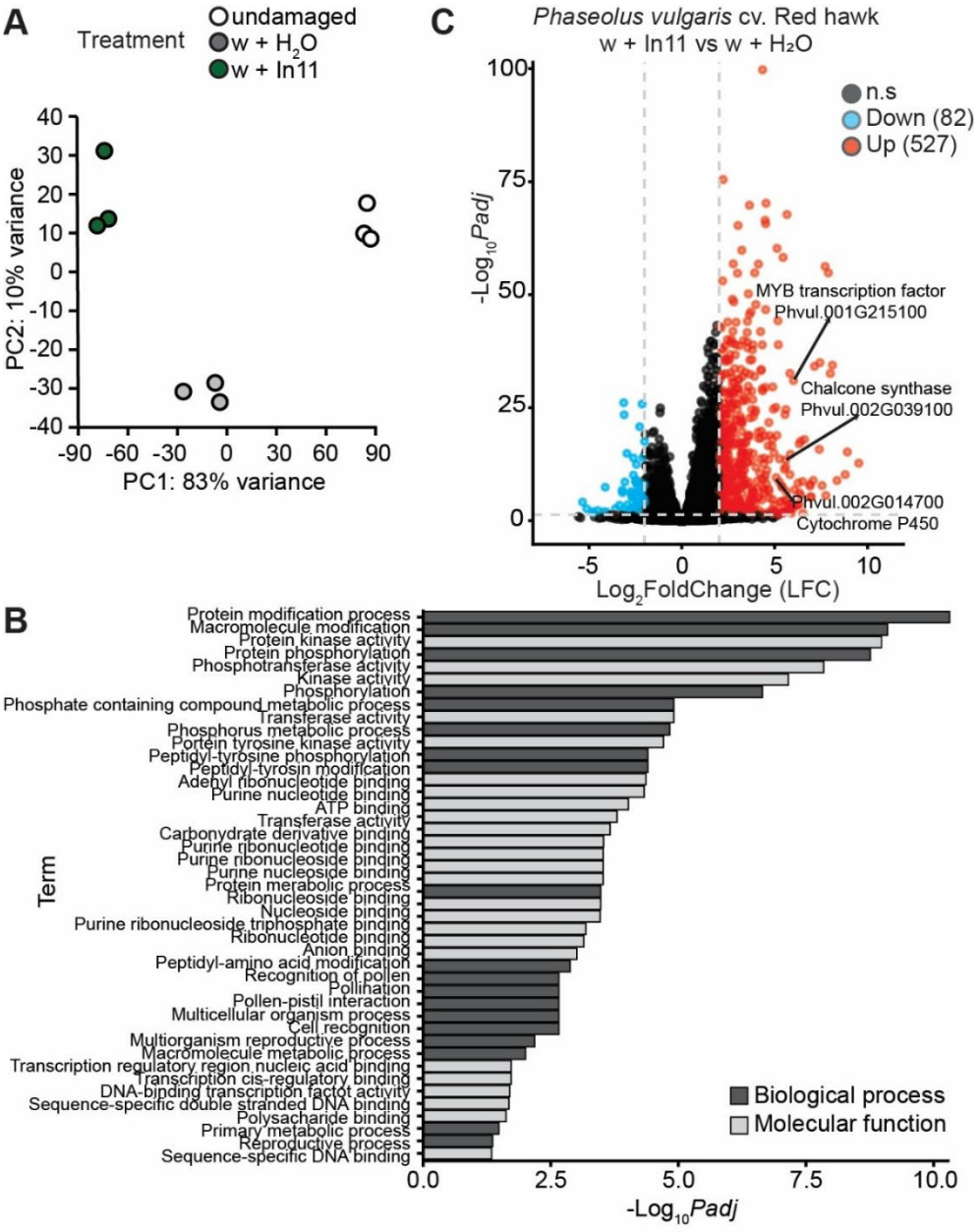

Fig. S3. RNAseq analysis of In11-treated common bean plants. (**A)** Principal component (PC) analysis of differentially expressed genes (DEGs) across all samples. (**B)** Volcano plot of DEGs 1 hour after wounding (w) + 1 µM In11 vs wounding alone (w + H2O) treatment in

Phaseolus vulgaris cv. Red Hawk. Selected marker genes *PvMYB*, *PvCHS*, and *PvCYP* are indicated. (**C)** Enriched Gene Ontology (GO) categories among In11-upregulated genes.

Chapter 3. Structural novelty in a cowpea protease inhibitor confers specific bioactivity against herbivores

3.1 Introduction

Plants defend against insect herbivores by recognizing herbivore-associated molecular patterns (HAMPs) and subsequently accumulating specific antiherbivore defenses. Among these, proteinase inhibitors (PIs) are a prominent example of inducible direct defenses that negatively impact caterpillar survival by inhibiting digestive proteases (Ryan, 1990). In tomato (*Solanum lycopersicum*) and potato (*S. tuberosum*), the herbivore-induced accumulation of the damage-associated molecular pattern (DAMP) systemin (Pearce et al., 1991) and subsequent recognition by the cell surface receptor kinase (RKs) systemin receptor 1 (SYR1) (Wang et al., 2018; Zhou et al., 2025) leads to the accumulation of Proteinase inhibitor I and II (PINs). Similarly, the recognition of In11 by the receptor-like protein (RLP) Inceptin Receptor (INR) in cowpea (*Vigna unguiculata*) (Steinbrenner et al., 2020) results in the upregulation of several members of an unrelated class of proteinase inhibitors termed Kunitz Trypsin Inhibitors (KTIs) (Guayazán Palacios et al., 2024; Schmelz et al., 2006; Steinbrenner et al., 2022), indicating that protease inhibition is a prevalent antiherbivore defense strategy in plants.

KTIs are members of the I3A family of serine protease inhibitors (Grosse-Holz & van der Hoorn, 2016; Rawlings et al., 2004), first characterized in soybean seeds (Kunitz, 1945) and crystallized in a complex with porcine trypsin (Blow et al., 1974; Song & Suh, 1998). These proteins are structurally conserved and functionally versatile targeting α-amylase, subtilisin-like proteases, PLCPs, trypsin, chymotrypsin and cathepsin D. Notably, the I3 family also includes miraculin, a taste-modifying glycoprotein found in miracle berry (*Synsepalum dulcificum*) (Brouwer et al., 1968; Kurihara & Beidler, 1968), which alters sweet perception in humans by binding the

Human sweet Taste Receptors hT1R2-hT1R3 (Koizumi et al., 2011), but has not been reported as a protease inhibitor.

KTIs typically consists of a β-trefoil fold stabilized by multiple disulfide bridges, which are important for protein stability. In most KTIs, two disulfide bonds link conserved cysteine residues, preserving the inhibitor's tertiary structure under denaturing conditions like high pH and protease-rich environments commonly found in the Lepidoptera digestive systems (Dow, 1992). The inhibitory function of KTIs is mediated through a canonical substrate-mimicry mechanism known as the standard or Laskowski mechanism (Laskowski & Kato, 1980). KTIs act as pseudo-substrates, binding tightly but reversibly to the active sites of target serine proteases. The interaction occurs primarily through a reactive site loop, which projects from the inhibitor surface and inserts into the enzyme's catalytic cleft. This loop contains a conserved P1 residue, typically a lysine (K) or arginine (R), that fits into the S1 specificity pocket of the protease. Hydrogen bonding and hydrophobic interactions stabilize the enzyme-inhibitor complex, effectively blocking substrate access without being cleaved, or in some cases, being cleaved and rapidly re-ligated in a non-productive turnover. The surrounding P2 and P1'–P4' positions further contribute to specificity and binding affinity.

The dynamic interaction between plant protease inhibitors and insect digestive proteases may represent an evolutionary arms race, analogous to the gene-for-gene model observed in plant-pathogen interactions, in which plants deploy inhibitors and insects evolve insensitive or novel proteases to escape inhibition (Jongsma & Bolter, 1997; Zhu-Salzman & Zeng, 2015) .This arms race suggests that inhibitor specificity and target adaptation are key factors shaping plant-insect interactions. In this study, we functionally characterized a subset of cowpea In11-induced KTIs using *Nicotiana benthamiana transient* expression as a heterologous system. We found that novel

cysteine residues are required for the species-specific antiherbivore function of VuKTI4 and identified candidate serine protease targets in *Spodoptera frugiperda* and *S. exigua* using a combination of AlphaFold-based multimer prediction and frass proteomics. Our study proposes a mechanism in which VuKTI4 escapes degradation in the midgut by inhibiting digestive serine proteases via highly stable decoy binding site, thus impairing digestion and therefore caterpillar growth.

3.2 Results

3.2.1 VuKTI4 is an In11 indued miraculin-like protein with novel cysteine residues.

To investigate the function of previously identified In11- induced Kunitz Trypsin Inhibitors (*KTI*), we first analyzed the KTI gene family across the tomato, Arabidopsis, miracle fruit, soybean, common bean, and cowpea genomes. We found that two of the strongest In11-induced genes, *Vigun04g045700* and *Vigun05g143300*, encode VuKTI4 and VuKTI5 and are members of a large gene family that is organized in two clades (**Fig. 1a**). Within clade II, VuKTI4 is a member of subclade a, and more closely related to miraculin, while *Vigun05g143300* is found in a legume-specific expansion within subclade b and more closely related to the soybean trypsin inhibitor (STI), a well-described antiherbivore defense (Sultana et al., 2023).

Owing to the role of disulfide bridges in the conserved beta-trefoil structure of canonical KTIs, we explored the number and location of cysteine residues in all the proteins. We found that all members of clade I and clade IIb had between 0 and 7 cysteines, although genes with 0 (*Vigun05g146100*) and 1 (*Glyma.19G074800*) are likely misassembled, while clade IIa had between 2 and 9 (**Fig. 1b**). As expected, proteins with more than four cysteines (102/120) had the two canonically conserved pairs. Additional disulfide-forming cysteines were found in two

insertions at the n- and c-termini of the protein, with the c-termini insertion located between two canonical cysteine residues and lost in the members of the legume-specific expansion in subclade IIb (see alignment file).Interestingly, the n-terminus insertion was unique to subclade IIa and seven out of nine members had a lysin or arginine residue between the two novel cysteines, except for *Vigun04g45700* and its soybean homolog. These results indicate that additional cysteine residues might contribute to overall protein stability or function to stabilize novel reactive loops with additional serine protease binding sites.

3.2.2 The species-specific antiherbivore activity of VuKTI4 is dependent on novel cysteine residues

To determine if In11-induced cowpea KTIs have antiherbivore functions, we conducted herbivory assays using the *N. benthamiana* transient protein expression system and *S. frugiperda* and *S. exigua* larvae. We found that second instar *S. frugiperda* larvae feeding on leaves expressing HA-tagged VuKTI4 gained ~25% less weight than those feeding on an empty vector (EV) or HA-tagged VuKTI5 (**Fig. 2a**), an In11-induced KTI with a typical cysteine content. Conversely, we observed a ~22% weight increase on *S. exigua* larvae feeding on leaves expressing VuKTI4-HA compared to the EV and VuKTI5-HA (**Fig. 2b**). In both species the effect of the inhibitor on larva weight gain was dependent on the novel cysteine residues as the mean weight of larvae feeding on the quadruple cysteine mutant VuKTI4Δ4-HA was not significantly different from EV or VuKTI5. Following this observation, we hypothesized that the strong antiherbivore function of VuKTI4 against *S. frugiperda* was due to increased resistance to cleavage or degradation in the midgut. To test this, we collected frass from larvae fed on leaves expressing the wild-type or quadruple mutant VuKTI4 and detected the presence of HA-tagged proteins using western blot. We detected the full-length (35 kDa) and lower molecular weight (25

kDa) protein in the frass of larvae feeding on VuKTI4-HA, while only the smaller version was present in the frass of those feeding on VuKTI4Δ4-HA, despite similar levels of expression in the leaf (**Fig. 2c**). Furthermore, we detected the VuKTI proteins in the frass of *S. exigua*, although the non-specific binding of the anti-HA antibody to other proteins in the frass hindered our ability to confidently assign low molecular weight bands. These results suggest that VuKTI4 is a bioactive and gut-stable species-specific antiherbivore protein.

3.2.3 AlphaFold Multimer (AFM) accurately predicts KTI-serine protease interactions

Due to the specific inhibitory action of KTIs against serine proteases, we focused on identifying candidate targets with homology to porcine pancreatic trypsin and bovine chymotrypsin. We found that trypsin and chymotrypsin-like serine proteases are a multigene family with 43 and 68 candidate proteins encoded in the *S. frugiperda* and *S. exigua* genomes, respectively (**Fig. 3a**).

To identify candidate VuKTI targets, we prioritized *S. frugiperda* genes expressed in in the midgut according to scRNA-seq data (Sun et al., 2025) and generated AFM models for every VuKTI-protease pair and ranked them according to their ipTM and pTM scores (**Fig. 3b**). First, we confirmed that the known complex of STI and porcine pancreatic trypsin (PTT) was correctly predicted by AFM (ipTM = 0.87, pTM =0.87) and placed the conserved Arg63 in the active site of the enzyme and interacted with the three catalytic residues Ser, His and Asp (**Fig. S1a**). The closely related VuKTI5 was also predicted to interact with PTT (ipTM = 0.91, pTM = 0.89) via the same contacts (**Fig. S1b)**. Surprisingly, VuKTI4 was not predicted to interact with PTT (ipTM = 0.17, pTM = 0.55) even though the canonical reactive site loop is present (**Fig. S1a**). Instead, the disulfide-stabilized novel loop was predicted to occupy the active site of PTT, but no hydrogen bonds were formed between the catalytic residues and the loop (**Fig. S1c**).

Furthermore, VuKTI5 was predicted to interact with 9 of the 17 expressed *S. frugiperda* trypsin/chymotrypsin-like serine proteases (red boxes) via the same mechanism, while AFM did not predict an interaction between VuKTI4 and any of the proteases (blue boxes) (**Fig. 3b**), although the novel loop was always inserted into the enzyme's active pocket. These results indicate that the interaction of individual canonical cowpea KTIs with digestive serine proteases might only weakly contribute to direct defenses via a known mechanism of inhibition, and that polymorphic gut-stable VuKTIs could be a primary defensive protein that inhibits unknown targets in a yet-to-be described manner.

3.2.4 Frass proteomics reveal differences in the protease content of KTI-fed *Spodoptera* sp caterpillars

Given that the KTI-protease interaction is highly stable, we sought to identify interacting proteases by immunoprecipitating HA-tagged KTI4 from the frass of *S. frugiperda* followed by trypsin digestion and proteomics. We successfully detected the full length and lower molecular weight version of VuKTI4 and VuKTI4Δ4-HA in the input (IN) and immunoprecipitated (IP) fraction of *S. frugiperda* and *S. exigua* (**Fig. 4a**). However, our optimization step (**Fig. 4b**) revealed that the clearest KTI bands were observed in a mildly reducing buffer (1 mM DTT) but, given the low amount of total protein observed by silver stain, further optimization might be required for efficient immunoprecipitation.

Nevertheless, our proteomics assay identified peptides corresponding to 141 proteins in the frass of *S. frugiperda*, five of which were uniquely found in the pulldown of the cysteine mutant, 51 in the wild-type, and 85 similarly accumulated in both in a single experiment. Four out of the 5 proteins uniquely found in the pulldown from the mutant were related to digestion according to their annotation as trypsin, aminopeptidases and nutrient sensing (enoyl-CoA hydratase).

Meanwhile, 18 proteins in the pulldown from the wild type KTI were associated with protein and carbohydrate metabolism and nutrient absorption and the other 33 with transcription, immunity and cuticle biosynthesis (**Table S1**). These results suggest that VuKTI4 might interfere with gut homeostasis by disrupting nutrient acquisition and transport, leading to reduced weight gain. Furthermore, the presence of transcriptional machinery in the frass suggests that the induced accumulation of inhibitor-resistant serine proteases or gut structural proteins might be contributing to reduced growth (Zhu-Salzman & Zeng, 2015).

For *S. exigua* we identified 168 proteins, four of which were consistently absent from the pulldown from the cysteine mutant and highly abundant in the wildtype across two independent experiments. These proteins were annotated as α-amylase, peptidase S1 domain-containing protein, proline-proline-glutamate (PPE) family protein, and REPAT, a family of proteins broadly characterized as insect effectors (Herrero et al., 2007; X. Wang et al., 2024; Yan et al., 2023). These results indicate that VuKTI4 might promote *S. exigua* growth by promoting carbohydrate and protein degradation and impairing other defenses in *N. benthamiana* via effectors.

3.4 Discussion

Together our experiments demonstrate that gut-stable In11-induced cowpea KTIs exhibit antiherbivore activity when expressed in a heterologous system. Unlike canonical KTIs, cysteine-rich VuKTI4 inhibited the growth of *S. frugiperda* but promoted *S. exigua* growth. Mutating these additional cysteines to the non-disulfide forming residue alanine (A) eliminated this effect and caterpillars of both species gained weight comparable to that of empty vector or VuKTI-fed larvae.

AFM protein complex predictions of VuKTI4 and *S. frugiperda* expressed serine proteases revealed that an exposed loop stabilized by the n-terminal cysteine pairs was occupying the catalytic pocket of the enzyme, although the complexes are unlikely according to the ipTM scores. Intriguingly, AFM correctly predicted the interaction between the conserved Arg residue of VuKTI5 and the active site of PTT but could not position the conserved Arg from VuKTI4.

Evidence from our proteomics data shows that, when in excess, trypsin cleaves VuKTI4 at several arginine and lysin residues (**Fig S2a**). This suggested to us that, although no hydrogen bonds are formed, the novel loop physically blocks the active site of gut serine proteases and acts as a "decoy" to neutralize the enzyme while avoiding cleavage, because there is no Arg or Lys residue between cysteines 34 and 39. Once proteases are inhibited, *S. frugiperda* might experience disrupted gut homeostasis and attempt to compensate by upregulating carbohydrate and protein metabolism. Although counterintuitive, the increased gut activity might hinder growth.

Lastly, this decoy site appears to be a rare mechanism to escape recognition as only two out of the 120 the KTIs in the gene family encode an insertion without active sites. Vigun04g048700.1, a homolog encoded next to *VuKTI4*, has a Lys in the novel loop and an Arg adjacent to it, and is likely an interactor of PTT (ipTM = 0.72, pTM = 0.67) and the *S. frugiperda* trypsin/chymotrypsin like encoded by LOC118278170 (ipTM = 0.79, pTM = 0.66). This indicated that the novel loop is a preferred contact point (vs the conserved Arg) (**Fig S2b**), and that mutations to the codons that encode amino acids in the novel loop are likely to be adaptive.

3.5 Materials and Methods

3.5.1 Phylogenetic analysis of Kunitz trypsin inhibitors (KTIs)

To identify members of the kunitz trypsin inhibitor (KTI) gene family, we used cowpea (*Vigna unguiculata*) VuKTI4 (Vigun04g045700.1) and soybean trypsin inhibitor (STI) (Glyma.08G341500.1) as queries to retrieve sequences from cowpea, common bean (*Phaseolus vulgaris*), soybean (*Glycine max*), tomato (*Solanum lycopersicum*) and Arabidopsis genomes available in Phytozome13 using the blast-align-tree (BAT v.0.1.1) pipeline developed by the Steinbrenner Lab (https://github.com/steinbrennerlab/blast-align-tree).

The resulting list of candidate genes was manually curated by removing pseudogenes, misassembled genes, and those lacking a kunitz motif (L,I,V,M)-X-D-X_2-G-X_2-(L,I,V,M)-X_5-Y-X-(L,I,V,M). The final KTI candidates were re-aligned using Clustal Omega (Sievers et al., 2011), and a maximum-likelihood phylogenetic tree was inferred using the IQ-TREE (Trifinopoulos et al., 2016) web server.

3.5.2 Molecular cloning of In11-induced KTIs

For plant protein expression, the full length coding sequence (without a stop codon) of In11-induced *VuKTI4* (Vigun04g045700.1) and *VuKTI5* (Vigun05g143300.1) were amplified from a 5' RACE cDNA library using the Q5 High-Fidelity DNA polymerase (NEB) and specific primers designed against the cowpea v1 genome available in Phytozome13 (Table S1) with added BpiI restriction enzyme recognition sites and overhangs compatible with Mo-Clo, and cloned into a level 0 cloning vector (pAGM1287, Addgene plasmid #47996). PCR products with the expected size were excised from a 1% agarose gel and purified using the Monarch® DNA Gel Extraction Kit Protocol (NEB #T1020) and quantified using a NanoDrop1000. The expression construct was assembled into a customized pGreenII (Hellens et al., 2000) with BsaI insertion sites by combining appropriate ratios of the 35S promoter (pICSL13001, Addgene plasmid #50265), the purified CDS products, a C-term 6xHA tag (pICSL50009, Addgene plasmid #50309) and nos

terminator (pICH41421, Addgene plasmid # 50339) following the recommended XL ligation protocol (Weber et al., 2011; Engler et al., 2014).

To generate VuKTI4 cysteine content variants, we designed DNA fragments of the *VuKTI4* CDS with added BpiI restriction enzyme recognition sites, overhangs and point mutations encoding alanines (A) at codons that originally encoded polymorphic cysteines 34 and 39, and the conserved cysteines 156 and 159. The resulting DNA fragment was synthesized by Twist Bioscience (San Francisco, CA) and cloned for plant expression as indicated above.

3.5.3 Transient expression of cowpea KTIs in *Nicotiana benthamiana*

To measure the effect of plant-expressed KTIs on caterpillar growth, we individually transformed pGreenII expression constructs indicated above into Agrobacterium (*Agrobacterium tumefacien*s GV3101). Individual strains were grown on Luria-Bertani (LB) medium supplemented with kanamycin, gentamicin, rifampin and tetracycline for 48 h and 1 mL of culture was centrifuged to recover the cell pellet. The pellets were resuspended in infiltration media (10 mM MES pH 5.6, 150 µM Acetosyringone, 10mM MgCl2) and incubated for 3 h. For plant expression, we adjusted the optical density (OD) of each strain to OD_{600}=0.45 and infiltrated two fully expanded leaves on 6-week-old *Nicotiana benthamiana* plants using a needleless syringe, taking care to infiltrate the entire leaf (~2 mL of culture per leaf). Plants with infiltrated leaves were kept under a 12 h light/dark cycle for 48 h to allow for protein expression. At 48 h, infiltrated leaves were excised from the plant at the petiole, placed in a humidity chamber constructed from a square petri dish with a layer of 1% agarose dissolved in water in the lid, and immediately infested with caterpillars for herbivory assays.

3.5.4 Herbivory assays and frass collection

First instar beet armyworm (*Spodoptera exigua*) and fall armyworm (*Spodoptera frugiperda*) larvae were obtained from Benzon Research Inc (Carlisle, PA). Upon receipt, larvae were incubated in their original packaging (artificial diet) under dark at 28ºC for 36-48 h to accelerate molting.

Second instar larvae were weighed using a digital scale (Mettler Toledo Inc), and randomly assigned to pre-assembled humidity chambers containing leaves expressing the wild type KTI proteins (VuKTI4-HA or VuKTI5-HA), the quadruple cysteine mutant (VuKTI4Δ4-HA) or an empty vector (EV). Each chamber received two larvae. The assembled chambers were placed under a custom imaging system with constant light, and larvae weights were recorded three and four days later. On the fourth day, all the frass in the chamber was collected and flash-frozen in liquid nitrogen alongside any leftover plant tissue for subsequent protein extraction and detection via western blot.

3.5.5 Western Blot

To detect the presence of the HA-tagged KTI proteins in leaf tissue and frass, total proteins were extracted using a 3X Laemmli buffer (50 mM Tris-Cl pH 6.8, 6% SDS, 30% glycerol, 16% β-mercaptoethanol, and 0.006% bromophenol blue sample buffer) and separated by SDS-PAGE in a 15% acrylamide gel before transfer to a nitrocellulose membrane. To detect the proteins, we use an anti-HA antibody (Millipore Sigma, #H3663) at 1:2000 diluted in 1% Tris-buffered saline with Tween-20 (TBST) buffer, with secondary antibody anti-mouse IgG-HRP conjugate (Sigma) at 1:10,000, followed by chemiluminescent visualization using a 1:10 mixture of SuperSignal West Pico PLUS /nano (Thermo Fisher, #34580). SimplyBlue™ SafeStain Protocol (Thermo Fisher) staining was used to verify equal loading.

3.5.6 Immunoprecipitation and proteomics

To identify candidate target proteases of VuKTI4, we characterized the frass proteome of the fall and beet armyworm after feeding on *N. benthamiana* leaves overexpressing VuKTI4 and the mutant version VuKTI4Δ4-HA. To this end, we took advantage of the HA-tagged VuKTIs and pulled down protein complexes using monoclonal anti-HA agarose beads (Millipore Sigma, #A2095). Briefly, total proteins were extracted from approximately 300 mg of frozen frass in 800 uL of extraction buffer (50 mM Tris pH 7.5, 150 nM NaCl, 10% glycerol, 1% NP-40, 1 mM DTT, 1mM PMSF, 1 XcOmplete™ Protease Inhibitor Cocktail). After centrifugation, the supernatant was recovered, mixed with 10 uL of washed anti-HA agarose beads, and incubated end-over-end at 4C for 30 minutes. The presence of immunoprecipitated proteins was confirmed by western blot as indicated above and Pierce™ Silver Stain (Thermo Fisher, #24612).

We used proteomics to identify candidate interactors of the HA-tagged KTIs among the immunoprecipitated proteins. We washed the anti-HA agarose beads with four changes of ammonium bicarbonate and performed overnight on-bead digestion using sequencing grade modified trypsin (Promega, #V5111). The next morning, we recovered the supernatant and washed the leftover beads with one more change of ammonium bicarbonate. The supernatant and wash were combined and shipped in dry ice to the UCDavis Proteomics core (Davis, CA) for identification using Data-Independent Acquisition (DAI) proteomics implemented in Spectronaut. The resulting peptides were compared against *Nicotiana benthamiana* v1 (Kurotani et al., 2023), and Uniprot *S. frugiperda* and *S. exigua* proteomes.

3.5.7 Alpha Fold Multimer (AFM)-assisted target protease discovery

To discover potential targets of the In11-induced KTIs we implemented a modified pipeline using AFM to discover interkingdom protein-protein interactions (Homma et al., 2023, 2024). Because of the putative function of KTIs, we focused on digestive serine proteases as candidate targets and identified genes in the genomes of the beet and fall armyworms encoding trypsin and chymotrypsin-like proteins using BAT v.0.1.1 with trypsin (PPT, Uniprot P00761) and bovine chymotrypsinogen A (CTRA, uniprot P00766) as queries. KTIs and candidate serine protease targets were first processed by removing the signaling peptide with SignalP 6.0 (Teufel et al., 2022). For multimer modeling, we used the batch-mode in the AlphaFold online server and provided a JSON file with all possible pairs of processed KTI and candidate serine protease. The well documented interaction between the soybean trypsin inhibitor (STI) and PPT (Blow et al., 1974) was used as a positive control, and all resulting multimers were ranked based on their ipTM and pTM scores (Abramson et al., 2024).

3.6 Figures

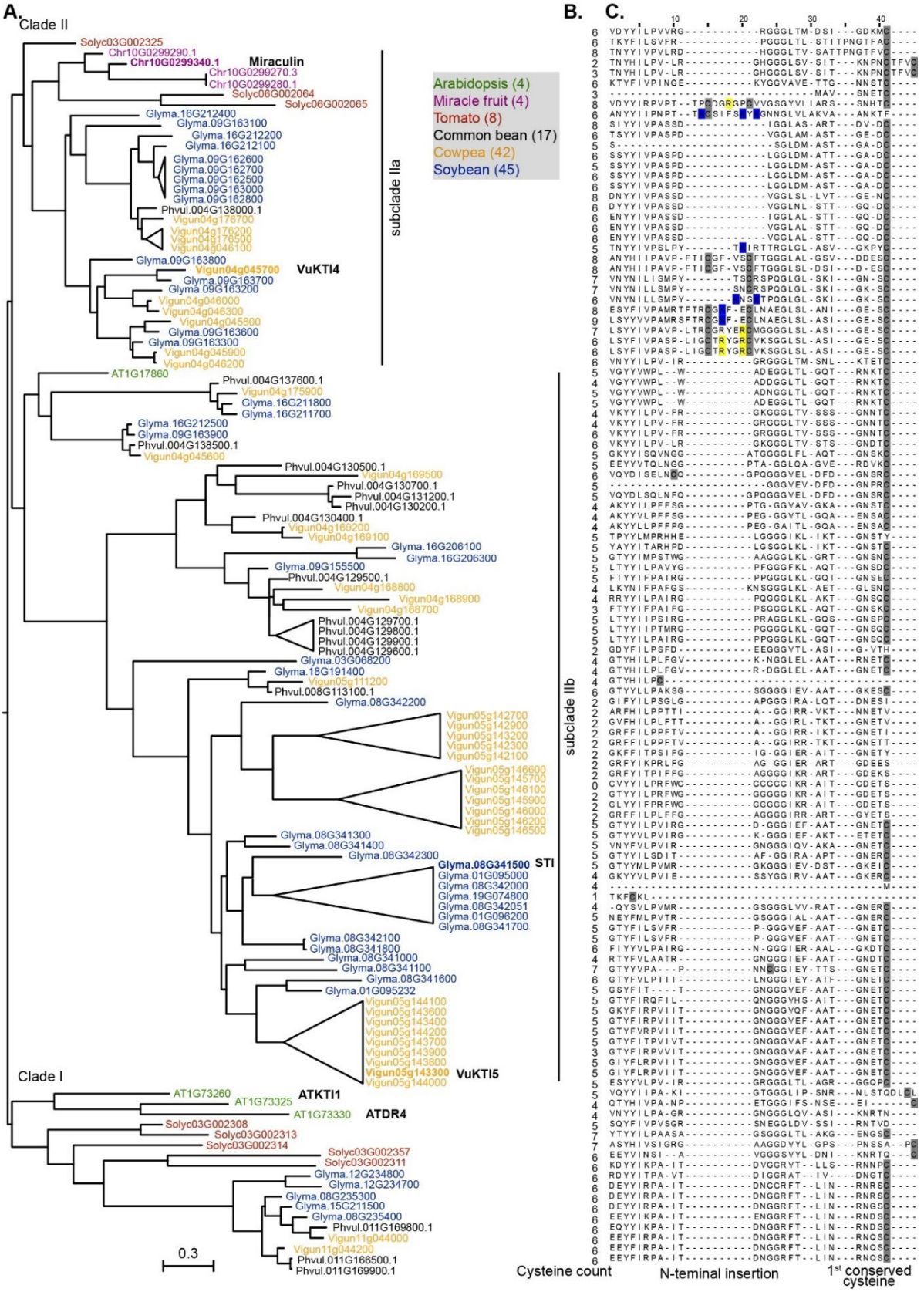
A.
Clade II
Solyc03G002325
Chr10G0299290.1
Chr10G0299340.1
Miraculin
Chr10G0299270.3
Chr10G0299280.1
Solyc06G002064
Solyc06G002065
Glyma.16G212400
Glyma.09G163100
Glyma.16G212200
Glyma.16G212100
Glyma.09G162600
Glyma.09G162700
Glyma.09G162500
Glyma.09G163000
Glyma.09G162800
Phvul.004G138000.1
Vigun04g176700
Vigun04g176200
Vigun04g176500
Vigun04g046100
Glyma.09G163800
Vigun04g045700
VuKTI4
Glyma.09G163700
Glyma.09G163200
Vigun04g046000
Vigun04g046300
Vigun04g045800
Glyma.09G163600
Glyma.09G163300
Vigun04g045900
Vigun04g046200
AT1G17860
Phvul.004G137600.1
Vigun04g175900
Glyma.16G211800
Glyma.16G211700
Glyma.16G212500
Glyma.09G163900
Phvul.004G138500.1
Vigun04g045600
Phvul.004G130500.1
Vigun04g169500
Phvul.004G130700.1
Phvul.004G131200.1
Phvul.004G130200.1
Phvul.004G130400.1
Vigun04g169200
Vigun04g169100
Glyma.16G206100
Glyma.16G206300
Glyma.09G155500
Phvul.004G129500.1
Vigun04g168800
Vigun04g168900
Vigun04g168700
Phvul.004G129700.1
Phvul.004G129800.1
Phvul.004G129900.1
Phvul.004G129600.1
Glyma.03G068200
Glyma.18G191400
Vigun05g111200
Phvul.008G113100.1
Glyma.08G342200
Vigun05g142700
Vigun05g142900
Vigun05g143200
Vigun05g142300
Vigun05g142100
Vigun05g146600
Vigun05g145700
Vigun05g146100
Vigun05g145900
Vigun05g146000
Vigun05g146200
Vigun05g146500
Glyma.08G341300
Glyma.08G341400
Glyma.08G342300
Glyma.08G341500
STI
Glyma.01G095000
Glyma.08G342000
Glyma.19G074800
Glyma.08G342051
Glyma.01G096200
Glyma.08G341700
Glyma.08G342100
Glyma.08G341800
Glyma.08G341000
Glyma.08G341100
Glyma.08G341600
Glyma.01G095232
Vigun05g144100
Vigun05g143600
Vigun05g143400
Vigun05g144200
Vigun05g143700
Vigun05g143900
Vigun05g143800
Vigun05g143300
VuKTI5
Vigun05g144000
Clade I
AT1G73260
AT1G73325
ATKTI1
AT1G73330
ATDR4
Solyc03G002308
Solyc03G002313
Solyc03G002314
Solyc03G002357
Solyc03G002311
Glyma.12G234800
Glyma.12G234700
Glyma.08G235300
Glyma.15G211500
Glyma.08G235400
Phvul.011G169800.1
Vigun11g044000
Vigun11g044200
Phvul.011G166500.1
Phvul.011G169900.1
0.3
subclade IIa
subclade IIb
Arabidopsis (4)
Miracle fruit (4)
Tomato (8)
Common bean (17)
Cowpea (42)
Soybean (45)
B.
C.
10
20
30
40
Cysteine count
N-teminal insertion
1st conserved cysteine

Figure 1. In11-induced VuKTI4 belongs to a structurally diverse legume-specific KTI family. **A.** Maximum likelihood tree depicting the phylogenetic relationship between Kunitz trypsin inhibitor (KTI) proteins from Arabidopsis, Miracle fruit, Tomato, Soybean, Common bean and Cowpea. In parentheses the number of KTI proteins identified in each species. In11-induced VuKTI4 (Vigun04g045700) and VuKTI5 (Vigun05g143300) are in bold. The scale bar indicates branch length as the mean number substitutions per site. **B.** Count of cysteine residues in each KTI protein. **C.** Protein alignment of the N-terminus region after the signaling peptide and the first conserved cysteine. The N-terminal insertion harboring additional cysteine residues (CYS34, CYS39) is indicated. Arginine (R, yellow) and lysine (K, blue) residues in the insertion, and cysteines (C, black) are highlighted.

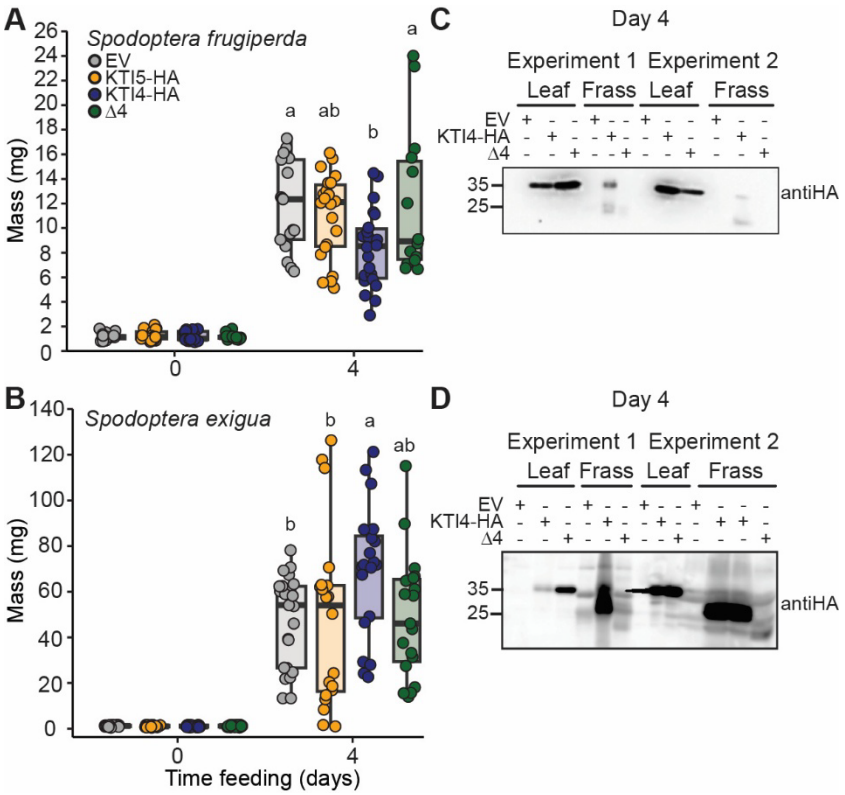

Figure 2. Antiherbivore activity and frass survival of VuKTI4. Caterpillar performance of **A.** *Spodoptera frugiperda* (n = 14-23, two independent experiments) and **B.** *Spodoptera exigua* (n=19-23, two independent experiments) 2nd instar larvae feeding on detached *Nicotiana benthamiana* overexpressing an empty vector (EV), *35S:KTI4-HA*, *35S:KTI5-HA*, and *35S:KTI4 Δ4-HA* and the silencing suppressor p19 for four days. VuKTI4 presence in leaves and frass collected from **C.** *S. frugiperda* and **D.** *S. exigua* after feeding on detached leaves for four days. Boxplots show the mean value and upper and lower quantiles, and each point represents individual larvae. Significant differences in caterpillar mass at day four were determined by ANOVA followed by a Tukey test ($\alpha = 0.05$).

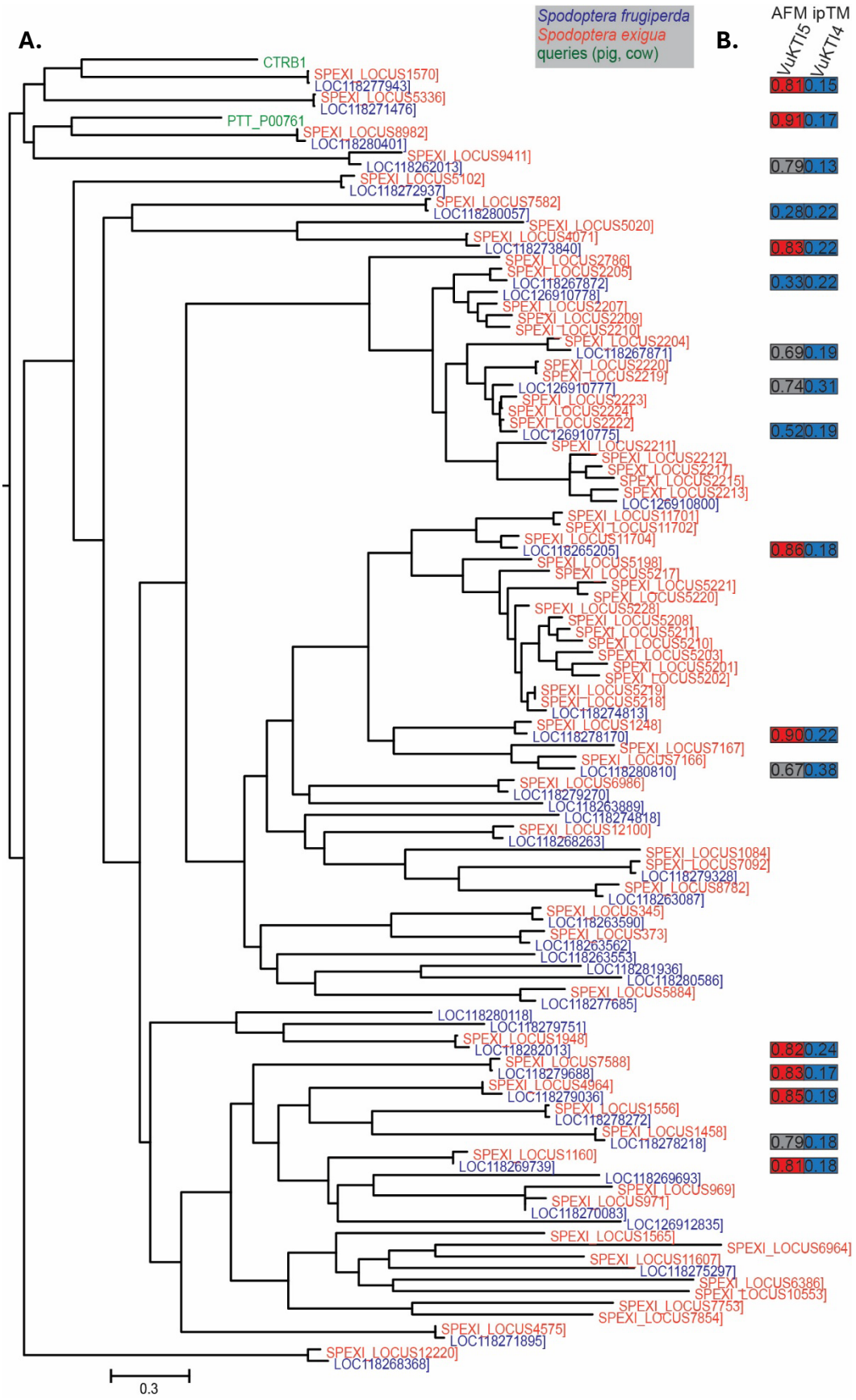
A.
Spodoptera frugiperda
Spodoptera exigua
queries (pig, cow)
B.
AFM ipTM
VuKTI5
VuKTI4
CTRB1
SPEXI_LOCUS1570]
LOC118277943]
SPEXI_LOCUS5336]
LOC118271476]
PTT_P00761
SPEXI_LOCUS8982]
LOC118280401]
SPEXI_LOCUS9411]
LOC118262013]
SPEXI_LOCUS5102]
LOC118272937]
SPEXI_LOCUS7582]
LOC118280057]
SPEXI_LOCUS5020]
SPEXI_LOCUS4071]
LOC118273840]
SPEXI_LOCUS2786]
SPEXI_LOCUS2205]
LOC118267872]
LOC126910778]
SPEXI_LOCUS2207]
SPEXI_LOCUS2209]
SPEXI_LOCUS2210]
SPEXI_LOCUS2204]
LOC118267871]
SPEXI_LOCUS2220]
SPEXI_LOCUS2219]
LOC126910777]
SPEXI_LOCUS2223]
SPEXI_LOCUS2224]
SPEXI_LOCUS2222]
LOC126910775]
SPEXI_LOCUS2211]
SPEXI_LOCUS2212]
SPEXI_LOCUS2217]
SPEXI_LOCUS2215]
SPEXI_LOCUS2213]
LOC126910800]
SPEXI_LOCUS11701]
SPEXI_LOCUS11702]
SPEXI_LOCUS11704]
LOC118265205]
SPEXI_LOCUS5198]
SPEXI_LOCUS5217]
SPEXI_LOCUS5221]
SPEXI_LOCUS5220]
SPEXI_LOCUS5228]
SPEXI_LOCUS5208]
SPEXI_LOCUS5211]
SPEXI_LOCUS5210]
SPEXI_LOCUS5203]
SPEXI_LOCUS5201]
SPEXI_LOCUS5202]
SPEXI_LOCUS5219]
SPEXI_LOCUS5218]
LOC118274813]
SPEXI_LOCUS1248]
LOC118278170]
SPEXI_LOCUS7167]
SPEXI_LOCUS7166]
LOC118280810]
SPEXI_LOCUS6986]
LOC118279270]
LOC118263889]
LOC118274818]
SPEXI_LOCUS12100]
LOC118268263]
SPEXI_LOCUS1084]
SPEXI_LOCUS7092]
LOC118279328]
SPEXI_LOCUS8782]
LOC118263087]
SPEXI_LOCUS345]
LOC118263590]
SPEXI_LOCUS373]
LOC118263562]
LOC118263553]
LOC118281936]
LOC118280586]
SPEXI_LOCUS5884]
LOC118277685]
LOC118280118]
LOC118279751]
SPEXI_LOCUS1948]
LOC118282013]
SPEXI_LOCUS7588]
LOC118279688]
SPEXI_LOCUS4964]
LOC118279036]
SPEXI_LOCUS1556]
LOC118278272]
SPEXI_LOCUS1458]
LOC118278218]
SPEXI_LOCUS1160]
LOC118269739]
LOC118269693]
SPEXI_LOCUS969]
SPEXI_LOCUS971]
LOC118270083]
LOC126912835]
SPEXI_LOCUS1565]
SPEXI_LOCUS6964]
SPEXI_LOCUS11607]
LOC118275297]
SPEXI_LOCUS6386]
SPEXI_LOCUS10553]
SPEXI_LOCUS7753]
SPEXI_LOCUS7854]
SPEXI_LOCUS4575]
LOC118271895]
SPEXI_LOCUS12220]
LOC118268368]
0.3
0.81 0.15
0.91 0.17
0.79 0.13
0.28 0.22
0.83 0.22
0.33 0.22
0.69 0.19
0.74 0.31
0.52 0.19
0.86 0.18
0.90 0.22
0.67 0.38
0.82 0.24
0.83 0.17
0.85 0.19
0.79 0.18
0.81 0.18

Figure 3. VuKTI4 is not predicted to interact with expressed trypsin/chymotrypsin-like homologs in *S. frugiperda*. A. Maximum-likelihood gene tree of *S. frugiperda* and *S. exigua* trypsin-like and chymotrypsin-like serine proteases (SP). Scale bar represents the rate of substitution. **B.** interface predicted template modeling (ipTM) scores for every KTI-SP pair. Red= high-quality prediction/likely interaction (>0.8), gray = gray zone (0.6-0.78), blue = failed prediction/unlikely interaction (<0.6).

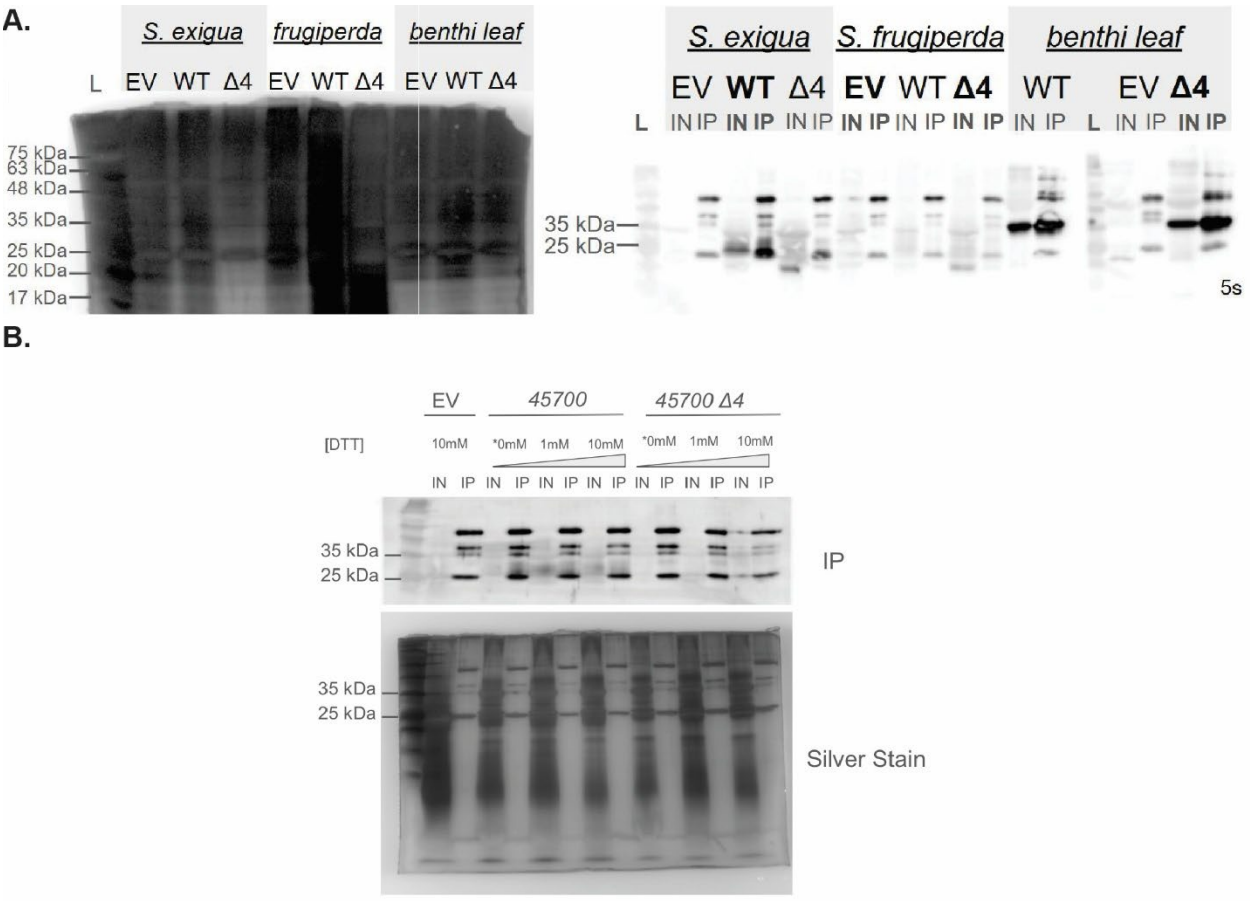

Figure 4. VuKTI4 is successfully immunoprecipitated under moderately reducing conditions from *S. frugiperda* and *S. exigua* frass. A. Sample silver stain of an SDS-PAGE gel of immunoprecipitated (IP) proteins (left) and anti-HA western blot of total input (IN) and IP proteins (right) from frass collected from larvae fed on *N. benthamiana* leaves expressing an empty vector (EV), VuKTI4-HA or VuKTI4Δ4-HA. **B.** Sample anti-HA western blot (top) and silver stain SDS-PAGE gel (bottom) from immunoprecipitated proteins from *S. exigua* frass during the optimization step under a range of concentrations of the reducing agent DTT.

3.7 Supplementary figures

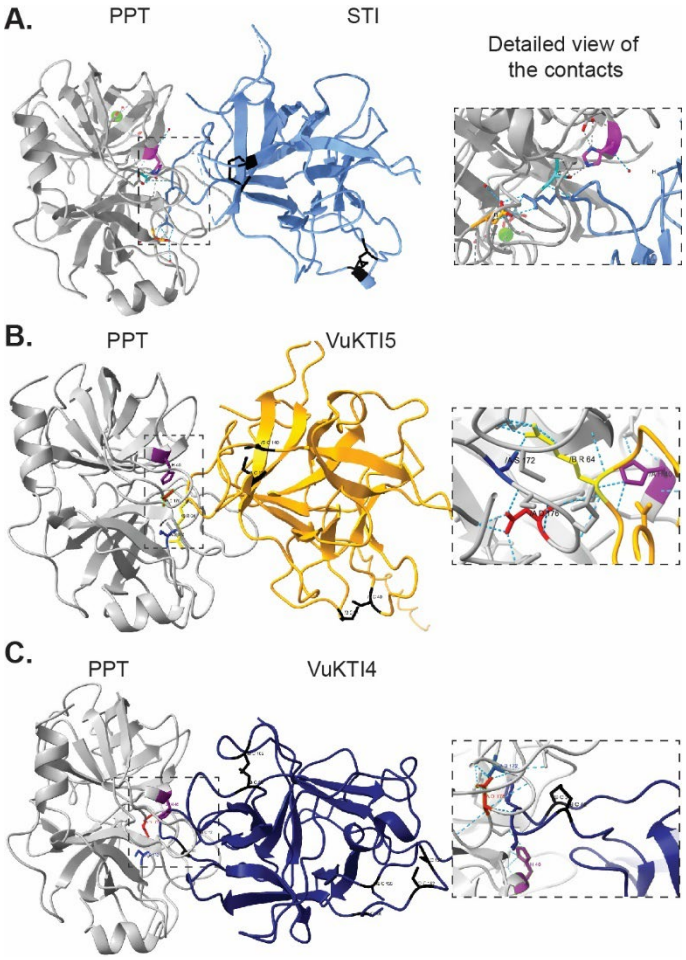

Figure S2. **AlphaFold-Multimer (AFM)-predicted complexes of porcine pancreatic trypsin (PTT) and KTIs**. Predicted complex of PPT and mature (no signal peptide) **A.** Soybean trypsin inhibitor (STI)**,** cowpea (*Vigna unguiculata*) Kunitz Trypsin Inhibitor **B.** VuKTI5 and **C.** VuKTI4. Conserved and polymorphic cysteine (C) residues in black. Conserved reactive arginine (R) in yellow. PTT catalytic triad residues histidine (H, purple), aspartic acid (D, red) and Serine (S, blue) with predicted hydrogen bonds.

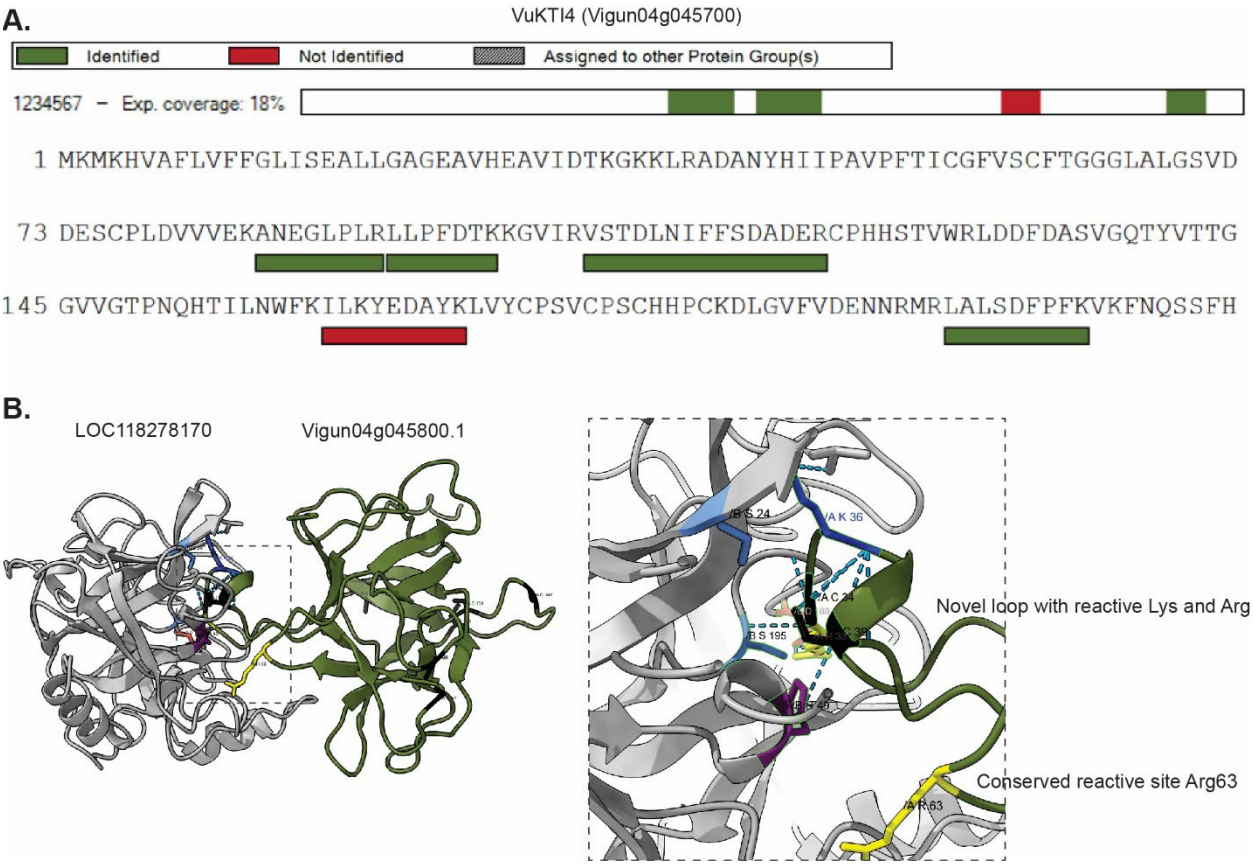

Figure S3. **VuKTI4 is cleaved by excess trypsin at conserved Arg 63 residue. A.** Sequence and location of peptides from VuKTI4 detected in *S. frugiperda* frass by proteomics **B.** Predicted complex of *S. frugiperda* trypsin/chymotrypsin-like (LOC118278170) and mature (no signal peptide) Vigun04g045800.1. Conserved and polymorphic cysteine (C) residues in black. Reactive arginine (R) and lysin (K) in yellow and cyan, respectively. Serine protease catalytic triad residues histidine (H, purple), aspartic acid (D, red) and Serine (S, blue) with predicted hydrogen bonds.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C. C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w

Adams, S., Grundy, J., Veflingstad, S. R., Dyer, N. P., Hannah, M. A., Ott, S., & Carré, I. A. (2018). Circadian control of abscisic acid biosynthesis and signalling pathways revealed by genome-wide analysis of LHY binding targets. *New Phytol.*, *220*(3), 893–907.

Alborn, H. T., Hansen, T. V, Jones, T. H., Bennett, D. C., Tumlinson, J. H., Schmelz, E. A., & Teal, P. E. A. (2007). Disulfooxy fatty acids from the American bird grasshopper Schistocerca americana, elicitors of plant volatiles. *Proc. Natl. Acad. Sci. U. S. A.*, *104*(32), 12976–12981.

Alborn, H. T., Turlings, T. C. J., Jones, T. H., Stenhagen G and Loughrin, J. H., & Tumlinson, J. H. (1997). An Elicitor of Plant Volatiles from Beet Armyworm Oral Secretion. *Science*.

Anderson, S. L., & Kay, S. A. (1995). Functional dissection of circadian clock- and phytochromeregulated transcription of the Arabidopsis CAB2 gene. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 1500–1504. https://doi.org/10.1073/pnas.92.5.1500

Arimura, G. I., Köpke, S., Kunert, M., Volpe, V., David, A., Brand, P., Dabrowska, P., Maffei, M. E., & Boland, W. (2008). Effects of feeding Spodoptera littoralis on lima bean leaves: IV. Diurnal and nocturnal damage differentially initiate plant volatile emission. *Plant Physiology*, *146*(3), 965–973. https://doi.org/10.1104/pp.107.111088

Benrey, B., Bustos-Segura, C., & Grof-Tisza, P. (2024). The mesoamerican milpa system: Traditional practices, sustainability, biodiversity, and pest control. *Biol. Control*, *198*(105637), 105637.

Bjornson, M., Pimprikar, P., Nürnberger, T., & Zipfel, C. (2021). The transcriptional landscape of Arabidopsis thaliana pattern-triggered immunity. *Nature Plants*, *7*(5), 579–586. https://doi.org/10.1038/s41477-021-00874-5

Blow, D. M., Janin, J., & Sweet, R. M. (1974). Mode of action of soybean trypsin inhibitor (Kunitz) as a model for specific protein-protein interactions. *Nature*, *249*, 54–57. https://doi.org/10.1038/249054a0

Boller, T., & Felix, G. (2009). A renaissance of elicitors: Perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual Review of Plant Biology*, *60*, 379–407. https://doi.org/10.1146/annurev.arplant.57.032905.105346

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*, 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R Journal*, *9*, 378–400. https://doi.org/10.32614/rj-2017-066

Brouwer, J. N., Van Der Wel, H., Francke, A., & Henning, G. J. (1968). Miraculin, the sweetness-inducing protein from miracle fruit. *Nature*, *220*, 373–374. https://doi.org/10.1038/220373a0

de Leone, M. J., Hernando, C. E., Romanowski, A., Careno, D. A., Soverna, A. F., Sun, H., Bologna, N. G., Vázquez, M., Schneeberger, K., & Yanovsky, M. J. (2020). Bacterial Infection Disrupts Clock Gene Expression to Attenuate Immune Responses. *Curr. Biol.*, *30*(9), 1740–1747.e6.

DeFalco, T. A., & Zipfel, C. (2021). Molecular mechanisms of early plant pattern-triggered immune signaling. *Mol. Cell*, *81*(17), 3449–3467.

Dicke, M., & Baldwin, I. T. (2010). The evolutionary context for herbivore-induced plant volatiles: beyond the "cry for help." In *Trends in Plant Science* (Vol. 15, pp. 167–175). https://doi.org/10.1016/j.tplants.2009.12.002

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21.

Dow, J. A. T. (1992). pH gradients in lepidopteran midgut. *Journal of Experimental Biology*, *172*, 355–375. https://doi.org/10.1242/jeb.172.1.355

Eberl, F., Fabisch, T., Luck, K., Köllner, T. G., Vogel, H., Gershenzon, J., & Unsicker, S. B. (2021). Poplar protease inhibitor expression differs in an herbivore specific manner. *BMC Plant Biology*, *21*. https://doi.org/10.1186/s12870-021-02936-4

Engler, C., Youles, M., Gruetzner, R., Ehnert, T. M., Werner, S., Jones, J. D. G., Patron, N. J., & Marillonnet, S. (2014). A Golden Gate modular cloning toolbox for plants. *ACS Synthetic Biology*, *3*(11), 839–843. https://doi.org/10.1021/sb4001504

Erb, M., & Reymond, P. (2019). Molecular Interactions Between Plants and Insect Herbivores. *Annu. Rev. Plant Biol*, *70*, 527–557. https://doi.org/10.1146/annurev-arplant-050718

Farinas, B., & Mas, P. (2011). Functional implication of the MYB transcription factor RVE8/LCL5 in the circadian control of histone acetylation. *Plant Journal*, *66*, 318–329. https://doi.org/10.1111/j.1365-313X.2011.04484.x

Felix, G., Duran, J. D., Volko, S., & Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant Journal*, *18*, 265–276. https://doi.org/10.1046/j.1365-313X.1999.00265.x

Felton, G. W., & Tumlinson, J. H. (2008). Plant-insect dialogs: complex interactions at the plant-insect interface. In *Current Opinion in Plant Biology* (Vol. 11, pp. 457–463). https://doi.org/10.1016/j.pbi.2008.07.001

Fenske, M. P., Hewett Hazelton, K. D., Hempton, A. K., Shim, J. S., Yamamoto, B. M., Riffell, J. A., & Imaizumi, T. (2015). Circadian clock gene LATE ELONGATED HYPOCOTYL directly regulates the timing of floral scent emission in Petunia. *Proc. Natl. Acad. Sci. U. S. A.*, *112*(31), 9775–9780.

Fraser, O. J. P., Cargill, S. J., & Spoel Steven H and van Ooijen, G. (2024). Crosstalk between salicylic acid signalling and the circadian clock promotes an effective immune response in plants. *Npj Biol Timing Sleep*, *1*(1), 1–9.

Gao, M., Zhang, C., & Lu, H. (2020). Coronatine is more potent than jasmonates in regulating Arabidopsis circadian clock. *Sci. Rep.*, *10*(1), 12862.

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. In *arXiv*. https://doi.org/10.48550/arXiv.1207.3907

Goodspeed, D., Chehab, E. W., Min-Venditti Amelia and Braam, J., & Covington, M. F. (2012). Arabidopsis synchronizes jasmonate-mediated defense with insect circadian behavior. *Proc. Natl. Acad. Sci. U. S. A.*, *109*(12), 4674–4677.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.*, *40*(Database issue), D1178–86.

Gouinguené, S., Pickett, J. A., Wadhams, L. J., Birkett, M. A., & Turlings, T. C. J. (2005). Antennal electrophysiological responses of three parasitic wasps to caterpillar-induced volatiles from maize (Zea mays mays), cotton (Gossypium herbaceum), and cowpea (Vigna unguiculata). *Journal of Chemical Ecology*, *31*, 1023–1038. https://doi.org/10.1007/s10886-005-4245-1

Grant, C. E., Bailey, T. L., & Noble, W. S. (2011). FIMO: Scanning for occurrences of a given motif. *Bioinformatics*, *27*(7), 1017–1018. https://doi.org/10.1093/bioinformatics/btr064

Green, T. R., & Ryan, C. A. (1972). Wound-Induced Proteinase Inhibitor in Plant Leaves: A Possible Defense Mechanism against Insects. *Science*, *175*(4023), 776–777.

Grof-Tisza, P., Muller, M. H., Gónzalez-Salas, R., Bustos-Segura, C., & Benrey, B. (2024). The Mesoamerican milpa agroecosystem fosters greater arthropod diversity compared to

monocultures. *Agriculture, Ecosystems and Environment*, *372*. https://doi.org/10.1016/j.agee.2024.109074

Grof-Tisza, P., Patiño Moreno, Y., Erb, C., Nobel, G., Clancy, M. V., & Benrey, B. (2025). Associational resistance in the milpa: herbivore-induced maize volatiles enhance extrafloral nectar-mediated defenses in common bean via shared parasitoids. *New Phytologist*. https://doi.org/10.1111/nph.70029

Grof-Tisza, P., Turlings, T. C. J., Bustos-Segura, C., & Benrey, B. (2024). Field evidence for the role of plant volatiles induced by caterpillar oral secretion in prey localization by predatory social wasps. *Biology Letters*, *20*. https://doi.org/10.1098/rsbl.2024.0384

Grosse-Holz, F. M., & van der Hoorn, R. A. L. (2016). Juggling jobs: Roles and mechanisms of multifunctional protease inhibitors in plants. *New Phytologist*, *210*, 794–807. https://doi.org/10.1111/nph.13839

Guayazán Palacios, N., Imaizumi, T., & Steinbrenner, A. D. (2024). The circadian clock regulates receptor-mediated immune responses to an herbivore-associated molecular pattern. *BioRxiv*. https://doi.org/10.1101/2024.11.06.622352

Harmer, S. L., Hogenesch, J. B., Straume, M., Chang, H. S., Han, B., Zhu, T., Wang, X., Kreps, J. A., & Kay, S. A. (2000). Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science*, *290*(5499), 2110–2113.

Heil, M. (2014). Herbivore-induced plant volatiles: Targets, perception and unanswered questions. In *New Phytologist* (Vol. 204, pp. 297–306). Blackwell Publishing Ltd. https://doi.org/10.1111/nph.12977

Hellens, R. P., Edwards, E. A., Leyland, N. R., Bean, S., & Mullineaux, P. M. (2000). pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. In *Plant Molecular Biology* (Vol. 42, Issue 6, pp. 819–832).

Homma, F., Huang, J., & van der Hoorn, R. A. L. (2023). AlphaFold-Multimer predicts cross-kingdom interactions at the plant-pathogen interface. *Nature Communications*, *14*. https://doi.org/10.1038/s41467-023-41721-9

Homma, F., Lyu, J., & van der Hoorn, R. A. L. (2024). Using AlphaFold Multimer to discover interkingdom protein–protein interactions. *Plant Journal*. https://doi.org/10.1111/tpj.16969

Howe, G. A., & Jander, G. (2008). Plant immunity to insect herbivores. In *Annual Review of Plant Biology* (Vol. 59, pp. 41–66). https://doi.org/10.1146/annurev.arplant.59.032607.092825

Hsu, P. Y., Devisetty, U. K., & Harmer, S. L. (2013). Accurate timekeeping is controlled by a cycling activator in Arabidopsis. *ELife*, *2*. https://doi.org/10.7554/eLife.00473

Huffaker, A., Pearce, G., Veyrat, N., Erb, M., Turlings, T. C. J., Sartor, R., Shen Zhouxin and Briggs, S. P., Vaughan, M. M., Alborn, H. T., Teal, P. E. A., & Schmelz, E. A. (2013). Plant elicitor peptides are conserved signals regulating direct and indirect antiherbivore defense. *Proc. Natl. Acad. Sci. U. S. A.*, *110*(14), 5707–5712.

Jongsma, M. A., & Bolter, C. (1997). The adaptation of insects to plant protease inhibitors. *Journal of Insect Physiology*, *43*, 885–895. https://doi.org/10.1016/S0022-1910(97)00040-1

Joo, Y., Goldberg, J. K., Chrétien, L. T. S., Kim, S.-G., Baldwin, I. T., & Schuman, M. C. (2019). The circadian clock contributes to diurnal patterns of plant indirect defense in nature: The plant clock supports timely indirect defense. *J. Integr. Plant Biol.*, *61*(8), 924–928.

Joo, Y., Schuman, M. C., Goldberg, J. K., Wissgott, A., Kim, S. G., & Baldwin, I. T. (2019). Herbivory elicits changes in green leaf volatile production via jasmonate signaling and the circadian clock. *Plant Cell and Environment*, *42*(3), 972–982. https://doi.org/10.1111/pce.13474

Kamioka, M., Takao, S., Suzuki, T., Taki, K., Higashiyama, T., Kinoshita, T., & Nakamichi, N. (2016). Direct repression of evening genes by CIRCADIAN CLOCK-ASSOCIATED1 in the Arabidopsis circadian clock. *Plant Cell*, *28*(3), 696–711.

Kappers, I. F., Aharoni, A., Van Herpen, T. W. J. M., Luckerhoff, L. L. P., Dicke, M., & Bouwmeester, H. J. (2005). Genetic engineering of terpenoid metabolism attracts bodyguards to Arabidopsis. *Science*, *309*, 2070–2072. https://doi.org/10.1126/science.1116232

Karban, R. (2011). The ecology and evolution of induced resistance against herbivores: Induced resistance against herbivores. *Funct. Ecol.*, *25*(2), 339–347.

Karban, R., & Baldwin, I. T. (1997). Induced Responses to Herbivory. In *University of Chicago Press*.

Karimi, M., Inzé, D., & Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.*, *7*(5), 193–195.

Katoh, K., Misawa, K., Kuma, K.-I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, *30*(14), 3059–3066.

Kessler, A., & Baldwin, I. T. (2001). Defensive function of herbivore-induced plant volatile emissions in nature. *Science*, *291*(5511), 2141–2144.

Kessler, A., & Heil, M. (2011). The multiple faces of indirect defences and their agents of natural selection. *Functional Ecology*, *25*, 348–357. https://doi.org/10.1111/j.1365-2435.2010.01818.x

Khan, Z. R., Midega, C. A. O., Bruce, T. J. A., Hooper, A. M., & Pickett, J. A. (2010). Exploiting phytochemicals for developing a “push-pull” crop protection strategy for cereal farmers in Africa. *J. Exp. Bot.*, *61*(15), 4185–4196.

Kim, S.-C., Edgeworth, K. N., Nusinow, D. A., & Wang, X. (2023). Circadian clock factors regulate the first condensation reaction of fatty acid synthesis in Arabidopsis. *Cell Rep.*, *42*(12), 113483.

Koizumi, A., Tsuchiya, A., Nakajima, K. I., Ito, K., Terada, T., Shimizu-Ibuka, A., Briand, L., Asakura, T., Misaka, T., & Abe, K. (2011). Human sweet taste receptor mediates acid-induced sweetness of miraculin. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 16819–16824. https://doi.org/10.1073/pnas.1016644108

Korneli, C., Danisman, S., & Staiger, D. (2014). Differential control of pre-invasive and post-invasive antibacterial defense by the Arabidopsis circadian clock. *Plant Cell Physiol.*, *55*(9), 1613–1622.

Kunitz, M. (1945). Crystallization of a trypsin inhibitor from soybean. *Science*, *101*, 668–669. https://doi.org/10.1126/science.101.2635.668

Kurihara, K., & Beidler, L. M. (1968). Taste-modifying protein from miracle fruit. *Science*, *161*, 1241–1243. https://doi.org/10.1126/science.161.3847.1241

Kurotani, K. I., Hirakawa, H., Shirasawa, K., Tanizawa, Y., Nakamura, Y., Isobe, S., & Notaguchi, M. (2023). Genome Sequence and Analysis of Nicotiana benthamiana, the Model Plant for Interactions between Organisms. *Plant and Cell Physiology*, *64*, 248–257. https://doi.org/10.1093/pcp/pcac168

Laskowski, M., & Kato, I. (1980). Protein inhibitors of proteinases. *Annual Review of Biochemistry*, *49*, 593–626. https://doi.org/10.1146/annurev.bi.49.070180.003113

Li, B., Meng, X., Shan, L., & He, P. (2016). Transcriptional regulation of pattern-triggered immunity in plants. *Cell Host Microbe*, *19*(5), 641–650.

Li, F., Li, W., Lin, Y. J., Pickett, J. A., Birkett, M. A., Wu, K., Wang, G., & Zhou, J. J. (2018). Expression of lima bean terpene synthases in rice enhances recruitment of a beneficial enemy of a major rice pest. *Plant Cell and Environment*, *41*, 111–120. https://doi.org/10.1111/pce.12959

Li, W., Wang, L., Zhou, F., Li, C., Ma, W., Chen, H., Wang, G., Pickett, J. A., Zhou, J. J., & Lin, Y. (2021). Overexpression of the homoterpene synthase gene, OsCYP92C21, increases emissions of volatiles mediating tritrophic interactions in rice. *Plant Cell and Environment*, *44*, 948–963. https://doi.org/10.1111/pce.13924

Li, Z., Bonaldi, K., Uribe, F., & Pruneda-Paz, J. L. (2018). A Localized Pseudomonas syringae Infection Triggers Systemic Clock Responses in Arabidopsis. *Current Biology*, *28*(4), 630-639.e4. https://doi.org/10.1016/j.cub.2018.01.001

Liang, T., Yu, S., Pan, Y., Wang, J., & Kay, S. A. (2024). The interplay between the circadian clock and abiotic stress responses mediated by ABF3 and CCA1/LHY. *Proc. Natl. Acad. Sci. U. S. A.*, *121*(7), e2316825121.

Ligterink, W., Kroj, T., zur Nieden, U., Hirt, H., & Scheel, D. (1997). Receptor-mediated activation of a MAP kinase in pathogen defense of plants. *Science*, *276*(5321), 2054–2057.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, *15*(12), 550.

Lu, H., McClung, C. R., & Zhang, C. (2017). Tick tock: Circadian regulation of plant innate immunity. *Annu. Rev. Phytopathol.*, *55*, 287–311.

Lu, S. X., Webb, C. J., Knowles, S. M., Kim, S. H. J., Wang, Z., & Tobin, E. M. (2012). CCA1 and ELF3 Interact in the control of hypocotyl length and flowering time in Arabidopsis. *Plant Physiol.*, *158*(2), 1079–1088.

Major, I. T., & Peter Constabel, C. (2008). Functional analysis of the kunitz trypsin inhibitor family in poplar reveals biochemical diversity and multiplicity in defense against herbivores. *Plant Physiology*, *146*(3), 888–903. https://doi.org/10.1104/pp.107.106229

Nagel, D. H., Doherty, C. J., Pruneda-Paz, J. L., Schmitz, R. J., Ecker, J. R., & Kay, S. A. (2015). Genome-wide identification of CCA1 targets uncovers an expanded clock network in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.*, *112*(34), E4802–10.

Ngou, B. P. M., Ding, P., & Jones, J. D. G. (2022). Thirty years of resistance: Zig-zag through the plant immune system. *Plant Cell*, *34*(5), 1447–1478.

Pearce, G., Strydom, D., Johnson, S., & Ryan, C. A. (1991). A Polypeptide from Tomato Leaves Induces Wound-Inducible Proteinase Inhibitor Proteins. *Science*, *253*, 895–898.

Poretsky, E., Dressano, K., Weckwerth, P., Ruiz, M., Char, S. N., Shi, D., Abagyan, R., Yang, B., & Huffaker, A. (2020). Differential activities of maize plant elicitor peptides as mediators of immune signaling and herbivore resistance. *Plant J.*, *104*(6), 1582–1602.

Rawat, R., Takahashi, N., Hsu, P. Y., Jones, M. A., Schwartz, J., Salemi, M. R., Phinney, B. S., & Harmer, S. L. (2011). REVEILLE8 and PSEUDO-REPONSE REGULATOR5 form a negative feedback loop within the arabidopsis circadian clock. *PLoS Genetics*, *7*. https://doi.org/10.1371/journal.pgen.1001350

Rawlings, N. D., Tolle, D. P., & Barrett, A. J. (2004). Evolutionary families of peptidase inhibitors. *Biochemical Journal*, *378*, 705–716. https://doi.org/10.1042/BJ20031825

Reymond, P., Bodenhausen, N., Van Poecke, R. M. P., Krishnamurthy, V., Dicke, M., & Farmer, E. E. (2004). A conserved transcript pattern in response to a specialist and a generalist herbivore. *Plant Cell*, *16*, 3132–3147. https://doi.org/10.1105/tpc.104.026120

Rodriguez-Saona, C., Kaplan, I., Braasch, J., Chinnasamy, D., & Williams, L. (2011). Field responses of predaceous arthropods to methyl salicylate: A meta-analysis and case study in cranberries. *Biological Control*, *59*, 294–303. https://doi.org/10.1016/j.biocontrol.2011.06.017

Ryan, C. A. (1990). Protease Inhibitors in Plants: Genes for Improving Defenses Against Insects and Pathogens. *Annual Review of Phytopathology*, *28*, 425–449. https://doi.org/10.1146/annurev.py.28.090190.002233

Schmelz, E. A., Alborn, H. T., Banchio, E., & Tumlinson, J. H. (2003). Quantitative relationships between induced jasmonic acid levels and volatile emission in Zea mays during Spodoptera exigua herbivory. *Planta*, *216*, 665–673. https://doi.org/10.1007/s00425-002-0898-y

Schmelz, E. A., Carroll, M. J., LeClere, S., Phipps, S. M., Meredith, J., Chourey, P. S., Alborn, H. T., & Teal, P. E. A. (2006). Fragments of ATP synthase mediate plant perception of insect attack. *Proc. Natl. Acad. Sci. U. S. A.*, *103*(23), 8894–8899.

Schmelz, E. A., Huffaker, A., Carroll, M. J., Alborn, H. T., Ali, J. G., & Teal, P. E. A. (2012). An amino acid substitution inhibits specialist herbivore production of an antagonist effector and recovers insect-induced plant defenses. *Plant Physiol.*, *160*(3), 1468–1478.

Schuman, M. C., & Baldwin, I. T. (2016). The Layers of Plant Responses to Insect Herbivores. *Annual Review of Entomology*, *61*, 373–394. https://doi.org/10.1146/annurev-ento-010715-023851

Snoeck, S., Abramson, B. W., Garcia, A. G. K., Egan, A. N., Michael, T. P., & Steinbrenner, A. D. (2022). Evolutionary gain and loss of a plant pattern-recognition receptor for HAMP recognition. *Elife*, *11*, e81050.

Snoeck, S., Garcia, A. G. K., & Steinbrenner, A. D. (2023). Plant Receptor-like proteins (RLPs): Structural features enabling versatile immune recognition. *Physiol. Mol. Plant Pathol.*, *125*(102004), 102004.

Snoeck, S., Guayazán-Palacios, N., & Steinbrenner, A. D. (2022). Molecular tug-of-war: Plant immune recognition of herbivory. *Plant Cell*, *34*(5), 1497–1513.

Song, H. K., & Suh, S. W. (1998). Kunitz-type soybean trypsin inhibitor revisited: Refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from Erythrina caffra and tissue-type plasminogen activator. *Journal of Molecular Biology*, *275*, 347–363. https://doi.org/10.1006/jmbi.1997.1469

Srinivasan, R., Tamò, M., & Malini, P. (2021). Emergence of Maruca vitrata as a Major Pest of Food Legumes and Evolution of Management Practices in Asia and Africa. *Annual Review of Entomology*, *66*, 141–161. https://doi.org/10.1146/annurev-ento-021220-084539

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313.

Steinbrenner, A. D., Muñoz-Amatriaín, M., Chaparro, A. F., Aguilar-Venegas, J. M., Lo Sassoum and Okuda, S., Glauser, G., Dongiovanni Julien and Shi, D., Hall, M., Crubaugh, D., Holton, N., Zipfel, C., Abagyan, R., Turlings, T. C. J., Close, T. J., Huffaker, A., & Schmelz, E. A. (2020). A receptor-like protein mediates plant immune responses to herbivore-associated molecular patterns. *Proc. Natl. Acad. Sci. U. S. A.*, *117*(49), 31510–31518.

Steinbrenner, A. D., Saldivar, E., Hodges, N., Guayazán-Palacios, N., Chaparro, A. F., & Schmelz, E. A. (2022). Signatures of plant defense response specificity mediated by herbivore-associated molecular patterns in legumes. *Plant Journal*, *110*(5), 1255–1270. https://doi.org/10.1111/tpj.15732

Sultana, M. S., Mazarei, M., Jurat-Fuentes, J. L., Hewezi, T., Millwood, R. J., & Stewart, C. N. (2023). Overexpression of soybean trypsin inhibitor genes decreases defoliation by corn earworm (Helicoverpa zea) in soybean (Glycine max) and Arabidopsis thaliana. *Frontiers in Plant Science*, *14*. https://doi.org/10.3389/fpls.2023.1129454

Sun, C., Shao, Y., & Iqbal, J. (2025). A comprehensive cell atlas of fall armyworm (Spodoptera frugiperda) larval gut and fat body via snRNA-Seq. *Scientific Data*, *12*, 250. https://doi.org/10.1038/s41597-025-04520-z

Sun, T., & Zhang, Y. (2022). MAP kinase cascades in plant development and immune signaling. *EMBO Rep.*, *23*(2), e53817.

Sun, Y., Wang, Y., Zhang, X., Chen, Z., Xia, Y., Wang, L., Sun, Y., Zhang, M., Xiao, Y., Han, Z., Wang, Y., & Chai, J. (2022). Plant receptor-like protein activation by a microbial glycoside hydrolase. *Nature*, *610*, 335–342. https://doi.org/10.1038/s41586-022-05214-x

Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., von Heijne, G., & Nielsen, H. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*, *40*, 1023–1025. https://doi.org/10.1038/s41587-021-01156-3

Turlings, T. C. J., & Erb, M. (2018). Tritrophic Interactions Mediated by Herbivore-Induced Plant Volatiles: Mechanisms, Ecological Relevance, and Application Potential. *Annual Review of Entomology*, *63*, 433–452. https://doi.org/10.1146/annurev-ento-020117-043507

Turlings, T. C. J., McCall, P. J., Alborn, H. T., & Tumlinson, J. H. (1993). An elicitor in caterpillar oral secretions that induces corn seedlings to emit chemical signals attractive to parasitic wasps. *J. Chem. Ecol.*, *19*(3), 411–425.

Wang, L., Einig, E., Almeida-Trapp, M., Albert, M., Fliegmann, J., Mithöfer, A., Kalbacher, H., & Felix, G. (2018). The systemin receptor SYR1 enhances resistance of tomato against herbivorous insects. *Nature Plants*, *4*, 152–156. https://doi.org/10.1038/s41477-018-0106-0

Wang, S., Steed, G., & Webb, A. A. R. (2022). Circadian entrainment in Arabidopsis. *Plant Physiology*, *190*, 981–993. https://doi.org/10.1093/plphys/kiac204

Wang, W., Barnaby, J. Y., Tada, Y., Li, H., Tör, M., Caldelari, D., Lee, D.-U., Fu, X.-D., & Dong, X. (2011). Timing of plant immune responses by a central circadian regulator. *Nature*, *470*(7332), 110–114.

Wang, Z. Y., Kenigsbuch, D., Sun, L., Harel, E., & Ong M S and Tobin, E. M. (1997). A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell*, *9*(4), 491–507.

Weber, E., Engler, C., Gruetzner, R., Werner, S., & Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE*, *6*(2). https://doi.org/10.1371/journal.pone.0016765

Weiss, J., Terry, M. I., Martos-Fuentes, M., Letourneux, L., Ruiz-Hernández, V., Fernández, J. A., & Egea-Cortines, M. (2018). Diel pattern of circadian clock and storage protein gene expression in leaves and during seed filling in cowpea (Vigna unguiculata). *BMC Plant Biol.*, *18*(1), 33.

Zeng, J., Ye, W., Hu, W., Jin, X., Kuai, P., Xiao, W., Jian, Y., Turlings, T. C. J., & Lou, Y. (2023). The N-terminal subunit of vitellogenin in planthopper eggs and saliva acts as a reliable elicitor that induces defenses in rice. *New Phytologist*, *238*, 1230–1244. https://doi.org/10.1111/nph.18791

Zhang, C., Xie, Q., Anderson, R. G., Ng, G., Seitz, N. C., Peterson, T., McClung, C. R., McDowell, J. M., Kong, D., Kwak, J. M., & Lu, H. (2013). Crosstalk between the Circadian Clock and Innate Immunity in Arabidopsis. *PLoS Pathogens*, *9*(6). https://doi.org/10.1371/journal.ppat.1003370

Zhang, C., Xie, Y., He, P., & Shan, L. (2024). Unlocking nature's defense: Plant pattern recognition receptors as guardians against pathogenic threats. *Mol. Plant. Microbe. Interact.*, *37*(2), 73–83.

Zhou, K., Wu, F., Deng, L., Xiao, Y., Yang, W., Zhao, J., Wang, Q., Chang, Z., Zhai, H., Sun, C., Han, H., Du, M., Chen, Q., Yan, J., Xin, P., Chu, J., Han, Z., Chai, J., Howe, G. A., … Li, C. (2025). Antagonistic systemin receptors integrate the activation and attenuation of systemic

wound signaling in tomato. *Developmental Cell*. https://doi.org/10.1016/j.devcel.2024.11.005

Zhou, S., & Jander, G. (2022). Molecular ecology of plant volatiles in interactions with insect herbivores. *Journal of Experimental Botany*, *73*, 449–462. https://doi.org/10.1093/jxb/erab413

Zhou, W., Brockmöller, T., Ling, Z., Omdahl, A., Baldwin, I. T., & Xu, S. (2016). Evolution of herbivore-induced early defense signaling was shaped by genome-wide duplications in Nicotiana. *Elife*, *5*, e19531.

Zhu-Salzman, K., & Zeng, R. (2015). Insect response to plant defensive protease inhibitors. *Annual Review of Entomology*, *60*, 233–252. https://doi.org/10.1146/annurev-ento-010814-020816

Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Springer New York. https://doi.org/10.1007/978-0-387-87458-6