UNIVERSITY OF CALIFORNIA SAN DIEGO

Probing layers of maize immunity through integration of genetic, transcriptomic and physiological approaches

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Biology

by

Elly Poretsky

Committee in charge:

Professor Alisa Huffaker, Chair
Professor Steven P. Briggs
Professor Wolfgang Busch
Professor Matthew D. Daugherty
Professor Trey Ideker

2021

The dissertation of Elly Poretsky is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

TABLE OF CONTENTS

Dissertation Approval Page iii

Table of Contents iv

List of Figures v

List of Supplemental Files vii

Acknowledgements viii

Vita x

Abstract of the Dissertation xi

Chapter 1. Comparative analysis of Pep- and FAC-induced antiherbivore responses in maize reveals a single QTL for FAC sensitivity 1

Chapter 2. Differential activities of maize plant elicitor peptides as mediators of immune signaling and herbivore resistance 65

Chapter 3. MutRank: an R shiny web-application for exploratory targeted mutual rank-based coexpression analyses integrated with user-provided supporting information ... 91

LIST OF FIGURES

Figure 1-1. Comparative analysis of ZmPep3- and Gln-18:3-induced VOC emission and early transcriptional changes 53

Figure 1-2. Early elicitor-induced transcriptional regulation of multiple components associated with antiherbivore responses 54

Figure 1-3. Screening *Zea mays* inbred lines for intact elicitor-induced defense responses but impaired FAC sensitivity 55

Figure 1-4. Genetic mapping for an FAC sensitivity locus based on Mo17 insensitivity to Gln-18:3 56

Figure 1-5. Fine-mapping the FAC sensitivity locus using the B73 x Mo17 NIL and IBM mapping populations 57

Figure 1-6. Fine-mapping the FAC sensitivity locus using a newly generated marker map for 105 IBM lines 58

Figure 1-7. A candidate genes within the FAC sensitivity locus denoted FACS is an ortholog of *Oryza sativa* OsLRR-RLK1 59

Figure 1-8. Transient expression in *N. benthamiana* of the FACS candidate gene cloned from B73 and Mo17 increases responsiveness to Gln-18:3 60

Figure 1-S1 Comparison of overlapping ZmPep3- and Gln-18:3-induced DEGs 61

Figure 1-S2. Genetic mapping sensitivity locus using the fold-change (FC) values of VOCs from three different biosynthetic pathways 62

Figure 1-S3. Alignment of the encoded amino acid FACS sequences of B73 and Mo17 63

Figure 1-S4. Native Gln-18:3 sensitivity in *N. benthamiana* is reduced in plants grown under diminished light intensity but heterologous expression of FACS increases responsiveness to Gln-18:3 64

Figure 2-1. Location and structure of maize PROPEP genes 68

Figure 2-2. The strength of *Zea mays* plant elicitor peptide (Pep) elicitor activity corresponds to the relative magnitude of induced changes in phytohormone levels 69

Figure 2-3. Gene expression changes elicited by *Zea mays* plant elicitor peptides (ZmPeps) correspond to the magnitude of induced changes in phytohormone levels and requires ethylene biosynthesis 72

Figure 2-4. Structure–function analysis of residues critical for *Zea mays* plant elicitor peptide 3 (ZmPep3) elicitor activity 73

Figure 2-5. ZmPep5a may act as an antagonist to negatively regulate responses induced by *Zea mays* plant elicitor peptides (ZmPeps) 75

Figure 2-6. ZmPEPR genes encode active receptors with differential endogenous expression in response to elicitor treatments 76

Figure 2-7. ZmPEPR CRISPR/Cas9-generated knockout lines implicate ZmPEPR1 as the primary contributor to *Zea mays* plant elicitor peptide (ZmPep)-mediated anti-herbivore defense responses 78

Figure 2-S1. Structure of maize PROPEP gene loci in the Nested Association Mapping (NAM) parent lines 86

Figure 2-S2. Sequence conservation of individual ZmPeps across NAM parent lines .. 87

Figure 2-S3. Relative volatile terpene-inducing activity for peptides from derived from the ZmPROPEP4.1 and ZmPROPEP4.2 duplicate precursor genes encoded in the B73 genome 88

Figure 2-S4. Relationship between relative levels of each phytohormone in leaves treated with water or ZmPeps 89

Figure 3-1. MutRank interface and workflow chart 94

Figure 3-2. Example workflow 1: validation of MutRank using a characterized biosynthetic pathway 96

Figure 3-3. Example workflow 2: using MutRank to predict enzymes in specialized metabolism 99

LIST OF SUPPLEMENTAL FILES

Supplemental File 1. Supplemental Tables 1-S1-11

Supplemental File 2. Supplemental Tables 2-S1-8

Supplemental File 3. Supplemental Tables 3-S1

Supplemental File 4. Supplemental Tables 3-S2-8

ACKNOWLEDGEMENTS

I would like to thank my family for the continuous support at every step of my education. Their support, patience and faith were an important factor for reaching this point. I would also like to thank all my friends throughout the years, that, each in their own way, contributed to the progression of my scientific path. A special thanks to Dr. Alyona Bobkova for offering endless support and guidance, from my very first steps at UCSD and all the way through my graduate education.

I would also like to thank members of the Huffaker and Schmelz group, in particular Dr. Philipp Weckwerth, Dr. Keini Dressano, Dr. Adam Steinbrenner, Dr. Yezhang Ding, Dr. Sowmya Poosapati and Mr. Miguel Ruiz for their continuous support, advice, mentoring and insightful discussions. I will fondly remember the long days in the lab, field and greenhouse. I learned a considerable amount about plants from the many conversations we had.

I would also like to thank Dr. Eric Schmelz that in addition to the intellectual discussions and advice throughout my graduate years, has also offered his technical expertise and access to lab instruments that were all indispensable.

Finally, I would like to express my sincere gratitude to my advisor, Dr. Alisa Huffaker for her dedication to my success as a scientist. Her mentorship was the perfect combination of actively participating and supporting the successful fulfillment of research goals while encouraging the pursuit of new research ideas. Her guidance contributed enormously to my development as a scientist and made a great impact on my future research interests.

Chapter 1, in part, is currently being prepared for submission for publication of the material. Poretsky, E, Ruiz, M, Ahmadian, N, Steinbrenner, AD, Dressano, K, Schmelz, EA, Huffaker, A. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material as it appears in The Plant Journal 2020. Poretsky, E., Dressano, K., Weckwerth, P., Ruiz, M., Char, S.N., Shi, D., Abagyan, R., Yang, B. and Huffaker, A. (2020) *The Plant Journal*. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material as it appears in PeerJ 2020. Poretsky, E. and Huffaker, A. (2020) *Peer J*. The dissertation author was the primary investigator and author of this paper.

VITA

2014 Bachelors of Science, The Open University of Israel

2015-2021 Graduate Researcher, University of California San Diego

2021 Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Poretsky, E., & Huffaker, A. (2020). MutRank: An R shiny web-application for exploratory targeted mutual rank-based coexpression analyses integrated with user-provided supporting information. PeerJ, 8, e10264.

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. The Plant Journal, tpj.15022.

Dressano, K., Weckwerth, P. R., Poretsky, E., Takahashi, Y., Villarreal, C., Shen, Z., Schroeder, J. I., Briggs, S. P., & Huffaker, A. (2020). Dynamic regulation of Pep-induced immunity through post-translational control of defence transcript splicing. Nature Plants, 6(8), 1008–1019.

Ding, Y., Weckwerth, P. R., Poretsky, E., Murphy, K. M., Sims, J., Saldivar, E., Christensen, S. A., Char, S. N., Yang, B., Tong, A., Shen, Z., Kremling, K. A., Buckler, E. S., Kono, T., Nelson, D. R., Bohlmann, J., Bakker, M. G., Vaughan, M. M., Khalil, A. S., … Huffaker, A. (2020). Genetic elucidation of interconnected antibiotic pathways mediating maize innate immunity. Nature Plants, 6(11), 1375–1388.

Ding, Y., Murphy, K. M., Poretsky, E., Mafu, S., Yang, B., Char, S. N., Christensen, S. A., Saldivar, E., Wu, M., Wang, Q., Ji, L., Schmitz, R. J., Kremling, K. A., Buckler, E. S., Shen, Z., Briggs, S. P., Bohlmann, J., Sher, A., Castro-Falcon, G., … Schmelz, E. A. (2019). Multiple genes recruited from hormone pathways partition maize diterpenoid defences. Nature Plants, 5(10), 1043–1056. https://doi.org/10.1038/s41477-019-0509-6.

Brandt, B., Munemasa, S., Wang, C., Nguyen, D., Yong, T., Yang, P. G., Poretsky, E., Belknap, T. F., Waadt, R., Alemán, F., & Schroeder, J. I. (2015). Calcium specificity signaling mechanisms in abscisic acid signal transduction in Arabidopsis guard cells. ELife, 4, e03599.

FIELDS OF STUDY

Major Field: Plant Biology

ABSTRACT OF THE DISSERTATION

Probing layers of maize immunity through integration of genetic, transcriptomic and physiological approaches

by

Elly Poretsky

Doctor of Philosophy in Biology

University of California San Diego, 2021

Professor Alisa Huffaker, Chair

To efficiently protect themselves against pests and disease, plants surveil for attacking organisms and upon recognition, activate protective inducible defenses. Here, I present my work on the regulation and function of maize inducible defenses by integrating genetic, transcriptomic and physiological approaches. This work elucidated mechanisms underlying three layers of the maize immune response, including: (1) A novel genetic locus associated with sensitivity to exogenous herbivore-associated elicitors of the fatty-acid amino-acid conjugate (FAC) family, (2) regulatory function of phytocytokines from the Plant elicitor peptide (Pep) family, and (3) biosynthesis of

antibiotic specialized metabolite defenses. Early maize signaling events triggered in the context of herbivory, were probed through comparative transcriptomic analyses upon treatment with ZmPeps and FACs, indicating a largely shared signaling pathway and identifying specific genes involved in antiherbivore defense. Genetic mapping using the Intermated B73 x Mo17 mapping population derived from B73, an FAC sensitive line, and Mo17, an FAC insensitive line, identified a single locus on chromosome 4 associated with FAC sensitivity that was further fine-mapped to a region containing 19 genes. A candidate gene within this region, FAC SENSITIVITY-ASSOCIATED (FACS), was expressed at significantly lower levels in the insensitive parent line, and heterologous expression of FACS increased FAC sensitivity in *Nicotiana benthamiana*, suggesting a role in regulation of FAC-induced responses. Work characterizing the maize ZmPep family led to several new insights into Pep signaling mechanisms: Maize Pep precursors (PROPEPs) were found to contain multiple nested active peptides, a phenomenon not previously observed for this family. Additionally, in contrast to Peps in Arabidopsis, individual maize Peps were found to have specific activities defined by the relative magnitude of elicited responses through rheostat-like tuning of phytohormone levels. Finally, peptide structure-function analysis and physiological assays identified ZmPep5a as a potential antagonist peptide. Finally, we report on the development of an R Shiny web-application that was developed to facilitate mutual rank-based coexpression analyses integrating user-provided supporting information. The utility of this user-friendly app was demonstrated through application to define two new biosynthetic pathways for maize terpenoid antibiotics.

Chapter 1: Comparative analyses of exogenous and endogenous antiherbivore elicitors enables a forward genetics approach to identify maize gene candidates mediating sensitivity to Herbivore Associated Molecular Patterns (HAMP)

ABSTRACT

Crop damage by herbivorous insects remains a significant contributor to annual yield reductions. Following attack, maize (*Zea mays*) responds to Herbivore Associated Molecular Patterns (HAMPs) and Damage Associated Molecular Patterns (DAMPs), activating dynamic direct and indirect anti-herbivore defense responses. To define underlying signaling processes, comparative analyses between Plant Elicitor Peptide (Pep) DAMPs and fatty acid-amino acid conjugate (FAC) HAMPs was conducted. Using RNA-seq to probe the early transcriptional changes following Pep and FAC treatment revealed quantitative differences in the strength of response but qualitative similarities, providing evidence for a shared signaling pathway. In further comparisons of FAC and Pep responses across diverse maize inbred lines, we identified Mo17 as part of a small subset inbred lines displaying selective FAC insensitivity. Genetic mapping for FAC sensitivity using the Intermated B73 x Mo17 mapping population identified a single locus on chromosome 4 associated with FAC sensitivity with multiple fine-mapping approaches narrowing the locus to 19 candidate genes. The top candidate gene was identified as a leucine-rich repeat receptor-like kinase (LRR-RLK), termed FAC Sensitivity (ZmFACS), that is orthologous to a rice gene previously associated with activation of induced responses to diverse Lepidoptera. Consistent with reduced sensitivity, ZmFACS expression was significantly lower in Mo17 as compared to B73. Transient heterologous expression of ZmFACS in Nicotiana benthamiana resulted in significantly increased

responses to FACs. Together, our results provide useful resources for studying early elicitor-induced antiherbivore responses in maize, and for better understanding gene candidates underlying the genetic basis of HAMP sensitivity in grain crops.

INTRODUCTION

Crop stress driven by insect pests and disease can cause 50% losses of total annual yield, with increased severity of environmental stresses expected to exacerbate the future losses (Chakraborty and Newton, 2011). Among the more damaging insect pests are lepidoptera in the family Noctuidae which include many *Spodoptera* species (Parra *et al.*, 2021). The fall armyworm (FAW; *Spodoptera frugiperda*) is highly polyphagous pest that attacks over 350 host plants across 76 plant families (Montezano *et al.*, 2018). Despite success as a generalist, FAW exhibit measurable specialization on grain crops, driving defoliation, seedling loss and introducing fungal pathogens contaminating grain with mycotoxins (Overton *et al.*, 2021). Together with native Heliothis spp. and Helicoverpa spp crop pests in the Americas, FAW has been partially controlled by transgenic stacking *Bacillus thuringensis* (Bt) cyotoxin-encoding genes in many crops (Shehryar *et al.*, 2020). However, despite the robust protection of stacked Bt genes, evidence is emerging from Brazil that FAW is evolving resistance to Bt-mediated crop protection (Horikoshi *et al.*, 2016). Furthermore, FAW has emerged as a formidable invasive pest. FAW was first detected in West Africa in 2016, but has now spread across the entire continent and causes billions of dollars in annual losses (Day *et al.*, 2017). Spreading further, FAW entered India in 2018, rapidly proliferated throughout Asia and is now in Australia (Overton *et al.*, 2021). Given the global challenge posed by FAW and

potential breakdown of current control measures, new knowledge of plant resistance mechanisms and control strategies are essential to reduce crop losses (Douglas, 2018).

As a major crop species attacked by diverse lepidoptera including *Spodoptera exigua* and FAW, maize (*Zea mays*) has been a leading research model for understanding plant responses to insect attack. For example, in what is now appreciated as a common phenomenon, indirect plant defense responses against lepidopteran pests were first described in maize (Turlings *et al.*, 1990). Volatile organic chemicals (VOC) emitted by leaves following *Spodoptera* herbivory can attract parasitoid wasps such as *Cotesia marginiventris,* which protect plants by parasitizing larvae (Turlings *et al.,* 1990). Herbivore-elicited volatiles are produced by young leaves, and while there is qualitative and quantitative variation among inbred lines, the blend is largely dominated by sesquiterpenes and monoterpenes along with fatty acid-derived green leafy volatiles (GLVs), indole and methyl anthranilate (Maffei, 2010). In addition to attracting parasitoids, indole and GLV components of the volatile blend act in interplant communication, priming defense responses in undamaged neighboring plants (Engelberth *et al.*, 2004; Erb *et al.*, 2015). Volatiles also mediate indirect maize defenses belowground; herbivory by Western corn rootworm (WCR; *Diabrotica virgifera*) larvae triggers emission of the sesquiterpene beta-caryophyllene, attracting entomophagous nematodes that prey on the larvae (Rasmann *et al.*, 2005). In addition to volatiles, maize also produces complex blends of directly protective chemicals and defensive proteins that vary with tissue, developmental stage and genetic background. Toxic and antifeedant chemicals include phenylpropanoids such as the silk-localized flavone glucoside toxin maysin (Waiss *et al.*,

1979; Casas *et al.*, 2016), numerous benzoxazinoid toxins such as dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside (DIMBOA-Glc) and 2-hydroxy-4,7-dimethoxy-1,4-benzoxazin-3-one glucoside (HDMBOA-Glc) (Oikawa *et al.*, 2004; Maag *et al.*, 2016; Wouters *et al.*, 2016), and acidic diterpenoids that have antifeedant activity (Schmelz *et al.*, 2011). While some maize defenses such as DIMBOA-Glc are constitutively present in young seedlings, the production of many defenses is upregulated by herbivory, enabling added protection against attack by minimizing production of costly defenses in the absence of herbivory (Erb, 2018; Fürstenberg-Hägg *et al.*, 2013; Mithöfer and Boland, 2012). Inducible responses may be variable across genotypes and environments, but the presence or absence of antiherbivore defenses is a major factor determining insect resistance (Chen *et al.*, 2009; Smith *et al.*, 2012).

Maize has also been a model for identification of molecules from insects that trigger protective responses. Maize response bioassays informed activity-guided fractionation efforts enabling discovery of the first precisely identified biochemicals from insect oral secretions (OS) that act as defense elicitors, termed herbivore-associated molecular patterns (HAMPs) (Alborn *et al.*, 1997; Felton and Tumlinson, 2008). Originally isolated from *Spodoptera exigua*, fatty acid-amino acid conjugates (FACs) are a family of molecules based on the conjugation of linolenic acid to either glutamine or glutamate in the insect (Alborn *et al.*, 1997; Pare *et al.*, 1998; Yoshinaga *et al.*, 2008; Lait *et al.*, 2003; Halitschke *et al.*, 2001). Among the naturally occurring FACs, 17-hydroxy *N*-linolenoyl L-glutamine (volicitin) *N*-linolenoyl L-glutamine (Gln-18:3) and are the mostly highly abundant and potent elicitors of foliar volatile emissions (Mori and Yoshinaga, 2011;

Schmelz *et al.*, 2009; Turlings *et al.*, 2000; Yoshinaga *et al.*, 2010; Yoshinaga *et al.*, 2008). FACs occur in diverse insects and play a nutritional role by increasing nitrogen assimilation efficiency in midgut tissues (Mori and Yoshinaga, 2011; Yoshinaga *et al.*, 2008; Yoshinaga *et al.*, 2007). FACs are potent defense elicitors in diverse plants, including maize, rice, soybean, Medicago and many solanaceous species (Grissett *et al.*, 2020; Turlings *et al.*, 2000; Shinya *et al.*, 2016; Wu and Baldwin, 2009). Although FACs have been the dominant HAMP studied in maize, additional insect-associated molecules also promote maize antiherbivore defenses. Disulfooxy fatty acid caeliferins isolated from the American bird grasshopper (*Schistocerca americana*) elicit defense responses in maize (Alborn *et al.*, 2007), as do yet unknown molecules in *Helicoverpa zea* frass and microbes associated with the insect digestive tract (Ray *et al.*, 2015; J., Wang *et al.*, 2018). Finally, plant hormones contained in both larval oral secretions and frass modulate maize immunity as well (Dafoe *et al.*, 2013; Acevedo *et al.*, 2019). While these diverse insect-associated molecules contribute to elicitation of maize defenses, maize is insensitive to numerous other oral cues from chewing insects such as Lepidopteran-produced inceptin, Helicoverpa-associated glucose oxidase and ATPases, and β-glucosidase from *Pieris brassicae* (Schmelz *et al.*, 2007; J., Wang *et al.*, 2018; Mattiacci *et al.*, 1995).

Maize signaling promoted by HAMPs is mediated and amplified by an array of endogenous signals (Schmelz *et al.*, 2003; Huffaker *et al.*, 2013; Poretsky *et al.*, 2020; Schmelz, 2015). Genarally, mechanical wounding of plant tissue leads to the release of damage-associated molecular patterns (DAMPs) including oligogalacturonic acid,

extracellular ATP, and peptides such as systemin and Plant Elicitor Peptides (Peps) (Huffaker *et al.*, 2006; Orozco-Cardenas and Ryan, 1999; Pearce *et al.*, 1991; Tanaka *et al.*, 2014). In addition to HAMPs, DAMPs further amplify wounding-mediated production of phytohormones including jasmonic acid (JA) and ethylene (ET) to regulate herbivore-associated defense responses (Diezel *et al.*, 2009; Erb *et al.*, 2012; Schmelz *et al.*, 2003; Shinya *et al.*, 2018). Additionally, rapid signaling cascades involving glutamate receptor-like proteins, MAP kinase (MAPK) cascades, Ca^{2+} influxes and bursts of reactive oxygen species serve to propagate immune signaling both spatially and temporally (Erb and Reymond, 2019). As amplifiers of maize immune signaling, maize encodes 13 ZmPeps contained in 6 precursor protein genes (Huffaker *et al.*, 2011; Huffaker *et al.*, 2013; Poretsky *et al.*, 2020). Peps in rice (OsPeps) have also been demonstrated to protect against herbivores, and act synergistically with HAMPs to generate stronger responses (Shinya *et al.*, 2018). While each ZmPep varies in the magnitude of elicited responses, they commonly promote JA and ET production, VOC emission, and accumulation of transcripts encoding proteinase inhibitors and other defense proteins. Among the ZmPep family, ZmPep3 is the most potent DAMP signal (Huffaker *et al.*, 2013; Poretsky *et al.*, 2020). ZmPeps are recognized by ZmPEPR1 and ZmPEPR2 receptors, and plants with lesions in ZmPEPR genes produce fewer volatiles and are less capable of generating a protective response against *Spodoptera* larvae after ZmPep treatment (Poretsky *et al.*, 2020). This is consistent with other studies demonstrating that impairments to wound and DAMP signaling commonly result in reduced herbivore resistance, and support functional roles for the interconnected signaling pathways (Onkokesung *et al.*, 2010; Orozco-Cardenas *et al.*, 1993; Poretsky *et al.*, 2020; Thaler *et al.*, 2002; L., Wang *et al.*, 2018).

Direct comparisons in maize of exogenous HAMPs and endogenous DAMPs, such as ZmPep3 and Gln-18:3, revealed striking overlap in the elicitation of defenses and protective responses against *Spodoptera* herbivores (Huffaker *et al.*, 2013). Given the highly similar activation of defenses, we sought to understand early maize responses to ZmPep3 and Gln-18:3 and comprehensively assess transcriptome wide overlap and response divergence. Transcriptional profiling showed that while ZmPep3 was a more potent signal, both ZmPep3 and Gln-18:3 promoted highly similar reprogramming responses at 2 hours, largely represented by transcripts encoding signaling proteins. We identify maize genes rapidly responding to both HAMP and DAMP signals that globally display over 70% overlap. Towards the goal of uncoupling the highly similar HAMP and DAMP responses, characterization ZmPep3 and Gln-18:3 sensitivity across diverse maize lines revealed defined inbreds specifically insensitive to Gln-18:3. Association mapping using the Intermated B73 x Mo17 (IBM) Recombinant Inbred Line (RIL) population to identification of a single locus specifically associated with response sensitivity to Gln-18:3, but not ZmPep3. This is the first time a reverse genetics approach in any plant has revealed a locus associated with sensitivity to FACs. Fine-mapping and characterization of this locus led to identification of an LRR-RLK gene, termed FAC Sensitivity (*ZmFACS*), as a predicted gene contributing to Gln-18:3 sensitivity. Using 2 diverse approaches, our work expands the current knowledge of defined maize genes involved in early signaling responses to defined HAMPs and DAMPs. Furthermore, we provide a long-sought path to uncoupling linked HAMP and DAMP responses.

RESULTS

Comparison of early HAMP- and DAMP-elicited transcriptional changes in maize

Comparative analyses of canonical defense responses against herbivores in maize and rice have revealed considerable connections between HAMP- and DAMP-elicited responses, consistent with shared signaling pathways (Huffaker *et al.*, 2013; Shinya *et al.*, 2018). To confirm that Gln-18:3 and ZmPep3 elicit anti-herbivore defenses in the maize B73 inbred similar to previous observations hybrid sweet-corn (Huffaker *et al.*, 2013), B73 seedlings were treated with water, ZmPep3 and Gln-18:3, and volatile organic compound (VOC) emission was measured 16 hours after elicitor treatment. In support of earlier findings, both ZmPep3 and Gln-18:3 treatments result in significantly higher levels of VOC emission to compared to water-treated samples (Fig. 1A). At equal concentrations ZmPep3 generates stronger induced volatile production than Gln-18:3 (Fig. 1A). To comprehensively compare early responses to Peps and FACs, B73 leaves were treated with H_2O, ZmPep3 and Gln-18:3 RNA-seq-based transcriptomes were generated for the 2-hour time point (Table S1). Analyses of the number of differentially expressed genes (DEGs) as compared to water-treated controls showed a similar pattern to that observed for elicitor-induced VOC emission, , namely a greater number of ZmPep3 elicited DEGs (1703) compared to Gln-18:3 elicited DEGs (358) (Fig. 1B, Table S2). While a quantitative difference in ZmPep3 and Gln-18:3 DEGs exists, Euler diagram analyses of the overlap between the combined up- and down-regulated DEGs reveals that Gln-18:3 responses display a high degree of overlap with ZmPep3 responses (Fig. 1C, Table S2). Specifically, 87% of all Gln-18:3 DEGs were also differentially expressed following

ZmPep3 treatment. In the context of up-regulated DEGs, Gln-18:3 elicited transcripts displayed 92.4% overlap with ZmPep3 responses (Fig. S1A). Further supporting largely shared processes, none of the upregulated DEGs in either treatment were downregulated in the other treatment, and vice versa (Fig. S1A).

To consider biological processes differentially regulated following ZmPep3 and Gln-18:3 treatment, both the maize Phytozome Gene Ontology (GO) and maize MAPMAN annotations were used (The Gene Ontology Consortium, 2019; Goodstein *et al.*, 2012; Thimm *et al.*, 2004). To visualize large-scale differences in the mean expression values of numerous enriched GO terms, a heatmap representing the mean expression values of the DEGs associated with each term was generated (Fig. 1D, Table S3). Among the GO terms enriched in the upregulated DEGs were terms associated with defense signaling and defense responses, including phytohormone and kinase signaling, specialized metabolism, response to wounding and defense to insects, that had the highest mean expression in the ZmPep3 treated leaves followed by the Gln-18:3 treated leaves (Fig. 1D, Table S3). Among the GO terms enriched in the downregulated DEGs were terms associated with growth and development, including red light signaling, gibberellic acid (GA) signaling and regulation of leaf morphogenesis, that had the lowest mean expression in the ZmPep3 treated leaves followed by Gln-18:3 treatment (Fig. 1D, Table S3). In contrast to GO terms, the maize MAPMAN bin annotation provides more specific gene groups and (Thimm *et al.*, 2004). Enrichment analysis revealed a diverse set of MAPMAN bins associated with ZmPep3- and Gln-18:3-upregulated genes, including PAMP-triggered immunity, MAPKs, indole biosynthesis, Guard cell S-type anion

channels SLAC anion channels, PIP/PIPL peptide signaling L-lectin receptors and transcription factors (TF)s in the myeloblastosis (MYB), WRKY domain (WRKY), basic helix–loop–helix (bHLH), TIFY and DREB families (Table S4). Enriched MAPMAN bins associated with ZmPep3- and Gln-18:3-downregulated genes included PHYTOCHROME B, cell cycle and ABA signaling, biosynthesis of gibberellic acid and brassinosteroids, and TFs from the MYB, C2H2, bHLH and Trihelix families (Table S4). Of the 51 enriched bins for ZmPep3-upregulated genes, only 13 contained more than 10 DEGs, which were associated with 6 different TF families, L- and G-lectin receptor families, and enzymes with peptidase, oxidoreductases, acyltransferases, glycosyltransferase and phosphotransferase activity (Fig. 1E, Table S4). Additionally, heatmap visualization of the ranked FPKM expression data of all ZmPep3-upregulated genes in these enriched bins showed that the majority of ZmPep3 upregulated DEGs also exhibited higher mean expression following Gln-18:3 treatment compared to the water treatment (Fig. 1E). While nearly 5-fold greater ZmPep3 DEGs occurred compared to Gln-18:3 DEGs, the ranked mean FPKM of all DEGs showed that for over 95% of upregulated DEGs, water treated samples had the lowest mean expression and that for over 95% of downregulated genes, water treated samples had the highest mean expression (Table S5). Overall trends in early transcriptional responses following ZmPep3 and Gln-18:3 treatment support highly similar regulation.

Identification of rapidly differentially expressed genes in response to HAMP and DAMP signals

A comparative analysis of the ZmPep3 and Gln-18:3 DEGs was used to probe transcriptional changes in gene groups and pathways with putative roles in regulation of antiherbivore defenses. MAPMAN pathway annotation and gene descriptions were used to group genes with shared functions (Thimm *et al.*, 2004). Groups included genes involved in signal transduction across membranes, MAPK signaling, phytohormone biosynthesis and signaling, transcription factors, cytoplasmic signaling, and antiherbivore defenses (Fig. 2, Table S6). After identifying the ZmPep3 DEGs in selected groups, a heatmap was generated illustrating the fold change values of the DEGs and including the number of DEGs compared to the total number of genes in each group (Fig. 2, Table S6). Genes that were also significantly differentially regulated following Gln-18:3 treatment were marked using thatched, cyan-colored boxes (Fig. 2, Table S6). Together this assembled a landscape of genes involved in signaling, including leucine-rich repeat receptor-like kinases (LRR-RLKs) and LRR-receptor-like proteins (LRR-RLPs), lectin receptors, glutamate-like receptors (GLR), respiratory burst oxidase homologs (RBOH), calcium-dependent protein kinases (CDPK), receptor-like cytoplasmic kinases (RLCK), genes involved in ubiquitylation, MAPK-associated genes and genes associated with multiple phytohormone signaling pathways. Among the genes involved in signaling, the lectin receptors had the highest proportion of DEGs compared to the group size (30%) while the LRR-RLK/RLP and ubiquitylation groups had among the highest proportion of DEGs split between the number of upregulated and downregulated DEGs (50% and 30% were upregulated, respectively). For bins involved in phytohormone biosynthesis and signaling, genes associated with JA and ET bins were upregulated, whereas the majority of those in GA, auxin, cytokinin (CK) and abscisic acid (ABA) groups were downregulated.

A large proportion of WRKY and TIFY TF groups were upregulated compared to the size of the group (32% and 52%, respectively), while the DEGs in the bHLH and MYB TF groups were split between upregulated and downregulated genes (61% and 76% were upregulated, respectively).). Groups of genes with either established or putative roles in antiherbivore defenses include genes involved in indole, terpenoid, flavonoid, and coumaroyl-CoA biosynthetic pathway, as well as trypsin inhibitors, glucosidases, peroxidases and proteases. Among the genes involved in defenses, the groups associated with indole and coumaroyl-CoA biosynthetic pathways had among the highest proportion of ZmPep3 DEGs compared to the size of the group (62% and 30%, respectively). The TIFY TF group and the coumaroyl-CoA biosynthetic pathway group had among the highest proportions of genes that were also upregulated by Gln-18:3 in comparison to ZmPep3 DEGs (61% and 62%, respectively).

Assessment of genetic variation in sensitivity to ZmPep3 and Gln-18:3

The large-scale overlap in the early transcriptional regulation following ZmPep3 and Gln-18:3 treatments supports the hypothesis that maize HAMPs and DAMPs activate antiherbivore defenses through highly similar signaling pathways. Towards the goal of identifying genes that uncouple ZmPep3 and Gln-18:3 responses, maize inbred lines were screened for differential responses to the two elicitors. Previous work on herbivore-induced VOCs revealed significant variation among genetically diverse maize lines (Degen *et al.*, 2004). We hypothesized that the observed response phenotype variation associated with herbivore-induced VOCs could be due to genetic impairments in either HAMP responsiveness, DAMP responsiveness or shared downstream signaling

components. To identify any genetic variability in Pep and FAC responses, a total of 27 maize inbred lines were screened for ZmPep3- and Gln-18:3-induced volatile sesquiterpene emission. This set included B73, Mo17 and W22 inbred lines, and parent lines of the Nested Association Mapping (NAM) population (not including CML52 due to poor germination), which collectively represent over 90% of maize genetic diversity (McMullen *et al.*, 2009). B73 was typical of examined inbred lines, emitting significantly more sesquiterpene volatiles after treatment with either ZmPep3 and Gln-18:3 as compared to water-treated leaves (Fig. 3A). Nineteen of the 27 lines responded similarly with increased sesquiterpene emissions to both treatments. Four inbred lines, HP301, CML333, MS71 and Ky21, did not emit significantly more sesquiterpenes in response to either ZmPep3 or Gln-18:3, suggesting the existence of potential mutations in shared signaling pathways (Fig. 3B). Importantly, four inbred lines, CML103, NC350, CML69 and Mo17, emitted significant sesquiterpene volatiles after ZmPep3 treatment but produced no response after Gln-18:3 treatment (Fig. 3C). The final response phenotype (Fig. 3C) is consistent with selective FAC insensitivity and the existence of a genetic basis that enable Pep and FAC responses to be uncoupled.

Differential responses in B73 and Mo17 enable forward genetics and the identification of an FAC sensitivity-associated locus

As an inbred line specifically insensitive to Gln-18:3, Mo17 was selected for further study due to the availability of established genetic resources with recombinant inbred lines (RILs), Near Isogenic Lines (NILs) and high-density genotypic markers (Eichten *et al.*, 2011; Lee *et al.*, 2002; Romay *et al.*, 2013). To confirm that Mo17 is specifically Gln-

18:3 insensitive, total VOCs were measured after treatment with water, ZmPep3 and Gln-18:3. As observed in the prior experiment, VOC emission was significantly increased after treatment with ZmPep3 in both B73 and Mo17, but only B73 emitted significantly more VOC after Gln-18:3 treatment (Fig. 4A). Because elicitor-induced VOC emission is a relatively late response, measured 16 hours after treatment, observed differences in Mo17 responses were occurring long after application of the initial signal. To assess whether Mo17 was insensitive to Gln-18:3 during earlier signaling events both inbred lines were treated for two hours with water, ZmPep3 and Gln-18:3 and analyzed for ET emission. Consistent with VOC emission, B73 emitted significantly more ET following both ZmPep3 and Gln-18:3 treatments, while Mo17 emitted significantly more ethylene following ZmPep3 treatment, but not following Gln-18:3 (Fig. 1B). Using both early and late markers for signal activation, our results support the hypothesis that Pep responsiveness in Mo17 is decoupled from FAC responsiveness and the existence of genetic variation in FAC-specific signaling components.

Based on the differential FAC sensitivity between B73 and Mo17, the Intermated B73-Mo17 (IBM) Recombinant Inbred Line (RIL) population (Lee *et al.*, 2002) was used for genetic mapping of FAC sensitivity (Lee *et al.*, 2002). Using 242 IBM-RILs, ZmPep3- and Gln-18:3-induced leaf VOCs were analyzed and compared to water-treated control plants. Association mapping was conducted based on the fold-change values of ZmPep3- and Gln-18:3-induced VOC emission as compared to water-treated leaves (Table S7). The single-nucleotide polymorphism (SNP)-based genetic marker map (B73 RefGen_V2) of the IBM-RILs (www.panzea.org, July 2012 All Zea GBS final build) was used for

association mapping. The general linear model (GLM) procedure in TASSEL 5 was used to calculate the statistical significance of SNP associations using the fold-change values of ZmPep3- and Gln-18:3-induced VOCs as traits (Romay *et al.*, 2013; Bradbury *et al.*, 2007). Association mapping revealed a single locus on chromosome 4 as significantly associated with total VOC emission specifically elicited by Gln-18:3 and not by ZmPep3 (Fig. 4C). The FAC sensitivity locus coordinates are based on adjusted *P*-values ($P < 0.05$; Boferroni-correction for multiple testing) and was defined as the region between B73 RefGen_V2 SNPs S4_237390439 and S4_238691014 (1.9Mbp), which correspond to the B73 RefGen_V4 region between *Zm00001d053820* and *Zm00001d053932* containing 77 genes (Fig. 4C, Table S8). Box-plot visualization of the ZmPep3 and Gln-18:3 VOC fold-change data split according to the allele identity at the most highly associated SNP (S4_237322925) confirmed that the FAC sensitivity locus was obtained due to Gln-18:3 sensitivity in the presence of the B73 allele and Gln-18:3 insensitivity in the presence of the Mo17 allele (Fig. 4D). To ensure that the Gln-18:3 specific association mapping result (Fig. 4C) was not driven by variation in a single VOC biosynthetic pathway, we investigated (*Z*)-3-hexenyl acetate, (*E*)-4,8-dimethyl-1,3,7-nonatriene (DMNT), and *E*-□-farnesene as separate mapping traits and obtained identical results consistent with a Mo17 lesion impacting early signal propagation (Fig. S2).

Narrowing the chromosome 4 locus associated with FAC sensitivity using Near Isogenic Lines

A second independent mapping population derived from the B73 and Mo17 inbred lines was used to provide additional support for the mapped FAC sensitivity locus. The

B73 x Mo17 NILs contain small introgression regions from of one of the parent genomes into the other nearly uniform genetic background (Eichten *et al.*, 2011). Two lines, b050 and b154, were identified as containing an introgression of the Mo17 genome into the B73 background within the FAC sensitivity locus, and were screened for ZmPep3- and Gln-18:3-elicited VOC production. Both lines emitted significantly more VOCs after treatment with ZmPep3, but only one line, b154, emitted significantly more VOCs after treatment with Gln-18:3 (Fig. 5A). These results provided additional support for the IBM-RIL derived FAC sensitivity locus (Fig. 4C) and narrowed the consideration of candidate genes by reducing the locus to 1.5 Mbp containing 54 genes based on B73 RefGen_V4 annotations between 242.2 Mbp and 243.7 Mpb (Fig. 5B).

Fine-mapping the FAC sensitivity locus using a newly generated IBM-RIL marker map

IBM-RILs provide high resolution resources to investigate the genetic basis of traits associated with development and disease, in part due to five generations of intermating before selfing, creating a high resolution RIL mapping population (Liu *et al.*, 2020; Lee *et al.*, 2002). Given considerable investments in the IBM-RIL population, higher-resolution genetic marker map are possible using genotype-by-sequencing (GBS) approaches to generate additional SNPs from DNA sequence data (Romay *et al.*, 2013). To further expand this direction, we obtained and analyzed the raw RNA-seq data (PRJNA179160) of a subset of 105 of the 302 IBM-RILs from a previous study on the regulation of gene expression in two-week old seedlings (Li *et al.*, 2013). The raw data was aligned to the B73 RefGen_V4 genome and used for variant calling. The resulting SNPs were filtered

followed by aggregation of all SNPs to a B73 and Mo17 gene-based marker map containing a total of 10,043 marker genes (Fig. 6A, Table S9). Association mapping was performed using the Gln-18:3-induced VOC fold-change data, similar to previous approaches (Fig. 4C) using a subset of 86 IBM-RIL lines (Table S7). After using the Bonferroni-correction for multiple testing and selecting marker genes with a significant threshold of $P < 0.05$, the FAC-sensitivity locus was fine-mapped to a greatly narrowed locus between 242.2 Mbp and 242.8 Mbp (B73 RefGen_V4) containing only 19 genes (Fig. 6B, Table S8). To assess the accuracy of the gene marker map at the FAC sensitivity locus, the allele counts of all SNPs in the region were aggregated and grouped based on the called genotype. These results indicated that higher allele counts matched the predicted SNP genotype, while counts of the alternative alleles were generally close to zero (Fig. 6C). The SNPs that were called as heterozygous were similarly distributed in both the B73 and Mo17 allele counts (Fig. 6C). To verify the fine-mapping results, 3 IBM-RILs for which a recombination event was detected within the fine-mapped region (B73 RefGen_V4: 242.2 Mbp and 242.8 Mbp) were tested for elicited VOC emission following treatment with water, ZmPep3 and Gln-18:3. In agreement with the revised IBM-RIL-based fine-mapping results, while all lines had significant ZmPep3-induced VOC, significant Gln-18:3-induced VOC were observed only when the genotype at the fine-mapped region was B73 (Fig. 6E). Among the 19 predicted genes within narrowed FAC sensitivity locus (Table S8), multiple genes exist that theoretically have the potential to impact signaling and transcriptionally mediated responses. These include genes predicted to encode proteins with DNA- or RNA-binding activity (*Zm00001d053855*, *Zm00001d053860*), ATP- or GTP-binding activity (*Zm00001d053857*, *Zm00001d053858*,

Zm00001d053861), RNA polymerase-related activity (*Zm00001d053872*, *Zm00001d053874*) and protein kinase superfamily proteins (Zm00001d053853, Zm00001d053876). Curiously, at the center the 19-gene locus exist 2 genes annotated as leucine-rich repeat receptor-like kinases (LRR-RLKs), namely *Zm00001d053866* and *Zm00001d053867*. The predicted *Zm00001d053866* sequence displays considerable physical overlap with the intact LRR-RK encoding gene *Zm00001d053867* suggesting a misannotation of *Zm00001d053866*. Given the physical genetic position and common role in early signal transduction events mediated by LRR-RLK family members (Zipfel Rev, other), we selected *Zm00001d053867,* termed *FAC SENSITIVITY-ASSOCIATED* (*ZmFACS*), as the top candidate for further assessment in mediating FAC sensitivity (Fig. 6F, Table S8).

ZmFACS is an ortholog of a rice receptor, OsLRR-RLK1, which positively regulates antiherbivore responses

Of the gene candidates present in the FAC sensitivity locus, *ZmFACS* exists as an orthologue of the rice gene *OsLRR-RLK1*, which was selected from a transcriptional profiling study and characterized as a positive regulator of rice antiherbivore responses to diverse lepidoptera attack and crude FAW OS (Fig. 7A) (Hu *et al.*, 2018; Zhou *et al.*, 2011). ZmFACS and OsLRR-RLK1 belong to the LRR-Xb subfamily of LRR-RLKs which also includes the phytosulfokine (PSKR1/2) and PSY1 receptors (PSY1R), with PSY1R being the more closely-related Arabidopsis LRR-RLK to ZmFACS and OsLRR-RLK1 (Fig. 7A) (Shiu *et al.*, 2004; Hu *et al.*, 2018). *OsLRR-RLK1* transcript accumulation significantly increases in rice seedlings in response to FAW OS (Hu *et al.*, 2018; Zhou *et al.*, 2011).

To understand whether ZmFACS is differentially expressed in maize upon FAC treatment, leaves of B73 and Mo17 plants were treated with water or Gln-18:3 and harvested after 0.5 and 1 hour. In B73, but not Mo17, *ZmFACS* transcripts displayed a significant two-fold increase in accumulation 1 hour after Gln-18:3 treatment (Fig. 7B). Moreover, basal *ZmFACS* expression levels are approximately 200-fold higher in B73 compared to Mo17 over analyzed treatments and time points (Fig. 7B). Given the difference in ZmFACS expression between the two inbred lines, genome sequence upstream and downstream of ZmFACS were compared. Comparison of the *ZmFACS* promoter sequences between B73 and Mo17 revealed large non-aligned regions, insertions of transposable element fragments of the DNA9 and Harbinger families as well as an expansion of hAT transposable element fragments (Fig. 4C)(Kohany *et al.*, 2006). In order to predict the existence of amino-acid (AA) sequence differences disrupting function, the aligned AA sequences of B73 ZmFACS and Mo17 ZmFACS were annotated through the identification of the signal peptides, LRRs, island domain, transmembrane domain, ATP binding site and the serine/threonine-protein kinase active site (Fig. S4)(Chen, 2021; Sievers *et al.*, 2011; Casas *et al.*, 2016; Kall *et al.*, 2007; Zhang *et al.*, 2020; Wang *et al.*, 2015; Mitchell *et al.*, 2019). Of the 13 total AA differences identified, none resulted in changes predicted to have large-scale impacts on protein function (Fig. S3). From these analyses, promoter sequence variation in Mo17 *ZmFACS* and dramatically reduced transcript abundance (Fig. 7) is more apparent than any potentially causal AA variation in the predicted ZmFACS proteins (Fig. S3).

Heterologous expression of ZmFACS proteins enhance response sensitivity to Gln-18:3 in tobacco (*Nicotiana benthamiana*)

To examine the role of the FACS candidate gene in promoting FAC sensitivity, *Nicotiana benthamiana* leaves were used as a heterologous protein expression system. As shown previously, multiple Solanaceous species, including *N. benthamiana* are naturally sensitive to FACs, and activate of antiherbivore defense responses upon FAC treatment (Grissett *et al.*, 2020; Xu *et al.*, 2015). Using ethylene emission to assess FAC responsiveness, *N. Benthamiana* leaves were infiltrated with carboxyl terminal yellow fluorescent protein (YFP)-tagged fusions of ZmFACS and *Arabidopsis thaliana* ELONGATION FACTOR-TU RECEPTOR (EFR) as a control (Zipfel *et al.*, 2006). Two days after infiltration the plants were treated with water, 1 µM elf18 or 1 µM Gln-18:3, and ethylene emission was measured after 2 h. Treatment of *N. Benthamiana* leaves with elf18, triggered a significant increase in ET production following the expression of EFR:YFP but not ZmFACS:YFP (Fig. S4A). In contrast, treatment with Gln-18:3 triggered significant increases in ET emission in both EFR:YFP and FACS:YFP expressing leaves (Fig. S4A). A positive Gln-18:3 result was anticipated given well established FAC elicited responses in *N. benthamiana* (Grissett et al., 2020). In an effort to leverage the experimental advantages of *N. benthamiana* as a rapid transient heterologous expression system, we considered different plant growth conditions. In contrast, Gln-18:3 treatments selectively triggered significant ET emission in the ZmFACS:YFP expressing plants (Fig. S4B). Putative native yet undescribed *N. benthamiana* protein(s) mediating FAC sensitivity predictably present under normal light conditions appear to have reduced function in control EFR:YFP expressing plants grown under diminished light. Protein

levels of EFR:YFP and ZmFACS:YFP were probed and visualized by western blotting demonstrating successful heterologous expression of both (Fig. S4C). Our results are consistent with ZmFACS:YFP heterologous expression increasing *N. benthamiana* responsiveness to FACs (Fig. S4B). In an effort to understand the causal basis of the Gln:18:3-elicited VOC association mapping result, *N. benthamiana* plants grown under diminished light intensity were used to compare Gln-18:3 responsiveness to both B73 ZmFACS and Mo17 ZmFACS. As a control, elf18 treatments triggered a significant increase in ET emission only in EFR:YFP expressing leaves (Fig. 8A). Only treatment with Gln-18:3 elicited significant ET emission in B73 ZmFACS:YFP and Mo17 ZmFACS:YFP expressing leaves (Fig. 8A), consistent with the hypothesis that ZmFACS from both B73 and Mo17 are functionally capable of increasing FAC response sensitivity in *N. benthamiana* (Fig. 8A). Protein levels of EFR:YFP, B73 ZmFACS:YFP and Mo17 ZmFACS:YFP were probed, visualized by western blotting, and displayed similar levels of heterologous expression (Fig. 8B).

DISCUSSION

Plants respond to many cues during herbivory including wound induced jasmonates, HAMPs and complex endogenous amplification signals which collectively activate regulatory mechanisms that promote protective antiherbivore defenses (Erb and Reymond, 2019). For nearly 30 years maize has been a model system for the study of induced plant responses elicited by insect herbivory, insect OS and defined HAMPs present in insect OS (Turlings *et al.*, 1990; Alborn *et al.*, 1997; Pare *et al.*, 1998; Turlings

et al., 1993). Subsequent efforts have expanded to an array of plant-insect models; however, while diverse FAC family HAMPs are known to commonly active plant defenses, many core components involving FAC mediated signal transduction remain to be identified (Halitschke *et al.*, 2001; Schmelz *et al.*, 2009; Wu and Baldwin, 2009; Erb and Reymond, 2019; Bonaventure *et al.*, 2011) . Previous research has revealed that like FACs, ZmPeps potently activate classical direct and indirect antiherbivore defenses in maize (Alborn *et al.*, 1997; Huffaker *et al.*, 2013; Poretsky *et al.*, 2020). The defensive biochemical outputs triggered by ZmPeps and FACs are increasingly well characterized; however, clear gene candidates for proximal mechanisms by which FAC outputs are activated remain unclear (Truitt *et al.*, 2004; Gilardoni *et al.*, 2011; Schmelz, 2015). Transcriptome analyses following biotic stress and treatments with defined elicitors provide comprehensive insights into the genetic regulation of plant physiological processes (Poretsky *et al.*, 2020; Zhou *et al.*, 2011; Gilardoni *et al.*, 2010; Tzin *et al.*, 2017; Tzin *et al.*, 2015; Heidel and Baldwin, 2004). In rice, transcriptional profiling following infestation with the caterpillar *Chilo suppressalis* identified WRKY and LRR-RLK genes involved in herbivore resistance (Tzin *et al.*, 2017; Hu *et al.*, 2018; Hu *et al.*, 2015). More recently, characterization of transcriptional reprogramming following treatment with diverse defense elicitors was used to identify a shared core of immune response genes in Arabidopsis and highlighted conserved roles for glutamate receptor-like calcium-permeable channels in general stress responses (Bjornson *et al.*, 2021). To define candidate regulatory genes and explore the degree of overlap between HAMP and DAMP signaling pathways, we employed both comparative transcriptional profiling and screens for genetic variation in maize responses to ZmPep3 and Gln-18:3.

In our present comparative analyses of early-elicited transcriptional changes in maize, we define a core set of 312 DEG genes common to defined HAMP and DAMP triggered immune processes (Fig. 1 and 2). Gene expression changes within two hours revealed significant quantitative differences yet a high degree of qualitative similarity between Gln-18:3 and ZmPep3 responses. While ZmPep3 treatment resulted in > 4-fold more DEGs than Gln-18:3, over 92% of positively regulated transcripts elicited by Gln-18:3 were shared by the ZmPep3 responses (Fig. S1). Moreover, 95% of the genes that were differentially regulated by ZmPep3 displayed greater average levels in Gln-18:3-treated samples compared to the water-treated controls (Table S6). Collectively our results support a high degree of overlap between ZmPep3 and Gln-18:3 mediated transcriptional responses in maize. From the shared DEGs, diverse regulators of ZmPep3- and Gln-18:3-induced maize responses have been identified that predominate in categories associated with signal transduction and transcriptional reprogramming (Fig. 2). A core set of 44 transcription factor genes rapidly responded to ZmPep3 and Gln-18:3, including transcription factors from the enrichment WRKY, TIFY, bHLH, MYB, ERF and DREB families (Table S6, Fig. 1C). While core transcription factors in the bHLH and WRKY families have demonstrated roles during herbivory (Schweizer *et al.*, 2013; Hu *et al.*, 2015; Li *et al.*, 2015), clearly a large array of transcription factors mediate extensive transcriptional changes underlying increases in protective defense metabolites and proteins following herbivory (Tzin *et al.*, 2017; Poretsky *et al.*, 2020; Erb and Reymond, 2019). Within transcription factor families, among the most DAMP associated DEG enrichments existed in the TIFY family (Fig. 2) which have established roles in the

regulation of jasmonate signaling (Chung and Howe, 2009). A classical feature of Gln-18:3 and ZmPep3 elicitation in maize is the rapid accumulation of jasmonates that co-incidently are readily observed after 2 hours of treatment (Schmelz *et al.*, 2003; Huffaker *et al.*, 2013). Elevated *TIFY* expression is consistent with jasmonate accumulation as a downstream signal mediating ZmPep3- and Gln-18:3-induced transcriptional changes (Erb and Reymond, 2019). Additionally, the expression of genes encoding numerous RLKs/RLPs/Lectin receptors, MAPK/MAP2K/MAP3K and calcium-dependent protein kinases (CDPK) were elevated, indicating specific networks likely to participate in phosphorylation cascades controlling response outputs (Fig. 2; Table S6). Transcripts for several maize orthologues of Arabidopsis genes encoding core components of wound and reactive oxygen species (ROS)-mediated signaling pathways also accumulated, for nicotinamide adenine dinucleotide phosphate (NADPH) oxidase family members termed Respiratory Burst Oxidase Homologs (RBOH)- and glutamate receptor like-encoding (GLR) genes (Mousavi *et al.*, 2013; Torres *et al.*, 2002). Many of these transcriptionally-regulated candidate genes are similar to genes mediating antiherbivore defense signaling in other species (Howe and Jander, 2008; Arimura *et al.*, 2005; Maffei *et al.*, 2007a; Maffei *et al.*, 2007b). Our transcriptomic study pinpoints discrete members within maize gene families for further characterization as predicted regulators of antiherbivore defense responses. These results also suggest that while ZmPep3 is a more potent elicitor in maize, Gln-18:3 and ZmPep3 largely share similar signaling components for antiherbivore response activation. Given the observed HAMP and DAMP overlap, the 1692 significant DEGs observed following ZmPep3 treatment represents a reasonable comprehensive view of early defense activation. However, not all genes of interest display large scale

transcriptional changes which further expands the challenge of considering 100's of candidate genes, which could potentially represent a new node in either HAMP or DAMP signaling (Dressano *et al.*, 2020).

The earliest insights into herbivore, OS and HAMP specific maize responses involved the analyses of elicited VOC production as indirect defenses (Turlings *et al.*, 1990; Turlings *et al.*, 1993; Alborn *et al.*, 1997). When measured over time, foliar sesquiterpenes dominate the late term profiles of herbivore and HAMP-elicited maize VOCs (Schmelz *et al.*, 2003; Turlings *et al.*, 1998). Using both ZmPep3- and Gln-18:3-induced production of volatile sesquiterpenes, we screened for genetic variation in diverse maize inbred lines and found that a majority of inbreds produced significant responses to both elicitors (Fig. 3). Four inbred lines, namely Ky21, HP301, Ms71 and CML333 did not emit statistically significant increases in sesquiterpenes to either signal. Our results were generally consistent with a previous study observing low terpene emission in these same lines upon elicitation with the synthetic 6-substituted indanoyl isoleucine conjugate analog of JA-isoleucine (Richter *et al.*, 2016). The low VOC response to three different elicitors indicates is consistent with either impaired defense signaling common to all three, or that elicited protective responses in these lines relies predominantly on nonvolatile defenses. We did not observe any inbred line to be responsive to Gln-18:3 yet specifically nonresponsive to ZmPep3. Collectively this supports the possible existence of defined maize inbreds, such as Ky21, with compromised signaling in shared pathways downstream of HAMPs, DAMPs and jasmonates (Fig. 3B) (Richter *et al.*, 2016). This possibility is intriguing as many core

signaling pathways are governed by gene family duplications and redundancies that create resiliency to single null mutations. For example, analyses of ZmPep receptor mutants, *Zmpepr1* and *Zmpepr2*, demonstrated that ZmPeps signal through both receptors (Poretsky *et al.*, 2020). Genetic redundancy displayed for both *ZmPROPEP* genes and *ZmPEPR* receptors is in concordance with the known role of Peps/PEPRs across diverse plant species as core amplifiers of signaling elicited by multiple inputs (Huffaker *et al.*, 2006; Yamaguchi *et al.*, 2010; Liu *et al.*, 2013; Tintor *et al.*, 2013; Ross *et al.*, 2014; Shinya *et al.*, 2016; Poretsky *et al.*, 2020). In contrast, several maize inbred lines were found to be specifically insensitive to Gln-18:3 as assigned by the elicited sesquiterpene volatile production assays (Fig. 3C). A reanalysis of Mo17 further demonstrated a selective deficiency to elicited ET production following Gln-18:3 treatment yet robust production after ZmPep3 treatment (Fig. 4B). Given that HAMP-elicited ET emission commonly occurs rapidly after elicitation and precedes significant VOC emission (Schmelz *et al.*, 2007; von Dahl *et al.*, 2007), we classified Mo17 as compromised in an early signaling node specifically acting downstream of Gln-18:3. We then choose a forward genetics approach to narrow the selection of candidate genes influencing HAMP signaling in maize.

To optimize data interpretation, we used a dual-input association mapping approach to ensure that the specific insensitivity of Mo17 to Gln-18:3 could be assessed in each IBM-RIL line compared to positive ZmPep3 responses. Using the fold-change in total elicited volatiles as a mapping trait, a single locus on chromosome 4 was identified to be significantly associated with sensitivity to Gln-18:3, but not to ZmPep3 (Fig. 4).

Further association mapping using individual volatile traits derived from distinct biosynthetic pathways demonstrated that Gln-18:3-induced production mapped to the same locus in each case and supported the hypothesis the underlying genetic difference was likely to effect signaling pathways rather than core VOC biosynthetic genes (Fig. S2). Following multiple NIL- and RIL-based fine mapping approaches which were uniquely enabled by comprehensive B73 and Mo17 community resources, the FAC sensitivity locus was successfully narrowed to a region containing 19 genes (Fig. 5 and 6). At the physical center of locus we identified the LRR-RLK signaling candidate gene ZmFACS (Fig. 6F). Surprisingly, ZmFACS exists as an ortholog of the gene *OsLRR-RLK1*, which was recently demonstrated to contribute to antiherbivore defenses in rice (Hu *et al.*, 2018). Transcript analyses demonstrated that expression of *OsLRR-RLK1* was induced in rice during sustained striped stem-borer (SSB, *Chilo suppresalis*) herbivory as well as following fall armyworm (FAW, *Spodoptera frugiperda*) OS application. Gene silencing of *OsLRR-RLK1* resulted in rice plants displaying diminished JA and ET production and decreased MAPK3/6 activation following SSB infestation (Hu *et al.*, 2018). SSB-induced increases in expression of several rice WRKY transcription factor genes were delayed in *OsLRR-RLK1*-silenced plants and reduced production of trypsin inhibitors was associated with decreased SSB resistance. In contrast, upon mechanical wounding alone *OsLRR-RLK1* silenced plants demonstrated no difference from wild type controls in JA/ET production, MAPK activation or defense gene expression (Hu *et al.*, 2018). A current hypothesis is that *OsLRR-RLK1* is specifically involved in regulating these outputs in response to an unknown family of HAMPs absent during simple mechanical damage. In the current maize study, ZmFACS emerged as an association mapping-derived candidate

gene linked to the specific insensitivity to a biochemically defined HAMP, namely Gln-18:3 (Pare *et al.*, 1998; Truitt *et al.*, 2004). Given the diverse approaches employed and convergent results obtained, it is likely that the ZmFACS and *OsLRR-RLK1* orthologs have similar roles in transduction and propagation of HAMP specific signals.

Precisely how ZmFACS mediates maize response sensitivity to Gln-18:3 remains to be demonstrated. Heterologous expression of ZmFACS in *N. benthamiana* can enhance elicited ET emission after Gln-18:3 application, supporting a role in promoting FAC responses (Fig. 8). In dicot models, the Arabidopsis homologs of ZmFACS and OsLRR-RLK1, AtPSKR1 and AtPSY1R, have been implicated in regulating pathogen resistance and wound responsiveness through upregulation of JA signaling in addition to their canonical role as the PSK and PSY1 receptors, respectively (Igarashi *et al.*, 2012; Mosher *et al.*, 2013; Mosher and Kemmerling, 2013; Shen and Diener, 2013). Receptors may mediate signaling through direct ligand binding or through activating or repressing pathway functions (Han *et al.*, 2014; Hohmann *et al.*, 2017; Smakowska-Luzan *et al.*, 2018). In a legume model for plant-herbivore interactions, cowpea (*Vigna unguiculata*) was recently leveraged in forward-genetic mapping approaches using the inceptin family HAMP peptides to uncover the LRR-RLP inceptin receptor (INR) (Steinbrenner *et al.*, 2020). Physical interactions between inceptin, INR and SERK co-receptor associations were supported by labeled-ligand binding assays. Importantly stable heterologous expression of INR in tobacco plants conferred both inceptin-induced responses and enhanced *Spodoptera* resistance (Steinbrenner *et al.*, 2020). Diverse receptors are increasingly implicated in regulating antiherbivore responses in diverse species through

a variety of mechanisms. Reverse-genetic approaches were used to partly characterize receptors from tobacco (*Nicotiana attenuata*) and rice (*O. sativa*) that regulate antiherbivore responses. The *N. attenuata* co-receptor NaSERK3/BAK1 was examined due to established roles for Arabidopsis orthologues in flagellin signaling, and when silenced, resulted in reduced wound- and OS-induced JA accumulation, but not reduced MPK activity or herbivore resistance, suggesting that NaBAK1 acts downstream of OS perception (Yang *et al.*, 2011). Because NaSERK1 was not silenced, it is possible that the wound- and OS-induced MPK activation were not reduced due to the presence of redundant SERK co-receptors, and that full activation of induced antiherbivore responses requires both NaSERK1 and NaSERK3 (Yang *et al.*, 2011; Ma *et al.*, 2016). HDS-ASSOCIATED RLK1 (HAK1) and HAK2 were identified through a sequence-mining approach to identify receptors responsive to fractionated *Spodoptera litura* oral secretions enriched in polysaccharides based on sequence homology to CERK receptors involved in recognition of chitin polysaccharides, (Uemura *et al.*, 2020). Arabidopsis *hak1* insertional knockout lines demonstrated reduced responses to an OS fraction containing polysaccharides, with the receptor-like cytoplasmic kinase (RLP) PBL27 identified as a HAK1-interacting coregulator. A *N. attenuata* receptor gene, *LECTIN RECEPTOR KINASE1* (NaLRK1) was among the most highly transcriptionally upregulated genes following FAC treatment in *N. attenuata* (Gilardoni *et al.*, 2010; Zhou *et al.*, 2011). Silencing of *NaLRK1* resulted in reduced accumulation of SA, but not JA, after OS treatment, suggesting a role in regulating a subset of responses downstream of OS perception (Gilardoni *et al.*, 2011).

Similar to most receptors previously identified as candidate regulators of antiherbivore defense responses, the precise role and molecular mechanism by which ZmFACS mediates sensitivity to Gln-18:3 remains to be determined. While heterologous expression of ZmFACS in *N. benthamiana* enhances sensitivity to Gln-18:3, additional experiments are required to determine if ZmFACS is necessary or sufficient for FAC sensitivity in maize. ZmFACS knockdown or defined CRISPR/Cas9 Zmfacs mutants must be generated in B73 to fully prove that ZmFACS is necessary for FAC sensitivity. Stable heterologous expression of FACS in non-responding plant species could be accomplished to determine if the gene is sufficient to confer sensitivity to Gln-18:3. While FACs have not yet been identified as a component of *Chilo suppressalis* oral secretions, FAC occur in the OS of at least 19 examined lepidoptera species (Yoshinaga *et al.*, 2010). Thus comparative testing of FACs sensitivity in *OsLRR-RLK1* silenced rice lines could be informative. More broadly, the evaluation of defense responses after FAC treatment in comparison to a range of elicitors, such as Peps or chitin, would inform whether *OsLRR-RLK1* functions in FAC signaling. In addition to these studies, experiments assessing whether ZmFACS directly binds Gln-18:3 as a ligand are of interest. Evidence for the existence of an FAC-binding protein was demonstrated using enriched plasma membrane preparations from hybrid maize (var. Delprim) and a radio-labeled [^{3}H]-FAC (Truitt *et al.*, 2004). A critical assessment of whether ZmFACS or OsLRR-RLK1 functions through direct interaction with FACs is predicted to prove challenging. Modern standards of proof are significant and recent critical analyses of the entire field states that not a single "*bona fide* HAMP/PRR pair has yet to be discovered" (Reymond, 2021). As the first challenge only modest radioisotope changes, such as [^{3}H]-FAC, have been reported to retain FAC

binding and biological activity. This is broadly consistent with lipid derived ligands that are comparatively intolerant to significant modification (J. Merkler and W. Leahy, 2018). The lack of free amide groups in Gln-18:3 make acridinium conjugate labeling similarly unfeasible (Steinbrenner et al., 2020). Biophysical ligand-binding assays could overcome this challenge (Sandoval and Santiago, 2020; Sharma and Russinova, 2018); however, within seconds of contacting the wounded leaf surface FACs undergo rapid lipoxygenase mediated modifications resulting active and inactive oxygenated derivatives (VanDoorn *et al.*, 2010). Plant mediated FAC modification creates an open question regarding the natural receptor ligands and could complicate the use of *in vitro* binding assays.

Our transcriptomic analyses of early maize responses to ZmPep3 and Gln-18:3 identified a targeted set of 312 shared transcriptionally co-regulated genes rich in signaling candidates. In an independent forward genetics approach we uncovered the candidate gene ZmFACS associated with maize FAC sensitivity. Previous studies have predominantly relied upon combinations of transcriptional profiling, candidate gene mutation and silencing approaches to highlight genes influencing herbivory signaling (Uemura *et al.*, 2020; Hu *et al.*, 2018, p.1; Hu *et al.*, 2015; Gilardoni *et al.*, 2011; Yang *et al.*, 2011, p.1). Our current study is the first to rely upon a comparatively unbiased forward genetics approach to identify a narrow locus and gene candidate mediating FAC-elicited response sensitivity in any plant species. Elicitor-induced antiherbivore signaling and defense activation have long been examined given the promise to naturally limit arthropod inflicted crop damage, but applications been constrained by our limited mechanistic understanding (Karban and Baldwin, 1997; Kessler and Baldwin, 2002). As a rather

unique model family of HAMPs, FACs have been closely examined by diverse research groups for a quarter of century and are established to both widely occur in lepidoptera herbivores and broadly elicit defenses responses in diverse monocot and dicot plants (Yoshinaga *et al.*, 2010; Halitschke *et al.*, 2001; Schmelz *et al.*, 2009; Grissett *et al.*, 2020; Shinya *et al.*, 2016). HAMP-induced plants were first discovered in the genus *Spodoptera* and are broadly present in lepidoptera that activate maize defenses (Ling *et al.*, 2021; Turlings *et al.*, 1993; Alborn *et al.*, 1997). An improved molecular understanding of how FAC responses are activated to directly or indirectly suppress *Spodoptera* pests has the continued potential to guide development of improved crop resistance to herbivores.

MATERIALS AND METHODS

Plant Growth Conditions

Zea mays plants were grown in BM2 soil (Berger Mixes) inside a greenhouse (12-h light, minimum of 300 µmol m-2s-1, and 12-h dark) at 24°C/28°C (night/day) temperature cycle. The plants were supplemented with a 18-18-21 Tomato Plant Food fertilizer (Miracle-Gro). *Nicotiana benthamiana* plants were grown inside a growth room under two different conditions, normal light intensity (150 µmol $m^{-2}s^{-1}$) and reduced light intensity (80 µmol $m^{-2}s^{-1}$), as specified. For plants grown under reduced light intensity condition the light intensity for was further reduced (10 µmol $m^{-2}s^{-1}$) by disconnecting the overhead lights on the shelf in the growth room immediately following agrobacterium infiltration and until elicitor treatment two days later. In both the normal light intensity and the reduced light intensity conditions, Gro Lite WS lamp (Interlectric Crop.) were included

as a supplementary light source and the plants experienced a 16-h light and 8-h dark cycle at 22°C. The plants were planted in BM2 soil (Berger Mixes) and supplemented with a 20-20-20 General Purpose fertilizer (Jack's Professional).

Elicitor treatment of plant leaves

The peptides elicitors ZmPep3 and elf18 were synthesized and purified as 23mers by Sigma-Aldrich. The FAC Gln-18:3 was obtained from Dr. Eric Schmelz . All elicitors were diluted in water to the concentrations indicated. For measurement of induced volatiles leaves were excised and put in a vial containing 1mL of the treatment solution for 16 hours or before volatile collection. For RNA-seq analysis whole leaves were excised, put in a vial containing 1mL of the treatment solution for 2 hours, after which 2 inches from the base of the leaves were cut and flash frozen in liquid nitrogen. For the analysis of transcript abundance qRT-PCR leaf tissue was treated by scratch application of 20uL treatment solution for the specified time before being flash frozen in liquid nitrogen. For measurement of induced ethylene in maize leaves were excised and put in a vial containing 1mL of the treatment solution for 1 hour before ethylene collection. Measurement of induced ethylene emission in *N. benthamiana* leaves was conducted using plants grown under two distinct growth conditions, normal light intensity and diminished light intensity, but the treatment procedure and ethylene collection was identical. For the treatment, leaves were infiltrated with Agrobacterium tumefaciens carrying constructs for the indicated YFP-fusion receptor genes two days before treatment. After two days the *N. benthamiana* leaves were infiltrated with treatment

solution, immediately cut with a cork borer and sealed in airtight tubes until ethylene collection.

Measurement of plant volatiles

Collection of volatile organic compounds (VOCs) from maize leaves was conducted by enclosing the leaves in glass tubes under light for 30 minutes and collecting head-space volatiles on a 50 mg Super Q (80/100 mesh; Alltech, https://www.alltech.com/). VOCs were eluted using methylene chloride, with the addition of nonyl acetate as an internal standard and analyzed by GC. Specific compounds were identified by comparing their retention times with those of pure standards. Ethylene emitted by leaves was measured by enclosing leaves in a plastic tube and collecting head-space volatiles for 2 hours. A syringe was used to inject and analyze head-space volatiles by GC using a standard curve, as previously described (Schmelz *et al.*, 2009).

RNA-seq Preparation and Analysis

Total RNA was isolated with the NucleoSpin RNA plant kit (Clontech), treated with the Turbo DNA-free kit (Ambion) and quantified using Qubit 2.0 fluorometer (Life Technologies). Preparation of RNA-seq data was done by Novogen Corporation. Briefly, insert size was checked on an Agilent 2100 Bioanalyzer (Agilent Technologies), and sequencing of the 250–300-base-pair (bp) insert cDNA library was conducted using Illumina HiSeq platform PE150. Filtering of raw reads was done by discarding reads with adaptor contaminations, reads with more than 10% of uncertain nucleotides and reads with more than 50% low quality nucleotides (base quality < 20). TopHap2 was used to

align reads to the Maize V4 genome with default parameters and mismatch parameter set to 2. HTSeq was used to analyze gene expression levels in union mode. DESeq was used for analysis of differential gene expression followed by calculation of p-values using the negative binomial distribution and adjusted using Benjamini–Hochberg procedure for False-Discovery Rate (FDR). Genes with Fold-change > 1 or < -1 and adjusted p-value < 0.05 were considered differentially expressed. GO term enrichment analysis was conducted with the Phytozome GO term annotation for Maize B73 V4 using the hypergeometric test to identify enriched GO terms with Benjamini–Hochberg FDR adjusted p-value < 0.05. The Maize B73 V4 MAPMAN bin annotations were used for bin enrichment analysis using the hypergeometric test to identify bins with Benjamini–Hochberg FDR adjusted p-value < 0.05.

Genetic mapping of the FAC sensitivity locus

In order to understand the genetic basis for FAC sensitivity in maize the NAM parent founds, B73, Mo17 and W22 were screened for elicitor induced volatile emission. ZmPep3 was used as a positive control to identify lines with decoupled Pep- and FAC-induced responses. Mo17 was among 4 maize inbred lines found to be insensitive to Gln-18:3 but not to ZmPep3. Based on the differential sensitivity of B73 and Mo17 to Gln-18:3, but not ZmPep3 we proceeded with screening the Intermated B73-Mo17 (IBM) genetic mapping population using volatile emission after treatment with water, ZmPep3 and Gln-18:3. Fold-change data for elicitor induced VOC emission was generated for 222 IBM RILs and used in TASSEL 5 to conduct the General Linear Model (GLM) association mapping. The imputed SNP marker HapMap file for the IBM mapping population was

obtained from Panzea and filtered for minimum allele count of 10% of the lines and minor allele frequency greater than 10% and below 90%. Significantly associated SNPs were assigned following the Bonferroni p-value correction procedure using a cutoff of adjusted $p < 0.05$.

Generating a SNP marker map based on RNA-seq data from 105 IBM lines

Raw FASTQ files for B73, Mo17 and 105 IBM lines were obtained from the NCBI PRJNA179160 study accessiond (Li *et al.*, 2013). The raw FASTQ files were first filtered using FASTP with default parameters (Chen *et al.*, 2018) and then aligned to the *Zea mays* B73v4 genome obtained from plant ensemble version 44 using the STAR RNA-seq aligner with adjusted parameters (outFilterMultimapNmax 10, outFilterMismatchNoverLmax 0.04, outFilterIntronMotifs RemoveNoncanonicalUnannotated, alignIntronMax 6000) (Dobin *et al.*, 2013). The BCFTools mpileup and call functions were used for SNP calling in all annotated CDS gene regions using the filtering parameters of DP ≥ 10 and QUAL=999 (Li, 2011). After the SNPs were called SNPs that had more than 10 lines with an allele read depth lower than 10 were removed. A final filtering step was conducted by removing all SNPs that disagreed with the SNPs called using the control B73 and the Mo17 inbred RNA-seq samples. Individual SNPs were assigned gene IDs based on their location and the genotypes of genes with the associated SNPs were called based on aggregated most common SNP genotype in each gene. The genotypes of 10,043 genes were called based on the most abundant SNP genotype in each gene.

RNA Isolation and Measurement of Transcript Abundance by qRT-PCR

Total RNA was isolated with TRIzol (Invitrogen) as per the manufacturer's protocol and treated with DNAse according to instructions (Life Technologies). M-MLV Reverse Transcriptase (Thermo Fisher Scientific) was used for cDNA synthesis with random decamer primers and 1µg of RNA. For the qRT-PCR 1µl of 2-fold diluted cDNA was added using SsoAdvanced(tm) Universal SYBR(R) Green Supermix (Bio-Rad) and a StepOne Real-Time PCR System (Applied Biosystems). Based on the threshold cycle (Ct) value, ΔCt was calculated relative to 60S ribosomal protein L17 (RPL17). Transcript abundance was calculated relative to its corresponding untreated control.

Plasmid Construction for Transient Assays and Immunoblotting in *Nicotiana benthamiana*

ZmFACS (Zm00001d053867) was amplified from genomic *Z. mays* B73 and Mo17 DNA and inserted into pENTR/D-TOPO vector (Invitrogen). Expression vectors of the corresponding ZmFACS entry vectors were generated through Gateway cloning into the destination vector pGWB441, to generate a C-terminal fusion with eYFP driven by the CaMV35S promoter (Nakagawa *et al.*, 2007). Expression vectors were transformed into Agrobacterium tumefaciens GV3101 (pMP90). A. tumefaciens carrying the expression vectors were infiltrated into *N. Benthamiana* leaves at OD600 of 0.8 for transient expression. For immunoblotting, 50mg of leaf tissue were ground in liquid nitrogen and homogenized in 100µL 2xSDS loading buffer (20% SDS was used in making the 5xSDS loading buffer, instead of 10%) for 5 minutes in 95°C. Western blotting was performed with α-GFP polyclonal (ThermoFisher) primary antibodies at 1:1,000 dilution and α-rabbit

(Sigma) secondary antibodies at 1:10,000 dilution. SuperSignal™ West Pico PLUS (ThermoFisher) chemiluminescent substrate was used for protein detection.

Phylogenetic Tree Construction

The Phytozome blastn tool was used to identify coding sequences (CDS) related to FACS from *Zea mays*, *Oryza sativa* and *Arabidopsis thaliana*. The CDS sequences of the putative genes were translated to amino acid sequences and the full sequences were aligned using MUSCLE with default parameters (Edgar, 2004). The best-fit model for the phylogenetic estimation was selected using ModelFinder (Kalyaanamoorthy *et al.*, 2017). A Maximum Likelihood (ML) phylogenetic tree was constructed using IQ-TREE under the selected best-fit model, with 1,000 bootstrap replications (Nguyen *et al.*, 2015). The tree was visualized and annotated using FigTree (http://github. Com/rambaut/figtree/).

SUPPLEMENTARY DATA:

Figure S1: Comparison of overlapping ZmPep3- and Gln-18:3-induced DEGs

Figure S2: Genetic mapping sensitivity locus using the fold-change (FC) values of VOCs from three different biosynthetic pathways

Figure S3: Alignment of the encoded amino acid FACS sequences of B73 and Mo17

Figure S4: Native Gln-18:3 sensitivity in *N. benthamiana* is reduced in plants grown under diminished light intensity but heterologous expression of FACS increases responsiveness to Gln-18:3

Table S1: Gene expression quantification of maize B73 leaves treated with water, ZmPep3 and Gln-18:3 2 hours after treatment

Table S2: ZmPep3- and Gln-18:3 induced differential gene expression (DEG) in maize B73 leaves 2 hours after treatment

Table S3: Analysis of ZmPep3- and Gln-18:3-induced DEGs GO term enrichment

Table S4: Analysis of ZmPep3- and Gln-18:3-induced DEGs MAPMAN bin enrichment

Table S5: Comparison of relative expression using the mean ranked FPKM of each gene

Table S6: List of DEGs and their log2(FC) data used for figure 2

Table S7: IBM VOC fold-change data used for QTL mapping

Table S8: Summary of the FAC sensitivity locus and finemapping results

Table S9: New IBM SNP genetic marker map using RNA-seq data

Table S10: Sequences of genes related to ZmFACS from *Zea mays*, *Oryza sativa* and *Arabidopsis thaliana* used for phylogeny tree construction

Table S11: List of primers used

AUTHOR CONTRIBUTIONS

AH conceived of research plans. EP, MR, NA, EAS and AH designed and performed experiments and analyzed the data. AS cloned maize receptor genes. KD prepared the samples for RNA-seq. EP, EAS and AH wrote the manuscript with input from all authors. AH agrees to serve as the author responsible for contact and ensures communication.

ACKNOWLEDGMENTS:

The authors are grateful to Philipp Weckwerth and Yezhang Ding for advice and discussions. This research was funded by This research was funded by NSF-IOS PBI CAREER #1943591, a Hellman Foundation Fellowship and UC San Diego Start-up funds to A. H., and by USDA NIFA AFRI #2018-67013-28125 to A. H. and E. A. S. E. P. was additionally funded by Cell and Molecular Genetics (CMG) Training Program at the University of California, San Diego.

Chapter 1, in part, is currently being prepared for submission for publication of the material. Poretsky, E, Ruiz, M, Ahmadian, N, Steinbrenner, AD, Dressano, K, Schmelz, EA, Huffaker, A. The dissertation author was the primary investigator and author of this paper.

REFERENCES:

Acevedo, F.E., Smith, P., Peiffer, M., Helms, A., Tooker, J. and Felton, G.W. (2019) Phytohormones in Fall Armyworm Saliva Modulate Defense Responses in Plants. J. Chem. Ecol., 45, 598–609.

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

Alborn, H.T., Turlings, T.C.J., Jones, T.H., Stenhagen, G., Loughrin, J.H. and Tumlinson, J.H. (1997) An Elicitor of Plant Volatiles from Beet Armyworm Oral Secretion. Science, 276, 945–949.

Arimura, G., Kost, C. and Boland, W. (2005) Herbivore-induced, indirect plant defences. Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids, 1734, 91–111.

Bjornson, M., Pimprikar, P., Nürnberger, T. and Zipfel, C. (2021) The transcriptional landscape of Arabidopsis thaliana pattern-triggered immunity. Nat. Plants. Available at: http://www.nature.com/articles/s41477-021-00874-5 [Accessed May 3, 2021].

Bonaventure, G., VanDoorn, A. and Baldwin, I.T. (2011) Herbivore-associated elicitors: FAC signaling and metabolism. Trends Plant Sci., 16, 294–299.

Bradbury, P.J., Zhang, Z., Kroon, D.E., Casstevens, T.M., Ramdoss, Y. and Buckler, E.S. (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics, 23, 2633–2635.

Casas, M.I., Falcone-Ferreyra, M.L., Jiang, N., Mejía-Guerra, M.K., Rodríguez, E., Wilson, T., Engelmeier, J., Casati, P. and Grotewold, E. (2016) Identification and Characterization of Maize salmon silks Genes Involved in Insecticidal Maysin Biosynthesis. Plant Cell, 28, 1297–1309.

Chakraborty, S. and Newton, A.C. (2011) Climate change, plant diseases and food security: an overview: Climate change and food security. Plant Pathol., 60, 2–14.

Chen, S., Zhou, Y., Chen, Y. and Gu, J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34, i884–i890.

Chen, T. (2021) Identification and characterization of the LRR repeats in plant LRR-RLKs. BMC Mol. Cell Biol., 22, 9.

Chen, Y., Ni, X. and Buntin, G.D. (2009) Physiological, Nutritional, and Biochemical Bases of Corn Resistance to Foliage-Feeding Fall Armyworm. J. Chem. Ecol., 35, 297–306.

Chung, H.S. and Howe, G.A. (2009) A Critical Role for the TIFY Motif in Repression of Jasmonate Signaling by a Stabilized Splice Variant of the JASMONATE ZIM-Domain Protein JAZ10 in Arabidopsis. Plant Cell, 21, 131–145.

Dafoe, N.J., Thomas, J.D., Shirk, P.D., Legaspi, M.E., Vaughan, M.M., Huffaker, A., Teal, P.E. and Schmelz, E.A. (2013) European Corn Borer (Ostrinia nubilalis) Induced Responses Enhance Susceptibility in Maize F. Marion-Poll, ed. PLoS ONE, 8, e73394.

Dahl, C.C. von, Winz, R.A., Halitschke, R., Kühnemann, F., Gase, K. and Baldwin, I.T. (2007) Tuning the herbivore-induced ethylene burst: the role of transcript accumulation and ethylene perception in Nicotiana attenuata: Ethylene biosynthesis during herbivory. Plant J., 51, 293–307.

Day, R., Abrahams, P., Bateman, M., Beale, T., Clottey, V., Cock, M., Colmenarez, Y., Corniani, N., Early, R., Godwin, J., Gomez, J., Moreno, P.G., Murphy, S.T., Oppong-Mensah, B., Phiri, N., Pratt, C., Silvestri, S. and Witt, A. (2017) Fall Armyworm: Impacts and Implications for Africa. outlook pest man, 28, 196–201.

Diezel, C., Dahl, C.C. von, Gaquerel, E. and Baldwin, I.T. (2009) Different Lepidopteran Elicitors Account for Cross-Talk in Herbivory-Induced Phytohormone Signaling. Plant Physiol., 150, 1576–1586.

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

Douglas, A.E. (2018) Strategies for Enhanced Crop Resistance to Insect Pests. Annu. Rev. Plant Biol., 69, 637–660.

Dressano, K., Weckwerth, P.R., Poretsky, E., Takahashi, Y., Villarreal, C., Shen, Z., Schroeder, J.I., Briggs, S.P. and Huffaker, A. (2020) Dynamic regulation of Pep-induced immunity through post-translational control of defence transcript splicing. Nat. Plants, 6, 1008–1019.

Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics, 5, 113.

Eichten, S.R., Foerster, J.M., Leon, N. de, Kai, Y., Yeh, C.-T., Liu, S., Jeddeloh, J.A., Schnable, P.S., Kaeppler, S.M. and Springer, N.M. (2011) B73-Mo17 Near-Isogenic Lines Demonstrate Dispersed Structural Variation in Maize. Plant Physiol., 156, 1679–1690.

Engelberth, J., Alborn, H.T., Schmelz, E.A. and Tumlinson, J.H. (2004) Airborne signals prime plants against insect herbivore attack. Proc. Natl. Acad. Sci., 101, 1781–1785.

Erb, M. (2018) Plant Defenses against Herbivory: Closing the Fitness Gap. Trends Plant Sci., 23, 187–194.

Erb, M., Meldau, S. and Howe, G.A. (2012) Role of phytohormones in insect-specific plant reactions. Trends Plant Sci., 17, 250–259.

Erb, M. and Reymond, P. (2019) Molecular Interactions Between Plants and Insect Herbivores. Annu. Rev. Plant Biol., 70, 527–557.

Erb, M., Veyrat, N., Robert, C.A.M., Xu, H., Frey, M., Ton, J. and Turlings, T.C.J. (2015) Indole is an essential herbivore-induced volatile priming signal in maize. Nat. Commun., 6, 6273.

Felton, G.W. and Tumlinson, J.H. (2008) Plant–insect dialogs: complex interactions at the plant–insect interface. Curr. Opin. Plant Biol., 11, 457–463.

Fürstenberg-Hägg, J., Zagrobelny, M. and Bak, S. (2013) Plant Defense against Insect Herbivores. Int. J. Mol. Sci., 14, 10242–10297.

The Gene Ontology Consortium (2019) The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res., 47, D330–D338.

Gilardoni, P.A., Hettenhausen, C., Baldwin, I.T. and Bonaventure, G. (2011) Nicotiana attenuata LECTIN RECEPTOR KINASE1 Suppresses the Insect-Mediated Inhibition of

Induced Defense Responses during Manduca sexta Herbivory. Plant Cell, 23, 3512–3532.

Gilardoni, P.A., Schuck, S., Jüngling, R., Rotter, B., Baldwin, I.T. and Bonaventure, G. (2010) SuperSAGE analysis of the Nicotiana attenuata transcriptome after fatty acid-amino acid elicitation (FAC): identification of early mediators of insect responses. BMC Plant Biol., 10, 66.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. and Rokhsar, D.S. (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Research, 40, D1178–D1186.

Grissett, L., Ali, A., Coble, A.-M., Logan, K., Washington, B., Mateson, A., McGee, K., Nkrumah, Y., Jacobus, L., Abraham, E., Hann, C., Bequette, C.J., Hind, S.R., Schmelz, E.A. and Stratmann, J.W. (2020) Survey of Sensitivity to Fatty Acid-Amino Acid Conjugates in the Solanaceae. J Chem Ecol, 46, 330–343.

Halitschke, R., Schittko, U., Pohnert, G., Boland, W. and Baldwin, I.T. (2001) Molecular Interactions between the Specialist Herbivore Manduca sexta (Lepidoptera, Sphingidae) and Its Natural Host Nicotiana attenuata . III. Fatty Acid-Amino Acid Conjugates in Herbivore Oral Secretions Are Necessary and Sufficient for Herbivore-Specific Plant Responses. Plant Physiol., 125, 711–717.

Han, Z., Sun, Y. and Chai, J. (2014) Structural insight into the activation of plant receptor kinases. Curr. Opin. Plant Biol., 20, 55–63.

Heidel, A.J. and Baldwin, I.T. (2004) Microarray analysis of salicylic acid- and jasmonic acid-signalling in responses of Nicotiana attenuata to attack by insects from multiple feeding guilds: Defence responses to herbivores from multiple feeding guilds. Plant Cell Environ., 27, 1362–1373.

Hohmann, U., Lau, K. and Hothorn, M. (2017) The Structural Basis of Ligand Perception and Signal Activation by Receptor Kinases. Annu. Rev. Plant Biol., 68, 109–137.

Horikoshi, R.J., Bernardi, D., Bernardi, O., Malaquias, J.B., Okuma, D.M., Miraldo, L.L., Amaral, F.S. de A. e and Omoto, C. (2016) Effective dominance of resistance of Spodoptera frugiperda to Bt maize and cotton varieties: implications for resistance management. Sci. Rep., 6, 34864.

Howe, G.A. and Jander, G. (2008) Plant Immunity to Insect Herbivores. Annu. Rev. Plant Biol., 59, 41–66.

Hu, L., Ye, M., Li, R., Zhang, T., Zhou, G., Wang, Q., Lu, J. and Lou, Y. (2015) The rice transcription factor WRKY53 suppresses herbivore-induced defenses by acting as a negative feedback modulator of map kinase activity. Plant Physiol., pp.01090.2015.

Hu, L., Ye, Meng, Kuai, P., Ye, Miaofen, Erb, M. and Lou, Y. (2018) OsLRR-RLK1, an early responsive leucine-rich repeat receptor-like kinase, initiates rice defense responses against a chewing herbivore. New Phytol., 219, 1097–1111.

Huffaker, A., Dafoe, N.J. and Schmelz, E.A. (2011) ZmPep1, an Ortholog of Arabidopsis Elicitor Peptide 1, Regulates Maize Innate Immunity and Enhances Disease Resistance. Plant Physiol., 155, 1325–1338.

Huffaker, A., Pearce, G. and Ryan, C.A. (2006) An endogenous peptide signal in Arabidopsis activates components of the innate immune response. Proc. Natl. Acad. Sci., 103, 10098–10103.

Huffaker, A., Pearce, G., Veyrat, N., Erb, M., Turlings, T.C.J., Sartor, R., Shen, Z., Briggs, S.P., Vaughan, M.M., Alborn, H.T., Teal, P.E.A. and Schmelz, E.A. (2013) Plant elicitor peptides are conserved signals regulating direct and indirect antiherbivore defense. Proceedings of the National Academy of Sciences, 110, 5707–5712.

Igarashi, D., Tsuda, K. and Katagiri, F. (2012) The peptide growth factor, phytosulfokine, attenuates pattern-triggered immunity: Attenuation of PTI by PSK. Plant J., 71, 194–204.

J. Merkler, D. and W. Leahy, J. (2018) Binding-based proteomic profiling and the fatty acid amides. Trends Res., 1. Available at: https://www.oatext.com/binding-based-proteomic-profiling-and-the-fatty-acid-amides.php [Accessed June 10, 2021].

Kall, L., Krogh, A. and Sonnhammer, E.L.L. (2007) Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. Nucleic Acids Res., 35, W429–W432.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Haeseler, A. von and Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods, 14, 587–589.

Karban, R. and Baldwin, I.T. (1997) Induced responses to herbivory, Chicago: University of Chicago Press.

Kessler, A. and Baldwin, I.T. (2002) P LANT R ESPONSES TO I NSECT H ERBIVORY : The Emerging Molecular Analysis. Annu. Rev. Plant Biol., 53, 299–328.

Kohany, O., Gentles, A.J., Hankus, L. and Jurka, J. (2006) Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics, 7, 474.

Lait, C.G., Alborn, H.T., Teal, P.E.A. and Tumlinson, J.H. (2003) Rapid biosynthesis of N-linolenoyl-L-glutamine, an elicitor of plant volatiles, by membrane-associated enzyme(s) in Manduca sexta. Proc. Natl. Acad. Sci., 100, 7027–7032.

Lee, M., Sharopova, N., Beavis, W.D., Grant, D., Katt, M., Blair, D. and Hallauer, A. (2002) Expanding the genetic map of maize with the intermated B73 × Mo17 (IBM) population. Plant Mol. Biol., 48, 453–461.

Li, H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, 27, 2987–2993.

Li, L., Petsch, K., Shimizu, R., Liu, S., Xu, W.W., Ying, K., Yu, J., Scanlon, M.J., Schnable, P.S., Timmermans, M.C.P., Springer, N.M. and Muehlbauer, G.J. (2013) Mendelian and Non-Mendelian Regulation of Gene Expression in Maize G. P. Copenhaver, ed. PLoS Genet, 9, e1003202.

Li, R., Zhang, J., Li, J., Zhou, G., Wang, Q., Bian, W., Erb, M. and Lou, Y. (2015) Prioritizing plant defence over growth through WRKY regulation facilitates infestation by non-target herbivores. eLife, 4, e04805.

Ling, X., Gu, S., Tian, C., Guo, H., Degen, T., Turlings, T.C.J., Ge, F. and Sun, Y. (2021) Differential Levels of Fatty Acid-Amino Acid Conjugates in the Oral Secretions of Lepidopteran Larvae Account for the Different Profiles of Volatiles. Pest Manag. Sci., ps.6417.

Liu, J., Fernie, A.R. and Yan, J. (2020) The Past, Present, and Future of Maize Improvement: Domestication, Genomics, and Functional Genomic Routes toward Crop Enhancement. Plant Commun., 1, 100010.

Liu, Z., Wu, Y., Yang, F., Zhang, Y., Chen, S., Xie, Q., Tian, X. and Zhou, J.-M. (2013) BIK1 interacts with PEPRs to mediate ethylene-induced immunity. Proc. Natl. Acad. Sci., 110, 6205–6210.

Ma, X., Xu, G., He, P. and Shan, L. (2016) SERKing Coreceptors for Receptors. Trends Plant Sci., 21, 1017–1033.

Maag, D., Köhler, A., Robert, C.A.M., Frey, M., Wolfender, J.-L., Turlings, T.C.J., Glauser, G. and Erb, M. (2016) Highly localized and persistent induction of Bx1 - dependent herbivore resistance factors in maize. Plant J., 88, 976–991.

Maffei, M.E. (2010) Sites of synthesis, biochemistry and functional role of plant volatiles. South Afr. J. Bot., 76, 612–631.

Maffei, M.E., Mithöfer, A. and Boland, W. (2007a) Before gene expression: early events in plant–insect interaction. Trends Plant Sci., 12, 310–316.

Maffei, M.E., Mithöfer, A. and Boland, W. (2007b) Insects feeding on plants: Rapid signals and responses preceding the induction of phytochemical release. Phytochemistry, 68, 2946–2959.

Mattiacci, L., Dicke, M. and Posthumus, M.A. (1995) β-Glucosidase: An Elicitor of Herbivore-Induced Plant Odor that Attracts Host-Searching Parasitic Wasps. Proc. Natl. Acad. Sci. U. S. A., 92, 2036–2040.

McMullen, M.D., Kresovich, S., Villeda, H.S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S., Thornsberry, J., Acharya, C., Bottoms, C., Brown, P., Browne, C., Eller, M., Guill, K., Harjes, C., Kroon, D., Lepak, N., Mitchell, S.E., Peterson, B., Pressoir, G., Romero, S., Rosas, M.O., Salvo, S., Yates, H., Hanson, M., Jones, E., Smith, S., Glaubitz, J.C., Goodman, M., Ware, D., Holland, J.B. and Buckler, E.S. (2009) Genetic Properties of the Maize Nested Association Mapping Population. Science, 325, 737–740.

Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.-Y., El-Gebali, S., Fraser, M.I., Gough, J., Haft, D.R., Huang, H., Letunic, I., Lopez, R., Luciani, A., Madeira, F., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Nuka, G., Orengo, C., Pandurangan, A.P., Paysan-Lafosse, T., Pesseat, S., Potter, S.C., Qureshi, M.A., Rawlings, N.D., Redaschi, N., Richardson, L.J., Rivoire, C., Salazar, G.A., Sangrador-Vegas, A., Sigrist, C.J.A., Sillitoe, I., Sutton, G.G., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Yong, S.-Y. and Finn, R.D. (2019) InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Research, 47, D351–D360.

Mithöfer, A. and Boland, W. (2012) Plant Defense Against Herbivores: Chemical Aspects. Annu. Rev. Plant Biol., 63, 431–450.

Montezano, D.G., Specht, A., Sosa-Gómez, D.R., Roque-Specht, V.F., Sousa-Silva, J.C., Paula-Moraes, S.V., Peterson, J.A. and Hunt, T.E. (2018) Host Plants of Spodoptera frugiperda (Lepidoptera: Noctuidae) in the Americas. Afr. Entomol., 26, 286–300.

Mori, N. and Yoshinaga, N. (2011) Function and evolutionary diversity of fatty acid amino acid conjugates in insects. J. Plant Interact., 6, 103–107.

Mosher, S. and Kemmerling, B. (2013) PSKR1 and PSY1R-mediated regulation of plant defense responses. Plant Signal. Behav., 8, e24119.

McMullen, M.D., Kresovich, S., Villeda, H.S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S., Thornsberry, J., Acharya, C., Bottoms, C., Brown, P., Browne, C., Eller, M., Guill, K., Harjes, C., Kroon, D., Lepak, N., Mitchell, S.E., Peterson, B., Pressoir, G., Romero, S., Rosas, M.O., Salvo, S., Yates, H., Hanson, M., Jones, E., Smith, S., Glaubitz, J.C., Goodman, M., Ware, D., Holland, J.B. and Buckler, E.S. (2009) Genetic Properties of the Maize Nested Association Mapping Population. Science, 325, 737–740.

Mousavi, S.A.R., Chauvin, A., Pascaud, F., Kellenberger, S. and Farmer, E.E. (2013) GLUTAMATE RECEPTOR-LIKE genes mediate leaf-to-leaf wound signalling. Nature, 500, 422–426.

Nakagawa, T., Suzuki, T., Murata, S., Nakamura, S., Hino, T., Maeo, K., Tabata, R., Kawai, T., Tanaka, K., Niwa, Y., Watanabe, Y., Nakamura, K., Kimura, T. and Ishiguro, S. (2007) Improved Gateway Binary Vectors: High-Performance Vectors for Creation of Fusion Constructs in Transgenic Analysis of Plants. Bioscience, Biotechnology, and Biochemistry, 71, 2095–2100.

Nguyen, L.-T., Schmidt, H.A., Haeseler, A. von and Minh, B.Q. (2015) IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol. Biol. Evol., 32, 268–274.

Oikawa, A., Ishihara, A., Tanaka, C., Mori, N., Tsuda, M. and Iwamura, H. (2004) Accumulation of HDMBOA-Glc is induced by biotic stresses prior to the release of MBOA in maize leaves. Phytochemistry, 65, 2995–3001.

Onkokesung, N., Gális, I., Dahl, C.C. von, Matsuoka, K., Saluz, H.-P. and Baldwin, I.T. (2010) Jasmonic Acid and Ethylene Modulate Local Responses to Wounding and Simulated Herbivory in Nicotiana attenuata Leaves. Plant Physiol., 153, 785–798.

Orozco-Cardenas, M., McGurl, B. and Ryan, C.A. (1993) Expression of an antisense prosystemin gene in tomato plants reduces resistance toward Manduca sexta larvae. Proc. Natl. Acad. Sci., 90, 8273–8276.

Orozco-Cardenas, M. and Ryan, C.A. (1999) Hydrogen peroxide is generated systemically in plant leaves by wounding and systemin via the octadecanoid pathway. Proc. Natl. Acad. Sci., 96, 6553–6557.

Overton, K., Maino, J., Day, R., Umina, P., Bett, B., Carnovale, D., Ekesi, S., Meagher, R. and Reynolds, O. (2021) Global crop impacts, yield losses and action thresholds for fall armyworm (Spodoptera frugiperda): A review. Crop Prot., 145, 105641.

Pare, P.W., Alborn, H.T. and Tumlinson, J.H. (1998) Concerted biosynthesis of an insect elicitor of plant volatiles. Proc. Natl. Acad. Sci., 95, 13971–13975.

Parra, J.R.P., Coelho, A., Cuervo-Rugno, J.B., Garcia, A.G., Andrade Moral, R. de, Specht, A. and Neto, D.D. (2021) Important pest species of the Spodoptera complex: Biology, thermal requirements and ecological zoning. J. Pest Sci. Available at: http://link.springer.com/10.1007/s10340-021-01365-4.

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

Poretsky, E., Dressano, K., Weckwerth, P., Ruiz, M., Char, S.N., Shi, D., Abagyan, R., Yang, B. and Huffaker, A. (2020) Differential activities of maize plant elicitor peptides as mediators of immune signaling and herbivore resistance. Plant J., tpj.15022.

Rasmann, S., Köllner, T.G., Degenhardt, J., Hiltpold, I., Toepfer, S., Kuhlmann, U., Gershenzon, J. and Turlings, T.C.J. (2005) Recruitment of entomopathogenic nematodes by insect-damaged maize roots. Nature, 434, 732–737.

Ray, S., Gaffor, I., Acevedo, F.E., Helms, A., Chuang, W.-P., Tooker, J., Felton, G.W. and Luthe, D.S. (2015) Maize Plants Recognize Herbivore-Associated Cues from Caterpillar Frass. J. Chem. Ecol., 41, 781–792.

Reymond, P. (2021) Receptor kinases in plant responses to herbivory. Curr. Opin. Biotechnol., 70, 143–150.

Richter, A., Schaff, C., Zhang, Z., Lipka, A.E., Tian, F., Köllner, T.G., Schnee, C., Preiß, S., Irmisch, S., Jander, G., Boland, W., Gershenzon, J., Buckler, E.S. and Degenhardt, J. (2016) Characterization of Biosynthetic Pathways for the Production of the Volatile Homoterpenes DMNT and TMTT in Zea mays. Plant Cell, 28, 2651–2665.

Romay, M.C., Millard, M.J., Glaubitz, J.C., Peiffer, J.A., Swarts, K.L., Casstevens, T.M., Elshire, R.J., Acharya, C.B., Mitchell, S.E., Flint-Garcia, S.A., McMullen, M.D., Holland, J.B., Buckler, E.S. and Gardner, C.A. (2013) Comprehensive genotyping of the USA national maize inbred seed bank. Genome Biol, 14, R55.

Ross, A., Yamada, K., Hiruma, K., Yamashita-Yamada, M., Lu, X., Takano, Y., Tsuda, K. and Saijo, Y. (2014) The Arabidopsis PEPR pathway couples local and systemic plant immunity. EMBO J., 33, 62–75.

Sandoval, P.J. and Santiago, J. (2020) In Vitro Analytical Approaches to Study Plant Ligand-Receptor Interactions. Plant Physiol., 182, 1697–1712.

Schmelz, E.A. (2015) Impacts of insect oral secretions on defoliation-induced plant defense. Curr. Opin. Insect Sci., 9, 7–15.

Schmelz, E.A., Alborn, H.T. and Tumlinson, J.H. (2003) Synergistic interactions between volicitin, jasmonic acid and ethylene mediate insect-induced volatile emission in Zea mays. Physiol. Plant., 117, 403–412.

Schmelz, E.A., Engelberth, J., Alborn, H.T., Tumlinson, J.H. and Teal, P.E.A. (2009) Phytohormone-based activity mapping of insect herbivore-produced elicitors. Proc. Natl. Acad. Sci., 106, 653–657.

Schmelz, E.A., Kaplan, F., Huffaker, A., Dafoe, N.J., Vaughan, M.M., Ni, X., Rocca, J.R., Alborn, H.T. and Teal, P.E. (2011) Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. Proc. Natl. Acad. Sci., 108, 5455–5460.

Schmelz, E.A., LeClere, S., Carroll, M.J., Alborn, H.T. and Teal, P.E.A. (2007) Cowpea Chloroplastic ATP Synthase Is the Source of Multiple Plant Defense Elicitors during Insect Herbivory. Plant Physiol., 144, 793–805.

Schweizer, F., Fernández-Calvo, P., Zander, M., Diez-Diaz, M., Fonseca, S., Glauser, G., Lewsey, M.G., Ecker, J.R., Solano, R. and Reymond, P. (2013) Arabidopsis Basic Helix-Loop-Helix Transcription Factors MYC2, MYC3, and MYC4 Regulate Glucosinolate Biosynthesis, Insect Performance, and Feeding Behavior. The Plant Cell, 25, 3117–3132.

Sharma, I. and Russinova, E. (2018) Probing Plant Receptor Kinase Functions with Labeled Ligands. Plant Cell Physiol., 59, 1520–1527.

Shehryar, K., Khan, R.S., Iqbal, A., Hussain, S.A., Imdad, S., Bibi, A., Hamayun, L. and Nakamura, I. (2020) Transgene Stacking as Effective Tool for Enhanced Disease Resistance in Plants. Mol. Biotechnol., 62, 1–7.

Shen, Y. and Diener, A.C. (2013) Arabidopsis thaliana RESISTANCE TO FUSARIUM OXYSPORUM 2 Implicates Tyrosine-Sulfated Peptide Signaling in Susceptibility and Resistance to Root Infection J. M. McDowell, ed. PLoS Genet., 9, e1003525.

Shinya, T., Hojo, Y., Desaki, Y., Christeller, J.T., Okada, K., Shibuya, N. and Galis, I. (2016) Modulation of plant defense responses to herbivores by simultaneous recognition of different herbivore-associated elicitors in rice. Sci. Rep., 6, 32537.

Shinya, T., Yasuda, S., Hyodo, K., Tani, R., Hojo, Y., Fujiwara, Y., Hiruma, K., Ishizaki, T., Fujita, Y., Saijo, Y. and Galis, I. (2018) Integration of danger peptide signals with herbivore-associated molecular pattern signaling amplifies anti-herbivore defense responses in rice. Plant J, 94, 626–637.

Shiu, S.-H., Karlowski, W.M., Pan, R., Tzeng, Y.-H., Mayer, K.F.X. and Li, W.-H. (2004) Comparative Analysis of the Receptor-Like Kinase Family in Arabidopsis and Rice. Plant Cell, 16, 1220–1234.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. and Higgins, D.G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol, 7, 539.

Smakowska-Luzan, E., Mott, G.A., Parys, K., Stegmann, M., Howton, T.C., Layeghifard, M., Neuhold, J., Lehner, A., Kong, J., Grünwald, K., Weinberger, N., Satbhai, S.B., Mayer, D., Busch, W., Madalinski, M., Stolt-Bergner, P., Provart, N.J., Mukhtar, M.S., Zipfel, C., Desveaux, D., Guttman, D.S. and Belkhadir, Y. (2018) An extracellular network of Arabidopsis leucine-rich repeat receptor kinases. Nature, 553, 342–346.

Smith, W.E.C., Shivaji, R., Williams, W.P., Luthe, D.S., Sandoya, G.V., Smith, C.L., Sparks, D.L. and Brown, A.E. (2012) A Maize Line Resistant to Herbivory Constitutively Releases (E)-β-Caryophyllene. J. Econ. Entomol., 105, 120–128.

Steinbrenner, A.D., Muñoz-Amatriaín, M., Chaparro, A.F., Aguilar-Venegas, J.M., Lo, S., Okuda, S., Glauser, G., Dongiovanni, J., Shi, D., Hall, M., Crubaugh, D., Holton, N.,

Zipfel, C., Abagyan, R., Turlings, T.C.J., Close, T.J., Huffaker, A. and Schmelz, E.A. (2020) A receptor-like protein mediates plant immune responses to herbivore-associated molecular patterns. Proc Natl Acad Sci USA, 202018415.

Tanaka, K., Choi, J., Cao, Y. and Stacey, G. (2014) Extracellular ATP acts as a damage-associated molecular pattern (DAMP) signal in plants. Front. Plant Sci., 5. Available at: http://journal.frontiersin.org/article/10.3389/fpls.2014.00446/abstract [Accessed July 11, 2020].

Thaler, J.S., Farag, M.A., Paré, P.W. and Dicke, M. (2002) Jasmonate-deficient plants have reduced direct and indirect defences against herbivores: Jasmonate-deficient plants. Ecol. Lett., 5, 764–774.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L.A., Rhee, S.Y. and Stitt, M. (2004) mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant Journal, 37, 914–939.

Tintor, N., Ross, A., Kanehara, K., Yamada, K., Fan, L., Kemmerling, B., Nurnberger, T., Tsuda, K. and Saijo, Y. (2013) Layered pattern receptor signaling via ethylene and endogenous elicitor peptides during Arabidopsis immunity to bacterial infection. Proc. Natl. Acad. Sci., 110, 6211–6216.

Torres, M.A., Dangl, J.L. and Jones, J.D.G. (2002) Arabidopsis gp91phox homologues AtrbohD and AtrbohF are required for accumulation of reactive oxygen intermediates in the plant defense response. Proc. Natl. Acad. Sci., 99, 517–522.

Truitt, C.L., Wei, H.-X. and Paré, P.W. (2004) A Plasma Membrane Protein from Zea mays Binds with the Herbivore Elicitor Volicitin. Plant Cell, 16, 523–532.

Turlings, T.C.J., Alborn, H.T., Loughrin, J.H. and Tumlinson, J.H. (2000) Volicitin, An Elicitor of Maize Volatiles in Oral Secretion of Spodoptera Exigua: Isolation and Bioactivity. J. Chem. Ecol., 26, 189–202.

Turlings, T.C.J., Lengwiler, U.B., Bernasconi, M.L. and Wechsler, D. (1998) Timing of induced volatile emissions in maize seedlings. Planta, 207, 146–152.

Turlings, T.C.J., McCall, P.J., Alborn, H.T. and 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, 411–425.

Turlings, T.C.J., Scheepmaker, J.W.A., Vet, L.E.M., Tumlinson, J.H. and Lewis, W.J. (1990) How contact foraging experiences affect preferences for host-related odors in the larval parasitoidCotesia marginiventris (Cresson) (Hymenoptera: Braconidae). J. Chem. Ecol., 16, 1577–1589.

Tzin, V., Fernandez-Pozo, N., Richter, A., Schmelz, E.A., Schoettner, M., Schäfer, M., Ahern, K.R., Meihls, L.N., Kaur, H., Huffaker, A., Mori, N., Degenhardt, J., Mueller, L.A. and Jander, G. (2015) Dynamic maize responses to aphid feeding are revealed by a time series of transcriptomic and metabolomic assays. Plant Physiol., pp.01039.2015.

Uemura, T., Hachisu, M., Desaki, Y., Ito, A., Hoshino, R., Sano, Y., Nozawa, A., Mujiono, K., Galis, I., Yoshida, A., Nemoto, K., Miura, S., Nishiyama, M., Nishiyama, C., Horito, S., Sawasaki, T. and Arimura, G. (2020) Soy and Arabidopsis receptor-like kinases respond to polysaccharide signals from Spodoptera species and mediate herbivore resistance. Commun Biol, 3, 224.

VanDoorn, A., Kallenbach, M., Borquez, A.A., Baldwin, I.T. and Bonaventure, G. (2010) Rapid modification of the insect elicitor N-linolenoyl-glutamate via a lipoxygenase-mediated mechanism on Nicotiana attenuata leaves. BMC Plant Biol., 10, 164.

Waiss, A.C., Chan, B.G., Elliger, C.A., Wiseman, B.R., McMillian, W.W., Widstrom, N.W., Zuber, M.S. and Keaster, A.J. (1979) Maysin, a Flavone Glycoside from Corn Silks with Antibiotic Activity Toward Corn Earworm13. J. Econ. Entomol., 72, 256–258.`

Wang, J., Li, H., Han, Z., Zhang, H., Wang, T., Lin, G., Chang, J., Yang, W. and Chai, J. (2015) Allosteric receptor activation by the plant peptide hormone phytosulfokine. Nature, 525, 265–268.

Wang, J., Yang, M., Song, Y., Acevedo, F.E., Hoover, K., Zeng, R. and Felton, G.W. (2018) Gut-Associated Bacteria of Helicoverpa zea Indirectly Trigger Plant Defenses in Maize. J. Chem. Ecol., 44, 690–699.

Wang, L., Einig, E., Almeida-Trapp, M., Albert, M., Fliegmann, J., Mithöfer, A., Kalbacher, H. and Felix, G. (2018) The systemin receptor SYR1 enhances resistance of tomato against herbivorous insects. Nat. Plants, 4, 152–156.

Wouters, F.C., Blanchette, B., Gershenzon, J. and Vassão, D.G. (2016) Plant defense and herbivore counter-defense: benzoxazinoids and insect herbivores. Phytochem. Rev., 15, 1127–1151.

Wu, J. and Baldwin, I.T. (2009) Herbivory-induced signalling in plants: perception and action. Plant Cell Environ., 32, 1161–1174.

Xu, S., Zhou, W., Pottinger, S. and Baldwin, I.T. (2015) Herbivore associated elicitor-induced defences are highly specific among closely related Nicotiana species. BMC Plant Biol., 15, 2.

Yamaguchi, Y., Huffaker, A., Bryan, A.C., Tax, F.E. and Ryan, C.A. (2010) PEPR2 Is a Second Receptor for the Pep1 and Pep2 Peptides and Contributes to Defense Responses in Arabidopsis. Plant Cell, 22, 508–522.

Yang, D.-H., Hettenhausen, C., Baldwin, I.T. and Wu, J. (2011) BAK1 regulates the accumulation of jasmonic acid and the levels of trypsin proteinase inhibitors in Nicotiana attenuata's responses to herbivory. J. Exp. Bot., 62, 641–652.

Yoshinaga, N., Aboshi, T., Abe, H., Nishida, R., Alborn, H.T., Tumlinson, J.H. and Mori, N. (2008) Active role of fatty acid amino acid conjugates in nitrogen metabolism in Spodoptera litura larvae. Proc. Natl. Acad. Sci., 105, 18058–18063.

Yoshinaga, N., Aboshi, T., Ishikawa, C., Fukui, M., Shimoda, M., Nishida, R., Lait, C.G., Tumlinson, J.H. and Mori, N. (2007) Fatty Acid Amides, Previously Identified in Caterpillars, Found in the Cricket Teleogryllus taiwanemma and Fruit Fly Drosophila melanogaster Larvae. J. Chem. Ecol., 33, 1376–1381.

Yoshinaga, N., Alborn, H.T., Nakanishi, T., Suckling, D.M., Nishida, R., Tumlinson, J.H. and Mori, N. (2010) Fatty Acid-amino Acid Conjugates Diversification in Lepidopteran Caterpillars. J. Chem. Ecol., 36, 319–325.

Zhang, W.-X., Pan, X. and Shen, H.-B. (2020) Signal-3L 3.0: Improving Signal Peptide Prediction through Combining Attention Deep Learning with Window-Based Scoring. J. Chem. Inf. Model., 60, 3679–3686.

Zhou, G., Wang, X., Yan, F., Wang, X., Li, R., Cheng, J. and Lou, Y. (2011) Genome-wide transcriptional changes and defence-related chemical profiling of rice in response to infestation by the rice striped stem borer Chilo suppressalis. Physiol. Plant., 143, 21–40.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T. and Felix, G. (2006) Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts Agrobacterium-Mediated Transformation. Cell, 125, 749–760.

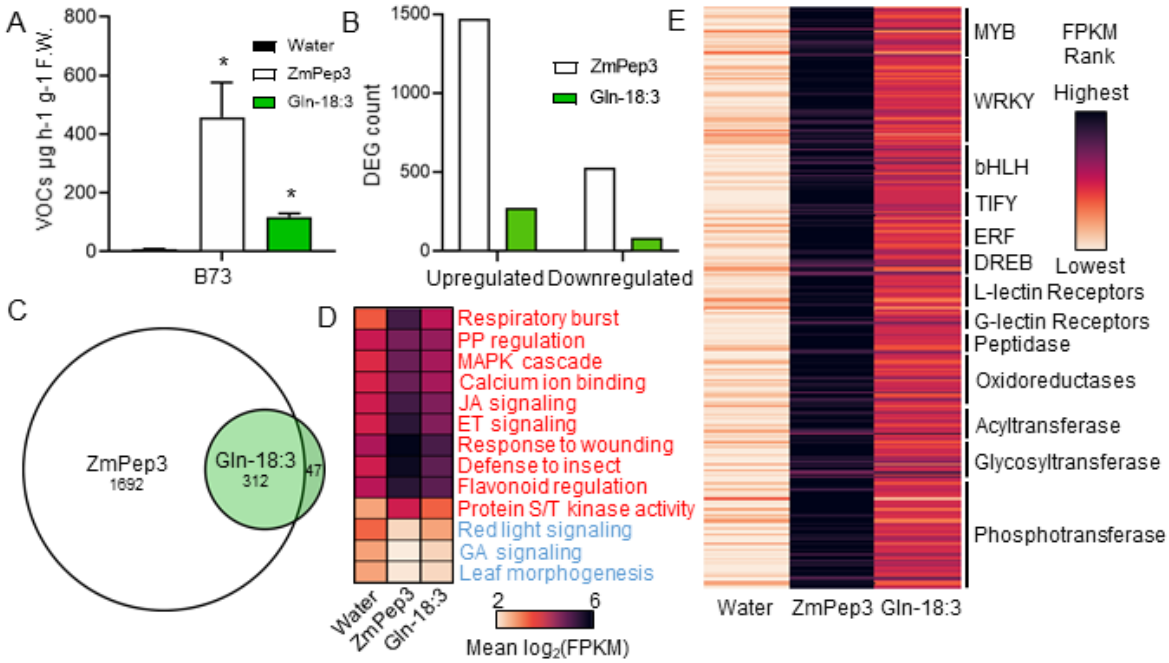


Figure 1-1. Comparative analysis of ZmPep3- and Gln-18:3-induced VOC emission and early transcriptional changes.

(A) Analysis of total elicitor-induced VOC emission 16 hours after treatment of B73 leaves with water, 5µM ZmPep3 or 5µM Gln-18:3. **(B)** Measurement of the number of differentially expressed genes (DEGs) following elicitor-induced 2 hour treatments measured by RNA-seq. **(C)** Euler diagram representing the overlap of the combined upregulated and downregulated DEGs. **(D)** Summary heatmap of the mean $\log_2$(FPKM) data of all ZmPep3- and Gln-18:3-induced DEGs in selected enriched GO terms. Enriched GO terms for upregulated DEGs are in a red font and downregulated DEGs are in a blue font. PP stands for phenylpropanoid **(E)** Summary heatmap of all ZmPep3 upregulated DEGs assigned to their enriched MAPMAN bins with more than 10 DEGs in respective bins. Row colors are based on the treatment mean of the ranked FPKM values for each gene. For all treatments shown n=4 and error bars represent SEM. Student t-tests (two-tailed distribution, unpaired), was used for detection of significant differences. Asterisks indicate significant difference with $P < 0.05$; ns, not significant.

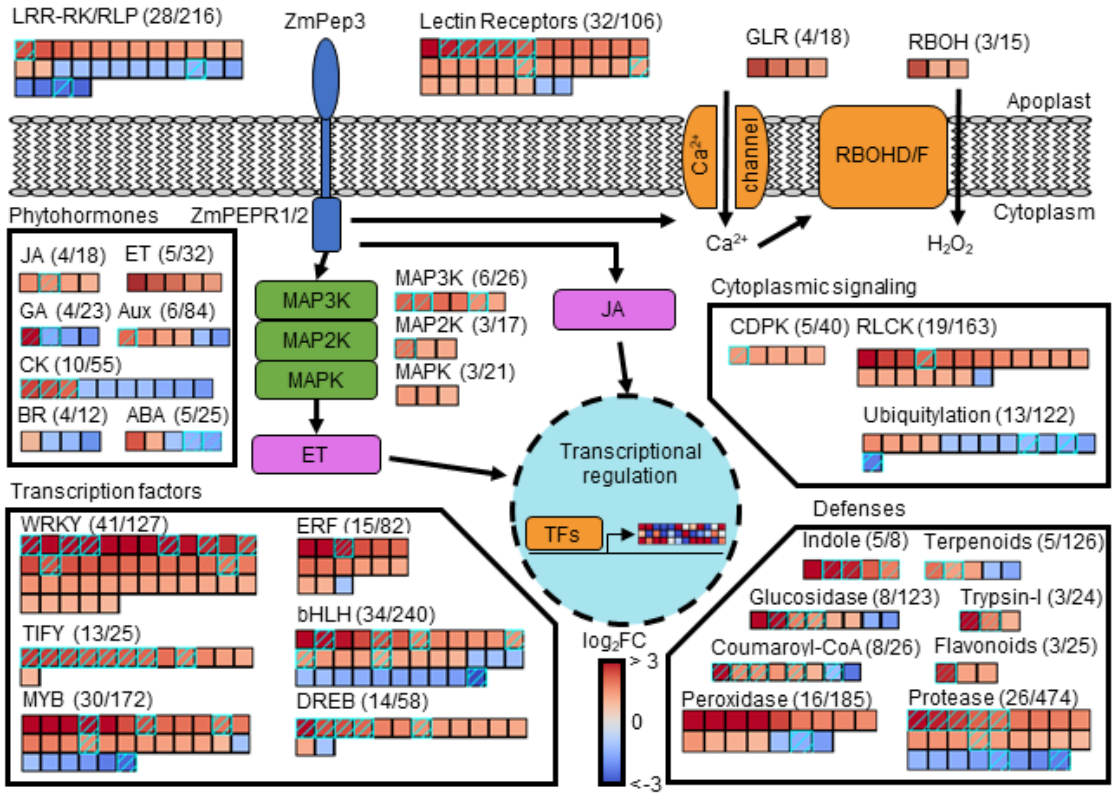


Figure 1-2. Early elicitor-induced transcriptional regulation of multiple components associated with antiherbivore responses.
Heatmaps of gene groups represent log_2 fold-change data from RNA-seq results of ZmPep3-induced differentially expressed genes and based the *Zea mays* V4 MAPMAN annotation. Hatched cyan boxes indicate genes that were also significantly differentially expressed following Gln-18:3 treatment. The selected gene groups were obtained either from the MAPMAN bin annotations or from searching selected keywords in the MAPMAN gene descriptions in the case of the glucosidases, peroxidases, proteases and trypsin-inhibitors (Trypsin-I) gene groups. In the parenthesis next to each group is the number of ZmPep3 treated DEGs compared to the size of the gene group. Abbreviations stand for: LRR-RK/ leucine-rich repeat receptor kinase/receptor like protein; GLR, glutamate-like receptors; RBOH, respiratory burst oxidase homolog; JA, jasmonic acid; ET, ethylene; GA, gibberellic acid; Aux, auxin; CK, cytokinins; BR, brassinosteroids; ABA, abscisic acid; MAPKs, mitogen activated protein kinase; CDPK, calcium-dependent protein kinase; RLCK, receptor-like cytoplasmic kinase; ERF, ethylene response factor; Trypsin-I, Trypsin inhibitor; TFs, transcription factors.

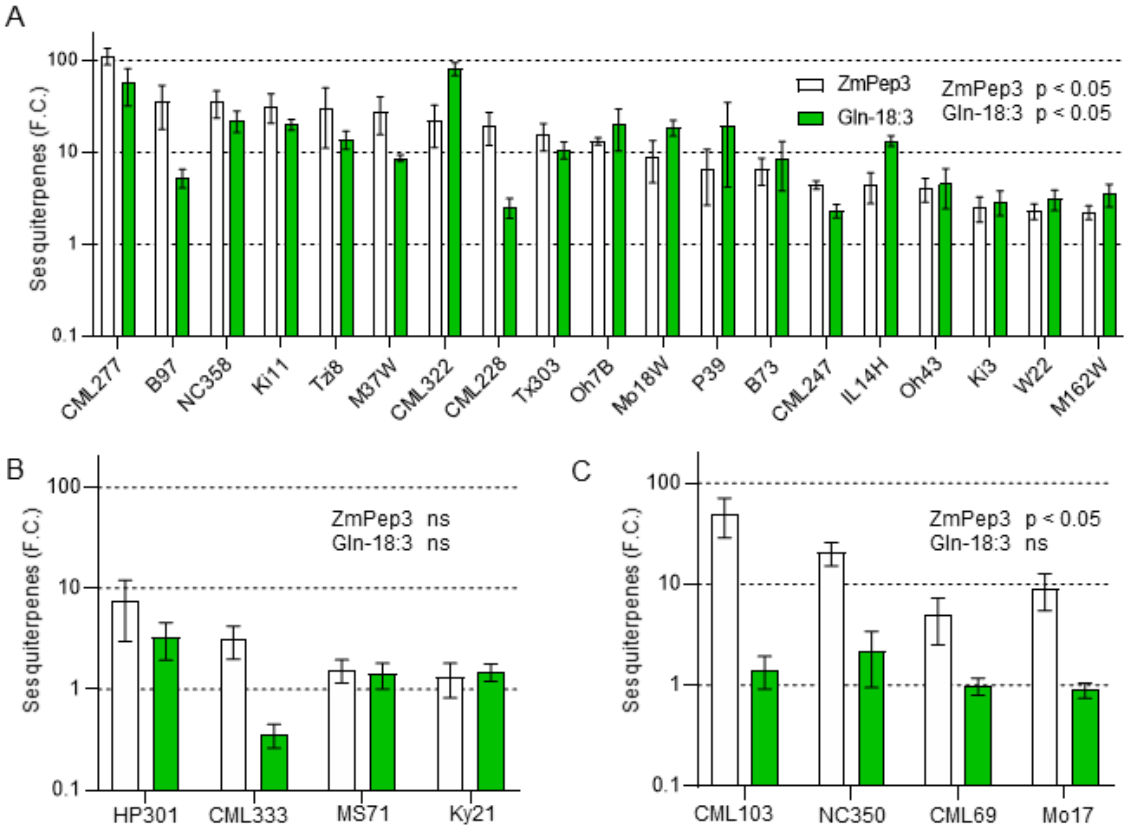


Figure 1-3. Screening *Zea mays* inbred lines for intact elicitor-induced defense responses but impaired FAC sensitivity.

A total of 27 *Zea mays* inbred lines, B73, Mo17, W22 and the NAM parents (not including CML52), were tested for elicitor-induced volatile sesquiterpene emission. Maize leaves were treated for 16 hours with water, 1μM ZmPep3 or 1μM Gln-18:3 and volatiles were analyzed by GC. Results are presented as elicitor-induced fold-change (F.C.) sesquiterpene emission compared to the water treatment, sesquiterpene defined as the sum of caryophyllene, α-bergamotene, β-farnesene and trans-nerolidol. For all treatments shown n=4 and error bars represent SEM. Significance of elicitor-induced sesquiterpene emission was determined by Student t-tests (two-tailed distribution, unpaired), for each line, and was used to divide the results into three categories: **(A)** Lines with significant elicitor-induced responses ($p < 0.05$), **(B)** lines where elicitor-induced responses are not significant (ns), and **(C)** lines with significant ZmPep3-induced responses but not significant FAC-induced responses.

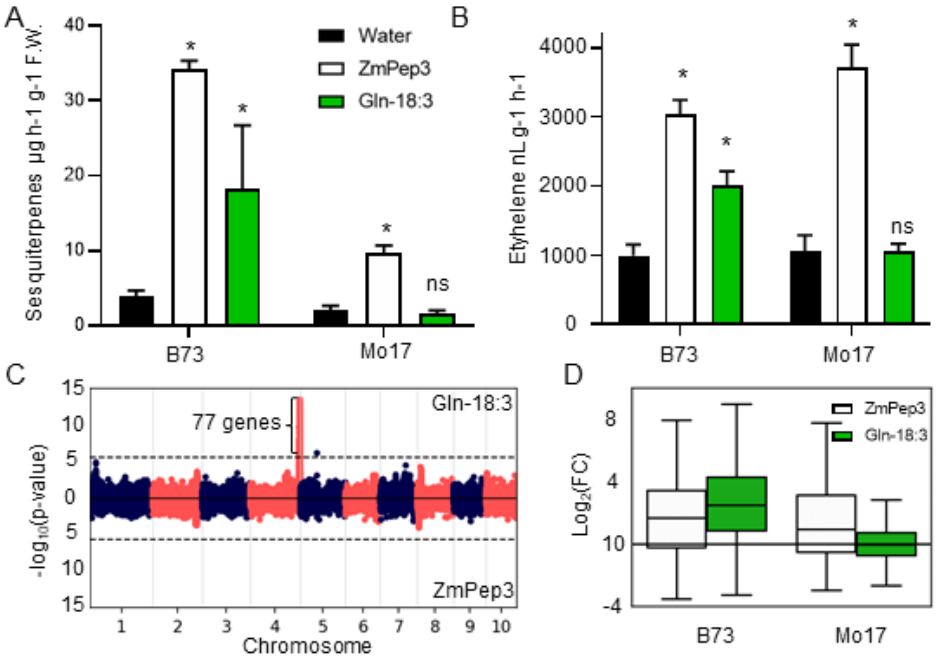


Figure 1-4. Genetic mapping for an FAC sensitivity locus based on Mo17 insensitivity to Gln-18:3.

(A-B) Analysis of total VOC emission 16 hours after treatment and **(B)** ethylene emission 2 hours after treatment of leaf 5 of B73 and Mo17 seedlings with water, 5µM ZmPep3 and 5µM Gln-18:3. **(C)** QTL mapping using the the imputed SNP marker map obtained from the Panzea project for the IBM mapping population with fold-change values from leaf 5 of 242 RILs treated with 1µM ZmPep3- and 1µM Gln-18:3-induced VOC in comparison to water treatment 16 hours after treatment. Fold-change values were calculated using the sum of the emitted sesquiterpene volatiles, defined as the sum of caryophyllene, α-bergamotene, β-farnesene and trans-nerolidol. **(D)** Box plots showing the distribution of elicitor-induced fold-change VOC emission at the top SNP associated with FAC sensitivity across the 222 IBM RILs. Dashed lines in manhattan plots indicate significance cutoff of $P < 0.05$ based on Bonferroni correction. For VOC treatment n=4 and for ethylene treatment n=6. Gene number are based on the B73v4 annotation. Error bars represent SEM. Student t-tests (two-tailed distribution, unpaired), was used for detection of significant differences. Asterisks indicate significant difference with $P < 0.05$; ns, not significant.

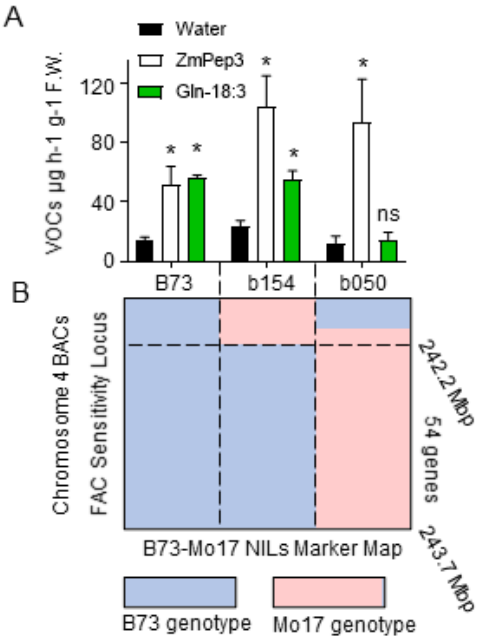


Figure 1-5. Fine-mapping the FAC sensitivity locus using the B73 x Mo17 NIL and IBM mapping populations.

(A) Total elicitor-induced VOC emission 16 hours after treatment of B73, b154 and b050 leaves with water, 5µM ZmPep3 or 5µM Gln-18:3. Total VOC is represented by the sum of hexenyl acetate, linalool, DMNT, caryophyllene, α-bergamotene, β-farnesene and trans-nerolidol. **(B)** Genetic map of the B73-Mo17 Near Isogenic Lines (NIL) mapping population shows the two lines, b153 and b050, containing an introgression of the Mo17 genotype into the B73 genetic background within the FAC Sensitivity Locus. Coordinates and gene number are based on the B73v4 annotation. For all treatments shown n=4 and error bars represent SEM. Student t-tests (two-tailed distribution, unpaired), was used for detection of significant differences. Asterisks indicate significant difference with $P < 0.05$; ns, not significant.

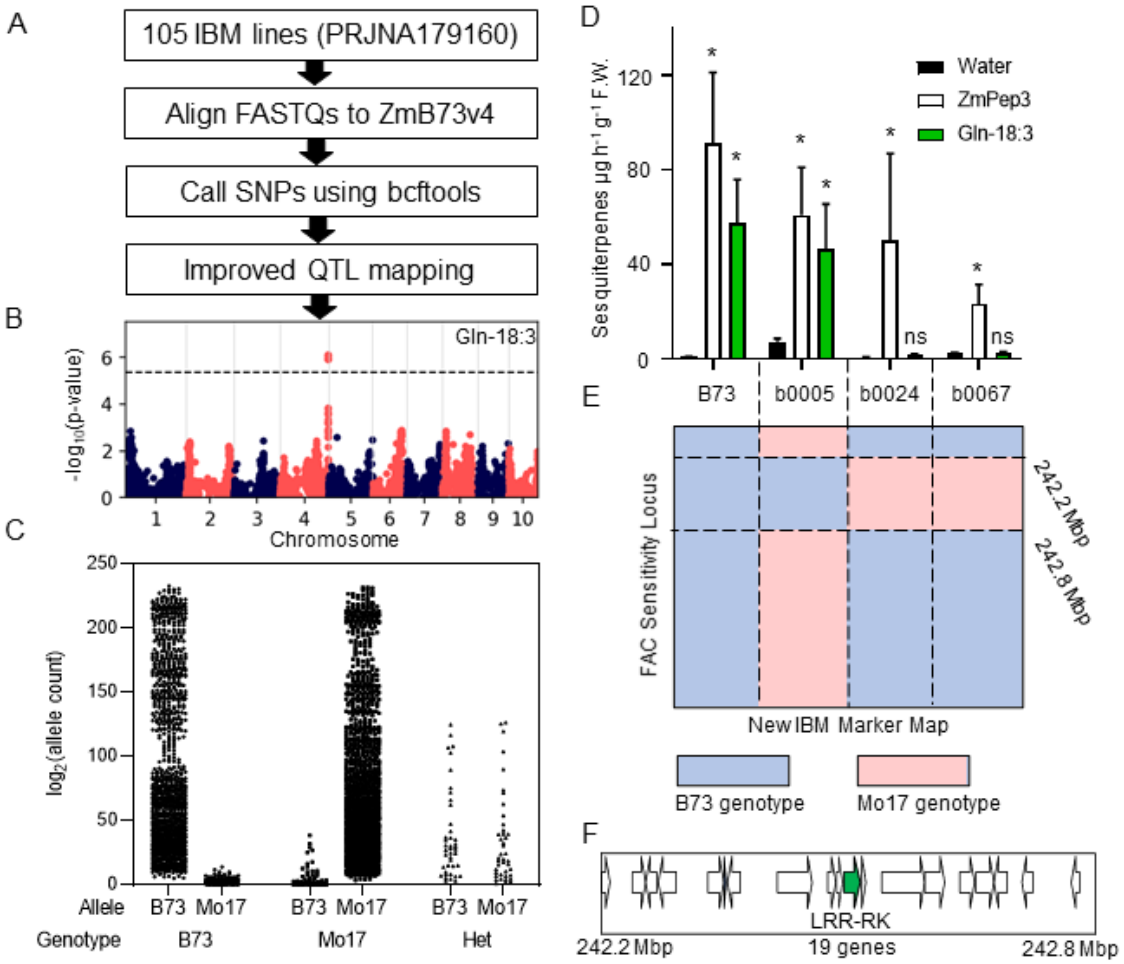


Figure 1-6. Fine-mapping the FAC sensitivity locus using a newly generated marker map for 105 IBM lines.

(A) Summarized workflow describing the procedure for generating the new IBM marker map using RNA-seq data from 105 IBM lines obtained from NCBI-SRA experiment PRJNA179160. **(B)** QTL mapping using the the new IBM marker map with fold-change values from leaf 5 of 86 RILs treated with 1µM ZmPep3- and 1µM Gln-18:3-induced VOC in comparison to water treatment 16 hours after treatment. This QTL mapping was done with the same data as the QTL mapping presented in Fig. 4C for overlapping lines. Dashed lines in manhattan plots indicate significance cutoff of p value < 0.05 based on Bonferroni correction. **(C)** Aggregated allele counts of all SNPs within the FAC sensitivity locus that were used for generating the new IBM marker map compared to the final assigned marker genotype. Het stands for heterozygous. **(D)** Elicitor-induced VOC emission, as measured by the sum of caryophyllene, α-bergamotene, β-farnesene, 16 hours after treatment of B73, b0005, b0024 and b0067 leaves with water, 5µM ZmPep3 or 5µM Gln-18:3 along the **(E)** new IBM genetic marker map of the FAC sensitivity locus. **(F)** The relative location of 19 genes present in the refined FAC sensitivity locus, LRR-RK gene highlighted in green. Coordinates and gene number are based on the B73v4 annotation. For all treatments shown n=3 and error bars represent SEM. Student t-tests (two-tailed distribution, unpaired), was used for detection of significant differences. Asterisks indicate significant difference with P < 0.05; ns, not significant.

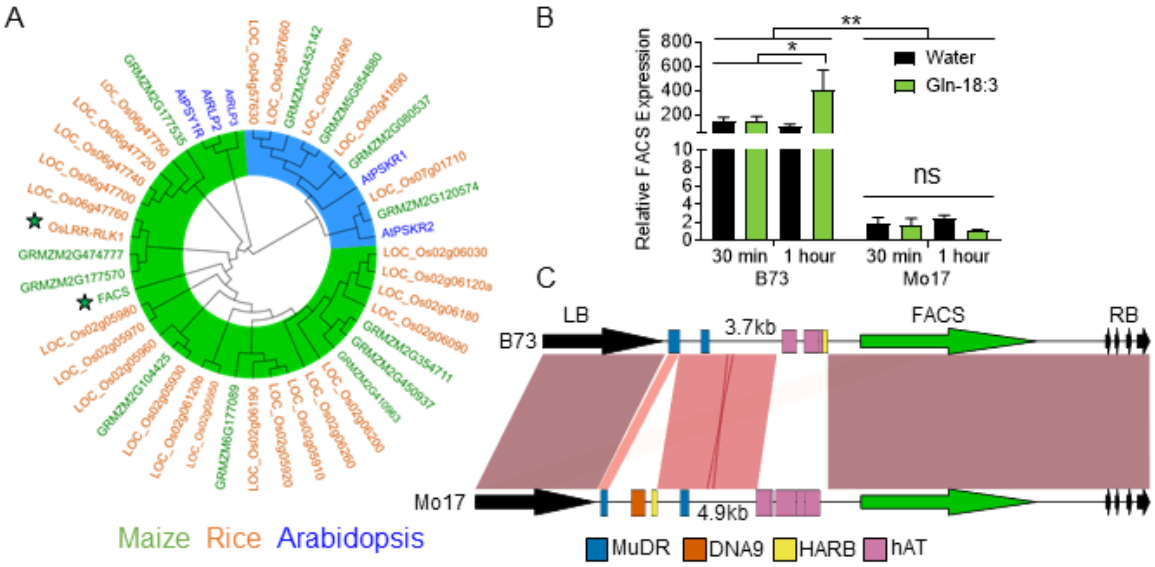


Figure 1-7. A candidate genes within the FAC sensitivity locus denoted FACS is an ortholog of *Oryza sativa* OsLRR-RLK1 necessary for anti-herbivory responses. (A) Phylogenetic tree generated from the protein sequences of genes related to FACS and OsLRR-RLK1 from *Arabidopsis thaliana* (blue font), *Oryza sativa* (orange font) and *Zea mays* (green font). Based on similarity to the *A. thaliana* genes, the phylogeny tree reveals two clades of AtPSKR1/2-related genes (highlighted in blue) and AtPSY1R-related genes (highlighted in green). FACS and OsLRR-RLK1 are designated by a green star. **(B)** Relative expression of FACS in leaf 5 of the maize B73 and Mo17 inbred lines treated with either water or Gln-18:3 for 0.5 hour and 1 hour determined by quantitative RT-PCR in comparison with the expression of the control gene Rpl17. **(C)** Comparative promoter sequence similarity visualization of the FACS (blue arrows) in B73 and Mo17 including the neighboring genes upstream (LB – Left border) and downstream (RB – Right border) from FACS for reference. The 3.7kb and 4.9kb sequences represent the genomic distance between the FACS start codon and the LB stop codon and were used for transposable element prediction using the Poaceae database at the Repbase Censor to identify a number of transposable element fragments from the MuDR family (blue), DNA9 (orange), Harbinger (HARB; yellow) and hAT (purple) families. For all treatments shown n=3 and error bars represent SEM. Student t-tests (two-tailed distribution, unpaired), was used for detection of significant differences. Asterisks indicate significant difference with $P < 0.05$; ns, not significant.

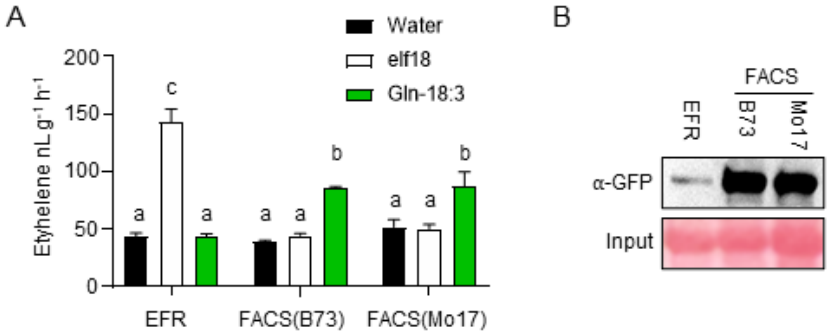


Figure 1-8. Transient expression in *N. benthamiana* of the FACS candidate gene cloned from B73 and Mo17 increases responsiveness to Gln-18:3.
(A) Analysis ethylene emission after 2 hours treatment with water, 1μM elf18 and 1μM Gln-18:3 in *N. benthamiana* leaves expressing EFR and FACS cloned from B73 and Mo17. **(B)** Detection of EFR and FACS cloned from B73 and Mo17 protein expression with α-GFP using a western blot. For all treatments shown n=4 and error bars represent SEM. Different letters represent significant differences (All ANOVAs followed by Tukey honestly significant difference (HSD), α = 0.05).

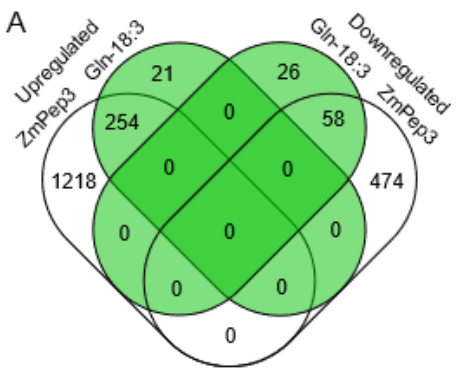


Figure 1-S1. Comparison of ZmPep3- and Gln-18:3-induced DEGs
(A) Venn diagram comparison of the upregulated and downregulated DEGs following ZmPep3 and Gln-18:3 treatments found using RNA-seq data after 2 hour treatment of maize B73 leaves. The elicitor-specific sets of upregulated and downregulated DEGs were used to identify enriched GO terms and Mapman bins using the hypergeometric test.

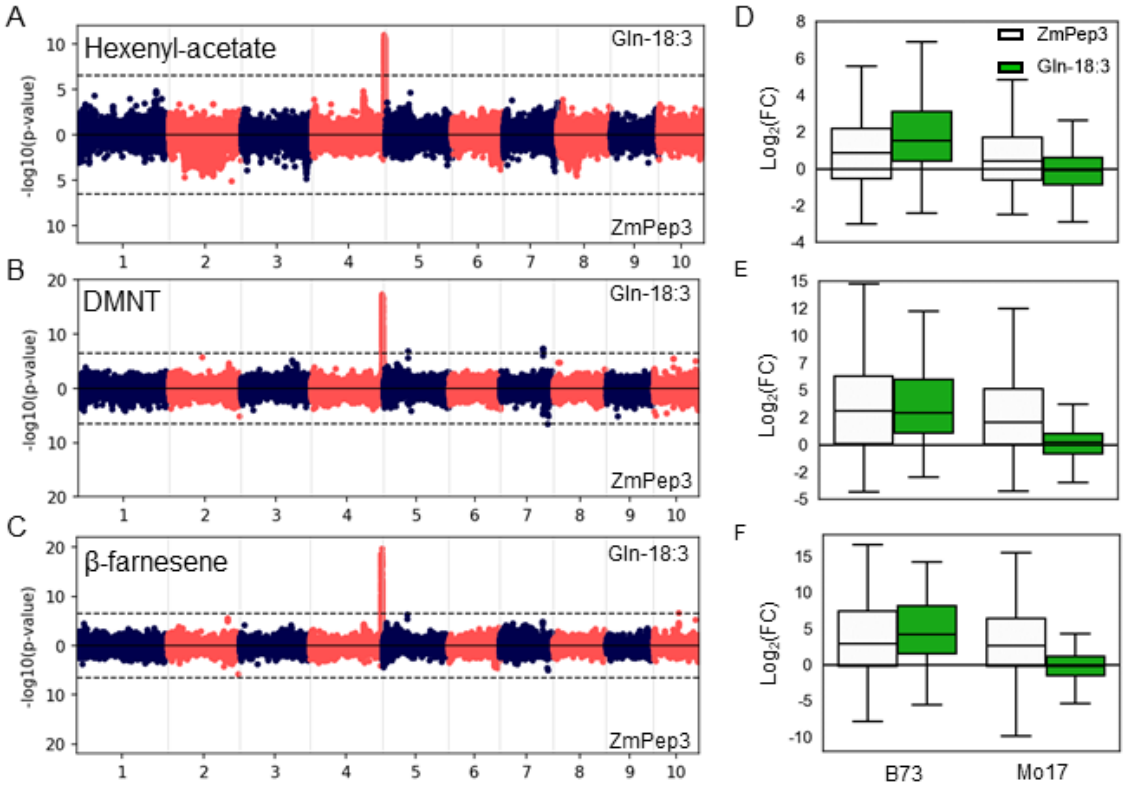


Figure 1-S2. Genetic mapping sensitivity locus using the fold-change (FC) values of VOCs from three different biosynthetic pathways.

QTL mapping using the imputed SNP marker map obtained from the Panzea project for the IBM mapping population with fold-change values from leaf 5 of 242 RILs treated with 1μM ZmPep3- and 1μM Gln-18:3-induced VOC in comparison to water treatment 16 hours after treatment. Fold-change values were calculated for individual VOCs from different biosynthetic pathways, including **(A)** hexenyl acetate, **(B)** DMNT and **(C)** β-farnesene. Box plots showing the distribution of elicitor-induced fold-change VOC emission at the top SNP associated with FAC sensitivity across the 242 IBM RILs, including **(D)** hexenyl acetate, **(E)** DMNT and **(F)** β-farnesene. Dashed lines in manhattan plots indicate significance cutoff of $P < 0.05$ based on Bonferroni correction.

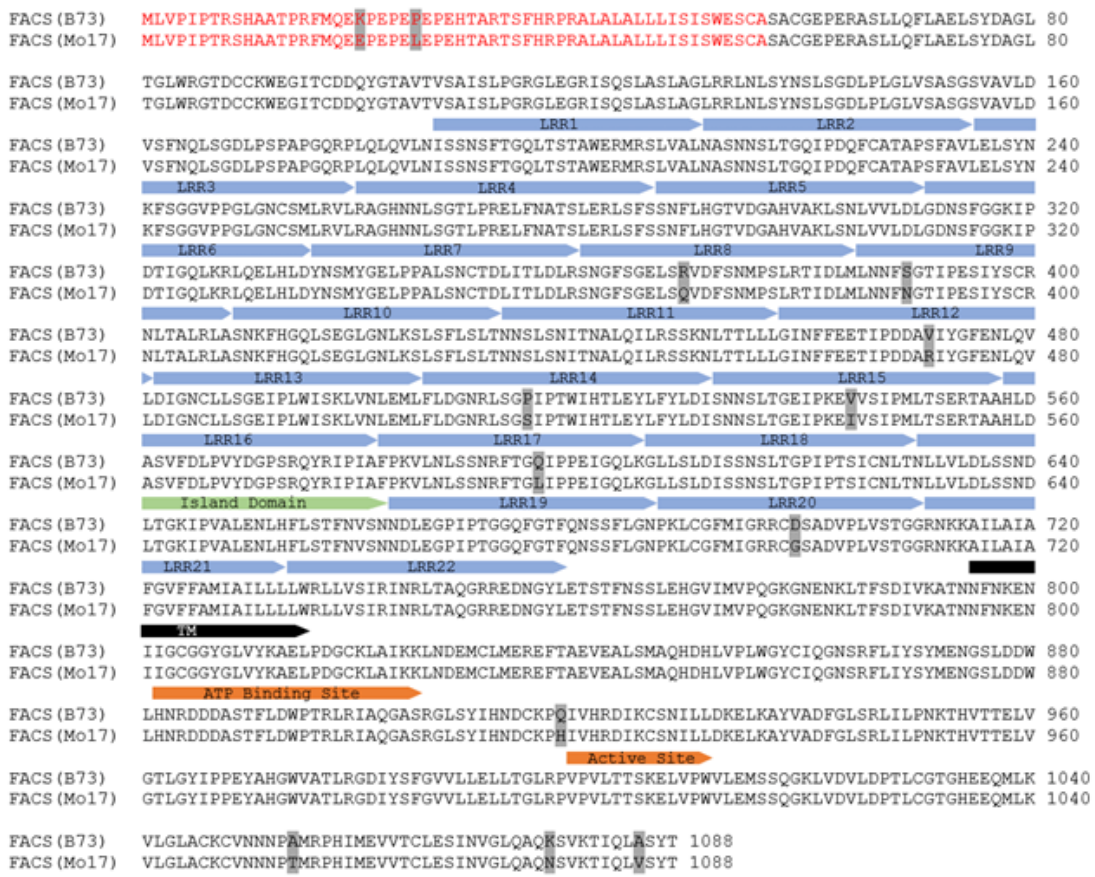


Figure 1-S3. Alignment of the encoded amino acid FACS sequences of B73 and Mo17

The amino acid sequences for FACS(B73) (Zm00001d053867) and FACS(Mo17) (Zm00014a001621) were aligned using Clustal Omega and visualized by highlighting in grey all matching amino-acids. The signal peptide, colored red, was predicted for both FACS(B73) and FACS(Mo17) using Signal-3L. The predicted LRR sequences, with a single predicted island domain between them, were manually annotated based on the consensus LRR motif sequence and labeled using a blue arrow below the matching LRR sequences. Transmembrane domain was predicted using Phobius and labeled using a black arrow below the matching sequence. Two predicted protein domains by InerPro using the amino-acid of FACS(B73) were included, the ATP binding site (IPR017441) and the Serine/threonine-protein kinase active site (IPR008271) and labeled using an orange arrows below the matching sequence.

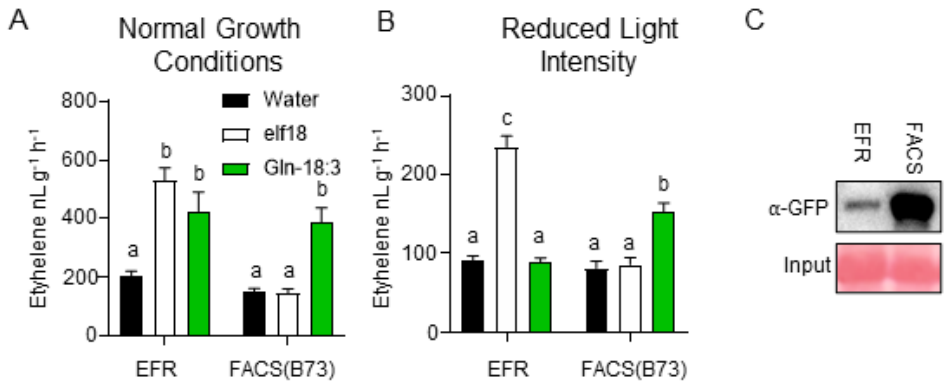


Figure 1-S4. Native Gln-18:3 sensitivity in *N. benthamiana* is reduced in plants grown under diminished light intensity but heterologous expression of FACS increases responsiveness to Gln-18:3.

(A) Analysis ethylene emission after 2 hours treatment with water, 1μM elf18 and 1μM Gln-18:3 in *N. benthamiana* leaves expressing EFR and FACS cloned from B73. Plants were grown under normal growth conditions before and after infiltration. **(B)** Analysis ethylene emission after 2 hours treatment with water, 1μM elf18 and 1μM Gln-18:3 in *N. benthamiana* leaves expressing EFR and FACS cloned from B73. Plants were grown under diminished light intensity before infiltration and left in darkness after infiltration for two days until treatment. **(C)** Detection of EFR and FACS cloned from B73 and Mo17 protein expression with α-GFP using a western blot. For all treatments shown n=4 and error bars represent SEM. Different letters represent significant differences (All ANOVAs followed by Tukey honestly significant difference (HSD), α = 0.05).

Chapter 2: Differential activities of maize plant elicitor peptides as mediators of immune signaling and herbivore resistance

the plant journal

The Plant Journal (2020) doi: 10.1111/tpj.15022

Differential activities of maize plant elicitor peptides as mediators of immune signaling and herbivore resistance

Elly Poretsky[1], Keini Dressano[1], Philipp Weckwerth[1], Miguel Ruiz[1], Si Nian Char[2], Da Shi[3], Ruben Abagyan[3], Bing Yang[2] and Alisa Huffaker[1,*]

[1]*Division of Biology, University of California San Diego, La Jolla, CA, USA,*
[2]*Division of Plant Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, USA, and*
[3]*Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA*

Received 16 April 2020; revised 13 August 2020; accepted 1 September 2020.
*For correspondence (e-mail ahuffaker@ucsd.edu).

SUMMARY

Plant elicitor peptides (Peps) are conserved regulators of defense responses and models for the study of damage-associated molecular pattern-induced immunity. Although present as multigene families in most species, the functional relevance of these multigene families remains largely undefined. While Arabidopsis Peps appear largely redundant in function, previous work examining Pep-induced responses in maize (Zm) implied specificity of function. To better define the function of individual ZmPeps and their cognate receptors (ZmPEPRs), activities were examined by assessing changes in defense-associated phytohormones, specialized metabolites and global gene expression patterns, in combination with heterologous expression assays and analyses of CRISPR/Cas9-generated knockout plants. Beyond simply delineating individual ZmPep and ZmPEPR activities, these experiments led to a number of new insights into Pep signaling mechanisms. ZmPROPEP and other poaceous precursors were found to contain multiple active Peps, a phenomenon not previously observed for this family. In all, seven new ZmPeps were identified and the peptides were found to have specific activities defined by the relative magnitude of their response output rather than by uniqueness. A striking correlation was observed between individual ZmPep-elicited changes in levels of jasmonic acid and ethylene and the magnitude of induced defense responses, indicating that ZmPeps may collectively regulate immune output through rheostat-like tuning of phytohormone levels. Peptide structure–function studies and ligand–receptor modeling revealed structural features critical to the function of ZmPeps and led to the identification of ZmPep5a as a potential antagonist peptide able to competitively inhibit the activity of other ZmPeps, a regulatory mechanism not previously observed for this family.

Keywords: *Zea mays*, peptide signaling, transcriptional response, signaling and hormones, plant–herbivore interactions, protein–protein interactions.

INTRODUCTION

Like all living organisms plants must protect themselves from attacking entities, and they have evolved sophisticated surveillance and signaling mechanisms to recognize and ward off attack. Molecular patterns from invading microbes, herbivores and parasites are perceived by pattern recognition receptors to activate immune responses involving kinase signaling and integral second messengers such as reactive oxygen species (ROS) and phytohormones (Zipfel, 2014; Hegenauer *et al.*, 2016; Mendy *et al.*, 2017; Qi *et al.*, 2017; Yu *et al.*, 2017b; Steinbrenner *et al.*, 2019). A myriad of endogenous molecules produced as a consequence of cellular damage have also emerged as critical signals to promote defense responses, including extracellular ATP, pyridine, glutamate, cell wall fragments and peptide hormones (McGurl *et al.*, 1992; Huffaker *et al.*, 2006; Mousavi *et al.*, 2013; Tanaka *et al.*, 2014; Toyota *et al.*, 2018; Wang *et al.*, 2019). Since the initial discovery that systemin is a regulator of solanaceous anti-herbivore defenses a large number of peptide families have been found to amplify plant immune responses (McGurl *et al.*, 1992; Yamaguchi and Huffaker, 2011; Ma *et al.*, 2014; Yu *et al.*, 2017b; Segonzac and Monaghan, 2019). While many of these peptides have only been identified in particular species, plant elicitor peptides (Peps) have emerged as widespread conserved signals across plant species, activating broadly protective immunity against microbial

pathogens, parasitic nematodes and arthropod herbivores (Huffaker *et al.*, 2011a, 2013; Trivilin *et al.*, 2014; Lee *et al.*, 2018; Ruiz *et al.*, 2018b; Shinya *et al.*, 2018; Zhang and Gleason, 2020).

Originally isolated biochemically from Arabidopsis, the 23-amino-acid peptide AtPep1 was found to be derived from the carboxyl terminus of a larger 92-amino-acid precursor protein, AtPROPEP1 (Huffaker *et al.*, 2006). Through comparisons of sequence homology, seven other Arabidopsis genes have been identified as *AtPROPEP1*-like, encoding precursors ranging in size from 80 to 109 amino acids. Although there is little direct amino acid identity across the eight precursors, each contains characteristic glutamate/lysine (EKE)-enriched motifs and a single AtPep1-like putative peptide signal domain in the carboxyl terminus region (Huffaker *et al.*, 2006). When synthesized, these putative peptide signals all demonstrated elicitor activity (Huffaker *et al.*, 2006; Huffaker and Ryan, 2007; Yamaguchi *et al.*, 2006, 2010; Bartels *et al.*, 2013). A search for homologous PROPEP sequences across plant species revealed many predicted orthologues, all containing 'EKE' motifs and putative bioactive peptides (Huffaker *et al.*, 2006, 2011a, 2013; Lori *et al.*, 2015). Although homology searches predicted PROPEP precursors, the putative active Pep sequences are widely variable, with only Gly17 being perfectly conserved, along with a His or Asn residue at the carboxyl terminus and an enrichment of Ser and Gly residues in the carboxyl half of the peptide (Huffaker *et al.*, 2006, 2013; Lori *et al.*, 2015). Despite this sequence variability, predicted PROPEPs/Peps have been demonstrated as functionally orthologous to Arabidopsis, regulating defense responses in fabaceous, rosaceous, solanaceous and poaceous plants (Huffaker *et al.*, 2011a, 2013; Trivilin *et al.*, 2014; Lee *et al.*, 2018; Ruiz *et al.*, 2018a; Shinya *et al.*, 2018; Zhang and Gleason, 2020).

Plant elicitor peptide signaling has been largely characterized in the model plant *Arabidopsis thaliana*. Arabidopsis Peps are liberated from PROPEP precursors in damaged cells through proteolytic processing by calcium-dependent metacaspases (Huffaker *et al.*, 2006; Huffaker and Ryan, 2007; Bartels *et al.*, 2013; Hander *et al.*, 2019; Shen *et al.*, 2019). Upon liberation from their precursors, AtPeps interact with two leucine-rich receptor kinases, PEP RECEPTOR1 and -2 (AtPEPR1 and AtPEPR2), which activate downstream signaling in complex with SOMATIC EMBRYO RECEPTOR KINASE (SERK) co-receptors and the membrane-associated cytoplasmic kinases BOTRYTIS-INDUCED KINASE 1 (BIK1) and BRASSINOSTEROID-SIGNALING KINASE 5 (BSK5) (Yamaguchi *et al.*, 2006, 2010; Krol *et al.*, 2010; Lu *et al.*, 2010; Postel *et al.*, 2010; Roux *et al.*, 2011; Liu *et al.*, 2013; Majhi *et al.*, 2019). Upon activation of the receptor complex, signaling occurs through the mitogen-activated protein (MAP) kinases MPK3/6/4/11, production of the phytohormones jasmonic acid (JA), ethylene (ET) and salicylic acid (SA) and of ROS generated by NADPH oxidase enzymes (Huffaker and Ryan, 2007; Bartels *et al.*, 2013; Huffaker *et al.*, 2013; Tintor *et al.*, 2013; Dressano *et al.*, 2019). This is accompanied by modulation of ion channels and transporters that increases calcium levels in the cytoplasm and reduces the export of protons to the apoplast (Huffaker *et al.*, 2006; Ma *et al.*, 2014). Together these activities alter patterns of gene expression and transcriptional splicing to activate broad anti-pathogen defense responses (Huffaker *et al.*, 2006; Ma *et al.*, 2014; Dressano *et al.*, 2019). Arabidopsis Peps have also been implicated in regulating abiotic stress responses, including starvation and enhanced tolerance of high-salt growth conditions (Gully *et al.*, 2015; Nakaminami *et al.*, 2018).

Characterization of expression patterns and functions of individual AtPeps and AtPEPRs has revealed largely overlapping activities (Huffaker and Ryan, 2007; Bartels *et al.*, 2013). When applied to leaves, all AtPeps promote MAP kinase activation, ET emission and accumulation of ROS, which has led to the conclusion that they may be largely redundant, although inducible expression of the precursor genes encoding PROPEP1, -2 and -3 in response to immune challenge indicates that these may be the dominant immunoregulatory signals (Huffaker and Ryan, 2007; Bartels *et al.*, 2013). Arabidopsis PEPR1 has been shown to bind AtPep1 6, and study of knockout lines indicates that it is the primary regulator of foliar AtPep signaling. Arabidopsis PEPR2 interacts only with AtPep1 and AtPep2, and while it plays a lesser role in aerial tissues it is strongly expressed in roots, where expression of 75% of AtPep1-modulated genes is fully dependent on AtPEPR2 (Yamaguchi *et al.*, 2010; Ma *et al.*, 2014).

Although AtPeps have thus far been observed to have largely redundant activities, studies in maize (*Zea mays*) indicate that this might not be the case for Peps in all species. Of the five ZmPeps identified previously, ZmPep3 was found to potently induce emission of herbivore-associated volatiles that serve as an indirect defense against lepidopteran pest insects by attracting *Cotesia marginiventris* wasps that parasitize *Spodoptera* species (Huffaker *et al.*, 2013, Huffaker, 2015). In contrast, ZmPep5 had no volatile-inducing activity, and the other three ZmPeps activated emission of varying levels in between. This indicated that although Peps are conserved regulators of immune responses across plant species there may be regulatory differences among species that are not discoverable in a single model system. To better understand the potential specificity and functional roles of ZmPeps and ZmPEPRs, we investigated their regulatory activities. In addition to defining function, this study revealed several surprising findings not previously seen in the Pep family, including precursors containing multiple active Peps, antagonist peptide activity and control of the magnitude of defense output through fine-tuned modulation of phytohormone levels.

RESULTS

Poaceous plants encode PROPEP precursors containing multiple Pep sequences

While five ZmPROPEP genes had previously been identified in maize through BLAST homology searches of the B73 genome sequence, improved genome annotation has revealed six genes encoding intact *ZmPROPEP* precursors and several pseudogenes (Figure 1a,b). The sixth PROPEP gene identified is a duplication of *ZmPROPEP4*, with the gene formerly referred to as *ZmPROPEP4*, now termed *ZmPROPEP4.1*, and the duplication termed *ZmPROPEP4.2* (Figure 1a,b). B73 also encodes three additional *ZmPROPEP4.1*-like pseudogenes, indicating a potentially high rate of gene duplication (Figure S1 in the online Supporting Information, Tables S1 and S2). The ZmPROPEP genes encode proteins larger than Arabidopsis PROPEPs, with maize precursors ranging between 142 and 199 amino acids in length. As in Arabidopsis, PROPEP genes share very little direct sequence identity and are primarily arrayed in clusters. *ZmPROPEP1* and *ZmPROPEP3* are tandemly encoded on chromosome 2. *ZmPROPEP4.1*, *ZmPROPEP4.2* and *ZmPROPEP5* are clustered together on chromosome 10 along with the truncated *ZmPROPEP4*-like pseudogenes. *ZmPROPEP2* is alone at a distal location on chromosome 10. Similar to the gene structure of AtPROPEP, some ZmPROPEP genes contain a single intron, including *ZmPROPEP1*, *ZmPROPEP2*, *ZmPROPEP3* and *ZmPROPEP5*, whereas the others lack introns, namely *ZmPROPEP4.1* and *ZmPROPEP4.2*. To understand whether the B73 gene family structure is representative of other maize varieties, we examined these loci in the newly released genome sequence for the nested-association mapping (NAM) parent lines as well as W22, Mo17, EP1, F7 and the *Z. mays* ssp. *mexicana* Zx-PI566673 (Portwood *et al.*, 2019). Aside from *ZmPROPEP4* genes, all PROPEP genes are present as single copies in each NAM parent, with the single exception of a *ZmPROPEP3* duplication in NC358 (Figure S1, Tables S1 and S2). The duplications of *ZmPROPEP4* observed in B73 occur in all NAM parent lines, with total duplications (including pseudogenes) ranging between two and six across inbreds, which we now term *ZmPROPEP4.1* through *ZmPROPEP4.6*. The majority of NAM parent lines contain three copies of *ZmPROPEP4* predicted to be functional, but individual inbreds can vary from a single copy (CML52, CML228) to five full copies (Mo17, B97, CML103).

A notable difference between maize PROPEP and Arabidopsis PROPEP genes is that several maize PROPEP genes appear to encode multiple potentially bioactive peptide signals. Prediction of bioactive peptide sequences within each precursor using BLAST homology searching and seeking motifs that match ZmPeps already demonstrated to be active (Huffaker *et al.*, 2011a, 2013) found that *ZmPROPEP4.1* and *ZmPROPEP4.2* each contained four putative peptide sequences whereas *ZmPROPEP5* contained two. Because Peps have previously been identified almost solely from the carboxyl terminus of PROPEP precursors, the peptide sequences from the carboxyl terminus of precursors containing multiple peptide sequences have been denoted as the 'a' peptides, for example ZmPep4.1a. Working backwards towards the amino terminus, subsequent predicted peptide sequences are designated by sequential letters: ZmPep4.1b, ZmPep4.1c and ZmPep4.1d (Figure 1b). Together, the six maize precursor genes contain a total of 13 ZmPeps (Figure 1a,b). Due to the fact that *ZmPROPEP4.2* shares 95.5% overall nucleotide sequence identity with *ZmPROPEP4.1*, ZmPep4.1a and ZmPep4.2a are identical in sequence and would be indistinguishable from one another once released from their respective precursors, leaving only 12 unique ZmPeps. The other *ZmPROPEP4.1*- and *ZmPROPEP4.2*-containing peptides are not identical but are highly similar: ZmPep4.2b has one conservative (ΔA4 to V) and one non-conservative substitution (ΔP6 to R) in the amino terminus of the peptide compared with ZmPep4.1b, whereas ZmPep4.2c has a single change, ΔH1 to T (all designated by underlined residues in the ZmPep4.2 sequences in Figure 1a). ZmPep4.2d contains one non-conservative (ΔR2 to W) and one conservative substitution (ΔK15 to R). To understand the conservation of predicted peptide sequences in maize varieties, each was examined across a set of 31 lines (Figure S2). The amino acid sequences of ZmPep1 and ZmPep5a are perfectly conserved across all lines. For ZmPep2 there is variation only in the first three amino acids, whereas for ZmPep3, ZmPep4a and ZmPep4b peptides the carboxyl terminus is conserved but the amino terminus has minor variation. ZmPep4c and ZmPep4d demonstrated more variation, but remain identical for the final seven amino acids. ZmPep5b is the only peptide with variable amino acids in the carboxyl terminus region.

Intriguingly, the propensity to encode PROPEPs containing multiple Pep sequences in a single precursor seems to be ubiquitous throughout poaceous plants. Of 42 precursors collectively identified in Phytozome databases, *Brachypodium distachyon*, *Oryza sativa*, *Panicum virgatum*, *Setaria italica*, *Sorghum bicolor* and *Z. mays*, 20 contained multiple putative bioactive Pep sequences (Figure 1b) (Goodstein *et al.*, 2012). These compound PROPEP precursors were disseminated across all species examined, and included between two and four Pep sequences: *B. distachyon* encodes four precursors with a total of five Peps, *O. sativa* encodes eight PROPEPs and a total of 15 Peps, *P. virgatum* contains eight precursors with 19 Peps, *S. bicolor* encodes eight PROPEPs and a total of 15 Peps and *S. italica* contains eight precursors with 12 Peps. A consensus motif generated from the 82 poaceous Pep sequences from these six species reveals a density of

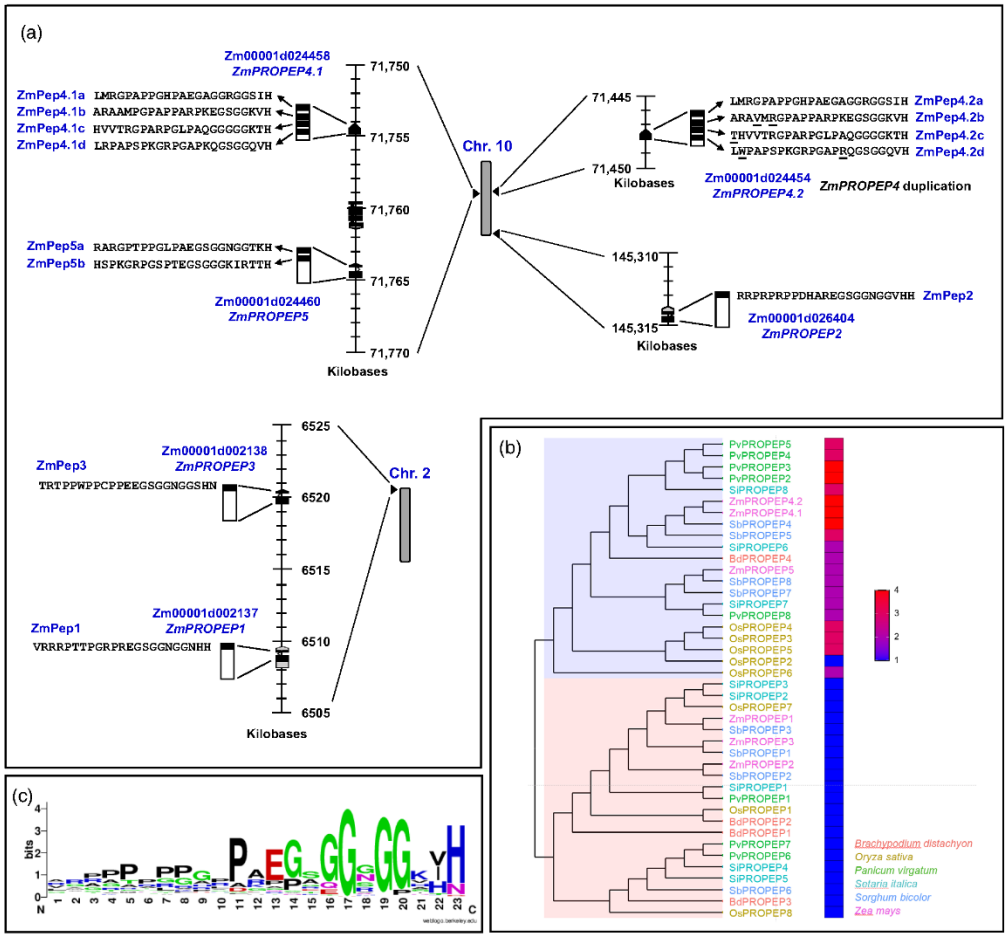


Figure 1. Location and structure of maize *PROPEP* genes.
(a) Location of *ZmPROPEP1* and *ZmPROPEP3* genes on maize B73 v4 chromosome 2, with the peptide-containing portion of each encoded protein indicated by black bars along with the amino acid sequence for each peptide and the location of *ZmPROPEP2*, *ZmPROPEP4.1*, *ZmPROPEP4.2* and *ZmPROPEP5* on chromosome 10. *ZmPROPEP4.1*, *ZmPROPEP4.2* and *ZmPROPEP5* each encode multiple peptides, with peptides most proximal to the carboxyl terminus designated 'a' peptides, followed, by b, c and d.
(b) Phylogenetic tree representing all identified PROPEPs in six different poaceous plants. The number of predicted peptides per precursor is represented by heat map values. Single- and multiple-plant elicitor peptide (Pep) PROPEPs separate into two distinctive clades.
(c) Logo file representing the consensus sequence for Peps from poaceous plants generated using the six species shown in the phylogenetic tree.

proline residues in the amino end of the peptide (Figure 1c). This is an interesting contrast to the high density of basic residues seen in the amino terminus of AtPeps (Huffaker *et al.*, 2006). Notably, of all amino acids, only Gly17 is perfectly conserved, a conservation that has been previously reported and demonstrated to be critical for both AtPep1 elicitor activity and association with the AtPEPR1 receptor (Pearce *et al.*, 2008; Tang *et al.*, 2015).

The magnitude of ZmPep-induced metabolic defense output correlates with relative phytohormone concentrations

While ZmPep1 and ZmPep3 have been found to have protective effects against pathogens and herbivores, the global activity of all ZmPeps has remained unexamined (Huffaker *et al.*, 2011a, 2013). To address this knowledge

gap, the activity of individual peptides was assessed. Moreover, given that ZmPROPEP4 and ZmPROPEP5 may simultaneously release all peptides contained in the precursor upon proteolytic processing, combinatorial treatments of these peptides (e.g. ZmPep4.1 All and ZmPep5 All) were also examined. In each case 5 nmol of peptide per leaf was applied, with combinatorial treatments composed of equimolar quantities of each peptide (1.25 nmol for each ZmPep4.1, and 2.5 nmol for each ZmPep5 peptide). As demonstrated previously, ZmPep3 is most potent in inducing herbivory-associated volatile terpene production, promoting increases of approximately 200-fold relative to leaves supplied with water (Figure 2a). Treatment with ZmPep1, ZmPep4.1a and the combined

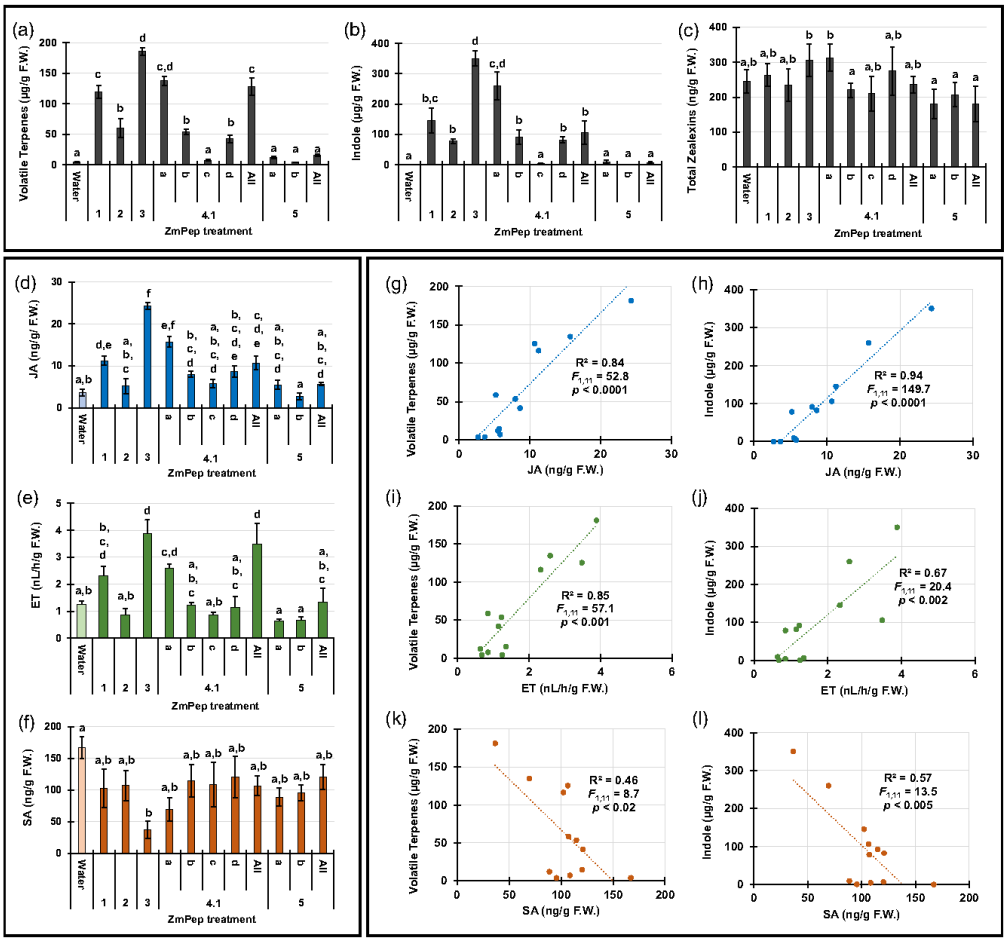


Figure 2. The strength of *Zea mays* plant elicitor peptide (Pep) elicitor activity corresponds to the relative magnitude of induced changes in phytohormone levels. Analysis of relative levels of defense hormones and metabolites in 21-day-old maize leaves post-treatment with either water or a 5 μM of ZmPeps. (a)–(c) Relative levels of defense metabolites extracted from leaf tissue: (a) indole levels; (b) total herbivore-associated volatile terpenes, defined as β-caryophyllene, *E*-β-farnesene, linalool, dimethylnonatriene (DMNT), trimethyltridecatetraene (TMTT) and α-bergamotene; (c) total zealexins. (d)–(f) Relative levels of defense-associated phytohormones post-treatment with either water or a 5 μM solution of ZmPeps: (d) jasmonic acid (JA) levels extracted after 12 h; (e) ethylene (ET) emitted from leaves after 2 h; (f) salicylic acid (SA) levels extracted after 12 h. (g)–(l) Relationship between relative levels of phytohormones and defense metabolites for each sample group: (g) JA versus indole; (h) JA versus herbivore-associated sesquiterpenes; (i) ET versus indole; (j) ET versus herbivore-associated sesquiterpenes; (k) SA versus indole; (l) SA versus herbivore-associated sesquiterpenes. For all experiments shown $n = 4 \pm$ SEM. Within plots, different letters (a–f) represent significant differences (one-way analysis of variance followed by Tukey's test corrections for multiple comparisons, $P < 0.05$).

The Plant Journal, (2020), doi: 10.1111/tpj.15022

ZmPep4.1 resulted in approximately 75% of the terpene levels observed after ZmPep3 treatment, whereas ZmPep2, ZmPep4.1b and ZmPep4.1d triggered accumulation of approximately one-third the terpene levels of ZmPep3. Significant increases in volatile terpenes were not observed in leaves treated with ZmPep4.1c, ZmPep5a, ZmPep5b or the combined ZmPep5 treatment. Combinatorial treatments with all ZmPep4.1 peptides resulted in the production of terpenes equivalent to the treatment with ZmPep4a, whereas the combination of all peptides from ZmPep5 did not increase terpene levels above those in water-treated leaves. To understand whether the highly similar peptides contained in the duplicated genes *ZmPROPEP4.1* and *ZmPROPEP4.2* were also similar in relative activity, we assessed induced volatile emission activities of the ZmPep4.1 and ZmPep4.2 peptides. No difference in activity was found for ZmPep4.1b versus ZmPep4.2b, or for the ZmPep4c and ZmPep4d peptides (Figure S3), and so for all subsequent experiments ZmPep4.1 peptides were selected for testing as representative of those encoded by *ZmPROPEP4* genes.

Indole, which is both emitted as a volatile priming signal and is a precursor to benzoxazinoid defense metabolites, was found to accumulate in leaves after ZmPep treatment as previously reported (Frey *et al.*, 1997; Huffaker *et al.*, 2013; Erb *et al.*, 2015). Compared with a water control, ZmPep3 was most potent when 5 nmol per leaf was applied, promoting a roughly 350-fold increase in indole accumulation after 12 h (Figure 2b). ZmPep4.1a was similarly active, resulting in an increase of roughly 250-fold, while ZmPep1, ZmPep2, ZmPep4.1b and ZmPep4.1d induced moderate levels of indole accumulation, approximately one-half to one-third the concentrations promoted by ZmPep3. A combined treatment using all four ZmPep4.1 peptides (1.25 nmol each, for a total application of 5 nmol) was also moderately active. In contrast, ZmPep4.1c, ZmPep5a, ZmPep5b and a combined ZmPep5 treatment did not result in significant indole accumulation. In contrast to the effects on indole and herbivory-associated volatile terpenes, ZmPep treatment did not strongly affect levels of zealexin anti-fungal acidic terpenoid defense metabolites (Figure 1c), supporting previous evidence indicating that these metabolic responses are not co-regulated (Huffaker *et al.*, 2011b; Schmelz *et al.*, 2011; Ding *et al.*, 2020).

To understand whether ZmPep-induced changes in relative levels of defensive phytohormones might contribute to the observed effects on indole and sesquiterpenes, production of JA, ET and SA was measured after ZmPep treatment. Strikingly, the pattern for relative ZmPep-induced increases in both JA and ET strongly resembled the patterns observed for both indole and sesquiterpenes. The ZmPep3 treatment caused the greatest accumulation of both hormones, with a fivefold change in JA and fourfold increase in ET compared with water-treated controls (Figure 2d,e). ZmPep1, ZmPep4.1a, and the combined ZmPep4.1 treatment resulted in an increase of approximately 50 60% in JA compared with ZmPep3. These treatments also triggered increased ET emission, with the combinatorial ZmPep4.1 treatment resulting in similar levels of ET emission to ZmPep3, while ZmPep1- and ZmPep4.1a-treated leaves emitted approximately 60% of the ZmPep3-induced levels of ET. Concentrations of JA in leaves treated with ZmPep4.1b and ZmPep4.1d were only moderately increased, to approximately double those of water-treated leaves, whereas JA in leaves treated with ZmPep2, ZmPep4.1c, ZmPep5a, ZmPep5b and combined ZmPep5 peptides did not significantly differ from water-treated controls.

Interestingly, although AtPeps were previously shown to simultaneously stimulate production of both JA and SA in Arabidopsis, ZmPeps appeared to have a potentially suppressive effect on accumulation of SA in maize (Huffaker *et al.*, 2013). Levels of SA in all ZmPep-treated leaves trended lower than levels in water-treated control leaves. Significantly for ZmPep3- and ZmPep4.1a-treated leaves, SA concentrations were reduced by approximately 75% and 60%, respectively, compared with water-treated leaves (Figure 2f). Just as ZmPep3 and ZmPep4.1a were most active in stimulating defense metabolism and JA/ET accumulation and emission, they are also most potent at suppressing SA production.

As predicted, given the similarities in patterns of ZmPep-induced increases in defense metabolites and in JA/ET production, a strong correlative relationship was observed between the concentrations of these phytohormones and the magnitude of defense metabolite accumulation. When concentrations of volatile terpenes and of indole for each treatment were plotted in relationship to JA levels, a linear positive correlation was apparent, with R^2 values of 0.84 ($F_{1,11} = 52.8$, $P < 0.0001$) and 0.94 ($F_{1,11} = 149.7$, $P < 0.0001$), respectively (Figure 2g,h). A similarly tight positive linear correlation was observed for ET emission and terpene accumulation, with R^2 equal to 0.85 ($F_{1,11} = 57.1$, $P < 0.0001$), while ET levels were somewhat less correlated with indole concentrations, having an R^2 of 0.67 ($F_{1,11} = 20.4$, $P < 0.002$) (Figure 2i,j). Correspondingly, levels of JA and ET demonstrate a linear positive correlation (Figure S4). Concentration of SA was loosely inversely correlated with terpene and indole levels, with a respective R^2 for each comparison of 0.46 ($F_{1,11} = 8.7$, $P < 0.02$) and 0.57 ($F_{1,11} = 13.5$, $P < 0.005$) (Figure 2k,l). Together, this evidence indicates that the magnitude of defense metabolite output in response to individual ZmPeps may be largely dependent on the relative increase in JA and ET.

Specificity of ZmPeps is defined by the relative magnitude of response rather than uniqueness

To better ascertain the global roles of each ZmPep, large-scale foliar gene expression changes 12 h post-treatment were profiled in comparison with water-treated controls

using the Agilent 44K maize microarray platform. Differentially expressed genes (DEGs) were defined as an increase or decrease of twofold or more with a false discovery rate (FDR)-corrected *P*-value of less than 0.05 (Table S4). As observed for metabolic outputs, ZmPeps had widely varying activity in stimulating gene expression changes. Each ZmPep treatment promoted distinct changes in expression, but these distinct changes signified varying expression of a single set of genes rather than differential expression of unique sets of genes. An Euler diagram representing upregulated DEGs appears almost as concentric rings, illustrating that less active ZmPeps induced expression of a small portion of the larger set that is differentially expressed in response to more active ZmPeps (Figure 3a). ZmPep3 was most active, triggering increased expression of 412 genes. Treatment with ZmPep1 or a ZmPep4.1 peptide blend triggered a subset of these DEGs, with only seven genes responsive to ZmPep1 or to the ZmPep4.1 blend not found in the ZmPep3 DEG set. ZmPep2 induced upregulation of a smaller subset of DEGs, all of which were also increased in response to the previous treatments. Genes for which expression was downregulated represented somewhat more specificity among treatment groups. ZmPep3 treatment again had the strongest effect, followed by treatment with ZmPep1 and the ZmPep4.1 blend (Figure 3b). Among the genes whose expression was suppressed by ZmPep1 and ZmPep4.1, 61 were distinct from the ZmPep3-responsive DEGs. ZmPep2 caused decreased expression for only 10 genes, all of which were also downregulated by ZmPep1, ZmPep3 and the ZmPep4.1 blend. Very few genes were differentially expressed in leaves treated with ZmPep5 peptides (Figure 3a,b).

To better visualize genes that were shared or unique among treatment groups, Venn webtools was used to identify the intersections between the DEGs in every treatment group. The intersections with the greatest number of genes were assigned into tiers representing the top five intersections for upregulated genes and the top four intersections for downregulated genes (Figure 3c, Table S5). Genes not contained in these tiers were summed and added to the unassigned group. As the most active peptide, ZmPep3 promoted expression of all five tiers of upregulated genes, whereas treatment with ZmPep4.1a resulted in expression of only four of the five tiers. In descending order of activity, ZmPep1 treatments activated expression of the lower three tiers, the ZmPep4.1 blend promoted expression of the lowest two tiers and ZmPep2 and ZmPep4.1b treatment activated only the lowest tier. The downregulated tiers were similarly expressed, with ZmPep3 promoting expression of all four tiers of downregulated genes. Interestingly, the combined ZmPep4.1 blend was the next most active at suppressing gene expression, resulting in downregulation of three of the four tiers. ZmPep1 downregulated expression of the genes in the lowest two tiers, and ZmPep4.1a affected only the lowest tier. To better understand the function of these gene tiers, AgriGO was used to calculate the Plant Gene Ontology (GO) Slim enrichment for DEGs within the tiers (Figure 3d, Table S6). A combination of all five tiers of upregulated genes was significantly enriched for response to biotic stimulus, response to extracellular stimulus, multi-organism process, metabolic process and carbohydrate metabolic process. Tiers 1:4 and 1:5, which included only genes upregulated by treatment with the most potent ZmPeps, were enriched specifically in cell death, catabolic processes and cell communication. For tiers representing genes with downregulated expression, all tiers were most enriched in photosynthesis and generation of precursor metabolites and energy, response to stress and response to abiotic stimulus.

Together, metabolic and gene expression profiling lead to similar conclusions: that there is a spectrum of activity among the ZmPeps, with ZmPep3, ZmPep1, ZmPep4.1a and combined ZmPep4.1 treatments acting as strong elicitors of immune responses, ZmPep5 peptides and ZmPep4.1c displaying little activity, and other ZmPeps having intermediate activity. Furthermore, while ZmPeps have differential activities, these differences are defined by the magnitude of defense output that each ZmPep promotes rather than a specificity in downstream responses distinct to each peptide. As seen for specialized defense metabolites, the relationship between the magnitude of changes in gene expression caused by treatment with any ZmPep correlates significantly with relative phytohormone levels. The number of genes increased or decreased in expression after ZmPep treatments showed a tight positive correlation with the relative levels of JA and ET produced (Figure 3e h). Conversely, the number of genes increased or decreased after ZmPep treatment negatively correlated with relative SA levels (Figure 3i,j). To test the requirement for ET biosynthesis in ZmPep-induced expression of defense genes, we examined defense marker gene expression in an ET-deficient line, a double mutant in ACC synthases 2 and 6 (*acs2*/*acs6*) (Young *et al.*, 2004; Zhou *et al.*, 2019). ZmPep-induced expression of genes encoding a serine protease inhibitor (*SerPIN*) and a terpene synthase responsible for the volatile sesquiterpene alpha-bergamotene (*TPS10*) was dramatically reduced in the *acs2*/*acs6* mutant line (Figure 3k,l), supporting a critical role for ET in ZmPep-regulated defenses.

Carboxyl terminus residues are essential for ZmPep activity

The differential activities of ZmPeps seem to be a function of their relative potency rather than specific regulation of distinct non-overlapping responses. Given that ZmPep3

The Plant Journal, (2020), doi: 10.1111/tpj.15022

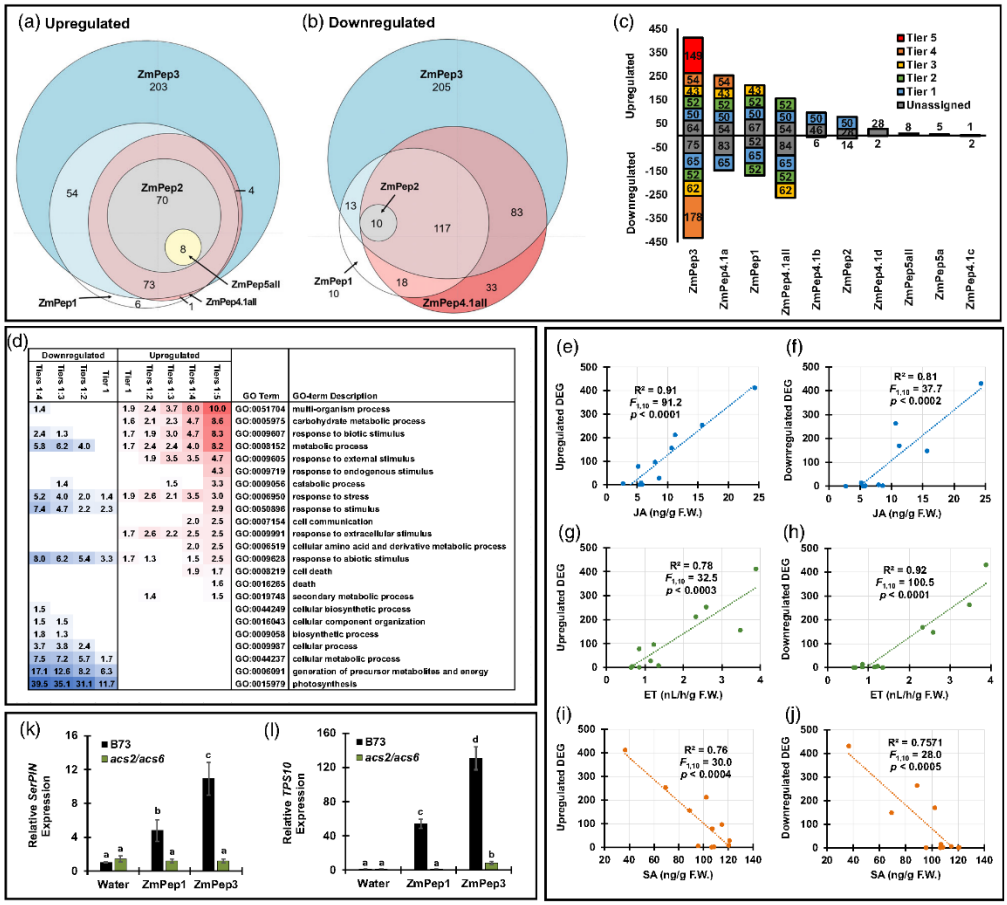


Figure 3. Gene expression changes elicited by *Zea mays* plant elicitor peptides (ZmPeps) correspond to the magnitude of induced changes in phytohormone levels and requires ethylene biosynthesis.

(a), (b) Euler diagrams representing the overlap in genes with increased (a) or decreased (b) expression in 21-day old leaves 12 h post-treatment with each ZmPep compared with water-treated controls as determined using the Agilent 44K maize microarray platform. (c) Differentially expressed genes (DEGs) categorized into tiers based on calculation of intersections between DEGs in each treatment using Venn webtools. Intersections with the greatest number of genes were assigned into tiers, and represent the top five intersections for upregulated genes and the top four intersections for downregulated genes. The number of genes represented in each tier is designated. (d) Enriched Gene Ontology (GO) terms for tiered ZmPep-responsive genes. Numbers represent the fold-enrichment over background, downregulated DEGs are represented in blue, upregulated in red. (e)–(j) Relationship between relative levels of phytohormones and number of upregulated or downregulated DEG for each sample group: (e) jasmonic acid (JA) versus upregulated DEGs; (f) JA versus downregulated DEGs; (g) ethylene (ET) versus upregulated DEGs; (h) ET versus downregulated DEGs; (i) salicylic acid (SA) versus upregulated DEGs; (j) SA versus downregulated DEGs. (k), (l) Relative expression of the defense marker gene *Serine Protease Inhibitor*(*SerPIN*) (k) and *Terpene Synthase 10*(*TPS10*) (l) 12 h post-treatment with water or ZmPeps in wild-type plants and the ET-deficient mutant plants *acs2*/*acs6*. Values represent the fold change in expression versus the water-treated wild-type (B73) samples after normalization against expression of the housekeeping gene *ZmRPL17*. Within plots, different letters (a–d) represent significant differences (one-way analysis of variance followed by Tukey's test corrections for multiple comparisons, $P < 0.05$). For all experiments shown, $n = 4 \pm$ SEM. F.W., fresh weight.

was found to be the most potent of the ZmPeps, structural properties contributing to its function were examined. Using a series of alanine-substituted and truncated ZmPep3 variants (Figure 4a), relative potency was explored using emission of herbivory-associated volatiles as a measurable output after treatment with 125 pmol solutions. The first eight amino acids of ZmPep3 were found to be non-essential for function, as truncated variants in which amino terminus residues were removed had activity similar to full-length ZmPep3 (Figure 4b). Truncation of the

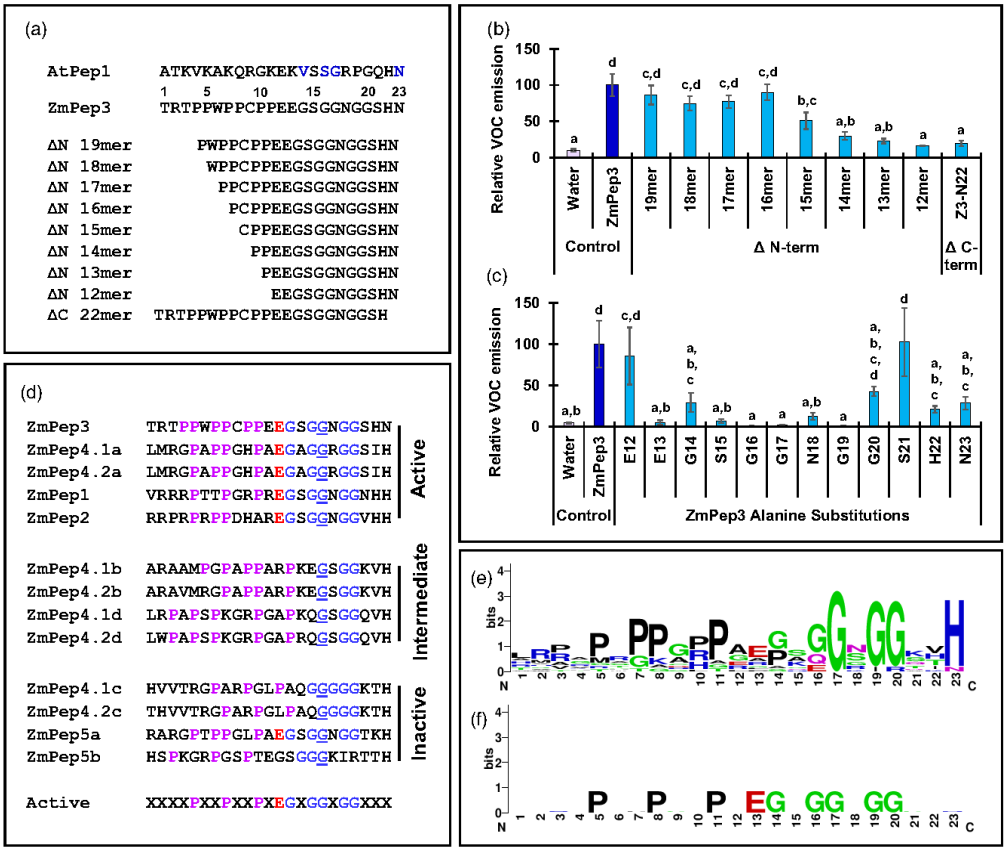


Figure 4. Structure–function analysis of residues critical for *Zea mays* plant elicitor peptide 3 (ZmPep3) elicitor activity.
(a) Alignment of Arabidopsis AtPep1 with ZmPep3. Residues found to be essential for AtPep1 activity are highlighted in blue. Truncated ZmPep3 variants used for structure–function study are shown below. (b) Relative herbivore-associated volatile organic compounds (VOCs) emitted by 21-day-old maize leaves 12 h post-treatment with water, or 125 pmol ZmPep3 or truncated ZmPep3 variants. The VOCs measured were indole, β-caryophyllene, *E*-β-farnesene, linalool, dimethylnonatriene (DMNT), trimethyltridecatetraene (TMTT) and α-bergamotene. (c) Relative VOCs emitted by leaves post-treatment with water, 125 pmol ZmPep3 or alanine-substituted ZmPep3 variants. (d) Classification of individual ZmPeps into functional categories, based on the strength of elicitor activity. Conserved prolines are designated in magenta, conserved glutamate residues in red and conserved glycine residues in blue. Gly17, the only conserved amino acid among Peps from all plant species, is underlined. A consensus sequence for the family of active peptides is designated at the bottom as 'active'. (e), (f) Logo files representing residue conservation among all ZmPeps (e), versus residue conservation among only those ZmPeps designated as 'active' (f). For experiments shown in (b) and (c), $n = 4 \pm$ SEM. Within plots, different letters (a–d) represent significant differences (one-way analysis of variance followed by Tukey's test corrections for multiple comparisons, $P < 0.05$).

ninth amino acid resulted in an approximately 50% loss of function compared with the full-length peptide, indicating that the carboxyl terminus 16mer is the minimum core sequence required for full activity. Successive amino terminus truncations had increasingly reduced activity, with apparent full inactivation occurring with the carboxyl terminus 12mer. Carboxyl terminus residues were confirmed as essential for ZmPep3 activity by the observation that removal of even a single residue from the carboxyl terminus abolished ZmPep3 activity.

To probe which carboxyl terminus residues were most critical for function, alanine substitutions across residues 12–23 were assayed (Figure 4c). Only two substitutions retained full activity compared with the wild-type (WT) peptide, E12ΔA and S21ΔA, with all others demonstrating a more than 50% reduction in volatile sesquiterpene emission. Alanine substitution at positions E13, G16, G17 and G19 completely abolished ZmPep3 activity. These results indicate that the core region required for ZmPep3 function spatially coincides with the region previously identified for

AtPep1, in which V13, S16, G17 and N23 (highlighted in blue in Figure 4a) were absolutely required for activity (Pearce *et al.*, 2008). Interestingly, despite the spatial conservation of core function for the Peps from both species, the side chains at the 13 and 16 positions differ between ZmPep3 and AtPep1 (E versus V, and G versus S). Activity of Peps has been shown to be family specific, i.e. poaceous Peps are active when treating poaceous plants, but not brassicas, and vice versa (Huffaker *et al.*, 2013). The requirement for differing residues at these key positions may in part explain why there is no cross-family reactivity of these functionally orthologous signals.

Amino acid variation underlies relative potency of ZmPeps

Based on the assessment of relative elicitor activity as evaluated through the magnitude of induced changes in levels of defense-associated phytohormones, metabolites and transcripts, the ZmPeps may be divided into three functional groups (Figure 4d): active, which includes ZmPep1, ZmPep2, ZmPep3 and ZmPep4.1a/ZmPep4.2a; intermediate, comprising ZmPep4.1b, ZmPep4.2b, ZmPep4.1d and ZmPep4.2d; and inactive, including ZmPep4.1c, ZmPep4.2c, ZmPep5a and ZmPep5d. A consensus sequence for all ZmPep sequences (Figure 4e) was generated for comparison with a consensus made only for ZmPeps classified as active (Figure 4f). Strikingly, only those ZmPeps categorized as active are perfectly conserved at the residues which were found to be critical for ZmPep3 activity as determined through alanine substitutions. Most notably, E13 is conserved in the active peptides but is not present in any of the others, with the exception of ZmPep5a.

ZmPep5a may act as an antagonist to negatively regulate ZmPep-induced responses

Interestingly, all of the critical residues found in the active ZmPeps are also conserved in ZmPep5a, which is inactive (Figure 2a e). We hypothesized that the lysine residue at position 22 in ZmPep5 might disrupt function by placing a relatively bulky positive charge at a site occupied by either a histidine or isoleucine in active ZmPeps (Figure 5a). To test whether modification of this site could reconstitute activity in ZmPep5a, a histidine residue as found in ZmPep3 at that position, was substituted. When this substitution was made, ZmPep5a-K22ΔH was found to be just as active as ZmPep3 in promoting both emission of herbivore-associated volatile organic compounds (VOCs) and increased expression of the defense gene encoding Terpene Synthase 10 (Figure 5b,c). Because other residues critical for ZmPep activity were perfectly conserved in ZmPep5a, we hypothesized that ZmPep5a was either unable to interact with ZmPEPR receptors or unable to activate signaling. If ZmPep5a were capable of interaction with ZmPEPRs without promoting signaling then it could potentially function as a competitive antagonist of signaling activation by active ZmPeps. To test whether ZmPep5a acts as a competitive antagonist, maize leaves were treated with 100 nM ZmPep5a in combination with varying concentrations of ZmPep3 ranging from 0 to 100 nM, and relative herbivore-associated VOC emission was measured (Figure 5d). Leaves co-treated with 100 nM each of ZmPep5a and ZmPep3 emitted VOCs at levels approximately 60% that of leaves treated with only 100 nM ZmPep3. When leaves were co-treated with 10 nM ZmPep3 in the presence of 100 nM ZmPep5a, relative VOC emission was less than 20% that of leaves supplied with 10 nM ZmPep3 alone. Homology modeling of ZmPEPR1 using the AtPep1/AtPEPR1 crystal structure (PDB ID 5GR8) as a template with peptide docking predicted that both ZmPep3 and ZmPep5a would interact with ZmPEPR1 (Figure 5e,f), with the carboxyl terminus of each peptide occupying a similar position (Tang *et al.*, 2015). For both peptides, intermolecular bonding was predicted with several of the same side chains from ZmPEPR1, including R455 and K525 (Figure 5g j). As predicted by these models, interaction of ZmPEPR1 with ZmPep5a at these sites would block potential interaction with ZmPep3. Together, these data demonstrate that ZmPep5a can reduce ZmPep3 signaling and has the potential to function *in vivo* as a negative regulator of ZmPep-induced immune responses through likely competitive antagonist activity.

ZmPep3 functionally associates with both predicted ZmPEPR receptors

Many species for which putative PEPR orthologues have been identified contain two or more PEPRs, and maize encodes two PEPRs. As might be expected, given that Peps are recognized within plant families but not between different families, PEPRs cluster together within families (Figure 6a, Table S7). Conservation mapping of leucine-rich repeat receptor families has been used to predict potential ligand-binding domains by assessing which residues are most conserved among related receptors within plant genera or families (Helft *et al.*, 2011). Using the 17 identified poaceous PEPR sequences, conservation mapping was applied to identify the potential binding face for poaceous Peps (Figure 6b). The region identified broadly matched potential sites of association between ZmPep3 and ZmPep5a with ZmPEPR1, as predicted using homology modeling and peptide docking.

Previous work has demonstrated that transient heterologous expression of ZmPEPR1 in *Nicotiana benthamiana* confers sensitivity to ZmPep1 as assessed by induced increases in ET (Lori *et al.*, 2015). To assess whether both predicted ZmPEPRs recognized ZmPep3, both were heterologously expressed in *N. benthamiana* as yellow fluorescent protein (YFP) fusion proteins, along with the brassicaceous pattern recognition receptor ELONGATION

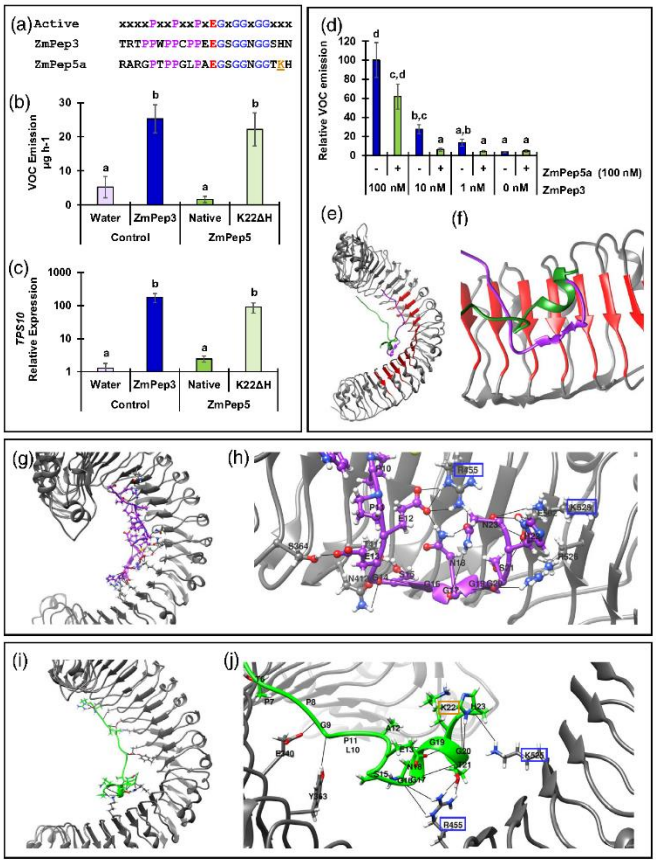


Figure 5. ZmPep5a may act as an antagonist to negatively regulate responses induced by *Zea mays* plant elicitor peptides (ZmPeps).
(a) Alignment of ZmPep3 and ZmPep5a with the consensus sequence for active ZmPeps. Conserved prolines are designated in magenta, conserved glutamate residues in red and conserved glycine residues in blue. Gly17, the only conserved amino acid among Peps from all plant species, is underlined. The lysine 22 residue in ZmPep5 that was subjected to site-directed mutagenesis is underlined and designated in orange. (b), (c) Relative emission of herbivore-associated volatile organic compounds (VOCs) (b) and expression of the marker gene *Terpene synthase 10*(*TPS10*) (c) in 21-day-old maize leaves 12 h post-treatment with water, or 5 μM solutions of ZmPep3, ZmPep5a or a ZmPep5a-K22ΔH variant. The VOCs measured were indole, β-caryophyllene, *E*-β-farnesene, linalool, dimethylnonatriene (DMNT), trimethyltridecatetraene (TMTT) and α-bergamotene. Relative gene expression values represent the fold change in expression versus water-treated leaves after normalization against expression of the housekeeping gene *ZmRPL17*. (d) Relative VOCs emitted by 21-day-old maize leaves 12 h post-treatment with water or with a titration of ZmPep3 solutions (1, 10 or 100 nM) in the presence or absence of 100 nM ZmPep5a. (e)–(j) Association of ZmPep with the receptor ZmPEPR predicted through homology modeling using the AtPepR1/AtPep1 crystal structure as a template for docking. General positioning (e) and close-up view (f) of predicted association for both ZmPep3 (purple) and ZmPep5a (green) with the ZmPEPR receptor. (g) Predicted intermolecular interactions between ZmPep3 and ZmPEPR1, with a close-up view in (h). Residues in ZmPEPR1 that are predicted to form hydrogen bonds with ZmPep3 residues found to be critical for elicitor activity are highlighted with blue boxes. (i) General predicted positioning of the association of ZmPep5a with the ZmPEPR receptor with possible intermolecular interactions visualized. (j) Close-up view of predicted intermolecular interactions between ZmPep5a and ZmPEPR1. Residues in ZmPEPR1 predicted to form hydrogen bonds with ZmPep5a are highlighted by blue boxes. The ZmPep5a Lys22 residue is designated by an orange box. For experiments shown in (b), (c) and (d) $n = 4 \pm$ SEM. Within plots, different letters (a–d) represent significant differences (one-way analysis of variance followed by Tukey's test corrections for multiple comparisons, $P < 0.05$).

FACTOR-TU RECEPTOR (EFR) as a positive control (Zipfel *et al.*, 2006). Treatment of EFR:YFP-expressing leaves with a 1 μM solution of the peptide ligand that binds to EFR, elf18, triggered increased ET emission within 2 h, whereas treatment with water or ZmPep3 did not (Figure 6c). In contrast, treatment of ZmPEPR1:YFP- or ZmPEPR2:YFP-expressing leaves with ZmPep3 promoted increased ET emission, while elf18-treated leaves emitted ET at levels similar to water-treated leaves. To further explore the functional interaction between ZmPep3 and the two ZmPEPR

receptors, ZmPep3-induced recruitment of co-receptors was assessed. For many leucine-rich repeat receptor kinases, including AtPEPRs, ligand-induced association of SERK co-receptors occurs within minutes (Postel *et al.*, 2010; Roux *et al.*, 2011). Co-expression of ZmPEPR:YFP constructs with two SERK:HA constructs (AtSERK1 and AtSERK3) facilitated co-immunoprecipitation assays to visualize ZmPEPR/SERK association. In water-treated leaves, very low levels of SERK:HA were observed to co-precipitate with ZmPEPR:YFP, indicating weak association between both ZmPEPR1 and ZmPEPR2 with SERK co-receptors (Figure 6d). However, 15 min after ZmPep3 treatment, the presence of SERK:HA in ZmPEPR:YFP co-immunoprecipitant was greatly increased for both ZmPEPR1 and ZmPEPR2 in association with both SERK1 and SERK3. Western blotting of crude extracts ensured that expression of all fusion proteins was stable across samples to rule out any false presumption of association or lack thereof based on variable expression. In sum, these experiments indicate that both ZmPEPR1 and ZmPEPR2 are activated by ZmPep3 to recruit SERK co-receptors and promote downstream signaling.

ZmPEPR1 is the primary receptor involved in ZmPep3-mediated anti-herbivore responses

In Arabidopsis, AtPEPR1 has been shown to be the dominant receptor required for AtPep responses in foliar tissues (Yamaguchi *et al.*, 2006, 2010; Krol *et al.*, 2010). Knockouts

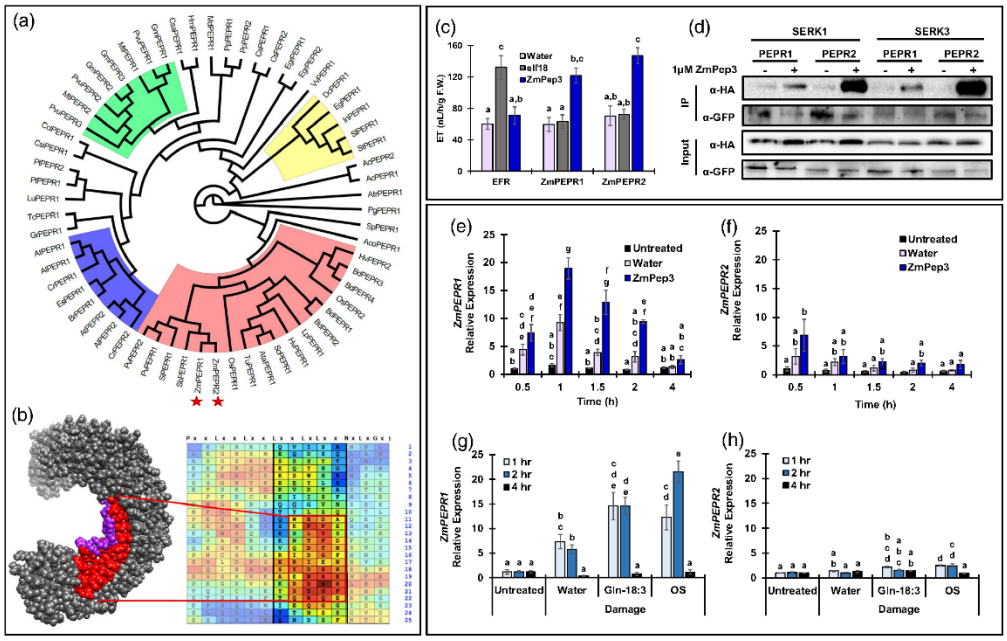


Figure 6. *ZmPEPR* genes encode active receptors with differential endogenous expression in response to elicitor treatments.
(a) Phylogenetic tree generated from predicted monocot, dicot and gymnosperm plant elicitor peptide receptor (PEPR) sequences. Red shading indicates poaceous plants, brassicaceous plants are in red, fabaceous plants in green and combined asterids are in yellow. Maize PEPRs are designated by a red star. (b) Comparison of predicted sites of interaction between ZmPep3 and ZmPEPR1 as predicted through molecular modeling to the ZmPEPR1 region containing the most conserved residues as determined through conservation mapping generated by analysis of 21 predicted poaceous PEPR sequences. (c) Relative ethylene (ET) emitted from *Nicotiana benthamiana* leaves expressing yellow fluorescent protein (YFP) fusions of either AtEFR, ZmPEPR1 or ZmPEPR2 2 h post-treatment with water or a 1 μM solution of either elf18 or ZmPep3. (d) Western blot of ligand-induced interaction between ZmPEPR:YFP and SERK:HA co-receptor fusion proteins co-expressed in *N. benthamiana* as determined through co-immunoprecipitation with anti-GFP beads 5 min after treatment with either water or 1 μM ZmPep3. For all samples both prior to immunoprecipitation (INPUT) and after immunoprecipitation (IP), ZmPEPR:YFP was detected using anti-GFP antibody and SERK:HA with anti-HA antibody. (e), (f) Relative expression of *ZmPEPR1* (e) and *ZmPEPR2* (f) in 21-day-old leaves treated with either water or 5 μM ZmPep3 as determined by quantitative (q)RT-PCR compared with expression of the *Rpl17* control gene. (g), (h) Relative expression of *ZmPEPR1* (g) and *ZmPEPR2* (h) in 21-day-old maize leaves treated with either water, 5 μM *N*-linolenoyl-L-glutamine (Gln-18:3), or *Spodoptera exigua* oral secretions (OS) as determined by qRT-PCR compared with expression of Zm*Rpl17*. In each case relative expression values represent the fold change in expression versus the untreated leaves after normalization against expression of the housekeeping gene *ZmRPL17*. For all experiments shown $n = 4 \pm$ SEM. Within plots, different letters (a–g) represent significant differences (one-way analysis of variance followed by Tukey's test corrections for multiple comparisons, $P < 0.05$).

in *AtPEPR1* have a stronger AtPep-insensitive phenotype than *AtPEPR2* knockouts, but a double knockout in both receptors is required for full loss of AtPep sensitivity (Yamaguchi *et al.*, 2010). To better understand the role of each ZmPEPR during the maize defense response we analyzed their expression patterns. Because the expression of genes encoding receptors is often rapidly upregulated in the presence of their ligand, we examined the expression of *ZmPEPR1* and *ZmPEPR2* in maize leaves scratch wounded and treated with either water or ZmPep3 (Figure 6e,f). Both *ZmPEPR1* and *ZmPEPR2* were wound-responsive, demonstrating increased expression with water treatment, but the magnitude of change was greater for *ZmPEPR1*, peaking at an approximately 10-fold increase relative to threefold for *ZmPEPR2* within 1 h. Furthermore, wound-inducible *ZmPEPR1* expression persisted longer than that of *ZmPEPR2*. Similarly, both *ZmPEPR1* and *ZmPEPR2* displayed increased expression after ZmPep3 treatment compared with both untreated and water-treated leaves, but the magnitude of change for *ZmPEPR1* was larger and more persistent. Expression of *ZmPEPR1* was also more responsive to treatment with herbivore-associated stimuli: leaves were scratch wounded and treated with either oral secretions (OS) from *Spodoptera exigua* or with *N*-linolenoyl-L-glutamine (Gln-18:3), a herbivore-associated molecular pattern (HAMP) found in oral secretions of lepidopteran insects (Alborn *et al.*, 1997). Expression of *ZmPEPR1* was again wound responsive and increased upon treatment with *S. exigua* OS or Gln-18:3 to levels more than 20-fold higher than basal expression (Figure 6g). *ZmPEPR2* was minimally responsive, increasing about twofold over basal expression levels in leaves treated with Gln-18:3 or *S. exigua* OS as compared with untreated leaves (Figure 6h).

To assess the relative contribution of each ZmPEPR to ZmPep3-mediated signaling and defense, CRISPR/Cas9 gene editing was used to generate disruptive frameshift mutations approximately 400 nucleotides after the start codon in both *ZmPEPR1* and *ZmPEPR2* genes. One homozygous line was obtained for *ZmPEPR1* containing a single-nucleotide insertion resulting in a frameshift that generated a stop codon at nucleotide 549 (Figure 7a). For *ZmPEPR2* a single homozygous line was obtained containing an 18-nucleotide out-of-frame insertion resulting in a premature stop codon at nucleotide 423 (Figure 7b). For comparison, a WT sibling segregant line was used. To examine whether disruption of either gene affected sensitivity to ZmPep3, relative emission of herbivore-associated VOCs was measured in excised leaves 12 h post-treatment with either water or ZmPep3. In the WT sibling line, ZmPep3 caused an approximately fourfold increase in VOC emission (Figure 7c). In contrast, in the *Zmpepr1* knockout line no ZmPep3-induced increase in VOC emission was observed (Figure 7d). The *Zmpepr2* knockout line retained ZmPep3-induced increased in VOC emission, but the magnitude of emission was reduced to approximately one-third that of the WT sibling line (Figure 7e). Interestingly, basal VOC emission from excised leaves treated with water was reduced in both *Zmpepr1* and *Zmpepr2* knockout lines compared with the WT sibling line.

Pre-treatment with ZmPep3 has been shown previously to promote a defense response capable of restricting *S. exigua* larval growth (Huffaker *et al.*, 2013). To assess whether *Zmpepr* knockout lines were compromised in ZmPep3-induced resistance to *S. exigua*, leaves of each line and the WT sibling were pre-treated with ZmPep3 for 48 h, after which individual second-instar larvae were allowed to feed on leaf disk samples. As observed previously, *S. exigua* larvae supplied with leaves from WT sibling plants pre-treated with ZmPep3 gained approximately one-third less mass than those on leaves pre-treated with water (Figure 7f). Notably, this ZmPep3-induced effect was lost in *Zmpepr1* knockout plants: when given leaf samples from *Zmpepr1* plants, *S. exigua* larval growth on ZmPep3-treated leaves was unchanged compared with that on water-treated leaves (Figure 7g). ZmPep3-induced decreases in larval mass gains still occurred in *Zmpepr2* knockout plants, but the relative decrease was significantly smaller than that observed in the WT sibling plants (Figure 7h). For all lines, no difference was observed in leaf area consumed. In sum, while both ZmPEPR1 and ZmPEPR2 are capable of functionally interacting with ZmPep3, expression patterns and knockout line phenotypes indicate that ZmPEPR1 probably plays a dominant role in mediating ZmPep3-activated foliar anti-herbivore defenses.

DISCUSSION

Plant elicitor peptides have been demonstrated to be conserved regulators of defense responses across diverse plant species and are multigene families in most species (Huffaker *et al.*, 2006, 2011a, 2013; Lori *et al.*, 2015; Lee *et al.*, 2018; Shinya *et al.*, 2018). However, the functional relevance of Peps as multigene families has remained largely undefined. Analysis of the Arabidopsis multigene PROPEP family showed that application of each AtPep triggered near-identical defense responses, leading to the conclusion that AtPeps were fundamentally functionally redundant, although tissue-specific expression patterns were hypothesized to confer some *in planta* specificity. In contrast, previous work examining ZmPep-induced responses in maize implied that ZmPeps might have some specificity of function, with some ZmPeps acting as potent regulators of herbivore-associated volatile emission and others having little effect (Huffaker *et al.*, 2013). To better define the function of individual ZmPeps, activities of each peptide were examined by assessing changes in defense-associated phytohormones, specialized metabolites and

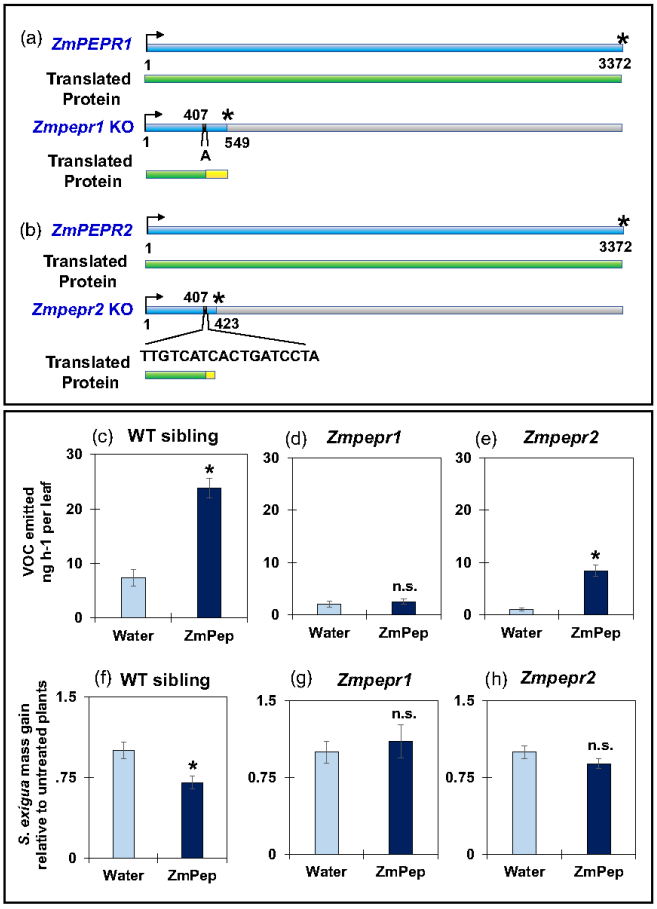


Figure 7. ZmPEPR CRISPR/Cas9-generated knockout lines implicate ZmPEPR1 as the primary contributor to *Zea mays* plant elicitor peptide (ZmPep)-mediated anti-herbivore defense responses.

(a), (b) Schematic of gene and translated proteins for native *ZmPEPR1* (a) and *ZmPEPR2* (b) versus CRISPR/Cas9-generated knockout line. Genes are shown in blue with the start codon marked by an arrow and the stop codon with a star. Inserted nucleotides in the *Zmpepr1* mutant line (a) and *Zmpepr2* (TTGTCATCACTGATCCTA) are designated below brackets. Proteins encoded by the genes are shown in green, with frameshifts in mutant line designated in yellow. (c)–(e) Herbivore-associated volatile organic compounds (VOCs) emitted from leaves of wild-type (WT) siblings of *Zmpepr1* and *Zmpepr2* CRISPR knockouts (c), *Zmpepr1* knockout plants (d) and *Zmpepr2* knockout plants (e) 12 h after treatment with water or 5 μM ZmPep3. The VOCs measured were indole, β-caryophyllene, *E*-β-farnesene, linalool, dimethylnonatriene (DMNT), trimethyltridecatetraene (TMTT) and α-bergamotene. For all, $n = 4 \pm$ SEM, with an asterisk (*) indicating a significant difference of $P < 0.05$ as determined by a pairwise Student's *t*-test within each graph. (f)–(h) Relative mass increase in second-instar *Spodoptera exigua* larvae allowed to feed for 24 h on leaf disks from WT sibling plants (F), *Zmpepr1* knockout plants (g) and *Zmpepr2* knockout plants (h) pre-treated with either water or 5 μM ZmPep3 for 48 h as compared with larvae on disks from untreated leaves. For all experiments shown $n = 20 \pm$ SEM, with an asterisk (*) indicating a significant difference of $P < 0.05$ as determined by a pairwise Student's *t*-test within each graph (n.s. designates not significant).

global gene expression patterns. Together, this has led to a number of new insights into how Peps function as signals, namely, that some PROPEP precursors harbor multiple Peps, that ZmPeps regulate defense-related phytohormone levels in a rheostat-like manner and that Peps can have antagonist-like negative regulatory activity.

Analysis of poaceous PROPEP precursors revealed that in addition to canonical precursors containing a single Pep some had additional internal Pep-like sequences, ranging between two and four putative Peps per precursor. In testing the four ZmPeps contained in the ZmPROPEP4.1 precursor, we found three of them to be active elicitors of

maize defense. Many animal peptide hormone precursors also contain multiple signals, including pro-opiomelanocortin (POMC) and proglucagon (Bell *et al.*, 1983; Navarro *et al.*, 2016). Interestingly, POMC contains four unique peptide hormones with differing functions: α-MSH, which regulates appetite and movement of melanin, ACTH, which stimulates glucocorticoid secretion, and β-endorphin and Met-enkephalin, endogenous opioids with widespread activities (Navarro *et al.*, 2016). In contrast, almost all plant peptide hormones are stored as single copies in their respective precursor proteins, with the exception of the hydroxyproline-rich systemins (HypSys) found in the Solanaceae (Pearce *et al.*, 2001; Ryan and Pearce, 2003). The proHypSys precursors contain between three and six copies of HypSys peptides, which through biochemical purification and isolation from plant tissues have been confirmed to be released as bioactive signals (Ryan and Pearce, 2003; Chen *et al.*, 2008). The HypSys peptides appear redundant rather than unique in function, and like Peps are regulators of defense responses, stimulating MAP kinase activity, extracellular alkalinization and accumulation of proteinase inhibitors (Pearce *et al.*, 2001; Ryan and Pearce, 2003). Compound storage of (semi)-redundant bioactive peptides in a single precursor could allow for increased signal amplification upon processing, and the finding that multiple families of immunoregulatory peptides have evolved this capability implies that it may be advantageous for rapid and robust amplification of an initial immune input.

Characterization of individual ZmPep activities demonstrated that their differential activity is defined by the relative magnitude of their response output rather than uniqueness. As the most potent elicitor, ZmPep3 stimulates production of the highest levels of both the phytohormones JA and ET and defense metabolites, as well as the greatest breadth of gene expression changes. Those ZmPeps that induced smaller increases in JA and ET triggered lower levels of defense metabolite accumulation and a narrower set of DEGs. These experiments revealed a striking direct linear relationship between relative levels of JA and ET and the degree of defense output, indicating that magnitude of response was essentially a function of phytohormone concentrations. Through their differential activities, ZmPeps have the potential to act as a rheostat system, controlling the levels of JA and ET as a means to achieve varying degrees of protective response. Rheostat-like regulation has been proposed as critical to the activity of other phytohormone-mediated processes such as auxin-regulated apical stem cell maintenance and protophloem differentiation in roots and ABA regulation of abiotic stress responses (Tischer *et al.*, 2017; Marhava *et al.*, 2018; Ma *et al.*, 2019).

Notably, genes differentially expressed in response to individual ZmPeps were not separate or even overlapping sets, but instead were almost entirely nested sets of varying sizes contained within one another. All genes differentially expressed in response to the less potent ZmPep2 were also differentially expressed in response to more potent Peps. Likewise, the set of genes upregulated in response to ZmPep3, the most active peptide, contained all but six of the genes that were upregulated by the other ZmPeps. In combination with the finding that the overall number of DEGs for each ZmPep was proportional to the relative increase in production of JA and ET, these gene sets could potentially be defined as tiers of those very sensitive to even small changes in JA/ET (represented in tier 1) versus those which are less sensitive (represented in tiers 4 and 5), requiring larger increases in phytohormone levels to trigger transcription. Altogether, the function of ZmPep family signals in fine-tuning hormone levels presents a unique working system for elicitation of gradations in JA/ET production that allows for delineation of phytohormone concentration-mediated response specificity. The ability to predictably adjust maize JA/ET concentrations through simple peptide application could have useful applications in research and practice.

How individual ZmPeps achieve varying activation of phytohormone production and defense output is unclear, but this could potentially occur through relative affinity of ligand receptor interactions due to ZmPep sequence variation. Through sequence-level comparisons of ZmPeps in combination with structure function analysis using truncated and alanine-substituted ZmPep3 variants, structural features associated with higher levels of activity were definable. The ZmPeps were readily categorized into three functional categories based on the magnitude of response output induced active, less active and inactive and although there is a large degree of variability in sequence between any given Pep, the consensus motif xxxPxxPxxPxEGxGGxGGxxx emerged for ZmPeps categorized as active. Moreover, specific residues contained in this motif, namely E13, G16, G17 and G19, were found to be essential for ZmPep3 activity in alanine substitution experiments. While less active or inactive ZmPeps might partially adhere to the consensus, with the exception of ZmPep5a, only active ZmPeps contain the full motif. Molecular modeling predicts hydrogen bonding between the side chain of ZmPep3 E13 and ZmPEPR1, and the freely rotating glycine residues appear to participate in a twisted conformation that, if accurate, could be more difficult to achieve with other amino acids at those positions. This parallels the finding that free rotation and compact packing of G17 in AtPep1 are essential for binding to and activating AtPEPR1 (Pearce *et al.*, 2008; Tang *et al.*, 2015).

ZmPep5a contained all the structural features characteristic of an active ZmPep but failed to induce any response in leaves. The lysine residue at position 22 was hypothesized to interfere with ZmPep5a receptor interactions and

disrupt function, which was supported by the observation that ZmPep5a was converted to an agonist with activity equivalent to ZmPep3 by a histidine substitution at this position. Interestingly, ZmPep5a was found to act as a competitive inhibitor of ZmPep3-induced responses, reducing the activity of ZmPep3 when co-administered in a dose-dependent manner. This suggests that ZmPep5a is able to compete for receptor binding with ZmPep3 but does not activate signaling, and may have an endogenous role as an antagonist and negative regulator of ZmPep responses. The binding of Peps to PEPRs recruits SERK co-receptors required for signaling activity, with Peps sandwiched at the interface between the two (Postel *et al.*, 2010; Roux *et al.*, 2011; Tang *et al.*, 2015). It may be that ZmPep5a is able to bind ZmPEPRs but that the K22 residue disrupts PEPR/SERK interactions to prevent downstream activation. Antagonist activity has not been previously observed for Pep signaling but does occur in other peptide hormone families. The epidermal patterning factor (EPF) peptides act through their receptor TOO MANY MOUTHS (TMM) to suppress ectopic stomatal development, whereas STOMAGEN, an EPF-like peptide acts as a negative regulator by directly binding to TMM and competitively inhibiting downstream signaling pathways (Sugano *et al.*, 2010). Similarly, two CEP peptides, CEP5 and XIP1, appear to have opposing activities dependent on the CEPR1 receptor in regulating lateral root formation (Roberts *et al.*, 2016). While none of the AtPeps act as antagonists, the apparent competitive inhibitory function of ZmPep5a found in maize suggests that in some species Peps may have evolved to become receptor antagonists.

In addition to providing new insights into mechanisms of regulation not previously associated with Peps, this work has identified a second maize PEPR, ZmPEPR2, confirmed the functional capability of both ZmPEPRs and defined ZmPEPR1 as the dominant receptor for ZmPep3-mediated anti-herbivore defense responses. In Arabidopsis, AtPEPR1 is the dominant receptor for foliar defenses while AtPEPR2 is implicated in root-specific processes (Yamaguchi *et al.*, 2010; Ma *et al.*, 2014). Similarly, although both ZmPEPRs confer sensitivity to ZmPep3 in a heterologous expression system, analyses of gene expression patterns and knockout line phenotypes supported ZmPEPR1 as the predominant contributor to ZmPep3-mediated insect resistance in maize leaves. An interesting question for future study is whether ZmPEPR2 is consistently a lesser contributor to ZmPep-mediated responses or if it is the primary contributor to ZmPep signaling in other tissues. Additionally, while the current study indicates that ZmPEPR1 contributes to ZmPep3-mediated defenses, detailed future study in the absence of exogenous peptide treatment will be required to understand the *in vivo* role of ZmPEPR1 with respect to herbivore resistance.

Our characterization of ZmPep signaling has demonstrated that while they are functionally orthologous to AtPeps as regulators of defense responses, there are a number of unique aspects particular to the ZmPep family, including two novel functions not previously ascribed to Pep signaling that could both increase immune amplification and allow for rapidly extinguishable signaling: evolution of precursor proteins containing multiple ZmPeps could facilitate more robust amplification upon immune elicitation through concerted release of several signaling peptides. In contrast, the action of ZmPep5a as a competitive inhibitor of ZmPep-induced responses could represent a mechanism to quench Pep signaling and prevent potentially detrimental overactivation of immunity. Although many general functional features are conserved between Pep signaling in Arabidopsis and maize, these findings also indicate that we cannot wholly extrapolate mechanisms from model systems as being perfectly conserved in crops, and vice versa. The Peps are a powerful tool for probing and manipulating immunity in diverse plants but there are likely to be unique features among species that will require empirical study for optimal understanding and deployment of the system.

EXPERIMENTAL PROCEDURES

Plant growth conditions

Zea mays plants were grown in BM2 soil (Berger Mixes, https://www.berger.ca/en/horticultural-products/) and supplemented with a 18-18-21 Tomato Plant Food fertilizer (Miracle-Gro, https://www.miraclegro.com/) inside a greenhouse (12-h light, minimum of 300 μmol $m^{-2}sec^{-1}$, and 12-h dark) in a 24°C/28°C (night/day) temperature cycle. *Nicotiana benthamiana* plants were grown in BM2 soil (Berger Mixes) and supplemented with a 20-20-20 General Purpose fertilizer (Jack's Professional, https://www.jrpeters.com/) inside a growth room (16-h light, 150 μmol $m^{-2}sec^{-1}$, and 8-h dark) at 22°C.

Identification and comparison of PROPEP sequences in multiple genomes

Individual B73 exon sequences of PROPEP1-5 were used as reference sequences to query the BLAST databases of different genomes using standalone blastn (default parameters). The blastn query results were used to generate a list of partial and full-length PROPEP sequences and genomic coordinates in the following reference genomes: B73 (NAM-5.0), W22 (NRGENE), Mo17 (CAU), EP1 (TUM), F7 (TUM), all NAM inbred lines and Zx-PI566673 (YAN) (Portwood *et al.*, 2019). All sequences were aligned to their respective B73 PROPEP coding sequences to determine predicted truncations and deleterious mutations and to extract all intact 23-amino-acid long Pep sequences. Neighboring PROPEP sequences were grouped into three loci and the genomic sequences of each locus in each genome were extracted and compared using the standalone blastn software (modified parameter: word_size 50) and the comparison results were visualized and annotated using genoPlotR (Guy *et al.*, 2010).

Peptide and elicitor treatment of leaves

Peptides for all comparative experiments were synthesized and purified as 23mers by Sigma-Aldrich (https://www.sigmaaldrich.c

om/). All peptide sequences are listed in Table S3. All peptides were diluted in water to the concentrations indicated for plant treatments to analyze transcript and metabolite abundance or for insect bioassays. Leaves were treated with solutions through either scratch application on intact plants or leaf excision, as previously described (Huffaker *et al.*, 2011a). Unless otherwise indicated, leaves were treated with 5 nmol of peptide per leaf. Tissue was harvested in liquid nitrogen for RNA and metabolite analysis at the time points indicated.

Measurement of hormones and defense metabolites

Levels of indole, JA, SA and terpenoid pools were measured using the previously described vapor phase extraction method coupled with GC/MS analysis (Schmelz *et al.*, 2004; Huffaker *et al.*, 2013). Indole levels were quantified by comparison with an external standard curve. Ethylene emitted by leaves was measured by GC using a standard curve, as previously described (Schmelz *et al.*, 2009). Volatiles were collected as described previously (Huffaker *et al.*, 2013). In short, to collect emitted volatiles, maize leaves were enclosed in glass tubes under light, and head-space volatiles were collected on 50 mg Super Q (80/100 mesh; Alltech, https://www.alltech.com/) for 30 min. Methylene chloride, with the addition of nonyl acetate as an internal standard, was used to elute the volatiles. Volatiles were analyzed by GC, and specific compounds were identified by comparing their retention times with those of pure standards.

Isolation of RNA and measurement of transcript abundance by quantitative RT-PCR

Total RNA was isolated with TRIzol (Invitrogen, https://www.thermofisher.com/us/en/home/brands/invitrogen.html) as per the manufacturer's protocol and cDNA synthesized using the RETROscript reverse transcriptase kit (Applied Biosystems, https://www.thermofisher.com/us/en/home/brands/applied-biosystems.html) and random decamer primers. Samples of cDNA were diluted twofold and analyzed by quantitative (q)RT-PCR using Power SYBR Green Master Mix (Applied Biosystems) with each primer at a concentration of 300 nM. Sequences for all primers used are contained in Table S8. Amplification was performed using the ABI 7300 sequence detection system (Applied Biosystems) following standard thermal profile conditions: 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min. Data were analyzed with SDS 1.3.1 software (Applied Biosystems). Relative expression levels were determined for four independent biological replicates; all reactions were run in triplicate. Threshold cycle (Ct) values were normalized to the housekeeping gene encoding ribosomal protein 17 (*Rpl17*) (Kirchberger *et al.*, 2007; Huffaker *et al.*, 2013). The abundance of each gene transcript was calculated relative to its corresponding untreated control with fold-change calculations performed with the Livak and Schmittgen method using the equation $2 - \Delta\Delta Ct$. Specificity of real-time PCR products was verified on 1% agarose gels (Livak and Schmittgen, 2001).

Measurement of transcript abundance by microarray and data analyses

Total RNA of leaves 12 h post-treatment with water or ZmPeps was prepared as described above with four biological replicates each. To eliminate DNA contamination, RNA was DNase treated with the TURBO DNA-free kit from Ambion (Applied Biosystems). Microarray sample preparation and analysis was performed by the University of Florida Interdisciplinary Center for Biotechnology Research (ICBR) gene expression core. For each sample, 1 μg of total RNA was provided to the ICBR gene expression core for cDNA synthesis, reverse transcription and biotin labeling using the Agilent Quick Amp Gene Expression Labeling kit and protocol (https://www.agilent.com/). This was followed by hybridization to the Agilent-016047 maize 44 K genome array, washing and scanning as per Agilent protocols. Data extraction was executed by the ICBR gene expression core and provided as CEL files. Raw data were subjected to quality control analysis using the ArrayQualityMetrics Bioconductor package, and three samples were removed for further analysis based on the outlier detection for distances between arrays metric (Kauffmann *et al.*, 2009). The limma Bioconductor package was used for background correction, followed by within-group quantile normalization using the qsmooth Bioconductor package (Ritchie *et al.*, 2015; Hicks *et al.*, 2018). Normalized data were processed using limma to fit each gene to a linear model, and the FDR was calculated using the Benjamini–Hochberg procedure. Annotation to the raw data was assigned using the Maize Microarray Annotation Database (Coetzer *et al.*, 2011). All data have been submitted to NCBI GEO and are available under the accession number GSE147439.

A list of DEGs was generated for every treatment by retaining all the genes that had an FDR *P*-value < 0.05 and log(fold change) > 1 for upregulated genes and < -1 for downregulated genes. To assess the degree of overlap between the intersections of DEGs, the R package eulerr was used to generate an Euler diagram (Larsson, 2019). Venn webtools was used to identify the intersections between DEGs for each treatment group (http://bioinformatics.psb.ugent.be/webtools/Venn/). The intersections with the highest number of genes were assigned as tiers, with the top five intersections for the upregulated genes and the top four intersections for the downregulated genes. Genes not contained in any tier were summed and added to the unassigned group. AgriGO was used the calculate the Plant GO Slim enrichment for the lists of DEGs within the tiers (Tian *et al.*, 2017).

Phylogenetic tree construction

Amino acid sequences of putative PROPEP and PEPR genes in the Poaceae were identified and aligned with MUSCLE using default parameters (Edgar, 2004). The aligned sequences were used to construct maximum likelihood (ML) phylogenetic trees using IQ-TREE according to the automatically selected optimal best-fit model and 1000 bootstrap replications (Nguyen *et al.*, 2015). The tree was visualized and annotated using the R package ggtree (Yu *et al.*, 2017a) and FigTree (Rambaut, available at http://tree.bio.ed.ac.uk/software/figtree).

Homology modeling of ZmPEPR1 and docking of ZmPeps

Using the AtPEPR1 crystal structure (PDB ID 5GR8) as a template, the ZmPEPR1 sequence was aligned through the zero end-gap global alignment (ZEGA) method with the Gonnet comparison matrix (Gonnet *et al.*, 1992; Abagyan and Batalov, 1997). Gap opening and extension penalties were set to 2.4 and 0.15, respectively. Using this alignment and template structure, a homology model of ZmPEPR1 was built with the ICM-Pro homology modeling tool and default parameters (Cardozo *et al.*, 1995). Biased probability Monte Carlo (BPMC) sampling was used to refine all side chains and insertions/deletions (Abagyan and Totrov, 1994). To define the docking region of ZmPEPR1, co-crystallized AtPep1 was used in the crystal structure template. A set of potential maps were generated for the docking region on a 0.5 Å three-dimensional grid, containing: (i) van der Waals interaction; (ii) electrostatic interaction; (iii) hydrogen bonds; and (iv) hydrophobic

potential grids. For each potential map, docking and scoring of ZmPeps was performed using a stochastic global energy optimization procedure in internal coordinates implemented in ICM-Pro 360 v.3.8-6a (Abagyan *et al.*, 1994), described as the following steps: (i) ZmPeps were sampled with the implicit solvation model to generate a series of starting conformations via the BPMC method, with each starting conformation placed into the docking region with four principal orientations. (ii) ZmPeps were sampled in the pre-calculated potential maps through BPMC sampling to optimize its positional and internal variables. (iii) After sampling, all top-ranking conformations were rescored using the ICM full atom scoring function and conformations were resorted by the docking score.

Plasmid construction and transient assays in *Nicotiana benthamiana*

ZmPEPR1 (Zm00001d050074) and *ZmPEPR2* (Zm00001d032116) were amplified from genomic *Z. mays* B73 DNA and inserted into pENTR/D-TOPO vector (Invitrogen). An amino terminus HA- or FLAG-tag was inserted into *ZmPEPR1* and *ZmPEPR2*, respectively, downstream of the signal peptide identified by SignalP, using a multistep PCR and ligation process. Expression vectors of the corresponding entry vectors of *ZmPEPR1* and *ZmPEPR2*, driven by a 35S promoter and tagged with YFP or 3× HA were generated through Gateway cloning of the entry clones into the destination vector pGWB441 or pGWB414, respectively (Nakagawa *et al.*, 2007). Expression vectors were transformed into *Agrobacterium tumefaciens* GV3101 (pMP90). *Agrobacterium tumefaciens* carrying the expression vectors was infiltrated into *N. benthamiana* leaves at OD_{600} of 0.8 for complementation assay and 0.3 for co-immunoprecipitation.

Immunoprecipitation and immunoblotting

Nicotiana benthamiana leaves were co-infiltrated with *A. tumefaciens* strains expressing either *AtSERK1:HA* or *AtSERK3:HA* and *ZmPEPR1:YFP* or *ZmPEPR2:YFP*. After 48 h, these leaves were infiltrated with a peptide solution for 5 min and immediately flash-frozen in liquid nitrogen. Frozen tissue was ground using a mortar and pestle. The frozen tissue was homogenized in 2 ml g^{-1} extraction buffer (50 mM TRIS pH 7.5, 150 mM NaCl, 2% Nonidet P-40, 1 mM DTT, 1× Roche Protease Inhibitor). Cell debris was removed by centrifugation (30 min, 15 000 ***g***), and the supernatant containing the solubilized membranes was incubated with 10 µl GFP-Trap® coupled to magnetic agarose beads (Chromotek, https://www.chromotek.com/) by end-over-end mixing at 4°C for 3 h. After washing with extraction buffer three times, the beads were resuspended in 50 µl of 2× SDS buffer for 5 min at 95°C.

Ethylene assay in *Nicotiana benthamiana*

Nicotiana benthamiana leaves were infiltrated with *A. tumefaciens* carrying constructs for the receptor-fusion proteins indicated. Two days after bacterial infiltration, the leaves were infiltrated with a 1 µM peptide solution and the infiltrated leaves were sealed in airtight tubes for 2 h to accumulate head-space volatiles. Ethylene accumulation was measured by collecting 1 ml of the head-space volatiles from each tube and analyzed by GC with flame-ionization detection in comparison with a standard curve as previously described (Schmelz *et al.*, 2009).

Generation of CRISPR/Cas9-mediated PEPR knockout lines

ZmPEPR1 and *ZmPEPR2* guide RNA (gRNA) target sites were selected using the B73 reference genome sequence and criteria as described (Brazelton *et al.*, 2015). Flanking regions with the target site at the middle were PCR-amplified from the maize genotype Hi-II and Sanger sequenced for accuracy of the genomic sequence including the gRNA complementary sequence. The gRNA gene was constructed in the intermediate vector and the expression cassette was mobilized through a gateway reaction into the Cas9-expressing binary vector for maize Hi-II transformation at the Iowa State University Plant Transformation Facility, as previously described (Char *et al.*, 2017). A total of 10 independent T_0 transgenic plants were obtained. To examine if the target gene sequence was edited, the PCR amplicons encompassing the gRNA target site from each plant were sequenced. Ultimately, for both *Zmpepr1* and *Zmpepr2*, one heterozygous mutant was obtained, both resulting from different plantlets derived from a single callus. The heterozygous mutant plants were self-pollinated, and the resultant progeny plants were genotyped to identify homozygous, heterozygous and WT segregant lines. Homozygous knockout plants and a WT sibling were selected for bioassays.

Insect herbivory bioassays

As per a previously published bioassay procedure, to observe the effects of ZmPep3-induced defenses on larval growth, intact leaves were scratch wounded and treated with peptide (Huffaker *et al.*, 2013). Forty-eight hours post-treatment, 2.5 cm leaf disks were harvested around each treatment site using a cork-borer and placed on a piece of filter paper in 12-well tissue culture plates. A single pre-weighed early second-instar *S. exigua* larvae (Benzon Research, https://www.benzonresearch.com/) was placed on each leaf disc, allowed to feed for 24 h and then reweighed.

Statistical analyses

Statistical analyses were conducted using JMP Pro 13.0 (SAS Institute Inc., https://www.sas.com/) and GraphPad Prism 8.0 (GraphPad Software, Inc., https://www.graphpad.com/). One-way analyses of variance were conducted to evaluate statistical differences. Tukey tests were used to correct for multiple comparisons between control and treatment groups. Student's *t*-tests (unpaired, two-tailed) were conducted for pairwise comparisons. *P*-values < 0.05 were considered significant.

ACKNOWLEDEGMENTS

The authors are grateful to Dr Eric A. Schmelz for providing seeds for the maize *acs2/acs6* double mutant and for helpful comments on this manuscript. This research was funded by NSF-IOS PBI CAREER #1943591, USDA NIFA AFRI #2018-67013-28125, a Hellman Foundation Fellowship and UC San Diego start-up funds to AH. KD was funded by a Ciencias sem Fronteiras/CNPq Fellowship (200260/2015-4). EP was funded by the Cell and Molecular Genetics (CMG) Training Program at the University of California, San Diego. BY was funded by NSF-IOS #1936492. RA and DS were funded by NIH R35 #GM131881.

AUTHOR CONTRIBUTIONS

AH conceived the research plans. EP, KD, MR, PW and AH designed and performed experiments and analyzed the data. DS and RA performed ligand receptor modeling analyses. BY, SNC, EP and AH designed gRNA constructs and generated the *Zmpepr1* and *Zmpepr2* maize mutants. EP, KD and AH wrote the manuscript with input from all authors. AH agrees to serve as the author responsible for contact and ensures communication.

CONFLICTS OF INTEREST

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

All relevant data can be found within the manuscript and its supporting materials. Additionally, all microarray data have been submitted to NCBI GEO and are available under the accession number GSE147439.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Structure of maize PROPEP gene loci in the nested association mapping parent lines.

Figure S2. Sequence conservation of individual *Zea mays* plant elicitor peptides across nested association mapping parent lines.

Figure S3. Relative volatile terpene-inducing activity for peptides from derived from the ZmPROPEP4.1 and ZmPROPEP4.2 duplicate precursor genes encoded in the B73 genome.

Figure S4. Relationship between relative levels of each phytohormone in leaves treated with water or *Zea mays* plant elicitor peptides.

Table S1. PROPEP blastn query and result sequences in *Zea mays* inbred lines

Table S2. Summary of PROPEPs identified in *Zea mays* inbred lines

Table S3. Poaceous PROPEP sequences with peptides noted and gene IDs

Table S4. Full microarray data for each sample group with statistical analyses, and annotation

Table S5. List of gene associated with each differentially expressed gene tier in microarray experiments

Table S6. List of genes associated with each Gene Ontology term in microarray experiments

Table S7. PEPR sequences with gene ID numbers

Table S8. List of primers used

REFERENCES

Abagyan, R. and Totrov, M. (1994) Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **235**, 983–1002.

Abagyan, R., Totrov, M. and Kuznetsov, D. (1994) ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Computat. Chemist.* **15**, 488–506.

Abagyan, R.A. and Batalov, S. (1997) Do aligned sequences share the same fold? *J. Mol. Biol.* **273**, 355–368.

Alborn, H.T., Turlings, T.C.J., Jones, T.H., Stenhagen, G., Loughrin, J.H. and Tumlinson, J.H. (1997) An Elicitor of Plant Volatiles from Beet Armyworm Oral Secretion ., **276**, 945–949.

Bartels, S., Lori, M., Mbengue, M., van Verk, M., Klauser, D., Hander, T., Boni, R., Robatzek, S. and Boller, T. (2013) The family of Peps and their precursors in Arabidopsis: differential expression and localization but similar induction of pattern-triggered immune responses. *J. Exp. Bot.* **64**, 5309–5321.

Bell, G.I., Santerre, R.F. and Mullenbach, G.T. (1983) Hamster preproglucagon contains the sequence of glucagon and two related peptides. *Nature*, **302**, 716–718.

Brazelton, V.A., Zarecor, S., Wright, D.A., Wang, Y., Liu, J., Chen, K., Yang, B. and Lawrence-Dill, C.J. (2015) A quick guide to CRISPR sgRNA design tools. *GM Crops Food*, **6**, 266–276.

Cardozo, T., Totrov, M. and Abagyan, R. (1995) Homology modeling by the ICM method. *Proteins*, **23**, 403–414.

Char S.N., Neelakandan A.K., Nahampun H. *et al.* (2017) An Agrobacterium-delivered CRISPR/Cas9 system for high-frequency targeted mutagenesis in maize. *Plant Biotechnology Journal*, **15**, 257–268. http://dx.doi.org/10.1111/pbi.12611

Chen, Y.C., Siems, W.F., Pearce, G. and Ryan, C.A. (2008) Six peptide wound signals derived from a single precursor protein in Ipomoea batatas leaves activate the expression of the defense gene sporamin. *J. Biol. Chem.* **283**, 11469–11476.

Coetzer, N., Myburg, A.A. and Berger, D.K. (2011) Maize microarray annotation database. *Plant Methods*, **7**, 31.

Ding Y., Murphy K.M., Poretsky E. *et al.* (2019) Multiple genes recruited from hormone pathways partition maize diterpenoid defences. *Nature Plants*, **5**, 1043–1056. http://dx.doi.org/10.1038/s41477-019-0509-6

Dressano K., Weckwerth P.R., Poretsky E. *et al.* (2020) Dynamic regulation of Pep-induced immunity through post-translational control of defence transcript splicing. *Nature Plants*, **6**, 1008–1019. http://dx.doi.org/10.1038/s41477-020-0724-1

Edgar, R.C.(2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, **5**, 113.

Erb, M., Veyrat, N., Robert, C.A., Xu, H., Frey, M., Ton, J. and Turlings, T.C. (2015) Indole is an essential herbivore-induced volatile priming signal in maize. *Nat Commun*, **6**, 6273.

Frey, M., Chomet, P., Glawischnig, E. *et al.* (1997) Analysis of a chemical plant defense mechanism in grasses. *Science*, **277**, 696–699.

Gonnet, G.H., Cohen, M.A. and Benner, S.A. (1992) Exhaustive matching of the entire protein sequence database. *Science*, **256**, 1443–1445.

Goodstein, D.M., Shu, S., Howson, R. *et al.* (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**, D1178–D1186.

Gully, K., Hander, T., Boller, T. and Bartels, S. (2015) Perception of Arabidopsis AtPep peptides, but not bacterial elicitors, accelerates starvation-induced senescence. *Front Plant Sci.* **6**, 14.

Guy, L., Kultima, J.R. and Andersson, S.G. (2010) genoPlotR: comparative gene and genome visualization in R. *Bioinformatics*, **26**, 2334–2335.

Hander, T., Fernandez-Fernandez, A.D., Kumpf, R.P. *et al.* (2019) Damage on plants activates Ca(2+)-dependent metacaspases for release of immunomodulatory peptides. *Science*, **363**.

Hegenauer, V., Furst, U., Kaiser, B., Smoker, M., Zipfel, C., Felix, G., Stahl, M. and Albert, M. (2016) Detection of the plant parasite Cuscuta reflexa by a tomato cell surface receptor. *Science*, **353**, 478–481.

Helft, L., Reddy, V., Chen, X., Koller, T., Federici, L., Fernandez-Recio, J., Gupta, R. and Bent, A. (2011) LRR conservation mapping to predict functional sites within protein leucine-rich repeat domains. *PLoS One*, **6**, e21614.

Hicks, S.C., Okrah, K., Paulson, J.N., Quackenbush, J., Irizarry, R.A. and Bravo, H.C. (2018) Smooth quantile normalization. *Biostatistics*, **19**, 185–198.

Huffaker, A., Dafoe, N.J. and Schmelz, E.A. (2011a) ZmPep1, an ortholog of Arabidopsis elicitor peptide 1, regulates maize innate immunity and enhances disease resistance. *Plant Physiol*, **155**, 1325–1338.

Huffaker, A., Kaplan, F., Vaughan, M.M., Dafoe, N.J., Ni, X., Rocca, J.R., Alborn, H.T., Teal, P.E. and Schmelz, E.A. (2011b) Novel acidic sesquiterpenoids constitute a dominant class of pathogen-induced phytoalexins in maize. *Plant Physiol.* **156**, 2082–2097.

Huffaker, A., Pearce, G. and Ryan, C.A. (2006) An endogenous peptide signal in Arabidopsis activates components of the innate immune response. *Proc. Natl. Acad. Sci. U S A*, **103**, 10098–10103.

Huffaker, A., Pearce, G., Veyrat, N. *et al.* (2013) Plant elicitor peptides are conserved signals regulating direct and indirect antiherbivore defense. *Proc. Natl. Acad. Sci. U S A*, **110**, 5707–5712.

Huffaker, A. and Ryan, C.A. (2007) Endogenous peptide defense signals in Arabidopsis differentially amplify signaling for the innate immune response. *Proc. Natl. Acad. Sci. U S A*, **104**, 10732–10736.

Huffaker, Alisa. (2015) Plant elicitor peptides in induced defense against insects. *Curr. Opin. Insect. Sci.* **9**, 44–50.

Kauffmann, A., Gentleman, R. and Huber, W. (2009) arrayQualityMetrics–a bioconductor package for quality assessment of microarray data. *Bioinformatics*, **25**, 415–416.

Kirchberger, S., Leroch, M., Huynen, M.A., Wahl, M., Neuhaus, H.E. and Tjaden, J. (2007) Molecular and biochemical analysis of the plastidic ADP-glucose transporter (ZmBT1) from Zea mays. *J. Biol. Chem.* **282**, 22481–22491.

Krol, E., Mentzel, T., Chinchilla, D. *et al.* (2010) Perception of the Arabidopsis danger signal peptide 1 involves the pattern recognition receptor AtPEPR1 and its close homologue AtPEPR2. *J. Biol. Chem.* **285**, 13471–13479.

Larsson, J. 2019. eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses. https://jolars.github.io/eulerr/

Lee, M.W., Huffaker, A., Crippen, D., Robbins, R.T. and Goggin, F.L. (2018) Plant elicitor peptides promote plant defences against nematodes in soybean. *Mol. Plant Pathol.* **19**, 858–869.

Liu, Z., Wu, Y., Yang, F., Zhang, Y., Chen, S., Xie, Q., Tian, X. and Zhou, J.M. (2013) BIK1 interacts with PEPRs to mediate ethylene-induced immunity. *Proc. Natl. Acad. Sci. U S A*, **110**, 6205–6210.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.

Lori, M., van Verk, M.C., Hander, T., Schatowitz, H., Klauser, D., Flury, P., Gehring, C.A., Boller, T. and Bartels, S. (2015) Evolutionary divergence of the plant elicitor peptides (Peps) and their receptors: interfamily incompatibility of perception but compatibility of downstream signalling. *J. Exp. Bot.* **66**, 5315–5325.

Lu, D., Wu, S., Gao, X., Zhang, Y., Shan, L. and He, P. (2010) A receptor-like cytoplasmic kinase, BIK1, associates with a flagellin receptor complex to initiate plant innate immunity. *Proc. Natl. Acad. Sci. U S A*, **107**, 496–501.

Ma, C., Guo, J., Kang, Y., Doman, K., Bryan, A.C., Tax, F.E., Yamaguchi, Y. and Qi, Z. (2014) AtPEPTIDE RECEPTOR2 mediates the AtPEPTIDE1-induced cytosolic Ca(2+) rise, which is required for the suppression of Glutamine Dumper gene expression in Arabidopsis roots. *J. Integr. Plant Biol.* **56**, 684–694.

Ma, Y., Miotk, A., Sutikovic, Z. *et al.* (2019) WUSCHEL acts as an auxin response rheostat to maintain apical stem cells in Arabidopsis. *Nat. Commun.* **10**, 5093.

Majhi, B.B., Sreeramulu, S. and Sessa, G. (2019) BRASSINOSTEROID-SIGNALING KINASE5 Associates with Immune Receptors and Is Required for Immune Responses. *Plant Physiol.* **180**, 1166–1184.

Marhava, P., Bassukas, A.E.L., Zourelidou, M. *et al.* (2018) A molecular rheostat adjusts auxin flux to promote root protophloem differentiation. *Nature*, **558**, 297–300.

McGurl, B., Pearce, G., Orozco-Cardenas, M. and Ryan, C.A. (1992) Structure, expression, and antisense inhibition of the systemin precursor gene. *Science*, **255**, 1570–1573.

Mendy, B., Wang'Ombe, M.W., Radakovic, Z.S., Holbein, J., Ilyas, M., Chopra, D., Holton, N., Zipfel, C., Grundler, F.M. and Siddique, S. (2017) Arabidopsis leucine-rich repeat receptor-like kinase NILR1 is required for induction of innate immunity to parasitic nematodes. *PLoS Pathog.* **13**, e1006284.

Mousavi, S.A., Chauvin, A., Pascaud, F., Kellenberger, S. and Farmer, E.E. (2013) GLUTAMATE RECEPTOR-LIKE genes mediate leaf-to-leaf wound signalling. *Nature*, **500**, 422–426.

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T. and Kimura, T. (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* **104**, 34–41.

Nakaminami, K., Okamoto, M., Higuchi-Takeuchi, M. *et al.* (2018) AtPep3 is a hormone-like peptide that plays a role in the salinity stress tolerance of plants. *Proc. Natl. Acad. Sci. U S A*, **115**, 5810–5815.

Navarro, S., Soletto, L., Puchol, S., Rotllant, J., Soengas, J.L. and Cerda-Reverter, J.M. (2016) 60 YEARS OF POMC: POMC: an evolutionary perspective. *J. Mol. Endocrinol.* **56**, T113–T118.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A. and Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274.

Pearce, G., Moura, D.S., Stratmann, J. and Ryan, C.A. (2001) Production of multiple plant hormones from a single polyprotein precursor. *Nature*, **411**, 817–820.

Pearce, G., Yamaguchi, Y., Munske, G. and Ryan, C.A. (2008) Structure-activity studies of AtPep1, a plant peptide signal involved in the innate immune response. *Peptides*, **29**, 2083–2089.

Portwood, J.L. 2nd, Woodhouse, M.R., Cannon, E.K. *et al.* (2019) MaizeGDB 2018: the maize multi-genome genetics and genomics database. *Nucleic Acids Res.* **47**, D1146–d1154.

Postel, S., Kufner, I., Beuter, C., Mazzotta, S., Schwedt, A., Borlotti, A., Halter, T., Kemmerling, B. and Nurnberger, T. (2010) The multifunctional leucine-rich repeat receptor kinase BAK1 is implicated in Arabidopsis development and immunity. *Eur. J. Cell Biol.* **89**, 169–174.

Qi, J., Wang, J., Gong, Z. and Zhou, J.M. (2017) Apoplastic ROS signaling in plant immunity. *Curr. Opin Plant Biol.* **38**, 92–100.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. and Smyth, G.K. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47.

Roberts, I., Smith, S., Stes, E. *et al.* (2016) CEP5 and XIP1/CEPR1 regulate lateral root initiation in Arabidopsis. *J. Exp. Bot.* **67**, 4889–4899.

Roux, M., Schwessinger, B., Albrecht, C., Chinchilla, D., Jones, A., Holton, N., Malinovsky, F.G., Tor, M., de Vries, S. and Zipfel, C. (2011) The Arabidopsis leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. *Plant Cell*, **23**, 2440–2455.

Ruiz, C., Nadal, A., Foix, L., Montesinos, L., Montesinos, E. and Pla, M. (2018a) Diversity of plant defense elicitor peptides within the Rosaceae. *BMC Genet.* **19**, 11.

Ruiz, C., Nadal, A., Montesinos, E. and Pla, M. (2018b) Novel Rosaceae plant elicitor peptides as sustainable tools to control Xanthomonas arboricola pv. pruni in Prunus spp. *Mol. Plant Pathol.* **19**, 418–431.

Ryan, C.A. and Pearce, G. (2003) Systemins: a functionally defined family of peptide signals that regulate defensive genes in Solanaceae species. *Proc. Natl. Acad. Sci. U S A*, **100**(Suppl 2), 14577–14580.

Schmelz, E.A., Engelberth, J., Alborn, H.T., Tumlinson, J.H. 3rdand Teal, P.E.,(2009) Phytohormone-based activity mapping of insect herbivore-produced elicitors. *Proc. Natl. Acad. Sci. U S A*, **106**, 653–657.

Schmelz, E.A., Engelberth, J., Tumlinson, J.H., Block, A. and Alborn, H.T. (2004) The use of vapor phase extraction in metabolic profiling of phytohormones and other metabolites. *Plant J.* **39**, 790–808.

Schmelz, E.A., Kaplan, F., Huffaker, A., Dafoe, N.J., Vaughan, M.M., Ni, X., Rocca, J.R., Alborn, H.T. and Teal, P.E. (2011) Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proc. Natl. Acad. Sci. U S A*, **108**, 5455–5460.

Segonzac, C. and Monaghan, J. (2019) Modulation of plant innate immune signaling by small peptides. *Curr. Opin Plant Biol.* **51**, 22–28.

Shen, W., Liu, J. and Li, J.F. (2019) Type-II Metacaspases Mediate the Processing of Plant Elicitor Peptides in Arabidopsis. *Mol. Plant*, **12**, 1524–1533.

Shinya, T., Yasuda, S., Hyodo, K. *et al.* (2018) Integration of danger peptide signals with herbivore-associated molecular pattern signaling amplifies anti-herbivore defense responses in rice. *Plant J.* **94**, 626–637.

Steinbrenner, A.D., Muñoz-Amatriaín, M. Muñoz-Amatriaín, M., *et al.* (2019) A receptor for herbivore-associated molecular patterns mediates plant immunity. https://doi.org/10.1101/679803.

Sugano, S.S., Shimada, T., Imai, Y., Okawa, K., Tamai, A., Mori, M. and Hara-Nishimura, I. (2010) Stomagen positively regulates stomatal density in Arabidopsis. *Nature*, **463**, 241–244.

Tanaka, K., Choi, J., Cao, Y. and Stacey, G. (2014) Extracellular ATP acts as a damage-associated molecular pattern (DAMP) signal in plants. *Front Plant Sci.* **5**, 446.

Tang, J., Han, Z., Sun, Y., Zhang, H., Gong, X. and Chai, J. (2015) Structural basis for recognition of an endogenous peptide by the plant receptor kinase PEPR1. *Cell Res.* **25**, 110–120.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W. and Su, Z. (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **45**, W122–w129.

Tintor, N., Ross, A., Kanehara, K., Yamada, K., Fan, L., Kemmerling, B., Nurnberger, T., Tsuda, K. and Saijo, Y. (2013) Layered pattern receptor signaling via ethylene and endogenous elicitor peptides during Arabidopsis immunity to bacterial infection. *Proc. Natl. Acad. Sci. U S A*, **110**, 6211–6216.

Tischer, S.V., Wunschel, C., Papacek, M., Kleigrewe, K., Hofmann, T., Christmann, A. and Grill, E. (2017) Combinatorial interaction network of abscisic acid receptors and coreceptors from Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U S A*, **114**, 10280–10285.

Toyota, M., Spencer, D., Sawai-Toyota, S., Jiaqi, W., Zhang, T., Koo, A.J., Howe, G.A. and Gilroy, S. (2018) Glutamate triggers long-distance, calcium-based plant defense signaling. *Science*, **361**, 1112–1115.

Trivilin, A.P., Hartke, S. and Moraes, M.G. (2014). Components of different signalling pathways regulated by a new orthologue of A t PROPEP 1 in tomato following infection by pathogens. *Plant Pathol.* **63**, 1110–1118.

Wang, C., Huang, X., Li, Q., Zhang, Y., Li, J.L. and Mou, Z. (2019) Extracellular pyridine nucleotides trigger plant systemic immunity through a lectin receptor kinase/BAK1 complex. *Nat. Commun.* **10**, 4810.

Yamaguchi, Y. and Huffaker, A. (2011) Endogenous peptide elicitors in higher plants. *Curr. Opin Plant Biol.* **14**, 351–357.

Yamaguchi, Y., Huffaker, A., Bryan, A.C., Tax, F.E. and Ryan, C.A. (2010) PEPR2 is a second receptor for the Pep1 and Pep2 peptides and contributes to defense responses in Arabidopsis. *Plant Cell*, **22**, 508–522.

Yamaguchi, Y., Pearce, G. and Ryan, C.A. (2006) The cell surface leucine-rich repeat receptor for AtPep1, an endogenous peptide elicitor in Arabidopsis, is functional in transgenic tobacco cells. *Proc. Natl. Acad Sci. U S A*, **103**, 10104–10109.

Young, T.E., Meeley, R.B. and Gallie, D.R. (2004) ACC synthase expression regulates leaf performance and drought tolerance in maize. *Plant J.* **40**, 813–825.

Yu, G., Smith, D.K., Zhu, H., Guan, Y. and Lam, T.-T.-Y. (2017a) ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Method Ecol Evolut*, **8**, 28–36.

Yu, X., Feng, B., He, P. and Shan, L. (2017b) From Chaos to harmony: responses and Signaling upon microbial pattern recognition. *Annu. Rev. Phytopathol.* **55**, 109–137.

Zhang, L. and Gleason, C. (2020) Enhancing potato resistance against root-knot nematodes using a plant-defence elicitor delivered by bacteria. *Nat. Plants*, **6**, 625–629.

Zhou, S., Zhang, Y.K., Kremling, K.A. *et al.* (2019) Ethylene signaling regulates natural variation in the abundance of antifungal acetylated diferuloylsucroses and Fusarium graminearum resistance in maize seedling roots. *New Phytol.* **221**, 2096–2111.

Zipfel, C. (2014) Plant pattern-recognition receptors. *Trends Immunol.* **35**, 345–351.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D., Boller, T. and Felix, G. (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell*, **125**, 749–760.

The Plant Journal, (2020), doi: 10.1111/tpj.15022

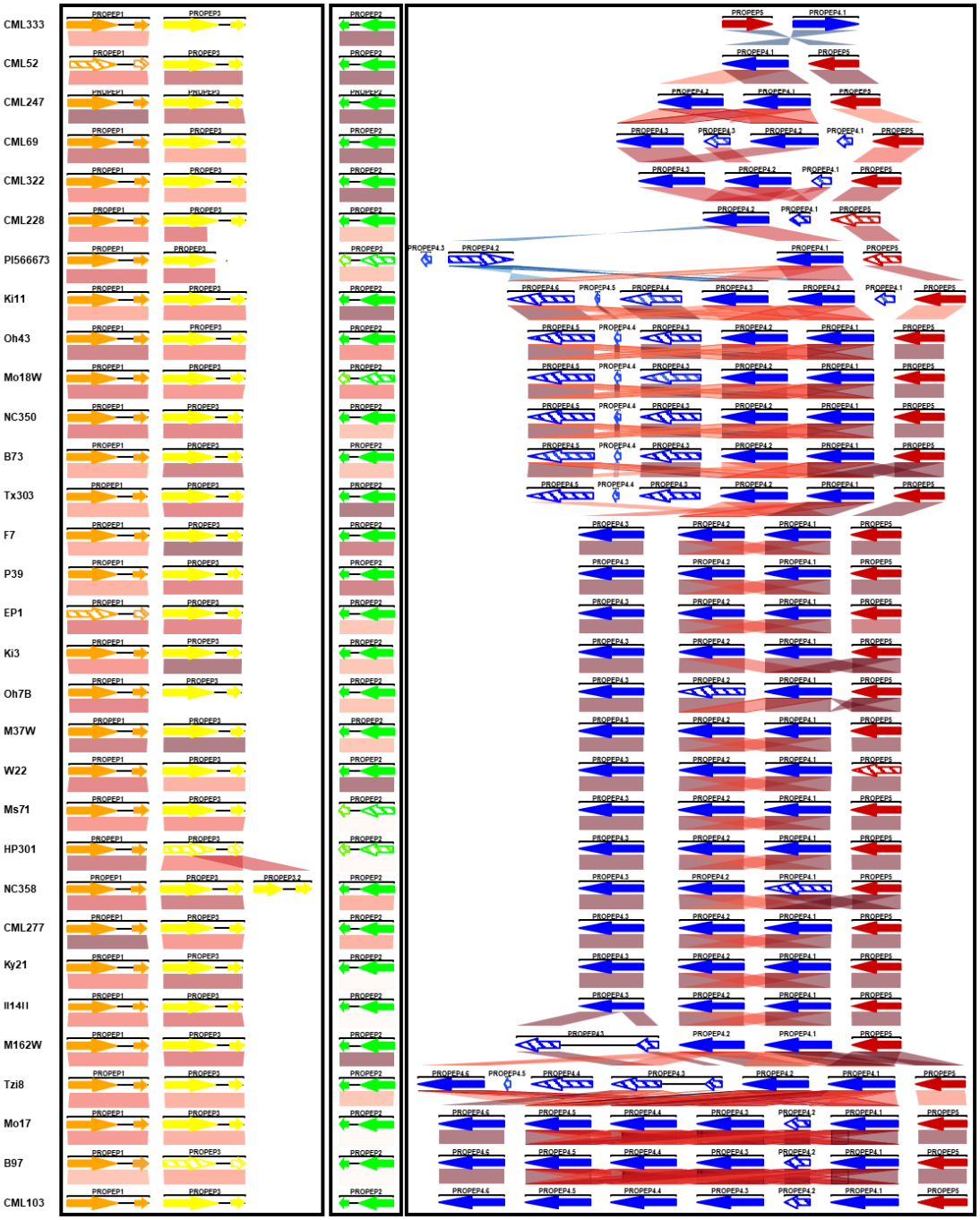


Figure S1. Structure of maize *PROPEP* gene loci in the Nested Association Mapping (NAM) parent lines. Copy number and organization of *ZmPROPEP* genes in each inbred is represented with ZmPROPEP1 designated in orange, *ZmPROPEP2* in green, *ZmPROPEP3* in yellow, *ZmPROPEP4* in blue and *ZmPROPEP5* in red. Genes with predicted deleterious mutations are denoted with stripes rather than solid colors.

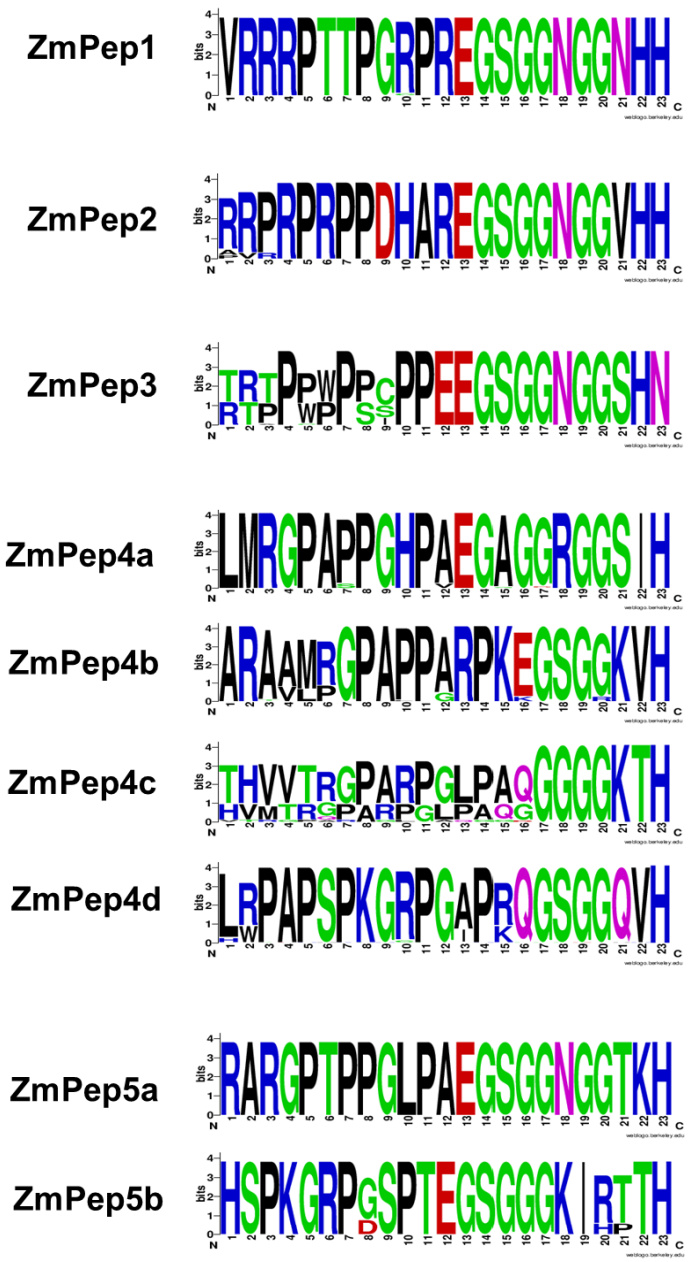


Figure S2. Sequence conservation of individual ZmPeps across NAM parent lines. Logo files represent the consensus sequence for Plant Elicitor Peptides from maize NAM parent lines generated using the sequences for genes identified in Figure S1.

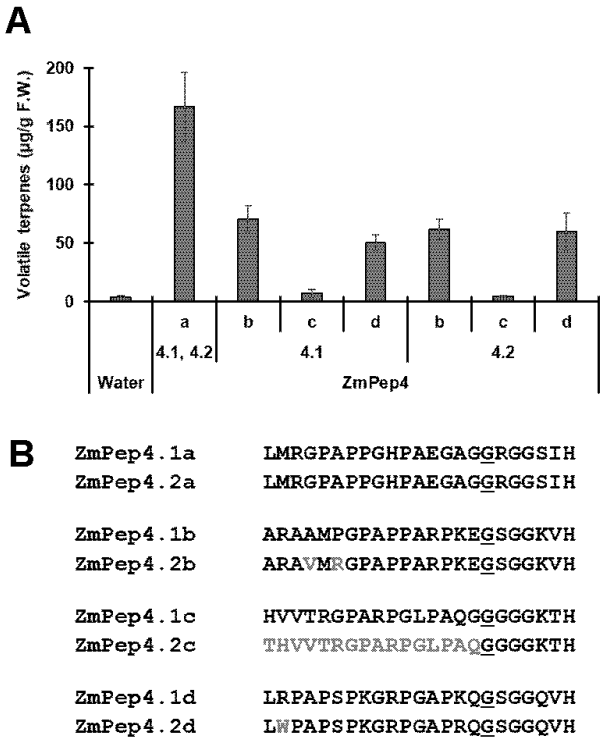


Figure S3. Relative volatile terpene-inducing activity for peptides from derived from the ZmPROPEP4.1 and ZmPROPEP4.2 duplicate precursor genes encoded in the B73 genome.
(A) Relative emission of herbivore-associated volatile terpenes defined as, β-caryophyllene, E-β-farnesene, linalool, dimethylnonatriene (DMNT), trimethyltridecatetraene (TMTT) and α-bergamotene in 21 day old maize leaves 12 h post-treatment with water, or 5 µM solutions of peptides. For all n = 4, error bars represent SEM.
(B) Sequence alignment of each ZmPep pair derived from the ZmPROPEP4.1 versus the ZmPROPEP4.2 precursor. Residues that differ between the two are highlighted in red text.

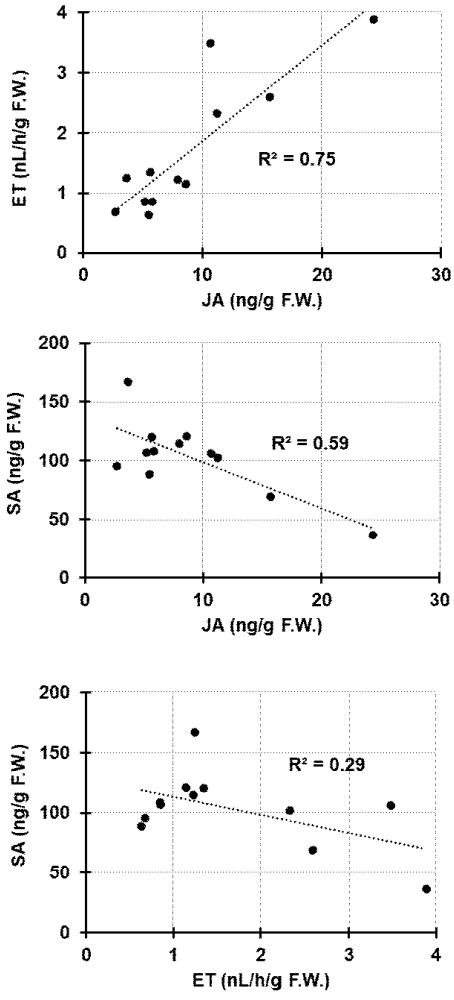


Figure S4. Relationship between relative levels of each phytohormone in leaves treated with water or ZmPeps

(A) JA versus ET

(B) JA versus SA

(C) ET versus SA

ACKNOWLEDGEMENTS

Chapter 2, in full, is a reprint of the material as it appears in The Plant Journal 2020. Poretsky, E., Dressano, K., Weckwerth, P., Ruiz, M., Char, S.N., Shi, D., Abagyan, R., Yang, B. and Huffaker, A. (2020) Differential activities of maize plant elicitor peptides as mediators of immune signaling and herbivore resistance. Plant J, tpj.15022. The dissertation author was the primary investigator and author of this paper.

Chapter 3: MutRank: an R shiny web-application for exploratory targeted mutual rank-based coexpression analyses integrated with user-provided supporting information

MutRank: an R shiny web-application for exploratory targeted mutual rank-based coexpression analyses integrated with user-provided supporting information

Elly Poretsky and Alisa Huffaker

Division of Biology, University of California, San Diego, La Jolla, CA, USA

ABSTRACT

The rapid assignment of genotypes to phenotypes has been a historically challenging process. The discovery of genes encoding biosynthetic pathway enzymes for defined plant specialized metabolites has been informed and accelerated by the detection of gene clusters. Unfortunately, biosynthetic pathway genes are commonly dispersed across chromosomes or reside in genes clusters that provide little predictive value. More reliably, transcript abundance of genes underlying biochemical pathways for plant specialized metabolites display significant coregulation. By rapidly identifying highly coexpressed transcripts, it is possible to efficiently narrow candidate genes encoding pathway enzymes and more easily predict both functions and functional associations. Mutual Rank (MR)-based coexpression analyses in plants accurately demonstrate functional associations for many specialized metabolic pathways; however, despite the clear predictive value of MR analyses, the application is uncommonly used to drive new pathway discoveries. Moreover, many coexpression databases aid in the prediction of both functional associations and gene functions, but lack customizability for refined hypothesis testing. To facilitate and speed flexible MR-based hypothesis testing, we developed MutRank, an R Shiny web-application for coexpression analyses. MutRank provides an intuitive graphical user interface with multiple customizable features that integrates user-provided data and supporting information suitable for personal computers. Tabular and graphical outputs facilitate the rapid analyses of both unbiased and user-defined coexpression results that accelerate gene function predictions. We highlight the recent utility of MR analyses for functional predictions and discoveries in defining two maize terpenoid antibiotic pathways. Beyond applications in biosynthetic pathway discovery, MutRank provides a simple, customizable and user-friendly interface to enable coexpression analyses relating to a breadth of plant biology inquiries. Data and code are available at GitHub: https://github.com/eporetsky/MutRank.

Submitted 25 August 2020
Accepted 7 October 2020
Published 9 November 2020

Corresponding author
Alisa Huffaker, ahuffaker@ucsd.edu

Academic editor
Shawn Gomez

Additional Information and Declarations can be found on page 11

DOI 10.7717/peerj.10264

Subjects Bioinformatics, Genetics, Genomics, Molecular Biology, Plant Science
Keywords Shiny, Transcriptomes, Coexpression analyses, Mutual rank, Gene function prediction, Functional association, Pathway discovery, Specialized metabolism, Customizable, Plant biology

How to cite this article Poretsky E, Huffaker A. 2020. MutRank: an R shiny web-application for exploratory targeted mutual rank-based coexpression analyses integrated with user-provided supporting information. **PeerJ 8:e10264 DOI 10.7717/peerj.10264**

INTRODUCTION

Visually-apparent biological complexity is greatly exceeded by the extreme diversity of specialized metabolites made by organisms for the mediation of essential biotic and abiotic interactions (*Dixon, 2001*; *Gershenzon & Dudareva, 2007*; *Pichersky & Lewinsohn, 2011*). In plants, the ability to identify and control the production of specialized metabolites has significant implications for human health and agriculture; however, efficient tools aiding in biosynthetic pathway discovery remain limited (*Dixon, 2001*; *Moghe & Kruse, 2018*). Clustering of plant specialized metabolism genes has historically been a useful, but not the sole, indicator of functional associations, and has accelerated the discovery of multiple specialized metabolite biosynthetic pathways (*Frey, 1997*; *Osbourn, 2010*; *Boutanaev et al., 2015*). For the discovery of non-clustered metabolic pathway genes, coexpression analyses have emerged as a powerful predictive tool. Genes in specialized metabolic pathways are often highly coregulated based on developmental, spatial, environmental and complex regulatory controls (*Schmelz et al., 2014*; *Lacchini & Goossens, 2020*). Genes that work together in functional specialized metabolic pathways are likely to require transcriptional coregulation and thus resulting patterns used to predict both functional associations and putative gene functions (*Chae et al., 2014*; *Wisecaver et al., 2017*). With increasingly affordable and accessible next generation sequencing technologies, new public and private custom large-scale transcriptomic datasets are routinely generated (*Zhou et al., 2020*). Studies in plants often generate hundreds and even thousands of transcriptomic samples from different genotypes, developmental stages, tissues and physiological conditions to understand traits of agronomic significance (*Sekhon et al., 2011*; *Stelpflug et al., 2016*; *Kremling et al., 2018*; *Machado et al., 2020*). Moreover, genomes and transcriptomes from thousands of plant species are expected to speed large-scale gene expression experiments in poorly understood models (*Twyford, 2018*; *One Thousand Plant Transcriptomes Initiative, 2019*). Public and lab-specific transcriptomic resources are far from static, instead they are continuously expanding and dynamic resources that require flexible tools for rapid and effective analyses.

Many databases and webtools, such as PLEXdb (*Dash et al., 2012*), Genevestigator (*Hruz et al., 2008*), PLANEX (*Yim et al., 2013*), CORNET (*De Bodt et al., 2010*, *2012*), ATTED-II (*Obayashi et al., 2018*), COXPRESdb (*Obayashi et al., 2012*), RiceFREND (*Sato et al., 2013*), ePlant (*Waese et al., 2017*) and STRING (*Szklarczyk et al., 2019*) have been developed to facilitate gene coexpression analyses. Coexpression analyses in studies and databases often use the Pearson's Correlation Coefficient (PCC) as a measure of coexpression. Mutual Rank (MR), the geometric mean of the ranked PCCs between a pair of genes, has been further proposed as an alternative measure of coexpression to PCC (*Obayashi & Kinoshita, 2009*). MR-based coexpression analyses provide better indication of functional associations and are more robust to inconsistencies caused by different microarray data processing methods compared to PCC-based coexpression analyses (*Obayashi & Kinoshita, 2009*). Collective findings have driven some coexpression databases to use MR as the primary measure of coexpression (*Obayashi et al., 2012*, *2018*; *Sato et al., 2013*). When the MR-and PCC-based coexpression databases of multiple plant

species from ATTED-II (*Obayashi et al., 2018*) were converted into coexpression networks and compared, the MR-based coexpression networks were more comparable than PCC-based coexpression networks across species, suggesting that MR-based coexpression networks accurately represent functional associations (*Wisecaver et al., 2017*). MR-based coexpression networks enabled the accurate prediction of clusters enriched for enzymes associated with validated plant specialized metabolic pathways (*Wisecaver et al., 2017*). *Wisecaver et al. (2017)* further demonstrate that MR analyses of transcripts are an improved and powerful tool for the functional prediction of unclustered biosynthetic pathway genes to serve as a springboard for hypothesis testing and validation.

While coexpression databases are useful, few enable flexible hypothesis testing and tool-based simplicity that integrates user-provided data and information. Data integration with coexpression results facilitates the meaningful interpretation of predicted functional associations and assignment of putative gene functions. For example, if a cytochrome P450 monoxygenase (CYP) is hypothesized to perform an oxidation step in a specific biosynthetic pathway, a user might ask "which of all possible *CYP* transcripts is most highly coexpressed with an established pathway gene?". More simply stated, any number of user-defined questions of targeted interest can be precisely examined. For any co-regulated process studied, the identification of 2–3 top candidates from a large gene family can greatly narrow efforts required for defined hypothesis testing and iterative re-testing. Towards this goal, we developed an R Shiny web-application, termed MutRank, to facilitate user control over both targeted and non-targeted MR-based coexpression analyses for rapid hypothesis testing. Using the R Shiny framework, we designed a flexible coexpression analysis platform that combines R packages to easily analyze and integrate user-provided expression data and information. Shiny web-applications are also advantageous for generating highly customizable and user-friendly interfaces that can run on most personal computers. In addition to identifying highly coexpressed genes in any user-provided dataset, MutRank automatically integrates supporting information such as gene annotations, differential-expression data, predicted protein domains and assigned Gene Ontology terms to provide useful tabular and graphical outputs as foundation for empirical hypothesis testing. Confirmed through diverse approaches, targeted and untargeted MR-based coexpression tools were recently leveraged to narrow gene candidates and accurately predict enzymes within multiple maize antibiotic biosynthetic pathways (*Ding et al., 2019*, *2020*). The goal of MutRank is to provide simple, customizable and readily accessible tools to speed research progress by using exploratory targeted coexpression analyses to predict gene functions and functional associations.

METHODS

Software packages and example supporting information used

MutRank was developed as a web application using the Shiny R package (1.4.0.2) (*Chang et al., 2020*) that creates the user interface and manages navigation across the different application components (Fig. 1A). It requires R (3.4.0) and Java (Version 8 Update 261) to be installed by the user, and the following R packages will be automatically installed: shiny (1.4.0.2) (*Chang et al., 2020*), hypergea (1.3.6) (*Boenn, 2018*), ontologyIndex (2.5)

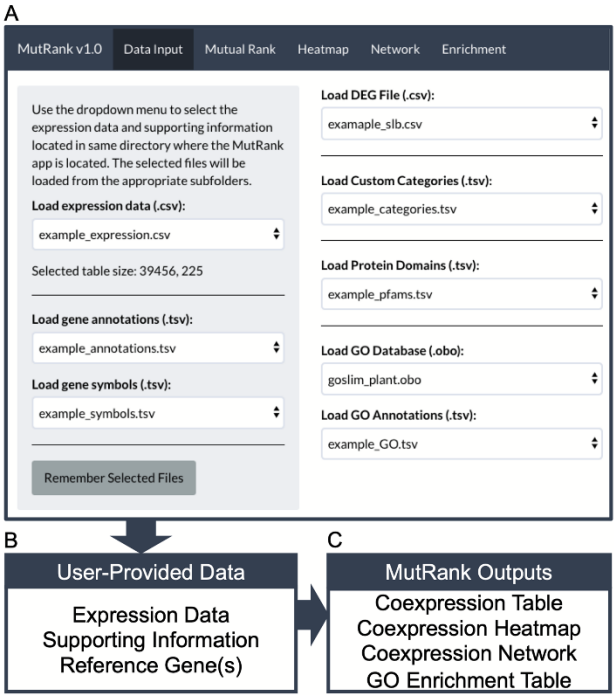


Figure 1 MutRank interface and workflow chart. (A) MutRank workflow starts at the Data Input tab at the top navigation bar that allows the selection of files to load and access different sections of MutRank. In the side panel users can select expression data files, gene descriptions and symbol annotations. In the main panel users can select additional supporting information which includes differential-expression data, custom categories, protein domains, and the Gene Ontology (GO) database along GO assignments. (B) With the user-provided expression data and integrated supporting information users can then select a single target reference gene or gene list to produce a (C) Mutual Rank-based coexpression table and to view the coexpression analysis results as a coexpression heatmap, coexpression network and a GO term enrichment table. Full-size DOI: 10.7717/peerj.10264/fig-1

(*Greene, Richardson & Turro, 2017*), reshape2 (1.4.3) (*Wickham, 2007*), RColorBrewer (1.1-2) (*Neuwirth, 2014*), data.table (1.12.8) (*Dowle & Srinivasan, 2020*), ggplot2 (3.3.0) (*Wickham, 2016*), visNetwork (2.0.9) (*Almende, Thieurmel & Robert, 2019*), igraph (1.2.4.2) (*Csardi & Nepusz, 2005*) and shinythemes (1.1.2) (*Chang, 2018*). To explain the features included in MutRank and to understand the required file structures we provide example expression data and supporting information. All the files used for examples are based on the *Zea mays* inbred B73 (RefGen_v3) genome annotation. The expression data is from the Expanded Maize Gene Expression Atlas (*Stelpflug et al., 2016*) (Fig. 1A; Table S1), gene annotations from the Phytozome database (*Goodstein et al., 2012*) (Fig. 1A; Table S2), and gene symbols from MaizeGDB (*Portwood et al., 2019*) (Fig. 1A; Table S3). Additional supporting information can be selected in the main panel (Fig. 1A).

As an example of analyzing a custom dataset, differential expression data was obtained for maize stems 24 hours after treatment with a fungal pathogen, specifically Southern leaf blight (SLB; *Cochliobolus heterostrophus*) (*Ding et al., 2019*) (Table S4). The predicted Pfam protein domain annotations and GO term assignments are derived from the Phytozome database (*Goodstein et al., 2012*) (Tables S5 and S6). The GO-basic and Plant-GO-Slim ontologies are from the GO Consortium (*Ashburner et al., 2000*; *The Gene Ontology Consortium, 2019*). Lists of maize terpene synthases (TPS) (*Ding et al., 2020*), cytochrome P450s (CYP) (*Ding et al., 2019*) and Pfam protein domains associated with specialized metabolism (SM) (*Wisecaver et al., 2017*) were used as categories to assign to coexpressed genes (Table S7).

Calculating mutual rank values

MutRank was developed as a user-friendly tool to quickly identify the most highly coexpressed genes based on MR values for any reference gene and expression dataset. One of the limitations of MutRank is that it does not calculate all pair-wise MR values. Unlike coexpression databases that pre-calculate all pair-wise MR values (*Obayashi et al., 2012*, *2018*; *Sato et al., 2013*), calculating all pair-wise MR values on the resources available on most personal computers is impractical. Instead, MutRank calculates all PCC values between the user-provided reference gene and all other genes to generate a limited list of genes with the highest PCC values (top 200 genes by default, maximum 1,000) for which it is feasible to calculate MR values. This practical trade-off between whole-genome and targeted coexpression analyses allows MutRank to rapidly complete calculations and to run on the resources of most personal computers. In addition to using a single reference gene, MutRank offers two additional methods for user-defined reference gene sets (Figs. 1B–2). The first method calculates the MR values between all genes in the reference gene set. The second method creates a novel compound reference gene from the average, sum, maximum or minimum expression values of the reference gene set. Using compound reference genes is important for capturing pan-genome patterns with key gene family members displaying highly variable expression across the analyzed germplasm (*Ding et al., 2020*).

Integrating user-provided supporting information

As an exploratory targeted coexpression analysis tool, MutRank integrates user-provided supporting information with the identified list of coexpressed genes (Fig. 1B). Users can provide gene annotations and symbols as easy-to-read information connected to the identified list of coexpressed genes. Additional supporting information in the form of lists of differentially-expressed genes, predicted Pfam domains and assigned Gene Ontology (GO) terms can be integrated with the coexpressed genes. Users can also define custom categories made from lists of genes, Pfam domains or GO terms. The goal of assigning a gene in the MR-based coexpression results as belonging to any of the categories is to have a noticeable indication that the gene is either present in the gene list or is assigned at least one of the Pfam protein domains or GO terms.

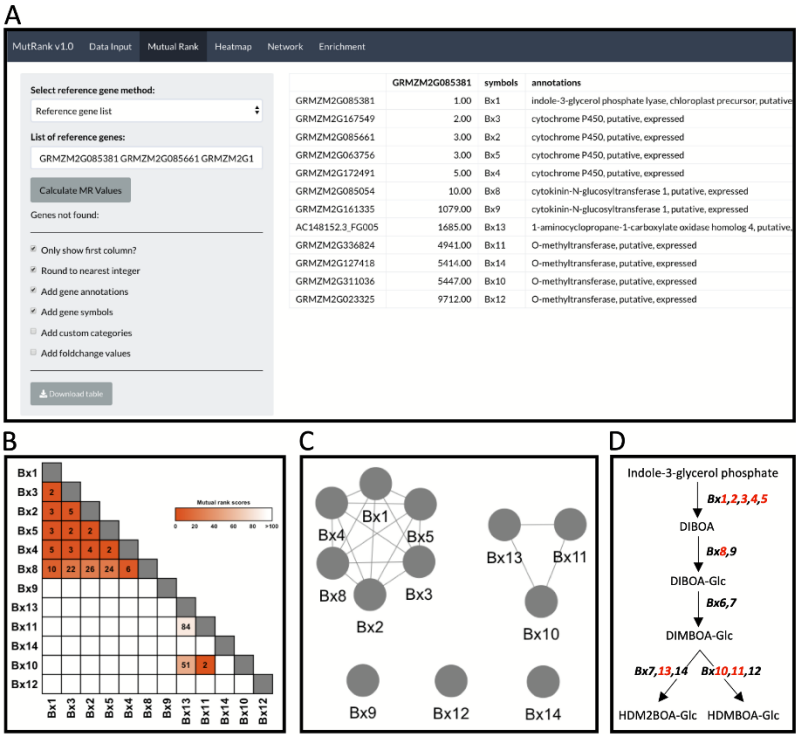


Figure 2 Example workflow 1: validation of MutRank using a characterized biosynthetic pathway. (A) In the Mutual Rank (MR) tab we used the reference gene list method with the characterized known enzymes in the benzoxazinoid (BX) biosynthetic pathway (*Bx1* to *Bx14*, note: *Bx6* and *Bx7* are absent from the example expression data) with default output, but excluding custom categories and fold-change values, to calculate the MR values and produce the MR coexpression table integrated with supporting information. The coexpression analysis results can be presented as a (B) coexpression heatmap and as a (C) coexpression network with an MR < 100 threshold for drawing edges between vertices showing two clusters of coexpressed genes. (D) Summarized diagram of the maize BX biosynthetic pathway with genes that were highly coexpressed designated in red.

Full-size DOI: 10.7717/peerj.10264/fig-2

Tabular and graphical outputs for coexpression analyses

The primary output is provided in the form of an MR coexpression table (Fig. 1C). User-provided supporting information can be automatically integrated into the table in separate columns for each of the coexpressed genes. The results from the MR coexpression table are used as the basis for three additional informative outputs: heatmap, network graph and a GO enrichment table (Fig. 1C). The heatmap, generated using ggplot2 (*Wickham, 2016*), provides an overview of the distribution of MR values among the top coexpressed genes. The R igraph package (*Csardi & Nepusz, 2005*) is used to convert the coexpression table into a coexpression network and to annotate the gene vertices with

user-provided data. The network graph visualization is produced with visNetwork package (*Almende, Thieurmel & Robert, 2019*) which allows the user to explore a dynamic network representation with supporting information. GO term enrichment is calculated using the hypergeometric test based on the GO database and all genes with MR values below a user-provided threshold (default MR < 100). The *P*-values are adjusted for false discovery rate and the results are presented in a separate table.

RESULTS

Example workflow 1: integrating coexpression analyses of genes encoding a specialized metabolic pathway with supporting information

In maize and other important grain crops, benzoxazinoids (BXs) are a highly-studied class of nitrogen-containing specialized metabolites with critical roles in plant protection against both herbivores and pathogens (*Frey, 1997*; *Meihls et al., 2013*; *Wouters et al., 2016*). Genes underlying early steps in the maize BX biosynthetic pathway, namely *Bx1* to *Bx8*, are consitutively expressed in seedlings and drive the production of 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside (DIMBOA-Glc). A majority of these genes, *Bx1* to *Bx5* and *Bx8*, are located together on chromosome 4 and represent the first biosynthetic gene cluster ever described in plants (*Frey, 1997*; *Dutartre, Hilliou & Feyereisen, 2012*). In contrast to largely constitutive production, the late stage BX pathway, namely *Bx10* to *Bx14* and encoded enzymes, display stress-inducible regulation resulting in the conversion of DIMBOA substrates to 2-(2-hydroxy-4,7-dimethoxy-1,4-benzoxazin-3-one)-β-D-glucopyranose (HDMBOA-Glc) and 2-(2-hydroxy-4,7,8-trimethoxy-1,4-benzoxazin-3-one)-β-D-glucopyranose (HDM_2BOA-Glc), which upon aglycone liberation (HDMBOA and HDM_2BOA) result in highly unstable bioactive molecules (*Maresh, Zhang & Lynn, 2006*; *Meihls et al., 2013*; *Wouters et al., 2016*). While displaying complex regulation of early- and late-stage *Bx* genes influenced by development and biotic stress (*Cambier, Hance & De Hoffmann, 2000*; *Wouters et al., 2016*), BX1 to BX14 collectively catalyze the production of multiple glucoside conjugates that can ultimately act as aglycone defenses (*Frey, 1997*; *Jonczyk et al., 2008*; *Meihls et al., 2013*; *Handrick et al., 2016*). The gene *Bx1* encodes an indole-3-glycerol phosphate lyase that cleaves indole-3-glycerolphosphate into free indole, acting as the first committed step in the pathway (*Frey, 1997*).

As an example to demonstrate both the power and remaining challenges of using Mutual Ranks to associate pathway genes to one another, we use the reference gene list method to investigate the coexpression of *Bx1* with other *Bx* pathway genes (Fig. 2A; Table S8). Users can select which supporting information to automatically integrate with the MR coexpression table generated (Fig. 2A). The final coexpression table includes columns with the MR values in reference to the first gene in the list (i.e., *Bx1*), gene symbols and gene annotations, and excludes the categories and fold-change columns (Fig. 2A). *Bx6* and *Bx7* were excluded from the coexpression analysis as they were not included in the expression dataset used for this analysis. The coexpression results in the table can be visualized as a coexpression heatmap that readily reveals the highly coexpressed gene

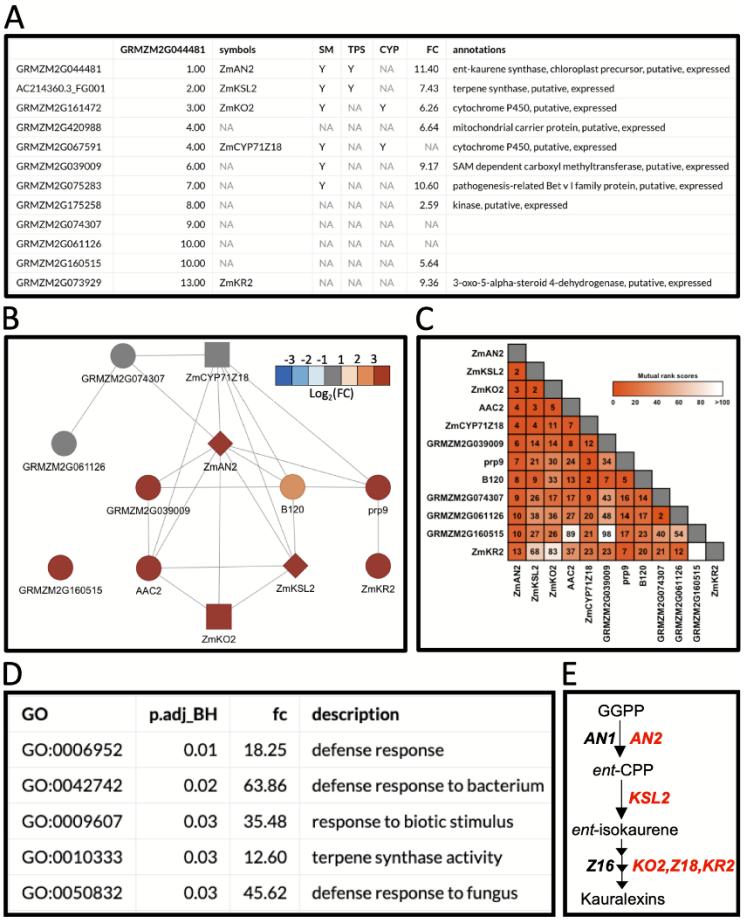
cluster of *Bx1* through *Bx5* and *Bx8*, as well as separate coexpression of *Bx10*, *Bx11* and *Bx13* with one another (Fig. 2B). Similar association patterns can also be observed using an interactive coexpression network with an MR < 100 threshold for drawing edges between vertices (Fig. 2C). Using the validated BX pathway as a simplistic MutRank example, we demonstrate the following: (1) the ease of observing strong co-expression of early *Bx* pathway genes, (2) the partial coexpression of late *Bx* pathway genes and (3) remaining challenges of bioinformatically-connecting complex pathways that display differential regulation of early and late steps (Fig. 2D) (*Meihls et al., 2013*; *Handrick et al., 2016*). Importantly, biosynthetic pathways function within the complex context of a living cell. The value in confirming established coexpression patterns is to first undertand how the user-defined dataset is performing. When compelling, these results then encourage further interrogation to address diverse hypotheses and complex surrounding processes, potentially identifying coexpressed transcription factors, transporters, or detoxification enzymes to investigate (*Lacchini & Goossens, 2020*).

Example workflow 2: using MutRank to predict enzymes in specialized metabolism

In the first example, we used BX-related defenses which have been studied in maize and other cereals for over 60 years (*Virtanen et al., 1955*; *Smissman, LaPidus & Beck, 1957*). More recently, maize diterpenoid pathways have been implicated in diverse protective roles providing fungal, insect and drought resistance (*Schmelz et al., 2011*; *Vaughan et al., 2015*; *Christensen et al., 2018*; *Ding et al., 2019*). Biosynthesis of protective *ent*-kaurane-related diterpenoids in maize, termed kauralexins, are mediated by multi-gene terpene synthase (TPS) and cytochrome P450 (CYP) families. Using MR-based coexpression analyses for discovery purposes (*Ding et al., 2019*) we examined one reference gene termed anther ear 2 (*ZmAN2*) (Table S8), that encodes an *ent*-copalyl diphosphate synthase (*ent*-CPS) responsible for the cyclization of geranylgeranyl diphosphate into bicyclic pathway precursor *ent*-copalyl diphosphate (*ent*-CPP) (*Harris et al., 2005*). Derived from two different genes encoding *ent*-CPS, *ent*-CPP is a key substrate shared by the kauralexin, dolabralexin and gibberellin biosynthetic pathways in maize (*Mafu et al., 2018*; *Ding et al., 2019*). Using *ZmAN2* as a reference gene, we calculated the non-targeted MR values between the top 200 coexpressed genes and integrated the supporting information (Fig. 3A). For simplification, we then selected the first 12 coexpressed genes and identified 1 TPS gene (Figs. 3A and 3B: diamond shaped vertex), a type I diterpene synthase: kaurene synthase-like 2 (*ZmKSL2*) and 2 CYP genes (Figs. 3A and 3B: square shaped vertices), *ZmCYP71Z18* and kaurene oxidase 2 (*ZmKO2*) that were highly coexpressed (Figs. 3A–3C). A GO-term enrichment analysis of the MR-based coexpression results using the GO-basic database revealed an enrichment of terms associated with defense responses and terpene synthesis (Fig. 3D). With candidates identified through similar MR-based coexpression relationships to those currently presented (Figs. 3A–3E), a recent study of kauralexin biosynthetic enzymes were systematically validated using genome wide association studies, heterologous enzyme co-expression assays, proteomics and characterization of defined genetic mutants (*Ding et al., 2019*). Two additional genes with

A

	GRMZM2G044481	symbols	SM	TPS	CYP	FC	annotations
GRMZM2G044481	1.00	ZmAN2	Y	Y	NA	11.40	ent-kaurene synthase, chloroplast precursor, putative, expressed
AC214360.3_FG001	2.00	ZmKSL2	Y	Y	NA	7.43	terpene synthase, putative, expressed
GRMZM2G161472	3.00	ZmKO2	Y	NA	Y	6.26	cytochrome P450, putative, expressed
GRMZM2G420988	4.00	NA	NA	NA	NA	6.64	mitochondrial carrier protein, putative, expressed
GRMZM2G067591	4.00	ZmCYP71Z18	Y	NA	Y	NA	cytochrome P450, putative, expressed
GRMZM2G039009	6.00	NA	Y	NA	NA	9.17	SAM dependent carboxyl methyltransferase, putative, expressed
GRMZM2G075283	7.00	NA	Y	NA	NA	10.60	pathogenesis-related Bet v I family protein, putative, expressed
GRMZM2G175258	8.00	NA	NA	NA	NA	2.59	kinase, putative, expressed
GRMZM2G074307	9.00	NA	NA	NA	NA	NA	
GRMZM2G061126	10.00	NA	NA	NA	NA	NA	
GRMZM2G160515	10.00	NA	NA	NA	NA	5.64	
GRMZM2G073929	13.00	ZmKR2	NA	NA	NA	9.36	3-oxo-5-alpha-steroid 4-dehydrogenase, putative, expressed



D

GO	p.adj_BH	fc	description
GO:0006952	0.01	18.25	defense response
GO:0042742	0.02	63.86	defense response to bacterium
GO:0009607	0.03	35.48	response to biotic stimulus
GO:0010333	0.03	12.60	terpene synthase activity
GO:0050832	0.03	45.62	defense response to fungus

Figure 3 Example workflow 2: using MutRank to predict enzymes in specialized metabolism. (A) Using the kauralexin biosynthetic gene *ANTHER EAR 2* (*AN2*) as a single reference gene, a Mutual Rank (MR)-based coexpression table was generated for the 200 most highly coexpressed genes (first 12 genes are shown) with the integrated supporting information. Using the first 12 genes in the list we generated a (B) coexpression network figure, with an MR < 10 threshold for drawing edges between vertices showing a cluster of coexpressed genes and a (C) coexpression heatmap. (A and B) Genes belonging to a category are denoted with "Y"; the categories used are SM (specialized metabolism), TPS (terpene synthases, diamond shape vertices) and CYP (cytochrome P450s, square shaped vertices). (A and B) Corresponding expression fold change (FC) increase 24 h after pathogen inoculation. (D) Results of the Gene Ontology (GO) term enrichment analysis using the GO-basic database and all genes with MR < 100 are over-represented for terms associated with biotic stress responses. *P*-values were calculated using a hypergeometric test and adjusted using the Bonferroni–Holm method. (E) Summarized diagram of the maize kauralexin biosynthetic pathway showing genes highly coexpressed with the reference gene *AN2* in red. Full-size DOI: 10.7717/peerj.10264/fig-3

defined roles in kauralexin biosynthesis that did not match any of the supporting information categories are the *ZmCYP71Z16* that is absent from the currently selected expression dataset and the coexpressed *kauralexin reductase2* (*ZmKR2*) encoding a 5α-steroid reductase that saturates B-series kauralexins (Figs. 3A–3C) (*Ding et al., 2019*). Together the combined use of MR analyses with biochemistry and defined genetic mutants defined roles for *ZmAn2*, *ZmKSL2*, *ZmKO2*, *ZmKR2*, *ZmCYP71Z18* and ZmCYP71Z16 in kauralexin biosynthesis and anti-pathogen defense enabling rapid assembly of the entire pathway (Fig. 3E) (*Ding et al., 2019*). Additional genes identified in the MR-based coexpression analysis encode predicted carrier proteins, pathogenesis-related proteins and kinases that might further contribute to the regulation and transport of diterpenoids (Figs. 3A–3C). In summary, straightforward MR analyses via MutRank provide a powerful starting point for defining networks surrounding specialized metabolism.

DISCUSSION

MutRank is a user-friendly and powerful tool for exploratory targeted gene coexpression analyses. MutRank enables the simple calculation of MR values for any reference gene or gene set from user-provided expression data. The Shiny web-application interface is ideal for combining MR-based coexpression analyses with useful R packages that produce informative tabular and graphical outputs. We implemented a number of features that allow users to integrate supporting information with the results of the coexpression analyses to facilitate prediction of putative gene functions. Example workflow 1 surveyed genes in the well-established maize BX biosynthetic pathway. Many of these genes were identified and characterized without the benefit of large-scale transcriptomic data (*Frey, 1997*). The lack of coexpression connections between early and late stage *Bx* biosynthetic genes (Figs. 2A–2D) likely provides a partial explaination for the relatively recent discovery of the terminal steps (*Meihls et al., 2013*; *Handrick et al., 2016*).

Public coexpression databases and tools, such as MutRank, provide intuitive user control over MR-based coexpression analyses to drive predictions and hypothesis testing of genes with currently unknown functions. Example workflow 2 was given as an example where MR-based coexpression analyses were used to guide recent hypothesis testing, and through a combination of diverse approaches, were demonstrated to correctly predict gene functions in the maize kauralexin antibiotic pathway (*Ding et al., 2019*). Importantly, we note here that custom use of further expression datasets were used to correctly predict the function of an additional kauralexin biosynthetic genes (*ZmCYP71Z16*) within the pathway using MR-based coexpression analyses (*Ding et al., 2019*). In *Ding et al. (2019)* the expression datasets were derived from the National Center for Biotechnology Information Sequence Read Archive project IDs SRP115041 and SRP011480. This esoteric detail speaks to an essentail point. Different MR coexpression patterns can be found in related datasets depending on sample size, plant growth conditions, genotypes used, tissue types, cell types, developemental age, presence or absence of biotic or aboitic stress and countless other factors important to the questions being examined. Given this, aggregate estimations of gene co-expression available on public websites typically fall short in facilitating elucidation of relationships of interest. Rapid progress requires flexible control

over the analyses of precise data subsets or of larger aggregated datasets for cross-comparison. MutRank allows for a large number of different datasets to be selected, and quickly analzyed and assessed. Most commonly, the search for meaningful coexpression relationships, whether of biosynthetic genes or for more complex regulatory processes, is a guided and highly iterative discovery process, relying on partial insights from related experimental systems. A common goal is to generate high-quality gene candidates for improved hypothesis testing that ultimately informs more expensive and time-consuming *in planta* analyses of defined mutants. As a further recent example, MR-based coexpression analyses were leveraged and played a key role in defining and disentangling a challenging 10-gene maize sesquiterpenoid antibiotic pathway partially sharing kauralexin biosynthetic genes (*Ding et al., 2020*). Research progress in plant specialized metabolism requires rapid, flexible and easy-to-use tools, through which diverse users of varying expertise levels can quickly compare results generated from public or customized user-provided datasets. We now routinely utilize MutRank as a rapid tool for exploratory targeted coexpression analyses facilitating the prediction of functional associations and putative gene functions. The goal of our current effort was to expand the ease and use of the R Shiny web-application tools to facilitate efforts of any biologists who seek to connect coregulated genes to important phenotypes.

CONCLUSION

The MutRank R Shiny web application provides an efficient, flexible and simple tool for conducting hypothesis-driven MR-based coexpression analyses. To enable rapid functional discovery, MutRank analyses are integrated with multiple customizable features for narrowing and prioritizing candidate genes and for hypothesis testing in predicted biochemical functions.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by NSF-IOS PBI CAREER #1943591 and USDA NIFA AFRI #2018-67013-28125. Elly Poretsky was additionally funded by the Cell and Molecular Genetics (CMG) Training Program at the University of California, San Diego. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
NSF-IOS PBI CAREER: #1943591.
USDA NIFA AFRI: #2018-67013-28125.
University of California, San Diego.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Elly Poretsky conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Alisa Huffaker conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

Data are available in the Supplemental Tables and at GitHub. Code is available at GitHub: https://github.com/eporetsky/MutRank.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.10264#supplemental-information.

REFERENCES

Almende BV, Thieurmel B, Robert T. 2019. visNetwork: network visualization using vis.js library. *Available at https://cran.r-project.org/web/packages/visNetwork/index.html.*

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. *Nature Genetics* **25(1)**:25–29 DOI 10.1038/75556.

Boenn M. 2018. Hypergea: hypergeometric tests. *Available at https://cran.r-project.org/web/packages/hypergea/index.html.*

Boutanaev AM, Moses T, Zi J, Nelson DR, Mugford ST, Peters RJ, Osbourn A. 2015. Investigation of terpene diversification across multiple sequenced plant genomes. *Proceedings of the National Academy of Sciences of the United States of America* **112(1)**:E81–E88 DOI 10.1073/pnas.1419547112.

Cambier V, Hance T, De Hoffmann E. 2000. Variation of DIMBOA and related compounds content in relation to the age and plant organ in maize. *Phytochemistry* **53(2)**:223–229 DOI 10.1016/S0031-9422(99)00498-7.

Chae L, Kim T, Nilo-Poyanco R, Rhee SY. 2014. Genomic signatures of specialized metabolism in plants. *Science* **344(6183)**:510–513 DOI 10.1126/science.1252076.

Chang W. 2018. Shinythemes: themes for shiny. *Available at https://cran.r-project.org/web/packages/shinythemes/index.html.*

Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J. 2020. Shiny: web application framework for R. *Available at https://shiny.rstudio.com/.*

Christensen SA, Sims J, Vaughan MM, Hunter C, Block A, Willett D, Alborn HT, Huffaker A, Schmelz EA. 2018. Commercial hybrids and mutant genotypes reveal complex protective roles for inducible terpenoid defenses in maize. *Journal of Experimental Botany* **69(7)**:1693–1705 DOI 10.1093/jxb/erx495.

Csardi G, Nepusz T. 2005. The igraph software package for complex network research. *Complex Systems* **1695(5)**:1–9.

Dash S, Van Hemert J, Hong L, Wise RP, Dickerson JA. 2012. PLEXdb: gene expression resources for plants and plant pathogens. *Nucleic Acids Research* **40(D1)**:D1194–D1201 DOI 10.1093/nar/gkr938.

De Bodt S, Carvajal D, Hollunder J, Van den Cruyce J, Movahedi S, Inzé D. 2010. CORNET: a user-friendly tool for data mining and integration. *Plant Physiology* **152(3)**:1167–1179 DOI 10.1104/pp.109.147215.

De Bodt S, Hollunder J, Nelissen H, Meulemeester N, Inzé D. 2012. CORNET 2.0: integrating plant coexpression, protein-protein interactions, regulatory interactions, gene associations and functional annotations. *New Phytologist* **195(3)**:707–720 DOI 10.1111/j.1469-8137.2012.04184.x.

Ding Y, Murphy KM, Poretsky E, Mafu S, Yang B, Char SN, Christensen SA, Saldivar E, Wu M, Wang Q, Ji L, Schmitz RJ, Kremling KA, Buckler ES, Shen Z, Briggs SP, Bohlmann J, Sher A, Castro-Falcon G, Hughes CC, Huffaker A, Zerbe P, Schmelz EA. 2019. Multiple genes recruited from hormone pathways partition maize diterpenoid defences. *Nature Plants* **5(10)**:1043–1056 DOI 10.1038/s41477-019-0509-6.

Ding Y, Weckwerth PR, Poretsky E, Murphy KM, Sims J, Saldivar E, Christensen SA, Char SN, Yang B, Tong A, Shen Z, Kremling KA, Buckler ES, Kono T, Nelson DR, Bohlmann J, Bakker MG, Vaughan MM, Khalil AS, Betsiashvili M, Briggs SP, Zerbe P, Schmelz EA, Huffaker A. 2020. Genetic elucidation of complex biochemical traits mediating maize innate immunity. *bioRxiv* DOI 10.1101/2020.03.04.977355.

Dixon RA. 2001. Natural products and plant disease resistance. *Nature* **411(6839)**:843–847 DOI 10.1038/35081178.

Dowle M, Srinivasan A. 2020. data.table: extension of 'data.frame'. *Available at https://cran.r-project.org/web/packages/data.table/index.html.*

Dutartre L, Hilliou F, Feyereisen R. 2012. Phylogenomics of the benzoxazinoid biosynthetic pathway of Poaceae: gene duplications and origin of the Bx cluster. *BMC Evolutionary Biology* **12(1)**:64 DOI 10.1186/1471-2148-12-64.

Frey M. 1997. Analysis of a chemical plant defense mechanism in grasses. *Science* **277(5326)**:696–699 DOI 10.1126/science.277.5326.696.

Gershenzon J, Dudareva N. 2007. The function of terpene natural products in the natural world. *Nature Chemical Biology* **3(7)**:408–414 DOI 10.1038/nchembio.2007.5.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* **40(D1)**:D1178–D1186 DOI 10.1093/nar/gkr944.

Greene D, Richardson S, Turro E. 2017. ontologyX: a suite of R packages for working with ontological data. *Bioinformatics* **33**:1104–1106 DOI 10.1093/bioinformatics/btw763.

Handrick V, Robert CAM, Ahern KR, Zhou S, Machado RAR, Maag D, Glauser G, Fernandez-Penny FE, Chandran JN, Rodgers-Melnick E, Schneider B, Buckler ES, Boland W, Gershenzon J, Jander G, Erb M, Köllner TG. 2016. Biosynthesis of 8-O-methylated benzoxazinoid defense compounds in maize. *Plant Cell* **28**:1682–1700 DOI 10.1105/tpc.16.00065.

Harris LJ, Saparno A, Johnston A, Prisic S, Xu M, Allard S, Kathiresan A, Ouellet T, Peters RJ. 2005. The Maize An2 gene is induced by Fusarium attack and encodes an ent-copalyl diphosphate synthase. *Plant Molecular Biology* **59(6)**:881–894 DOI 10.1007/s11103-005-1674-8.

Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P. 2008. Genevestigator V3: a reference expression database for the

meta-analysis of transcriptomes. *Advances in Bioinformatics* **2008**:1–5 DOI 10.1155/2008/420747.

Jonczyk R, Schmidt H, Osterrieder A, Fiesselmann A, Schullehner K, Haslbeck M, Sicker D, Hofmann D, Yalpani N, Simmons C, Frey M, Gierl A. 2008. Elucidation of the final reactions of DIMBOA-glucoside biosynthesis in maize: characterization of *Bx6* and *Bx7*. *Plant Physiology* **146(3)**:1053–1063 DOI 10.1104/pp.107.111237.

Kremling KAG, Chen S-Y, Su M-H, Lepak NK, Romay MC, Swarts KL, Lu F, Lorant A, Bradbury PJ, Buckler ES. 2018. Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature* **555(7697)**:520–523 DOI 10.1038/nature25966.

Lacchini E, Goossens A. 2020. Combinatorial control of plant specialized metabolism: mechanisms, functions, and consequences. *Annual Review of Cell and Developmental Biology* **36(1)**:1–23 DOI 10.1146/annurev-cellbio-011620-031429.

Machado FB, Moharana KC, Almeida-Silva F, Gazara RK, Pedrosa-Silva F, Coelho FS, Grativol C, Venancio TM. 2020. Systematic analysis of 1298 RNA-seq samples and construction of a comprehensive soybean (*Glycine max*) expression atlas. *Plant Journal* **103(5)**:1894–1909 DOI 10.1111/tpj.14850.

Mafu S, Ding Y, Murphy KM, Yaacoobi O, Addison JB, Wang Q, Shen Z, Briggs SP, Bohlmann J, Castro-Falcon G, Hughes CC, Betsiashvili M, Huffaker A, Schmelz EA, Zerbe P. 2018. Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize. *Plant Physiology* **176(4)**:2677–2690 DOI 10.1104/pp.17.01351.

Maresh J, Zhang J, Lynn DG. 2006. The innate immunity of maize and the dynamic chemical strategies regulating two-component signal transduction in *Agrobacterium tumefaciens*. *ACS Chemical Biology* **1(3)**:165–175 DOI 10.1021/cb600051w.

Meihls LN, Handrick V, Glauser G, Barbier H, Kaur H, Haribal MM, Lipka AE, Gershenzon J, Buckler ES, Erb M, Kollner TG, Jander G. 2013. Natural variation in maize aphid resistance is associated with 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one glucoside methyltransferase activity. *Plant Cell* **25(6)**:2341–2355 DOI 10.1105/tpc.113.112409.

Moghe GD, Kruse LH. 2018. The study of plant specialized metabolism: challenges and prospects in the genomics era. *American Journal of Botany* **105(6)**:959–962 DOI 10.1002/ajb2.1101.

Neuwirth E. 2014. RColorBrewer: colorBrewer palettes. *Available at https://cran.r-project.org/web/packages/RColorBrewer/index.html*.

Obayashi T, Aoki Y, Tadaka S, Kagaya Y, Kinoshita K. 2018. ATTED-II in 2018: a plant coexpression database based on investigation of the statistical property of the mutual rank index. *Plant and Cell Physiology* **59(2)**:440 DOI 10.1093/pcp/pcx209.

Obayashi T, Kinoshita K. 2009. Rank of correlation coefficient as a comparable measure for biological significance of gene coexpression. *DNA Research* **16(5)**:249–260 DOI 10.1093/dnares/dsp016.

Obayashi T, Okamura Y, Ito S, Tadaka S, Motoike IN, Kinoshita K. 2012. COXPRESdb: a database of comparative gene coexpression networks of eleven species for mammals. *Nucleic Acids Research* **41(D1)**:D1014–D1020 DOI 10.1093/nar/gks1014.

One Thousand Plant Transcriptomes Initiative. 2019. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**:679–685 DOI 10.1038/s41586-019-1693-2.

Osbourn A. 2010. Gene clusters for secondary metabolic pathways: an emerging theme in plant biology—Figure 1. *Plant Physiology* **154(2)**:531–535 DOI 10.1104/pp.110.161315.

Pichersky E, Lewinsohn E. 2011. Convergent evolution in plant specialized metabolism. *Annual Review of Plant Biology* **62(1)**:549–566 DOI 10.1146/annurev-arplant-042110-103814.

Portwood JL, Woodhouse MR, Cannon EK, Gardiner JM, Harper LC, Schaeffer ML, Walsh JR, Sen TZ, Cho KT, Schott DA, Braun BL, Dietze M, Dunfee B, Elsik CG, Manchanda N, Coe E, Sachs M, Stinard P, Tolbert J, Zimmerman S, Andorf CM. 2019. MaizeGDB 2018: the maize multi-genome genetics and genomics database. *Nucleic Acids Research* **47(D1)**:D1146–D1154 DOI 10.1093/nar/gky1046.

Sato Y, Namiki N, Takehisa H, Kamatsuki K, Minami H, Ikawa H, Ohyanagi H, Sugimoto K, Itoh J-I, Antonio BA, Nagamura Y. 2013. RiceFREND: a platform for retrieving coexpressed gene networks in rice. *Nucleic Acids Research* **41(D1)**:D1214–D1221 DOI 10.1093/nar/gks1122.

Schmelz EA, Huffaker A, Sims JW, Christensen SA, Lu X, Okada K, Peters RJ. 2014. Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant Journal* **79(4)**:659–678 DOI 10.1111/tpj.12436.

Schmelz EA, Kaplan F, Huffaker A, Dafoe NJ, Vaughan MM, Ni X, Rocca JR, Alborn HT, Teal PE. 2011. Identity, regulation, and activity of inducible diterpenoid phytoalexins in maize. *Proceedings of the National Academy of Sciences of the United States of America* **108(13)**:5455–5460 DOI 10.1073/pnas.1014714108.

Sekhon RS, Lin H, Childs KL, Hansey CN, Buell CR, De Leon N, Kaeppler SM. 2011. Genome-wide atlas of transcription during maize development: maize gene atlas. *Plant Journal* **66(4)**:553–563 DOI 10.1111/j.1365-313X.2011.04527.x.

Smissman E, LaPidus J, Beck S. 1957. Notes—corn plant resistance factor. *Journal of Organic Chemistry* **22(2)**:220 DOI 10.1021/jo01353a036.

Stelpflug SC, Sekhon RS, Vaillancourt B, Hirsch CN, Buell CR, De Leon N, Kaeppler SM. 2016. An expanded maize gene expression atlas based on RNA sequencing and its use to explore root development. *Plant Genome* **9(1)**:1–16 DOI 10.3835/plantgenome2015.04.0025.

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Von Mering C. 2019. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research* **47(D1)**:D607–D613 DOI 10.1093/nar/gky1131.

The Gene Ontology Consortium. 2019. The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Research* **47(D1)**:D330–D338 DOI 10.1093/nar/gky1055.

Twyford AD. 2018. The road to 10,000 plant genomes. *Nature Plants* **4(6)**:312–313 DOI 10.1038/s41477-018-0165-2.

Vaughan MM, Christensen S, Schmelz EA, Huffaker A, Mcauslane HJ, Alborn HT, Romero M, Allen LH, Teal PEA. 2015. Accumulation of terpenoid phytoalexins in maize roots is associated with drought tolerance: maize root phytoalexins play a role in drought tolerance. *Plant, Cell & Environment* **38(11)**:2195–2207 DOI 10.1111/pce.12482.

Virtanen AI, Hietala PK, Lundén R, Prydz H. 1955. 2(3)-Benzoxazolinone, an anti-fusarium factor in rye seedlings. *Acta Chemica Scandinavica* **9**:1543–1544 DOI 10.3891/acta.chem.scand.09-1543b.

Waese J, Fan J, Pasha A, Yu H, Fucile G, Shi R, Cumming M, Kelley LA, Sternberg MJ, Krishnakumar V, Ferlanti E, Miller J, Town C, Stuerzlinger W, Provart NJ. 2017. ePlant: visualizing and exploring multiple levels of data for hypothesis generation in plant biology. *Plant Cell* **29(8)**:1806–1821 DOI 10.1105/tpc.17.00073.

Wickham H. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* **21(12)**:1–20 DOI 10.18637/jss.v021.i12.

Wickham H. 2016. *ggplot2: elegant graphics for data analysis.* New York: Springer.

Wisecaver JH, Borowsky AT, Tzin V, Jander G, Kliebenstein DJ, Rokas A. 2017. A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *Plant Cell* **29(5)**:944–959 DOI 10.1105/tpc.17.00009.

Wouters FC, Blanchette B, Gershenzon J, Vassão DG. 2016. Plant defense and herbivore counter-defense: benzoxazinoids and insect herbivores. *Phytochemistry Reviews* **15(6)**:1127–1151 DOI 10.1007/s11101-016-9481-1.

Yim WC, Yu Y, Song K, Jang CS, Lee B-M. 2013. PLANEX: the plant co-expression database. *BMC Plant Biology* **13(1)**:1–9 DOI 10.1186/1471-2229-13-83.

Zhou P, Li Z, Magnusson E, Gomez Cano F, Crisp PA, Noshay JM, Grotewold E, Hirsch CN, Briggs SP, Springer NM. 2020. Meta gene regulatory networks in maize highlight functionally relevant regulatory interactions. *Plant Cell* **32(5)**:1377–1396 DOI 10.1105/tpc.20.00080.

ACKNOWLEDGEMENTS

Chapter 3, in full, is a reprint of the material as it appears in PeerJ 2020. Poretsky, E. and Huffaker, A. (2020) MutRank: an R shiny web-application for exploratory targeted mutual rank-based coexpression analyses integrated with user-provided supporting information. PeerJ, 8, e10264. The dissertation author was the primary investigator and author of this paper.