UNIVERSITY OF CALIFORNIA, SAN DIEGO

The ethylene transcriptional response mediated by ETHYLENE INSENSITIVE3 in *Arabidopsis thaliana*

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

in

Biology

by

Katherine Noelani Chang

Committee in charge:

Professor Joseph R. Ecker, Chair
Professor Stephen P. Briggs
Professor Joanne Chory
Professor Trey Ideker
Professor Amy Kiger

2011

The Dissertation of Katherine Noelani Chang is approved, and it is acceptable in quality and form for publication on microfilm or electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2011

DEDICATION

This dissertation is dedicated to JEC and CLD

because this work would not have been possible without their support.

EPIGRAPH

Tug on anything at all and you'll find it

connected to everything else in the universe.

John Muir

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Joseph Ecker for giving me the opportunity to work with him. He has been an incredible resource and has challenged me to think critically and develop a broader scientific perspective. I thank my committee members, Joanne Chory, Steve Briggs, Trey Ideker, and Amy Kiger for their advice, support and patience. The guidance of my committee has been key to the completion of this work.

Joe provided me access to high technology research and a great atmosphere to work. Members of his lab are excellent and diverse researchers and I have learned a lot from each of them. I would like to thank current members Hong Qiao, Bob Schmitz, Junshi Yazaki, Ryan Lister, Joe Nery, Ronan O'Malley, Mary Galli, Cesar Barragan, and Mark Urich, and past members Brian Gregory, Ramlah Nehring, Hai Li, Liz Colangelo, and Paul Shinn. Discussions with various members of the Salk Plant Biology department have also been very helpful. I appreciate all that Tsegaye Dabi has done to keep the Plant Biology department at Salk functioning. Kim Emerson and Nancy Benson have been indispensible, helping with all things administrative-related. Students Jerry Chan, Michael Ho, and volunteer Abdullah Jamali have each contributed to the larger scope of my work.

My bioinformatician collaborators Gary Hon, Dwight Kuo, and Mattia Pelizzola, have contributed to this work significantly. On the experimental side, Bob Schmitz and Mark Urich completed the RNA-Seq experiments Joe Nery completed all the Illumina sequencing. Hai Li performed the genetic and physiological characterization studies

of HLS1 and its homologs. Hongwei Guo generated the EIN3 antibody used in this study.

This work has been funded by an NIH CMG training grant, NSF Plant Systems Biology IGERT training grant, The Gates Millennium Foundation, The USA Funds Foundation, and a Department of Energy grant to Joseph Ecker. The NSF Plant Systems Biology IGERT grant was awarded to Julian Schroeder and Steve Briggs.

I would like to acknowledge Cathy Pugh and Tom Tomp at the UCSD Division of Biology, for the smorgasbord of graduate student related issues they have never failed to iron out. Alyona Bobkova has also been very helpful with the NSF PSB IGERT program.

Finally, I thank my family and friends who have given me their unconditional support throughout these years. My mother and sister have been there for me wherever I go and whatever I choose to do. Chris Dupont has cheered me up numerous times with burnt steak and made me laugh about life. Gerald Pao has been my mentor and one of my best friends. Kendra Hogan, Dwight Kuo, Zack Smith, and Mari Manuia, have all made this experience much more enjoyable.

Chapter One, in full, consists of the following manuscript to be submitted as a review in the format of the journal Current Opinion in Plant Biology.

Chang KN, Ecker JR. Unraveling the role of ethylene in plant hormone crosstalk.

Chapter Two, in full, consists of the following manuscript in preparation in the format of the journal *Science* as a report.

Chang KN, Li H, Hon G, Pelizzola M, Schmitz RJ, Urich M, Kuo P. Dwight, Nery J, Qiao H, Ideker T, Ecker JR. Dynamic transcription factor binding reveals role of EIN3 in plant hormone crosstalk.

I was the primary researcher and author for both manuscripts and Joseph R. Ecker directed and supervised the research that forms the basis for these chapters. Hai Li contributed the genetic and phenotypic analysis in Chapter Two.

VITA

EDUCATION

2011	Ph.D., Biology, University of California, San Diego
2003	M.S., Biological and Environmental Engineering, Cornell University
2001	B.S., Biosystems Engineering with High Honors, University of Hawai`i, Manoa
1997	High School Diploma, Valedictorian, Waimea High School, Kaua`i

WORK EXPERIENCE

2003-2004	Biotechnology Fellow, Department of Agricultural and Biological Engineering, University of California, Davis
2001	Engineering Intern, Roy F. Western, Inc.
1999-2001	Research Assistant, Marine Bioproducts Center, University of Hawai`i, Manoa
1998	Research Assistant, Department of Biosystems Engineering, University of Hawai`i, Manoa

PUBLICATIONS

Chang KN, Pelizzola M, Kuo DP, Qiao H, Ideker T, Ecker JR. Temporal transcription factor binding reveals role of EIN3 in plant hormone crosstalk. Manuscript in preparation for submission to Science.

Boutrot F, Segonzac C, Chang KN, Qiao H, Ecker JR, Zipfel C, Rathjen JP. 2010. Direct transcriptional control of the Arabidopsis immune receptor FLS2 by the ethylene-dependent transcription factors EIN3 and EIL1. PNAS 107(32):14502-7.

Qiao H, Chang KN, Yazaki J, Ecker JR. 2009. Interplay between ethylene, ETP1/ETP2 F-box proteins, and degradation of EIN2 triggers ethylene responses in Arabidopsis. Genes Dev 23(4):512-21.

Chang, KN, Holtmeier, NL, Spanswick, RS, and Ahner, BA. 2001. Phytochelatin- and Glutathione-mediated Cadmium Transport in Tonoplast Vesicles Isolated from Oat (*Avena sativa*) Roots. Master thesis, Cornell University.

Sylva T, Kinoshita C, Romano R, Toma M, Tsang S, Chang K. 2003. Bioremediation of Petroleum Impacted Soils from Investigation-derived wastes. Remediation Journal. Autumn: 79-90.

PRESENTATIONS/MEETINGS

Chang KN, Pelizzola M, Kuo DP, Qiao H, Ideker T, Ecker JR. 2010. Temporal gradation of EIN3 transcription factor binding regulates the ethylene transcriptional response. Gordon Research Conference in Plant Molecular Biology. Invited speaker.

Chang KN, Kuo DP, Ideker T, Ecker JR. 2008. Using ChIP-Sequencing to understand the ethylene transcriptional regulatory response mediated by EIN3. University of California, San Diego, Plant Systems Biology NSF IGERT Symposium. Invited speaker.

Chang KN, Gregory BD, Yazaki J, Lister R, Chen H, Ecker JR. 2008. Rapid alteration in DNA methylation associated with response to ethylene in Arabidopsis thaliana. Gordon Research Conference in Plant Molecular Biology.

Zapala MA, Hsiao G, Chang KN, Schork NJ. 2005. Evol-genomic: a software tool for translational genomics. American Society of Human Genetics Annual Meeting.

Spanswick R, Ahner BA, Chang KN. 2004. Nutritional dependence of cadmium/phytochelatin tonoplast transport. 13th International Workshop on Plant Membrane Biology.

Chang KN, Holtmeier NL, Spanswick RM, Ahner BA. 2003. Inducible transport of Cd-phytochelatin complexes. American Society of Plant Biologists.

ABSTRACT OF THE DISSERTATION

The ethylene transcriptional response mediated by

ETHYLENE INSENSITVE3 in *Arabidopsis thaliana*

by

Katherine Noelani Chang

Doctor of Philosophy in Biology

University of California, San Diego, 2011

Professor Joseph R. Ecker, Chair

The sessile nature of plants necessitates a phenotypic plasticity that enables plants to respond to changes in environment throughout growth and development. Orchestration of molecular components at various levels in a spatial and temporal dimension is required. However, how signals are integrated to produce a specific

response, e.g. how a simple hydrocarbon, the plant hormone ethylene, can cause a diverse set of morphological response, remains elusive. To better understand the integration of signals involved in generating these phenotypes, we characterized the transcriptional regulation of the ethylene response by identifying key protein-DNA interactions and transcriptional profiles in a temporal manner.

The plant hormone ethylene regulates numerous growth and developmental processes in plants, including stem cell division, differential cell growth, stress response, response to pathogens, germination, senescence, fruit ripening and the triple response. The triple response, a decrease in cellular elongation, increase in radial swelling, and an exaggerated apical hook, has been the hallmark of the ethylene response in dark grown seedlings. To determine the molecular mechanisms of the ethylene response, we characterized the dynamic ethylene transcriptional response in etiolated seedlings. We identified targets of the master regulator of the ethylene signaling pathway, ETHYLENE INSENSITIVE3 (EIN3), using temporal chromatin immunoprecipitation sequencing (ChIP-Seq) and transcript sequencing (mRNA-Seq).

First we identified the minimum complement of genes required for the ethylene transcriptional response in etiolated seedlings. Then we characterized targets of EIN3 based on their transcriptional state as well as their function in relation of the ethylene signaling pathway and response. We found a singular EIN3 binding pattern that increased upon ethylene treatment, which did not correspond to target gene transcription. Binding of EIN3 established feedforward transcriptional cascades, feedback circuitry of the ethylene signaling pathway, and interconnections between hormone response pathways at a multitude of levels, e.g. hormone biosynthesis,

signal reception, signaling, and transcriptional response. Finally, we showed that mutants in a gene family targeted by EIN3 exhibit hormone response phenotypes in several developmental stages, thus demonstrating the integral role EIN3 plays in the orchestration of hormone crosstalk in plant growth and development.

CHAPTER 1:

INTRODUCTION

ETHYLENE IN PLANT GROWTH AND DEVELOPMENT

Ethylene (chemical formula, C_2H_4) is a gaseous phytohormone regulating plant growth and development. The identification of ethylene as a plant hormone began in the 19th century, during the time coal gas was used to light streetlamps. Observations that trees located near streetlamps senesced earlier than others lead to the discovery that burning coal produces ethylene. In 1901, Neljubov found ethylene to be the active component in a gas affecting the direction of pea seedling growth in the dark *(1)*. Later Gane demonstrated that ethylene was a produced in ripening fruit *(2)*.

Ethylene is responsible for a broad range of biotic and abiotic responses in plants. Because of its roles in senescence, fruit ripening and pathogen responses *(3)*, ethylene has a major impact in agriculture. Ethylene stimulates fruit ripening in climacteric plants (*e.g.* tomatoes, bananas, apples) *(4)*. The ethylene response can be chemically manipulated to increase yields of fruits such as pineapples and tomatoes. On the other hand, from field to consumer, ethylene must be kept at a minimum to reduce fruit spoilage.

The biosynthesis and signaling of ethylene is relatively conserved in plants and has been studied in brassica, tobacco, *Medicago*, maize, soybean, tomato, red pepper, persimmon, melon, cucumber, lettuce, grapevine, carrot, cotton, rice, petunia, peach, rubber tree, beech, moss, nitrogen-fixing bacteria, and more. Numerous reports exist detailing ethylene biosynthesis as well as the involvement of ethylene in relation to environmental stimuli. Parallels can be drawn between orthologs, for example, a protein interacting with the ethylene receptors in *Arabidopsis thaliana* was found to cause dramatically decreased rates of ripening in tomatoes *(5, 6)*.

Throughout the lifetime of a plant, ethylene orchestrates many key biological processes: seed germination, root hair development, root stem cell division, root nodulation, abiotic and biotic stress, leaf abscission, sex determination, flower senescence, and fruit ripening. In seedling germination, ethylene maintains the curvature of the apical hook to protect the stem cells in the shoot. When exogenous ethylene is applied to *Arabidopsis* seedlings grown in the dark, they exhibit what is known as the triple response, displaying an exaggerated apical hook, increase in radial swelling, and decrease in cellular elongation *(7)*. This hallmark of the ethylene response has been used in phenotypic screens and subsequent genetic analysis to identify key components in the ethylene signaling pathway *(1, 7)*. Aberrant ethylene response phenotypes that were discovered include ethylene overproduction, ethylene insensitivity, hypersensitivity, and constitutive ethylene response *(7)*. Isolation, cloning, and characterization of mutants hypersensitive or insensitive to ethylene have contributed to our current understanding of the ethylene signaling pathway, as reviewed in *(8, 9)*. In the last decade, phenotypic screens for weak ethylene insensitive (wei) or enhanced ethylene response (eer) mutants have been performed to identify novel components involved in the ethylene signaling pathway *(10-12)*.

ETHYLENE BIOSYNTHESIS

Most plants are capable of synthesizing ethylene from S-adenosyl-methionine (S-AdoMet) in a three-step process. First, the Yang cycle of methionine biosynthesis feeds into the production of ethylene as methionine is converted to S-AdoMet using the enzyme S-AdoMet synthase, which requires adenosine triphosphate (ATP) *(13)*. Second, 1-aminocyclopropane-1-carboxylic acid (ACC) is generated from S-AdoMet

using ACC synthase. The conversion of S-AdoMet to ACC is the first committed step to the production of ethylene and is the rate-limiting reaction of ethylene biosynthesis. Nine ACS genes exist in *Arabidopsis* (ACS1-2, ACS4-9, ACS11) and they are regulated by environmental stimuli that trigger ethylene biosynthesis, including wounding/herbivory, pathogen infection, UV-B, and the presence of other hormones such as auxin, cytokinin, and brassinosteroid, reviewed in *(14)*. ACS regulation at the both at the transcriptional (for auxin) or translational (for cytokinin) levels have been found to occur*(15, 16)*.The third step of ethylene biosynthesis is the production of ethylene from ACC with ACC oxidase, as reviewed in *(14)*.

Regulation of ethylene biosynthesis has been examined through the study of the ethylene overproducer (*eto*) mutants. *eto* mutants are affected in the stability of active ACS enzyme, and rapid regulation of ethylene biosynthesis in response to plant pathogens, herbivory, or other environmental changes occurs through the protein accumulation of ACS *(17)*.

ETHYLENE SIGNAL PERCEPTION

The regulation of the ethylene response occurs in a negative manner by the ethylene receptors and its interactor, CONSTITUTIVE TRIPLE RESPONSE1 (CTR1) (Fig. 1). Gaseous ethylene diffuses through the plant cell wall and plasma membrane, and is bound and perceived by a family of partially redundant receptor histidine kinases located in the endoplasmic reticulum (ER), ETHYLENE RECEPTOR1 (ETR1), ETHYLENE RECEPTOR2 (ETR2), ETHYLENE RESPONSE SENSOR1 (ERS1), ETHYLENE RESPONSE SENSOR2 (ERS2), and ETHYLENE INSENSITIVE4 (EIN4) *(18-22)*. These receptors share homology to bacterial two

component histidine kinases and require a copper cofactor to bind ethylene *(19, 23, 24)*. Copper is transported by RAN1 (RESPONSIVE-TO-ANTAGONIST1), a protein similar to the copper transporter P-type ATPase, and incorporated into the receptors *(25, 26)*. Ethylene binding of the receptors occurs in the N-terminus transmembrane domain *(27)*.

The ethylene receptors share similarities and differences in the structure and functions of their protein domains. Each receptor contains an N-terminal transmembrane domain that binds ethylene, a GAF domain that mediates interactions between the receptors, and a kinase domain responsible for the interaction with the downstream signaling component, CTR1. ETR1, ETR2, and EIN4 have an additional receiver domain which has been reported to be required for the growth recovery upon ethylene treatment and subsequent removal of ethylene gas *(28)*.

The ethylene receptors were classified into two families based on their sequence and phylogeny. Type I receptors (ETR1, ERS1) have three hydrophobic transmembrane domains and contain a functional histidine kinase domain. Type II (ETR2, ERS2, EIN4) receptors have four hydrophobic transmembrane domains and degenerate histidine kinase domains with *in vitro* serine threonine kinase activity *(29)*. The ethylene receptors have been shown to heterodimerize and have been isolated as large protein complexes *(30, 31)*. The formation of these complexes occurs in both constitutive and ethylene-induced manner. Quintuple mutant analysis has enabled the discovery that the gain of function *ers1-1* mutant requires *ETR1, ETR2, ERS2, EIN4* (http://www.sippe.ac.cn/e_ktzz_wenqg.asp).

Researchers are still investigating the role and importance of the kinase and receiver domains in the ethylene receptors. The kinase activity of the receptors

remains of dubious significance. Mutations in the kinase domains of the ethylene receptors reveal that the kinase domain may be involved in growth-related phenotypes mediated by ethylene *(32)*. Although it has been reported that ethylene represses ETR1 histidine kinase activity *(33)* and the kinase domain has been shown to interact with CTR1, the kinase domain is not necessary for ethylene signaling *(32, 34, 35)*. Similarly, the receiver domain is not necessary for the receptors to repress the ethylene response *(32)*.

Another key player in the perception of ethylene was recently identified and characterized. REVERSION-TO-ETHYLENE SENSITIVITY1 (RTE1) is a membrane protein localized to the ER and Golgi apparatus *(36-41)*. *RTE1* was found to suppress ethylene insensitivity of *etr1-2* and is a positive regulator of ETR1. RTE1 has a high affinity to ETR1 and its interaction is specifically with ETR1 and not the other ethylene receptors. ETR1 receptor complexes are RTE1-regulated in mechanism that remains elusive. The function of RTE1 has been difficult to determine because RTE1 does not contain any canonical protein domains and the functions of its orthologs in plants, animals, and protists are unknown. Recently, it has been reported that RTE1 is not involved in the growth recovery upon ethylene removal, but does play a role in ethylene-stimulated bending *(35)*.

The ethylene receptors interact with a Raf-like mitogen activated protein kinase kinase kinase (MAPKKK), CTR1, which is colocalized in the ER. Specifically, the N-terminus of CTR1 interacts with the kinase domain of ETR1 and ERS1 *(42-44)*. This interaction is not dependent on the receptor histidine kinase activity. CTR1 was determined to be downstream of the ethylene receptors, but upstream of the first positive ethylene signaling gene, ETHYLENE INSENSITIVE2 (EIN2), using epistatic

analysis. It has been hypothesized that CTR1 triggers a MAPK cascade, resulting in the negative regulation of the ethylene signaling pathway. However, experimental evidence has been inconclusive in demonstrating the downstream MAPKK and MAPK components required for the ethylene response. When ethylene gas is not present, the interaction between the receptors and CTR1 or CTR1 itself represses the ethylene signaling pathway.

ETHYLENE SIGNAL TRANSDUCTION

In the presence of ethylene gas, ethylene binds to the receptors via a copper cofactor and the repression of the pathway is relieved, allowing EIN2 to signal in an unknown manner to a transcription factor in the nucleus, ETHYLENE INSENSITIVE3 (EIN3) (Fig. 1). EIN2 is a 12 domain transmembrane protein located in the ER *(45, 46)*, similar to Nramp metal transporters *(47)*. Despite its annotation as a putative metal transporter, EIN2 does not have the capability to transport metals and its molecular function remains elusive.

EIN2 is regulated in a ubiquitin-mediated proteasome dependent manner by F-box proteins EIN2 TARGETING PROTEIN1 (ETP1) and 2 (ETP2) *(48)*. ETP1 and ETP2 transcription is induced by ethylene, thus ethylene regulates the accumulation of EIN2. Recent studies have shown that the C-terminus of EIN2 is cleaved and is subsequently translocated to the nucleus, where it may function in downstream signaling *(45)*. The cleavage and subsequent translocation of the C-terminus of EIN2 is ethylene dependent. Qiao and Ecker recently found that the C-terminus of EIN2 interacts with a nuclear localized protein, which in turn, interacts with the downstream transcription factor, ETHYLENE INSENSITIVE3 (EIN3) *(45)*. Although it has yet to be

demonstrated, the nuclear interaction between EIN2 and EIN3 may provide the mechanism in which the C-terminus of EIN2 transmits the ethylene signal to EIN3.

The C-terminus of EIN2 signals to EIN3 and its homolog ETHYLENE INSENSITIVE3-LIKE1 (EIL1), transcription factors located in the nucleus *(49)*. The protein levels of the downstream nuclear-localized transcription factor EIN3 are dependent on the presence of EIN2, and the exact mechanism of regulation has not yet been elucidated. Previous studies have shown that EIN3 is necessary and sufficient for the ethylene response. Like EIN2, EIN3 protein stability is also regulated in an ubiquitin-mediated proteasome-dependent manner through F-box proteins, here, EBF1 and EBF2 (EIN3 BINDING F-BOX) *(50-55)*. Previous studies have shown that EBF2 and EBF1 mRNA is regulated in an unknown manner by *ETHYLENE-INSENSITIVE5 (EIN5) (50, 54)*, and that EBF1 and EBF2 are regulated differentially at the transcript level by ethylene *(54)*. Recently it has been shown that EBF1 and EBF2 protein stability is regulated by ethylene, specifically dependent on EIN2 *(55)*, and that the degradation of EBF1 and EBF2 most likely controls the levels of EIN3 and EIL1. EIN3 and EIL1 bind the promoter regions of transcription factors ETHYLENE RESPONSE FACTOR1 (ERF1) and ETHYLENE RESPONSE DNA-BINDING FACTOR1 (EDF1), initiating a transcriptional cascade resulting in the variety of ethylene mediated physiological and morphological responses *(56, 57)*.

Several other genes may play a role in the ethylene signaling pathway, but their functions and mechanisms remain elusive. For example, double mutants in *ETHYLENE INSENSITVE 6 (EIN6)* and *ENHANCER of ETHYLENE INSENSITIVITY (EEN)* are deficient in EIN3 accumulation *(50)*. EIN6 is a jumonji-class transcription factor that has been implicated in flowering, but otherwise, its function in the ethylene

response is unknown *(58)*. The expression of *EIN6* is not ethylene-inducible suggesting that EIN6 protein stability may be regulated by ethylene. Mutants in *ein6* alone display an ethylene insensitive root. The identity of EEN has not been determined, and the manner in which it enhances the ethylene response phenotype of EIN6, though interesting, is also unknown.

Other mutants have been found to exhibit enhanced ethylene responses in different conditions and plant organs, but their function in relation to the ethylene signaling pathway is unknown. Interpretations of epistatic analysis do not apply for enhanced/hypersensitive phenotypes, therefore making it difficult to place these genes in the ethylene signaling pathway, if indeed they play a central role in ethylene signaling. At best, mutation analysis can only determine whether these genes depend on the ethylene signaling pathway to produce an ethylene response phenotype. ENHANCED ETHYLENE RESPONSE (EER) mutants that have been identified include *eer1, eer2, eer4,* and *eer5. EER1*, a protein phosphatase 2A regulatory subunit A (PP2A) was found to act downstream of ETR1 and EIN2 because mutations in *ETR1* and *EIN2* masked the *eer1* phenotype *(11)*. *EER2* mutants displayed an enhanced ethylene response in the light, but its function has not been determined *(59)*. Mutants in *EER4,* a nuclear-localized TFIID-interacting transcription factor, exhibited a decrease in ethylene response genes such as *ERF1*, but how *EER4* functions in the ethylene signaling pathway was not clearly established *(60)*. *EER5* was identified as a putative proteasome COP9 initiation factor PCI/PINT-associated module domain-containing protein and was found to interact with EIN2, but *eer5* mutants were not impaired in EIN3 accumulation *(12)*.

EMERGING QUESTIONS IN ETHYLENE SIGNALING

Despite its importance in plant physiology, homeostasis, and development, several mechanisms of ethylene signaling remain unclear. Recent conflicting evidence suggests that there is a CTR1 independent ethylene signaling pathway *(33)* (http://www.sippe.ac.cn/e_ktzz_wenqg.asp), however this has yet to be substantiated. In addition, the majority of the genetic data indicates that ethylene signaling occurs in a largely linear fashion. Another area of ethylene signaling that is under contention is the identification of MKKs and MPKs, if any, involved in ethylene signaling. One study presented *in vitro* protoplast data that MPK3/6 are responsible for phosphorylating EIN3 *(61)*, while another clearly shows that these MPKs are involved in ethylene biosynthesis *(62)*. Recently the previously demonstrated importance of MKK9 in ethylene signaling has been negated *(55)*. Finally, downstream ethylene signaling with respect to the transcriptional response has not been well studied. Understanding how ethylene signaling manifests in many different morphological phenotypes is significant, as modules of ethylene responses can eventually be manipulated to produce a specific morphotype.

Thus, the most significant questions of the ethylene signaling pathway that need to be addressed are: 1) What, if any, are the components of the kinase cascade downstream of CTR1 and upstream of EIN2? 2) What is the molecular function of EIN2 and how does EIN2 transmit the ethylene signal to EIN3? 3) How does ethylene result in many different morphological phenotypes? What are the molecular bases of these phenotypes?

ETHYLENE TRANSCRIPTIONAL RESPONSE

This study focuses on the last question, as the mechanism(s) in which ethylene induces a myriad of different responses is unknown. It has been previously shown that EIN3 protein accumulation determines whether the ethylene response occurs and is the rate-limiting step of the ethylene response *(50-52)(63)(63)*. As mentioned previously, the protein levels of EIN3 are regulated in a 26S proteasome-dependent manner by F-box proteins EBF1 and EBF2. EIN3 accumulates upon ethylene treatment and is stabilized at one hour of ethylene treatment *(50)*. The half-life of EIN3 protein is less than 30 minutes and EIN3 is rapidly degraded in the absence of ethylene. Overexpression of EIN3 and EIL1 but not other paralogs (e.g. EIL2, EIL3) confer a constitutive ethylene response phenotype *(49)*. Overexpression of EIN3 in the mutant background of an upstream component in the ethylene signaling pathway (*ein2)* also results in a constitutive ethylene response phenotype, demonstrating that EIN3 is sufficient for the ethylene response. Double mutant plants *ein3-1 eil1-1* are refractory to the second prolonged phase of the ethylene response *(10, 64, 65)*. Conversely, overexpression of only the C-terminus of *EIN2*, as well as *EIN3* or *EIL1* but not other homologs confer a constitutive triple response phenotype *(49)*.

EIN3 and EIL1 have distinct roles in the ethylene response based on the genetic and biochemical analyses of An et al. *(55)*. Mutants in *EBF1* and *EBF2* were crossed with *EIN3* and *EIL1* mutants and the ability to suppress the constitutive triple response phenotypes of the *ebf1 ebf2* mutants was assessed. *ein3,* but not *eil1* suppressed the *ebf1 ebf2* mutant etiolated seedling phenotypes. EIL1 was shown to function in the adult plant, regulating stem elongation and leaf expansion, while EIN3 regulated the ethylene response in etiolated seedlings.

EIN3 has been shown to bind the primary ethylene response element (PERE) as a homodimer in the promoter region of the transcription factor ERF1 (ETHYLENE RESPONSE FACTOR1) and EDF1 (ETHYLENE RESPONSE DNA-BINDING FACTOR1) *(66, 67)*. Solano et al. (1998) found that *EIN3* is necessary and sufficient for *ERF1* expression. ERF1, in turn, binds other ethylene inducible genes with GCC box promoter elements (*e.g. basic-chitinase, PDF1.2*). *ERF1* is responsible for a subset of ethylene responses as its overexpression leads to a partial triple response and causes a decrease in one of the ethylene responsive genes, *HLS1* (HOOKLESS1) in an *ein3* background suggesting that it acts downstream of *EIN3 (56)*. ERF1 is part of the ERF family of transcription factors thought to activate the expression of secondary ethylene response genes, including pathogen defense genes through interaction with the jasmonic acid pathway *(68)*. Because the ethylene transcriptional cascade culminates in a variety of defense and growth responses and *ERF1* is responsible for a subset of these responses, EIN3 is thought to be the master regulator of the transcriptional cascade initiated by ethylene.

Previous microarray experiments revealed that hundreds of genes (3-7% of the genome) have altered expression profiles in response to ethylene *(10, 69, 70)*. Yet, the targets of EIN3 remained elusive because known targets of EIN3 are transcription factors themselves (*e.g. ERF1, EDF1*), and it was difficult to distinguish direct targets from those further downstream in the ethylene transcriptional program. At the time of this study, only the promoter regions of *ERF1* and *EDF1* were shown bound by EIN3 *(66)*. In the last three years, several studies identified new targets of EIN3, and a total of seven have been reported (Fig. 2). Yanagisawa and Konishi demonstrated that EIN3 binds the promoter region of the F-box protein, EBF2,

directing the protein degradation of EIN3 itself *(53)*. Through electromobility shift assays, *SALICYLIC ACID INDUCTION DEFICENT2 (SID2) (71)* and *PROTOCHLOROPHYLLIDE OXIDOREDUCTASE A and B (PORA/B) (72)* were demonstrated to be targets of EIN3. Chen et al. found that EIN3 and EIL1 negatively regulate SID2, an enzyme necessary for the pathogen-induced biosynthesis of salicylic acid (SA). This was the first report of EIN3/EIL1 acting as a repressor, as all other studies reported induction of EIN3/EIL1 target genes. Another set of enzymes was targeted for transcription by EIN3/EIL1; PORA/B enzymes are involved in chlorophyll synthesis during the deetiolation of seedlings.

A plant immunity related gene, a leucine-rich repeat receptor kinase, *FLAGELLIN-SENSITIVE2 (FLS2*), was targeted by EIN3/EIL1. Boutrot et al. initially found that *ein2* mutants were impaired in regulated immune responses. FLS2 is involved in the plant immune response because of its recognition of the bacterial peptide flg22. FLS2 was also found to be transcriptionally regulated by ethylene *(73)*. The data presented in this thesis lead to the determination that FLS2 was a target of EIN3. Subsequent chromatin immunoprecipitation (ChIP) experiments followed by quantitative polymerase chain reaction (QPCR) further supported that FLS2 was a direct target of EIN3.

EIN3 Protein-DNA Binding Domain Characteristics

The family of *EIN3/EIL* transcription factors is novel. EIN3 is a member of a family of five sequence-specific DNA binding proteins including EILs (EIN3-LIKE) 1-5 *(65)*. EIN3 contains four basic regions in the N-terminus, as well as an acidic and proline rich region. The N-terminus is highly conserved among EIN3 and the EILs and

is sufficient for DNA binding. The C terminus contains an Asn-rich region and is required for EIN3 to interact with EBF2 *(50)*.

The structure of the DNA binding domain of the most distant EIN3 relative in the EIN3 family, EIL3, was determined using nuclear magnetic resonance (NMR) imaging (Yamasaki et al. 2005). The EIL3 DNA binding domain is composed of mainly alpha class protein folds and have similarity to the LEM/SAP HeH motif, a helix-extended loop-helix arranged in parallel helices *(74)*. Unfortunately the EIL3 binding domain does not have homology to other well-characterized transcription factor binding domains and how its paralog EIN3 might possibly bind to DNA in the presence of ethylene is unknown. One caveat of the NMR study is that EIL3 was previously shown to be unable to recognize the *ERF1* promoter region *in vitro (56)*.

Interestingly, the well-known human cancer gene BREAST CANCER1 (BRCA1) and its homologs and orthologs contain an EIN3-like DNA binding domain (http://pfam.sanger.ac.uk/). Non-sequence specific DNA binding has been demonstrated for BRCA1 *in vitro (75)*, unlike the sequence-specific binding of EIN3 previously determined. The majority of the EIN3-like proteins exist in eukaryotes. EIN3-like protein domains exist in plants such as rice, tobacco, mung bean, tomato, banana, carnation, and orchid (Fig. 3). Based on the sequence phylogeny, metazoan EIN3-like proteins are BRCA1-like and are an out-group from the plant EIN3-like proteins. *Arabidopsis thaliana* does contain BRCA1 orthologs, however to our knowledge these do not have an EIN3-like protein DNA binding domain.

EIN3 Transcription Factor Activity

Previous attempts to determine whether EIN3 binding is activated by its phosphorylation were inconclusive. EIN3 has been reported to be phosphorylated at two sites *in vitro*, the T174 site was reported to stabilize EIN3, while the T592 was suggested to play a role in EIN3 degradation *(61)*. However, as mentioned previously the MKK and MPKs shown by Yoo et al. to phosphorylate EIN3, were not required for the ethylene response *(55)*. In addition, available phosphoproteomes do not contain peptides of EIN3 (http://phosphat.mpimp-golm.mpg.de/phosphat.html, http://gator.masc-proteomics.org) *(45, 76)*. It is possible that the tissues queried for EIN3 phosphorylation affect the detection of phosphorylated EIN3, however, the importance of EIN3 phosphorylation with respect to transcription factor activity has yet to be demonstrated.

HORMONE CROSSTALK

ChIP-Seq and mRNA-Seq will enable the genome-wide identification of the minimum complement of the ethylene transcriptional response mediated by EIN3. In addition to determining genes that are involved in the ethylene response in etiolated seedlings, we will most likely observe connections between ethylene and other hormones necessary for plant growth and development. Many biological processes are regulated by more than one hormone, and a plethora of hormone connections have been discovered throughout plant growth and development, in various tissue or cell types. Spatiotemporal regulation of growth and development as well as responses to environmental stimuli requires the coordination of a multitude of signal inputs and outputs for a variety of hormone pathways.

Studies scoring mutants with a phenotype relevant to one hormone (e.g. hypocotyl or root length, germination rate, lateral root growth, flowering time) often reveals phenotypes in relevant to another hormone. However, crosstalk in the strictest sense, or the sharing of biosynthesis, perception, or signaling components between two hormones, does not occur in plants and a recent review also mentions that crosstalk is likely to occur at the transcriptional regulatory network level *(77)*.

A microarray analyses revealed that different hormone pathways regulate the same biological processes, but that the transcriptional programs that are responsible for this regulation are not shared *(78)*. Nemhauser and colleagues attempted to identify a core growth regulation module because growth is regulated by many hormones, including brassinosteroids, gibberellin, auxin, and ethylene. In addition mutants in these pathways exhibit aberrant growth phenotypes. However, Nemhauser et al. found that hormones regulate the same biological processes (e.g. cellular elongation) by nonoverlapping transcriptional responses *(78)*. Although a major caveat to this study is that the spatiotemporal resolution of the dataset was limiting, the results suggest that plants have many ways of regulating growth using a suite of hormones. The massive permutations of the pathways in growth regulation are indeed necessary when a sessile organism must adapt to its environment.

Although the transcriptional regulatory networks that may or may not be shared by hormone pathways have not been determined, there is overwhelming evidence for hormone connections, as discussed below. The points of interconnection between ethylene and auxin (IAA), cytokinin (CK), methyl jasmonate (MJ), gibberellin (GA), salicylic acid (SA), brassinosteroid (BR), and abscisic acid (ABA) have been reported, as well as interactions of ethylene and light, nitric oxide, hydrogen peroxide,

and other small molecules. The significance of these interactions must eventually be placed into the context of gene regulation, in terms of transcription, translation, post-translation, or metabolite production. Here, we focus on the interaction of ethylene and auxin and ethylene and methyl jasmonate.

ETHYLENE AND AUXIN

The connections between ethylene and auxin have a long history and have been shown to occur at many levels. An early study reported that ethylene biosynthesis is regulated by auxin. Auxin was found to induce the expression of ACS, the rate-limiting enzyme of ethylene biosynthesis *(79)*. The original screens for ethylene triple response mutants identified two auxin transport mutants, *aux1* and *eir1 (7)*. A ethylene response mutant that did not display the exaggerated apical hook phenotype, *hookless1* (*hls1*), was eventually found to regulate the protein levels of an auxin signaling gene, AUXIN RESPONSE FACTOR2 (ARF2) *(80)*.

Other auxin mutants exhibit an insensitive or weak insensitive (wei) ethylene response and *aux1, eir1, axr1,* and *hls1* auxin mutants displayed tissue specific ethylene phenotypes. Mutants in positive regulators of the ethylene signaling pathway, EIN2 and EIN3, also have increased cellular elongation in the hypocotyl and root in the etiolated seedling. Mutants of key players in the auxin pathway at the levels of biosynthesis (*taa1*, *asa1, asb1*), transport (*aux1, eir1*), perception (*tir1*), signaling (*axr1*) display aberrant ethylene response phenotypes *(7, 10, 81)*.

In general, ethylene and auxin both inhibit growth in the seedling. A study examining the interdependency of the ethylene and auxin pathway revealed that auxin is required for the ethylene response *(81)*. Stepanova and colleagues analyzed

ethylene and auxin mutants for hypocotyl and root growth phenotypes as well as the distribution of signal in ethylene and auxin response reporter lines (EBS:GUS and DR5:GUS, respectively) under auxin and ethylene treatment. Auxin mutants *aux1, eir1, axr1,* and *tir1* displayed ethylene insensitivity in the root. However, ethylene mutants *ein2, ein3*, and *eil1* displayed hypocotyl and root lengths similar to wildtype in the presence of auxin/IAA. This suggested that a basal level of auxin is required for the ethylene response in roots, but that ethylene is not required for the auxin response in hypocotyls or roots. The expression of the auxin reporter construct in roots revealed that the precursor to ethylene, ACC, increased the auxin response in the root, expanding the auxin response to the transition zone. This root expression pattern correlated to that observed in EBS:GUS lines, expression occurred in the root transition zone in the presence of ACC in wildtype seedlings but not in the auxin transport mutant, *aux1 (81)*. Additional microarray studies in root tissue further supported the role of auxin-mediated ethylene response controlling the growth of roots.

Stepanova and colleagues further dissected the ethylene and auxin response with the microarray studies, analyzing changes in gene expression under ethylene and auxin treatment in the wildtype and ethylene *(ein2*) and auxin (*aux1*) mutants under auxin and ethylene treatment, respectively. Interactions between ethylene and auxin were reduced to three types, 1) one hormone sensitizing cells for response to another, 2) one hormone triggering the biosynthesis or signaling of another, or 3) both hormones targeting the same transcription modules independently. One of the most important conclusions of this study is that hormone dependency appears to be directional in the case of ethylene and auxin. Basal levels of auxin are required for the

ethylene response, but the converse is not true. The current model is that ethylene increases auxin to regulate growth *(82-84)*.

ETHYLENE AND METHYL JASMONATE

Many examples of ethylene and methyl jasmonate hormone crosstalk have been reported, and those studying the plant defense response have found that interactions between hormone signaling pathways dictate the set of defenses that are activated in response to a specific environmental stimuli or developmental regimen.

In general, ethylene and methyl jasmonate work synergistically in the plant pathogen response against necrotrophic pathogens and mutants that are deficient in ethylene or jasmonic acid signaling confer susceptibility to fungal pathogens. However, negative interactions between ethylene and methyl jasmonate pathways do exist, and it is the interplay between these pathways that determines the stress response of a plant. Ethylene and methyl jasmonate induce and coregulate defense related genes such as *ERF1*, *PLANT DEFENSIN 1.2* (*PDF1.2*), *BASIC CHITINASE* (b-CHI), β-1,3-glucanase and other pathogen related (PR) genes *(68)*. ERF1 was specifically shown to be regulated by both ethylene and jasmonic acid signaling, as overexpression of ERF1 was sufficient to rescue pathogen deficiency of ethylene signaling (*ein2)* or jasmonic acid signaling (*coi1*) mutants *(68)*. ERF1 and another AP2/ERF transcription factor ORA59 were previously shown to induce *PDF1.2*, and a recent study has shown that ORA59 binds two GCC boxes to transcriptionally activate PDF1.2 *(85)*. Other genes involved in the regulation of transcription may also have a role in the regulation of ethylene and methyl jasmonate crosstalk. For example, a histone deacetylase, *HDA19*, was also regulated by ethylene and methyl

jasmonate. Although HDA19 was shown to localize in the nucleus *(86)*, its mechanistic function in relation to these hormone responses is unknown.

SYSTEMATIC ANALYSIS OF HORMONE CROSSTALK USING GENOME-WIDE DATASETS

Due to the complex nature of the interconnections between plant hormone pathways, large datasets are useful in identifying key players (hubs) of hormone crosstalk. The multitude of connections between hormone response pathways must be approached in a systematic manner, using a network approach. Building protein-protein interaction (PPI) networks, protein modification networks, and transcriptional regulatory networks of protein-DNA interaction (PDI) will help us identify the density and topology of the interactions between hormone biosynthesis, signaling, and transcription.

PROTEIN-PROTEIN INTERACTIONS IN *ARABIDOPSIS THALIANA*

Literature curation of PPI studies (e.g. yeast two hybrid (Y2H), bimolecular fluorescent complementation (BiFC), coimmunoprecipitation) revealed a high false positive rate in available PPI data. Recently the Arabidopsis Interactome Mapping Consortium used a high throughput Y2H study to identify protein-protein interactions in Arabidopsis. High throughput nucleic acid programmable protein microarrays (HT-NAPPA) are currently being used to provide a complementary protein-protein interaction dataset. HT-NAPPA can query the binding of proteins that are membrane bound in their native condition, as well as proteins that have auto-activation activity in Y2H studies.

Preliminary experiments have determined that >90 proteins in addition to EBF1 and EBF2, interact with the master regulator of the ethylene response, EIN3 *(87)*. These proteins are likely coregulator candidates. More studies mapping the interactions between proteins are necessary to help us understand the extent of hormone crosstalk.

PROTEIN-DNA INTERACTIONS IN *ARABIDOPSIS THALIANA*

Cis-regulatory elements, transcription factor motifs, and protein-DNA interactions can be found at The Arabidopsis Gene Regulatory Information Server (AGRIS, http://arabidopsis.med.ohio-state.edu/). To date, genome-wide *in vivo* transcription factor studies have only been performed for 15 transcription factors, including: floral identity transcription factors SEPELLATA3, APETALA1, and FLC; hormone response transcription factors, BES1, BZR1, EIN3; as well as others TGA2 and LFY1, AGL15, HY5, GL3, AtbHLH15, WRKY53, GL1, E2F. A protein-DNA interaction (PDI) network was generated with the data at AGRIS and from the literature. An increase of ChIP-Seq studies will help us to map all possible PDIs and later determine which subsets are significant for the regulation of a specific process in a cell type.

Challenges in chromatin immunoprecipitation

Chromatin immunoprecipitation has been used to study protein-DNA associations *in vivo (88-90)*. Proteins are first cross-linked to DNA using formaldehyde or other cross-linkers to retain the *in vivo* structure of protein-DNA complexes. Transcription factor binding sites or loci with specific histone modifications

are identified using traditional immunoprecipitation techniques. Large-scale ChIP was first performed using microarrays by Ren and colleagues *(91)*. With the advent of cost effective high throughput sequencing, this technique has gained widespread use, as it possible to query whole genomes for loci of transcription factor or other protein (e.g. RNA polymerase II) binding or specific histone modifications (e.g. histone H3K4me3, H3K9me2) *(92)*. Again, the main advantage of this technique is that it enables us to generate an *in vivo* snapshot of protein-DNA binding under various experimental conditions at high resolution.

One major disadvantage of this technique is that the success of these experiments depends largely on the efficacy of the antibody to the transcription factor or histone modification of interest. The non-specificity of antibodies may convolute/obfuscate the signal to noise ratio and thus prevent the identification of true transcription factor targets or loci containing specific histone modifications. Also, the availability of antibodies is often an issue, as many companies generate these antibodies in individual rabbits, mice, goats, and bleed the animal when done. The specificity and sensitivity of native antibodies vary quite immensely, and once a working lot of antibody is depleted, it may prove difficult to repeat experiments. Monoclonal antibodies, which in general are more reproducibly produced, do yield different numbers of protein-DNA binding sites in comparison to polyclonal antibodies *(93)*. In general, it appears that polyclonal antibodies are not saturated as quickly as monoclonal antibodies.

Additional challenges that may prove difficult to address include the low abundance of the protein of interest or transient protein-DNA interactions. Genetic manipulation to address the abundance of the protein or increasing the strength of

crosslinking can result in increasing the false positives of this assay. Finally, this technique can also prove extremely challenging in general, as the protein of interest must be bound to specific fragments of DNA in a differential manner. Therefore, for a successful genome-wide ChIP experiment, the protein of interest must be relatively abundant or induced upon a specific stimulus and the antibody used in the ChIP should be very specific, sensitive, and of sufficient titer.

Tagging the transcription factor of interest using an epitope and using transgenic organisms for chromatin immunoprecipitation is favored as this negates the generation of many different native antibodies and may possibly address the low abundance of proteins of interest. However, the major disadvantage of this method is that it is relatively unknown whether the native transcriptional state of the organism is somehow altered in the generation of the transgenic lines and proper controls must be performed to show that this is not the case. Transforming an organism lacking the protein of interest with a construct containing the fusion protein of interest is necessary to establish that the tagged protein functions similarly to the native protein. Many epitope tags have been used for chromatin immunoprecipitation in mammalian systems, including GFP/YFP, FLAG, HA. Some researchers are moving towards a non-antibody based chromatin immunoprecipitation using a Halo-tag system (Promega) involving capture of a tagged protein using a covalent linkage of the Halo ligand on beads to the Halo-tagged protein.

ChIP-Sequencing (ChIP-Seq) improves the dynamic range issues encountered by genomic amplification of ChIP samples required for ChIP-chip, as the former procedure foregoes the need for amplification. The recent availability of high throughput sequencing machines (*e.g.* Illumina GAII, HiSeq, and ABI SOLiD) has

made ChIP sequencing feasible. ChIP sequencing is also advantageous in comparison to ChIP-chip because DNA binding motifs can be recovered from the sequencing data with confidence and high resolution *(92, 94-96)*.

Genome-wide chromatin immunoprecipitation studies in Arabidopsis thaliana

Less than ten studies reporting the use of genome-wide chromatin immunoprecipitation in plants exist. However, ChIP-PCR for specific loci is commonly used in numerous studies examining the histone modifications of genomic loci or in a few studies of transcription factor DNA binding. One of the reasons that high throughput chromatin immunoprecipitation studies are lacking is the aforementioned caveat of chromatin immunoprecipitation relating to the availability of specific and sensitive native antibodies. In addition to the limited number of plant specific native antibodies available, ChIP in plants has an added difficulty in comparison animals because plants have a cell wall. Plant cell walls contain a high abundance of starch and this may interfere with efficient nuclei isolation. In order to access the chromatin in a plant cell, the rigid cell wall must be broken and the nuclei isolated from the other organelles and cytosol. The starches should be removed as they complicate the nuclei purification and decrease chromatin yield. Circumventing the issues with the plant cell wall proves difficult, as liquid culture of plant cells is not a natural system; researchers are wary of the artifacts that may occur from protoplasting plant cells. Plant physiology does not lend itself easily to tissue culture (with the exception of suspension cell culture), and issues in the consistency between ChIP experiments often arise.

Therefore, the number of high throughput chromatin immunoprecipitation studies in plant pale in comparison to that of in mammalian or other eukaryotic systems (e.g. yeast, drosophila, mouse and human). Several chromatin immunoprecipitation tiling microarray studies have been published *(97-100)*, although some of these have relatively poor resolution and many false positives *(98, 101)*. Recently, two studies have successfully demonstrated the use of ChIP-Seq in floral tissue, identifying the targets of transcription factors key in the establishment of floral identity *(102, 103)*. Although the first chromatin immunoprecipitation sequencing study was reported in 2008, more than two years ago, the lack of plant protein-DNA studies employing ChIP in comparison to non-plant genome-wide ChIP studies reveals the technical challenges facing plant biologists. To our knowledge, only a few plant biology laboratories have been able to successfully master this technique *(104)*.

CONCLUSIONS

Ethylene is involved in a myriad of plant growth and developmental processes and is especially agriculturally relevant because of its role in fruit ripening. However, there are several emerging questions in ethylene signaling and we do not understand how ethylene regulates many different morphological responses. Because the transcription factors EIN3 and EIL1 are necessary and sufficient to mediate most if not all ethylene responses, identification of their targets should help us further understand ethylene transcriptional regulation. A recent genetic study found that EIN3, not EIL1, was the master regulator of the ethylene transcriptional response in etiolated seedlings *(55)*, and that EIN3 and EIL1 have distinct functions in the ethylene response. Therefore, the initial characterization of the transcriptional

program of ethylene responses was most amenable to study by identifying targets of EIN3.

To characterize the ethylene transcriptional response mediated by EIN3, we decided to use a genome-wide approach, chromatin immunoprecipitation followed by high throughput sequencing (ChIP-Seq). The main advantage of ChIP-Seq is that it is *in vivo* experiment capable of mapping global protein-DNA interactions. Prior to this study, a native polyclonal antibody was designed to recognize the non-DNA binding part of EIN3 (residues 349-581 at the C-terminus) and found to detect EIN3 specifically *(50)*. Corresponding genome-wide expression profiles (mRNA-Seq) under ethylene response conditions enable the determination of whether the EIN3 targets were direct/indirect, activated/repressed. In addition, the correlation of ChIP-Seq and mRNA-Seq data may be used to identify the minimum complement required for the ethylene transcriptional response. Finally, integration of the protein-DNA interactions, protein-protein interactions, and the ethylene transcriptional response can serve to generate a coherent picture of the global transcriptional profile in response to ethylene signaling.

The interconnection of ethylene and other hormone responses require a mapping of protein-protein and protein-DNA interactions, as it is likely that a subset of these are relevant for a specific morphological response, either to an environmental stimuli or developmental regimen. Genome-wide studies coupled with specific genetic and phenotypic studies will help us gain an understanding of the significance and extent of plant hormone crosstalk.

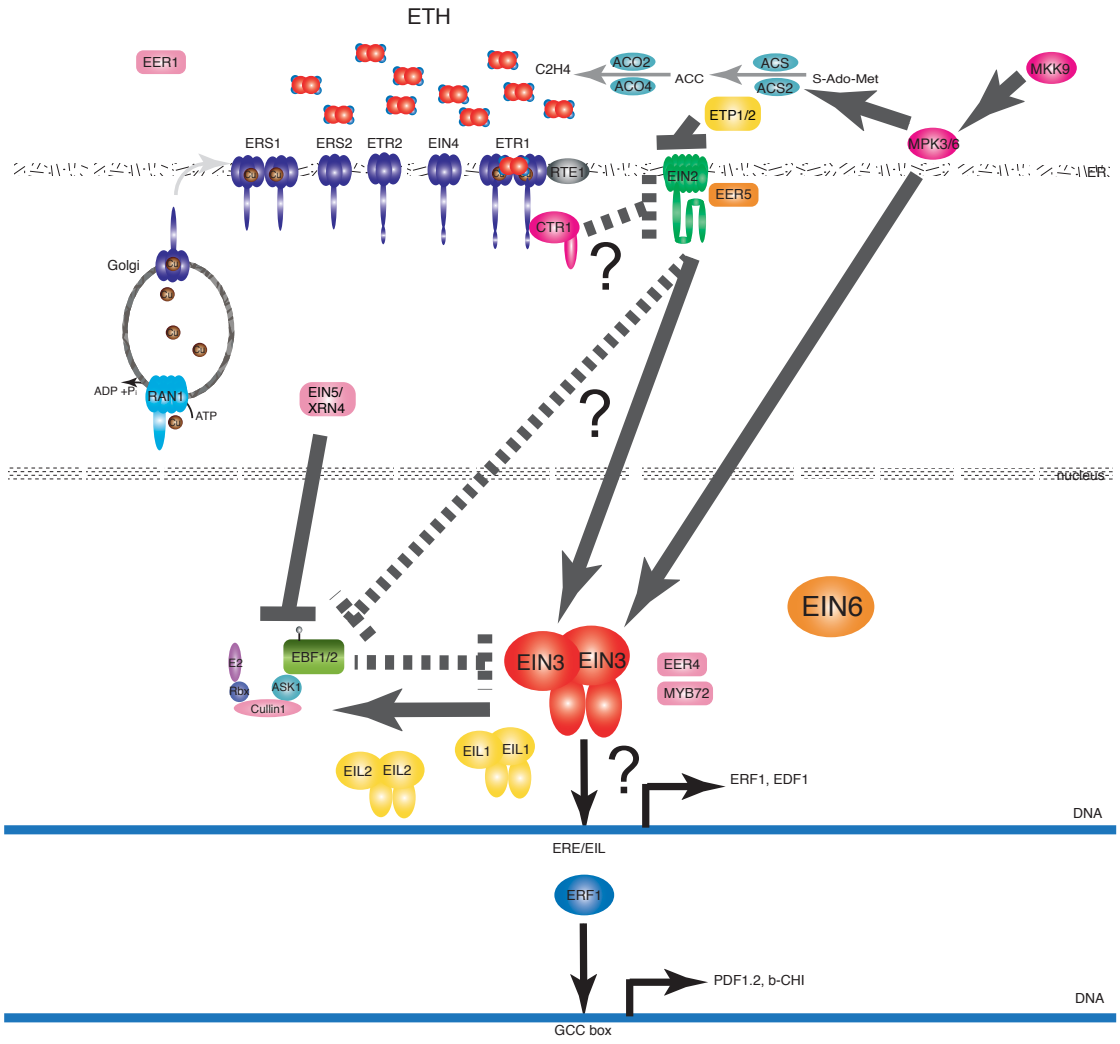


Fig. 1. Current model of ethylene signaling.

A

```
ERF1:                 ---------TCAAGATACATGCCCCCTTGAATCC----
EDF1:                 ---------AAATAGTGCATCCCCCCTTTGATCA----
EDF2:                 ---------TTTTAGTTCATCCCCCTTTGAAGCA----
EDF3:                 ---------GTTAAGTTCATTCCCCCTGATGTTA----
EDF4:                 ---------GTGAAGTTCATCCCCCCTGAGTTTT----
EBF2:                 ------ACTCTTAAGTTCAATGTATACACGCGAAAAAG
SK1-7:                -------GCGCCTCCATCATGCCGGCCACATCCCG---
SK2-3:                AGACCCTAAACATGCCTCATTAGCTATA----------
EIN3_ChIP-Seq:        -----------AASATKCRT-----------------
TEIL:                 ------------AGRTWCRT-----------------
PORA_B:               ---------------TACAT-----------------
PORB:                 ---------------TTCAAA----------------
SID2:                 --------------RTWCRT-----------------
FLS2_Distal:          ------------AGATGCAT-----------------
FLS2_Proximal:        ------------ACATACAT-----------------
```

B

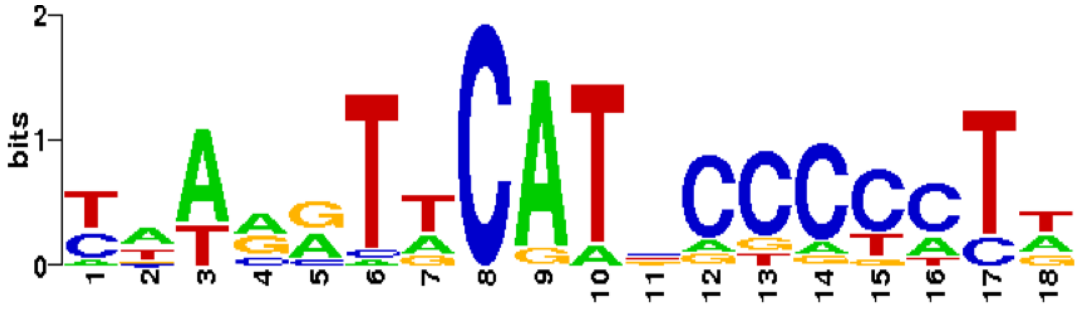


Fig. 2. EIN3 DNA binding motif. (**A)** DNA binding motif alignment of known EIN3 targets. (B) Weblogo representation of alignment.

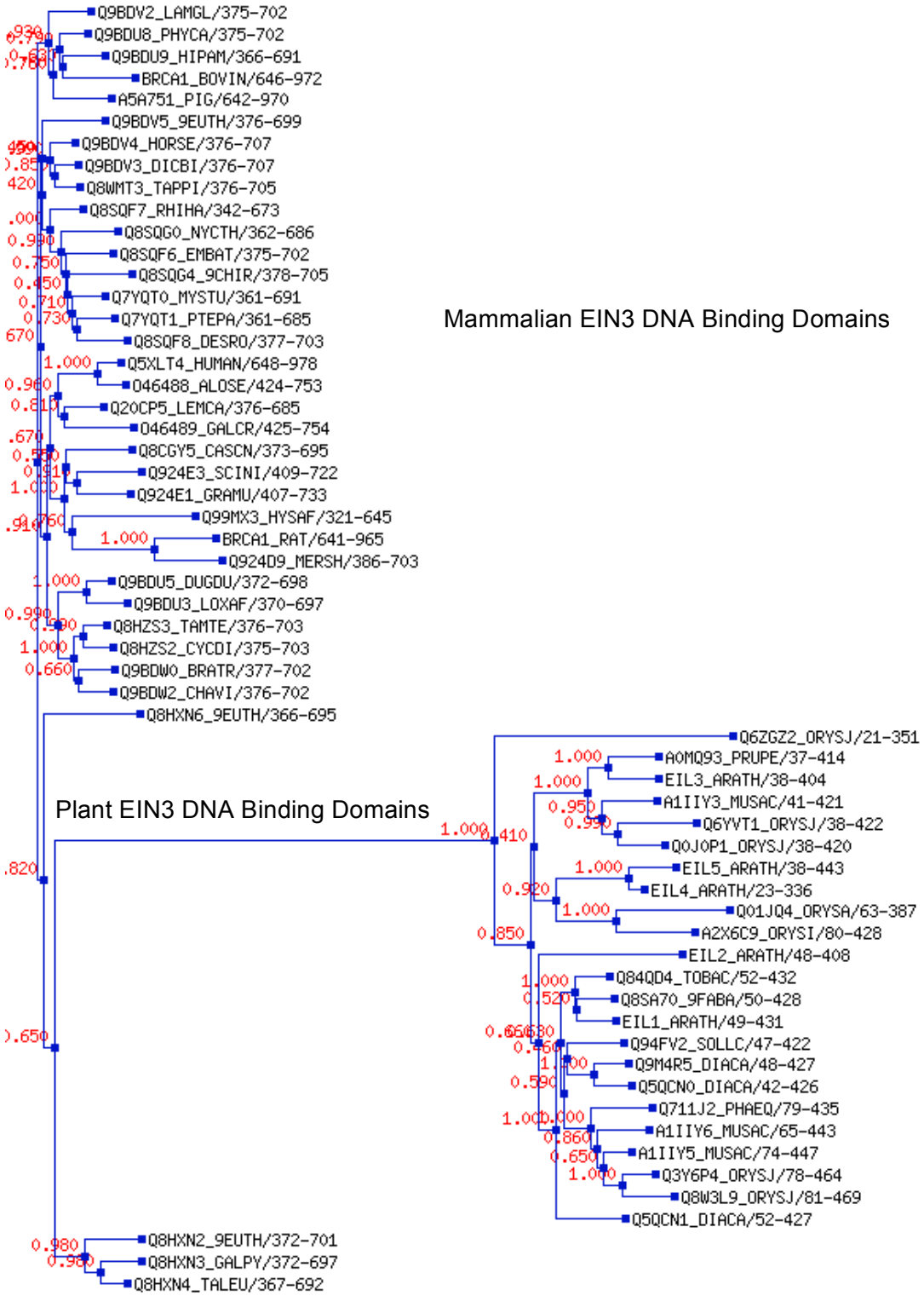


Fig. 3. EIN3 DNA binding domain phylogeny. Phylogenetic tree from Pfam database (http://pfam.sanger.ac.uk/family/PF04873). Viridiplantae out group at the right.

Chapter One, in full, consists of the following manuscript to be submitted as a review in the format of the journal Current Opinion in Plant Biology.

Chang KN, Ecker JR. Unraveling the role of ethylene in plant hormone crosstalk.

I was the primary researcher and author for this manuscript and Joseph R. Ecker directed and supervised the research that formed the basis of this chapter.

CHAPTER 2:

DYNAMIC TRANSCRIPTION FACTOR BINDING REVEALS ROLE OF EIN3 IN PLANT HORMONE CROSSTALK

ABSTRACT

The plant hormone ethylene regulates a multitude of growth and developmental processes, however the temporal control and organization of the ethylene signaling network remains elusive. We characterized the dynamic ethylene transcriptional response by identifying targets of the master regulator of the ethylene signaling pathway, ETHYLENE INSENSITIVE3 (EIN3), using temporal chromatin immunoprecipitation sequencing and transcript sequencing. EIN3 binding establishes downstream transcriptional cascades, feedback circuitry of the ethylene signaling pathway, and interconnections between hormone response pathways. Finally, we show that mutants in a gene family targeted by EIN3 exhibit hormone response phenotypes in several developmental stages, demonstrating the integral role EIN3 plays in the orchestration of hormone crosstalk during plant growth and development.

INTRODUCTION

Ethylene, a simple hydrocarbon gas, regulates many biological processes in plants including stem cell division, differential cell growth, stress and pathogen responses, senescence, fruit ripening, and seedling germination. Despite its importance, we lack a comprehensive understanding of how ethylene mediates this myriad of morphological responses. The dynamic nature of the ethylene response, a rapid growth inhibition independent of the master transcriptional regulator ETHYLENE INSENSITIVE3 (EIN3), followed by an EIN3-dependent sustained growth inhibition, calls for a temporal study of the ethylene response *(64)*. EIN3 protein accumulates in an ethylene-dependent manner and is necessary and sufficient for the ethylene response *(50)*. Although hundreds of ethylene response genes have been identified,

distinguishing direct targets from those further downstream in the ethylene transcriptional program is difficult because targets of EIN3 are also transcription factors (*e.g. ETHYLENE RESPONSE FACTOR1* (*ERF1*). Therefore, to understand the dynamics of the EIN3-mediated ethylene transcriptional response, we performed a genome-wide study of ethylene-induced EIN3 protein-DNA interactions using chromatin immunoprecipitation followed by sequencing (ChIP-Seq) and simultaneously determined the repertoire of target genes that were transcriptionally regulated by ethylene (mRNA-Seq).

RESULTS AND DISCUSSION

We performed the EIN3 ChIP-Seq and mRNA-Seq in three-day-old dark grown seedlings during a timecourse of ethylene gas treatment (fig. S1, Table S1, S2). Temporal ChIP-Seq enabled the stringent identification of EIN3 binding regions. We identified 1460 EIN3 binding regions in the *Arabidopsis* genome, associated with 1314 genes (Table S3). Genes associated with EIN3 binding regions are referred to as EIN3 targets.

The majority of EIN3 binding occurred near the transcriptional start sites (TSS), in promoters or in 5' untranslated regions (UTRs) (61%) (fig. S1). Sequences of EIN3 binding regions were significantly enriched in the consensus TEIL motif ($P < 10^{-87}$), and *de novo* motif analysis identified the canonical EIN3 motif (fig. S3). Four of seven known EIN3 targets were identified *(56, 71-73, 105)* (Fig. 1A, fig. S3); those not identified in our study have been shown to function in chlorophyll or pathogen-induced salicylic acid biosynthesis, processes that do not occur in etiolated seedlings. One example of a known target of EIN3 is shown (Fig. 1A). *ETHYLENE BINDING*

FACTOR2 (EBF2), an F-box protein that directs the proteolysis of EIN3 itself, exhibits ethylene-induced transcription and EIN3 binding (Fig. 1A, fig. S4). Two additional binding sites in the *EBF2* promoter were also identified (fig. S4).

We then evaluated EIN3 targets with respect to changes in their steady-state transcriptional levels upon ethylene treatment. We first focused on EIN3 targets transcriptionally regulated by ethylene (EIN3 targets, ethylene-Regulated = EIN3-R). These targets were likely to represent the minimum complement of ethylene response (Fig. 1B). Although genes were both induced and repressed by ethylene (Fig 1C, fig. S6) (t-test p-value ≤ 0.05, fold-difference cutoff ≥ 50%), the majority of EIN3-R (85%) were up-regulated by ethylene, confirming the role of EIN3 as an activator (Fig. 1B, 1C). Consistent with kinetics of EIN3-dependent growth inhibition *(106)*, the ethylene transcriptional response was activated and sustained after one hour of ethylene gas treatment (Fig. 1D). Because many EIN3-R were transcription factors, conjecture that EIN3 initiates a transcriptional cascade is supported (fig. S5) *(107)*.

Numerous studies have reported that transcription factor binding does not necessarily coincide with changes in transcription *(108)*, especially for master regulators targeting other transcription factors resulting in transcriptional cascades. As expected, a subset (29%) of the EIN3 targets were ethylene-regulated (EIN3-R), while the majority of EIN3 targets were not ethylene-regulated (EIN3-NR) (67%), and a negligible amount of target transcripts (4%) were below detection (EIN3-ND) (RPKM < 1) (Fig. 2A, Table. S2).

Although all EIN3 targets share a similar temporal pattern of binding -- induction of EIN3 binding by ethylene -- this binding pattern does not correlate to the steady-state mRNA levels of these targets (Fig. 2B). Because the majority of EIN3

targets were EIN3-NR, we asked if there were distinguishing features that differentiate EIN3 targets that show a transcriptional response to ethylene versus those that did not. EIN3 targets regulated by ethylene (EIN3-R) exhibited more EIN3 binding near the TSS in comparison to EIN3-NR or EIN3-ND, a trend that was consistent over the duration of ethylene treatment (Fig. 2C). Specific gene families were enriched in EIN3-R and EIN3-NR groups (p-value < 0.05). EIN3-R was uniquely enriched in genes in BZR, TIFY, and bHLH transcription factor families (Fig. 2D).

We then focused on all EIN3 targets and asked whether they have a specific functional role (Table S4). The repertoire of EIN3 targets that included hormone-related genes could be classified into three functional categories based on what is known of EIN3 and the ethylene response. EIN3 targets included downstream effectors of the ethylene response, ethylene signaling components, and genes involved in hormone crosstalk (Fig. 3). Several negative regulators of the ethylene signaling pathway *(8)* are targets of EIN3 (Fig. 3A), suggesting that EIN3 is involved in feedback regulation. Support of the functional role of EIN3 in the negative regulation of the ethylene signaling pathway can be found in the literature, as CONSTITUTIVE TRIPLE RESPONSE (CTR1) *(44)*, ETHYLENE RECEPTOR2 (ETR2) *(109)* protein stability is ethylene-induced.

Hormone crosstalk genes have also been previously identified as targets of flowering transcription factors targets *(102, 103)*. In our study, the EIN3 target dataset contains more than twice the proportion of hormone genes than in the genome (46%, Hypergeometric p-value = 10^{-96}) *(78, 110, 111)*. Many genes are involved in more than one hormone response, confirming extensive hormone crosstalk in *Arabidopsis* (Fig. 3B, 3C). However, analysis of the topology of the hormone response network of

protein-protein interactions (PPIs) *(112)* and protein-DNA interactions (PDI)s *(113)* revealed low connectivity (Fig. 3C) *(114)*, possibly due to the lack of available data. Because many hormones act in concert to regulate various biological processes, crosstalk is expected, and indeed EIN3 targets hormone pathways at multiple levels (Fig. 3D). EIN3-NR genes may be gated by other hormones such as auxin, as ethylene is not required for the auxin response *(81)*.

Although we have shown extensive hormone crosstalk with ethylene occurs through EIN3-protein DNA interactions, as evidenced by the over-representation of EIN3 targets with the hormone annotation, the importance and significance of EIN3 binding in hormone crosstalk must be demonstrated. To explore the specific role of EIN3 in hormone crosstalk, we examined an EIN3 target known to be involved in ethylene-auxin hormone crosstalk, *HOOKLESS1 (HLS1)*. *HLS1* links ethylene and light signaling with auxin during differential growth of the apical hook and HLS1 protein is necessary for the accumulation of AUXIN RESPONSE FACTOR2 (ARF2) DNA-binding protein *(80, 115)*.

Close examination of our data revealed that EIN3 targets *HLS1* and two of its most similar homologs *HOOKLESS1-LIKE HOMOLOG1/2* (*HLH1/2*), implicating a mechanism for ethylene-auxin crosstalk (Fig. 4A, fig. S8)**.** Binding of EIN3 to the promoters of *HLS1*, *HLH1*, *HLH2* increased upon ethylene treatment (fig. S8). EIN3 binding of *HLS1* and *HLH2* promoters was specific and significantly decreased in the *ein3-1* mutant (Fig. 4B). In addition, EIN3 binding sites in the *HLS1/HLH2* promoters were similar to known EIN3 motifs (fig. S8). The functional significance of the EIN3 binding site in the ethylene response was supported by a previous study that identified two allelic mutations in the *HLS1* promoter sufficient to yield a weak

hookless phenotype *(116)*. Although the existence of a GCC box motif in the *HLS1* promoter lead to the conjecture that ERF1 binds the promoter of *HLS1 (116)*, EIN3 binding does not negate ERF1 binding of the *HLS1* and *HLH* promoters. However because ERF1 mRNA is induced by EIN3, it is likely ERF1 regulation of *HLS1*/*HLH*s would also be regulated by EIN3.

The steady-state transcriptional levels of the *HLS1* gene family members vary upon ethylene treatment. *HLS1* is up-regulated by ethylene (EIN3-R), *HLH2* is expressed but unaffected by ethylene (EIN3-NR), *HLH1/3* are not expressed in etiolated seedlings (EIN3-ND) (Fig. 4A, lower panel). EIN3 targets *HLS1* and its homologs, and *HLS1* is ethylene-induced, but the question remains whether the ethylene-induction of *HLS1* is dependent on EIN3. Previous studies have shown that induction of HLS1 expression by ethylene is dependent on functional EIN2, and *HLS1* mRNA was absent in *ein2-1* as examined by northern blot *(116)*. *HLS1* steady-state transcript levels were also significantly reduced in the *ein3-1 eil1-1* mutant (data not shown). Taken together, our results in conjunction with the previous data suggest that ethylene and auxin crosstalk of HLS1 is mediated by EIN3. Although the ethylene response is not necessary for basal auxin responses as shown by mutant and reporter line analyses in roots *(81)*, ethylene most likely plays a significant role in ethylene-auxin crosstalk, and we found 18% of auxin response genes *(117)* were EIN3 targets (Hypergeometric *p*-value = 10^{-15}).

Because *HLS1* acts downstream of EIN3 and the *hookless* phenotype is repressed by a mutation in a key player in the auxin response, ARF2, but not upstream components of the ethylene signaling pathway *(118)*, we systematically generated double, triple and quadruple mutants for the *HLS1* gene family to further

understand their role in relation to ethylene-auxin hormone crosstalk. The phenotypes observed support the role of the *HLS1* gene family in auxin regulated plant growth and development (fig. S9-12). We observed severe defects in the embryonic patterning in mutants as evidenced by the dramatic appearance of an additional organ in the apical part of the *quadruple* mutant embryo (Fig. 4C, fig. S10). Abnormal etiolated seedling phenotypes were also evident in *hls1hlh* mutants, only combinations with the *hls1* mutant exhibited an absence of the apical hook, suggesting *HLS1* was the major player in the apical hook formation and maintenance (Fig. 4D). It is logical that *hls1* results in an etiolated seedling phenotype *HLS1* because it is the only ethylene-regulated EIN3 target in its gene family. Adult *hls1hlh1* containing mutants displayed a dwarf phenotype, similar to the auxin mutant *axr1 (119)* (Fig. 4E). Floral morphology of the *hls1hlh1* containing mutants also suggested defects in the concerted regulation of hormone response; mutants displayed two stigmas atop a gynoecium, similar to the *arf3/ettin* mutant floral phenotype *(120)*. A detailed analysis of *HLS1* and *HLH* combination mutants are described in the Supplemental Online Material.

Redundancy of EIN3 targets identified through the ChIP-Seq timecourse analysis lead us to identify the mechanism of the interconnection between the ethylene and auxin pathway through *HLS1*. Only the EIN3-R target in the HLS1 gene family exhibited an ethylene response phenotype in etiolated seedlings. It is possible that the regulation of *HLS1* by EIN3 occurs in a differential manner or through various *HLS1* family members in specific stages in plant growth and development, thus a simple pattern of temporal regulation easily translates to one that is complex. The functional redundancy of EIN3 was not addressed in this study and it is unknown

whether EIN3 homologs may function to also add levels of complexity to ethylene hormone crosstalk.

CONCLUSIONS

A singular temporally increasing pattern of transcription factor binding exists for the master regulator of the ethylene response, EIN3. In a minority of cases ethylene-regulated EIN3 targets seem to behave in a manner consistent with a canonical view of transcription, in which a bound activator elicits transcription upon physical recruitment to DNA. We discovered that EIN3 targeted many genes involved in hormone responses including downstream effectors of the ethylene response, negative regulators of the ethylene signaling pathway, and genes involved in hormone crosstalk. Finally we characterized a family of genes involved in ethylene-auxin hormone crosstalk with various states of transcription, and resulting genetic analysis revealed a multitude of auxin-related phenotypes in numerous stages of development.

Fig. 1. Dynamics of ethylene-induced EIN3 binding and transcription supports the role of EIN3 as an activator of the ethylene response. **(A)** Ethylene treatment results in an increase of EIN3 binding in three regions of the EBF2 promoter, corresponding to an increase in steady-state mRNA levels. Binding and transcription levels are indicated by reads per kilobase per million reads in sample (RPKM). **(B)** Patterns of EIN3 binding and expression of ethylene-regulated targets are strikingly evident over a timecourse of ethylene gas treatment. EIN3 binding increases with ethylene treatment to a maximum at 4 hours of ethylene treatment for all targets. Each line in the heatmap represents the RPKM value for the representative EIN3 binding site (left panel) and transcript (right panel). **(C)** (Upper panel) Equivalent numbers of genes are up- and down-regulated upon ethylene treatment. (Lower panel) The majority of EIN3 targets that are differentially expressed upon ethylene treatment are up-regulated. **(D)** The ethylene transcriptional response was activated and after one hour of ethylene gas treatment.

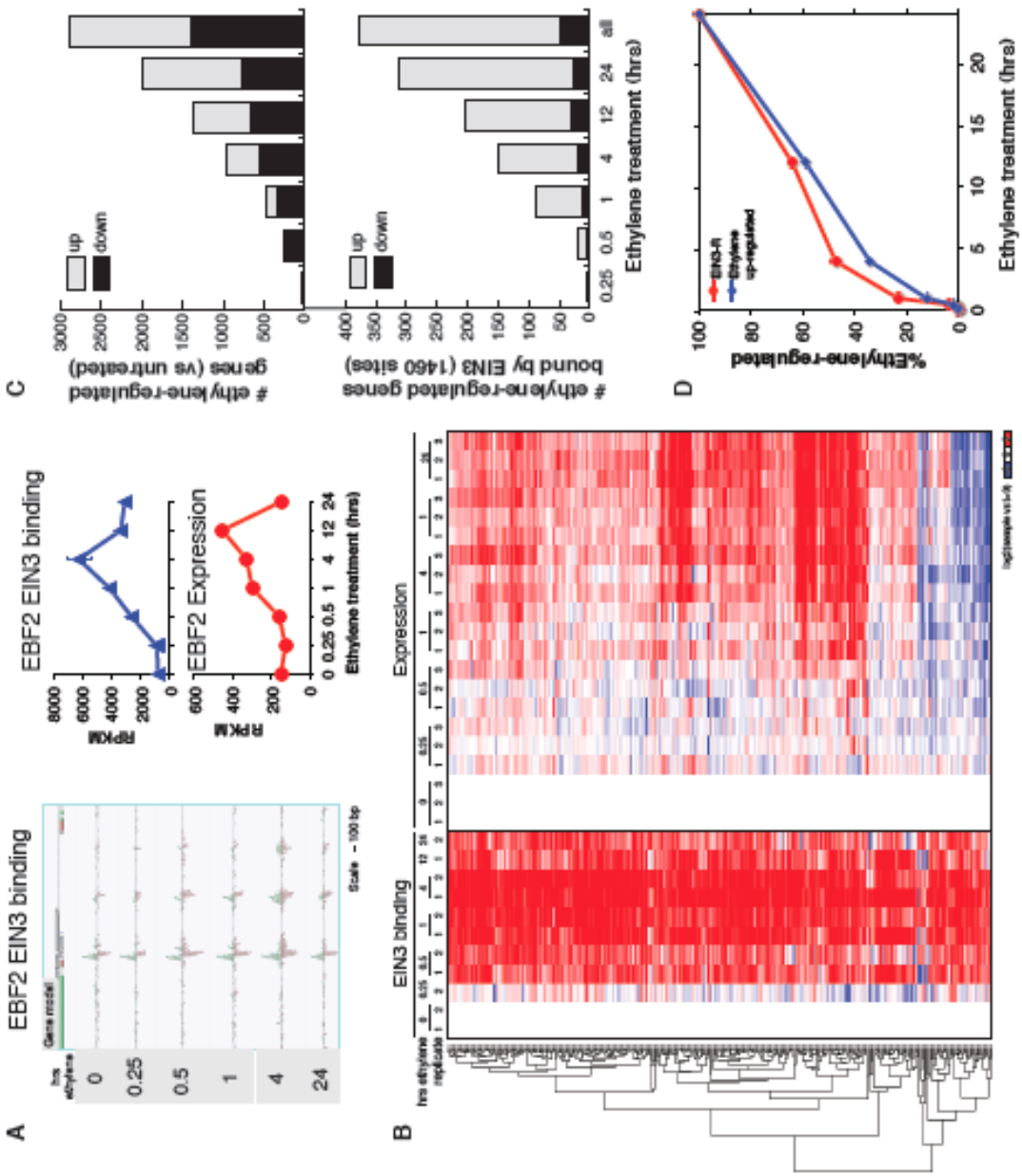

A
EBF2 EIN3 binding
Gene model
Scale - 100 bp
0
0.25
0.5
1
4
24
EBF2 EIN3 binding
RPKM
8000
6000
4000
2000
0
EBF2 Expression
RPKM
600
400
200
0
0 0.25 0.5 1 4 12 24
Ethylene treatment (hrs)
B
EIN3 binding
Expression
C
ethylene-regulated genes (vs untreated)
3000
2500
2000
1500
1000
500
0
up
down
ethylene-regulated genes bound by EIN3 (1460 sites)
400
350
300
250
200
150
100
50
0
up
down
0.25 0.5 1 4 12 24 all
Ethylene treatment (hrs)
D
%Ethylene-regulated
100
80
60
40
20
0
EIN3-R
Ethylene up-regulated
0 5 10 15 20
Ethylene treatment (hrs)

Fig. 2. Singular EIN3 binding dynamic exists for varied transcriptional regulation of targets. **(A, B)** Many EIN3 targets are not ethylene-regulated, yet exhibit a progression of EIN3 binding upon ethylene treatment. **(C)** Ethylene-regulated EIN3 targets (EIN3-R) exhibit increased binding at transcription start sites upon ethylene treatment (black arrows) in comparison to EIN3-NR or EIN3-ND. Each boxplot represents the distribution of EIN3 ChIP-Seq RPKMs near the TSS. **(D)** Distribution of gene families among EIN3-R and EIN3-NR targets reveal unique over-representation of gene families.

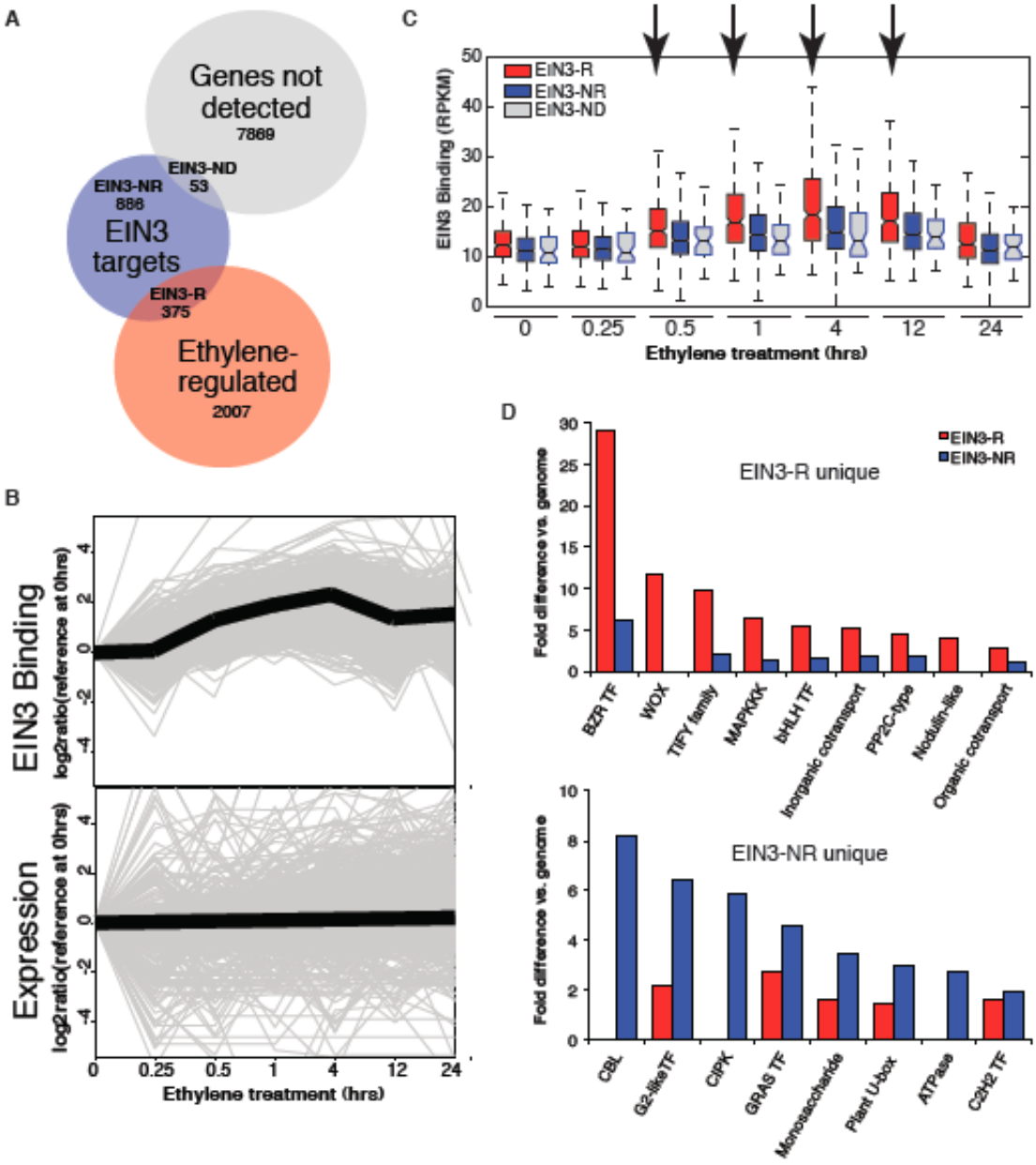

A
Genes not detected
7869
EIN3-ND
53
EIN3-NR
886
EIN3 targets
EIN3-R
375
Ethylene-regulated
2007
B
EIN3 Binding
log2ratio(reference at 0hrs)
Expression
log2ratio(reference at 0hrs)
0 0.25 0.5 1 4 12 24
Ethylene treatment (hrs)
C
EIN3 Binding (RPKM)
EIN3-R
EIN3-NR
EIN3-ND
0 0.25 0.5 1 4 12 24
Ethylene treatment (hrs)
D
EIN3-R unique
EIN3-R
EIN3-NR
Fold difference vs. genome
BZR TF
WOX
TIFY family
MAPKKK
bHLH TF
Inorganic cotransport
PP2C-type
Nodulin-like
Organic cotransport
EIN3-NR unique
Fold difference vs. genome
CBL
G2-like TF
CIPK
GRAS TF
Monosaccharide
Plant U-box
ATPase
C2H2 TF

Fig. 3. Functional classification of EIN3 targets reveals genes involved in hormone responses. **(A)** Feedback (ethylene signaling components) of the ethylene response and feedforward (downstream effectors, bottom). Known EIN3 targets are noted by asterisks; all other EIN3 targets were discovered by this study. **(B, C)** EIN3 targets are involved in hormone crosstalk. Node color represents hormone annotation, as indicated in B; large nodes are EIN3 targets. Blue edges represent protein-protein interactions (PPI) and black edges are protein-DNA interactions (PDI). **(D)** EIN3-mediated ethylene crosstalk occurs at many different levels. PPIs are from the Arabidopsis Interactome Mapping Consortium, and EIN3 PDIs are from this study.

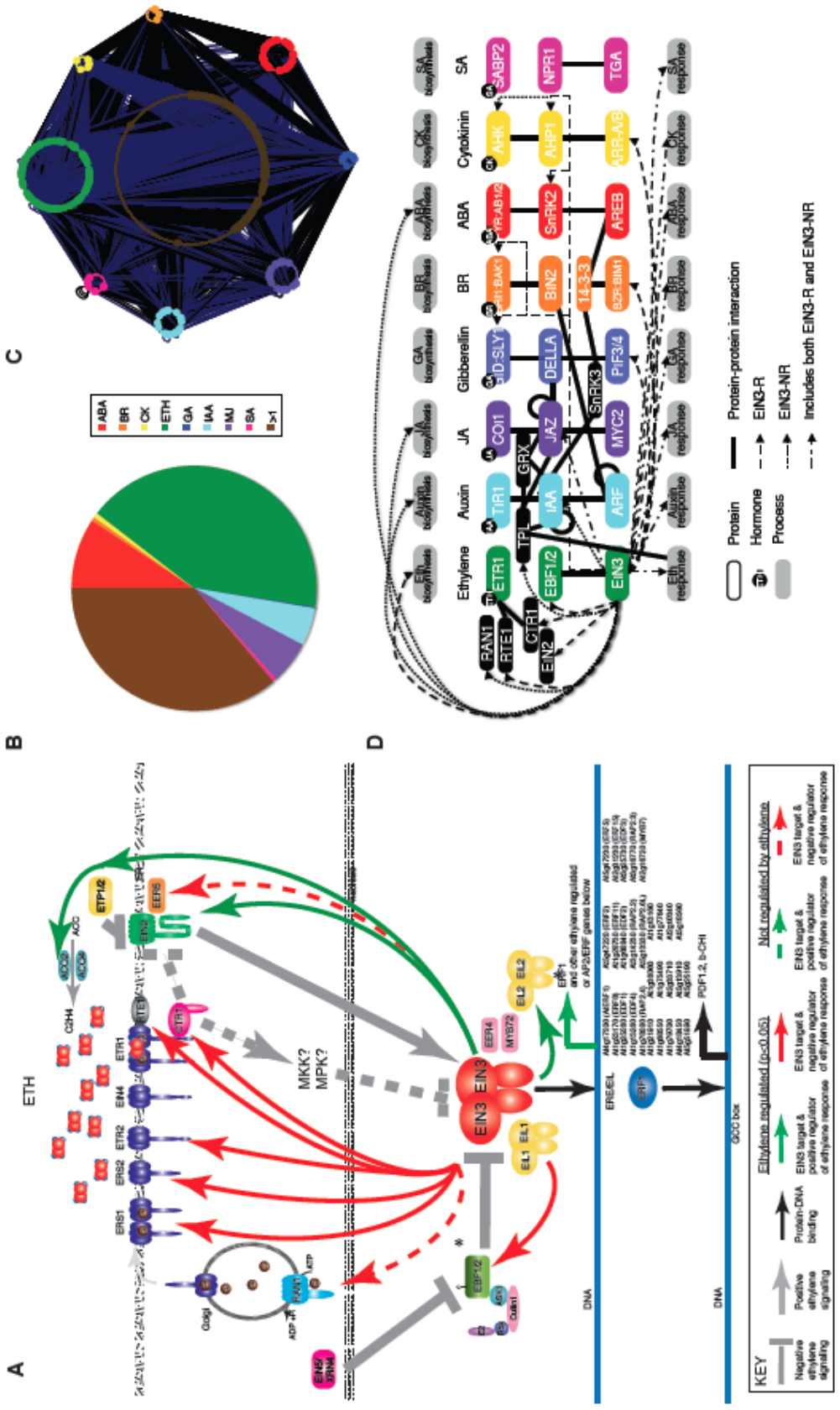
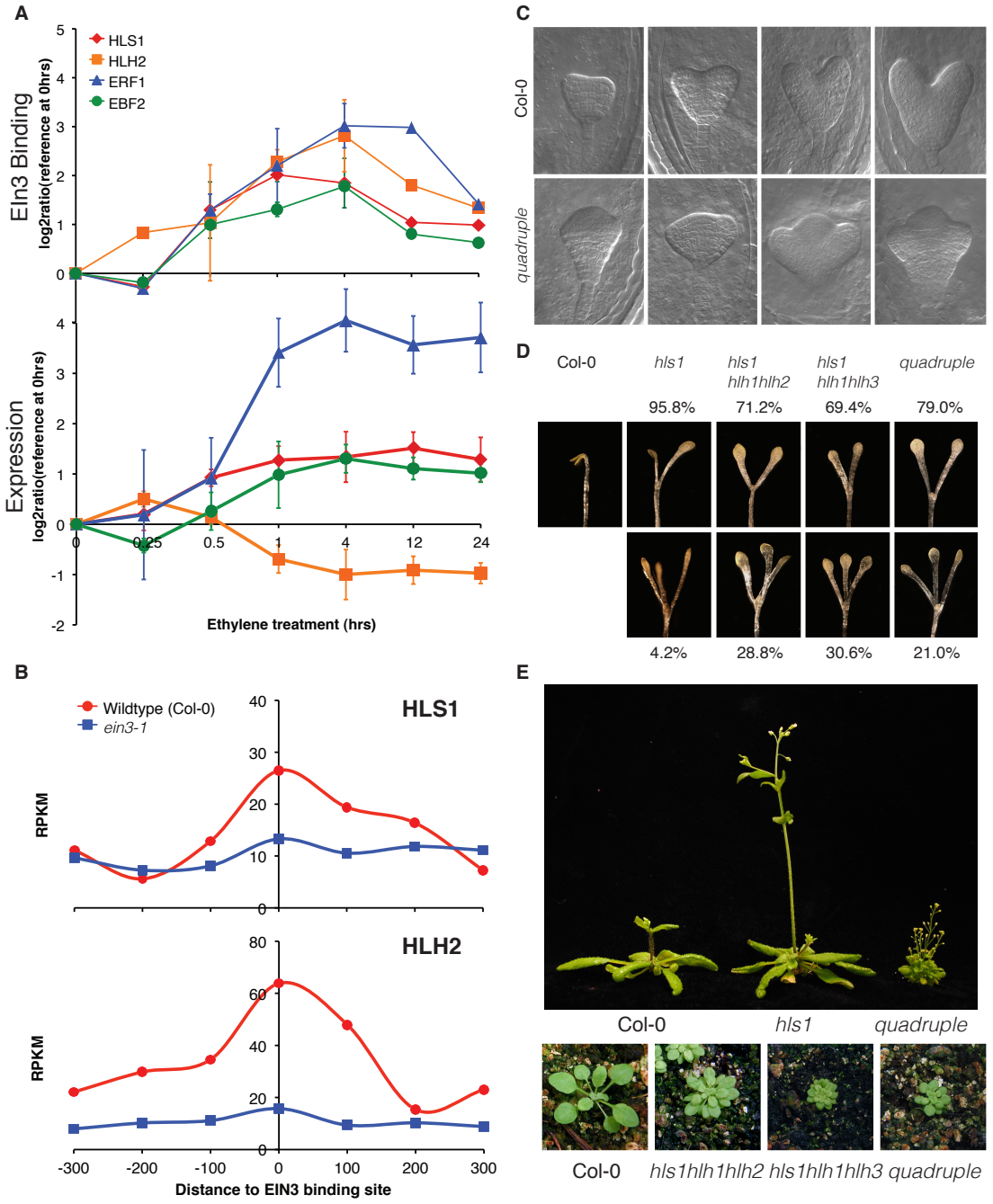

A
B
C
D
ETH
MKK?
MPK?
EIN3
EIL1
EBF1/2
ERF1
PDF1.2, b-CHI
DNA
GCC box
KEY
Negative ethylene signaling
Positive ethylene signaling
Protein-DNA binding
Ethylene regulated (p<0.05)
Not regulated by ethylene
EIN3 target & positive regulator of ethylene response
EIN3 target & negative regulator of ethylene response
ABA
BR
CK
ETH
GA
IAA
JA
SA
>1
Ethylene
Auxin
JA
Gibberellin
BR
ABA
Cytokinin
SA
ETR1
TIR1
COI1
GID-SLY1
BRI1-BAK1
PYR-ABI1/2
AHK
SABP2
CTR1
TPL
GRX
JAZ
DELLA
BIN2
SnRK2
AHP1
NPR1
EIN2
SnRK3
14-3-3
TGA
EBF1/2
EIN3
MYC2
PIF3/4
BZR-BIM1
ARR-A/B
AREB
RAN1
RTE1
Protein
Hormone
Process
Protein-protein interaction
EIN3-R
EIN3-NR
Includes both EIN3-R and EIN3-NR

Fig. 4. EIN3 binding is the mechanism of *HLS1* ethylene-auxin hormone crosstalk. **(A)** (Top panel) EIN3 targets *HLS1* and two of its three homologs. Temporal EIN3 binding and expression patterns are shown with known EIN3 targets as a control. *HLH1* and *HLH3* are not expressed in etiolated seedlings and not shown. Binding of *HLH1* is shown in fig. S8. **(B)** Binding of EIN3 to *HLS1/HLH2* promoters is dependent on presence of EIN3. Enrichment is decreased in the mutant *ein3-1* background. **(C, D, E)** Mutations in *HLS1* and its homologs reveal severe growth and developmental defects. **(C)** HLS1 gene family has a role in embryo patterning. Phase contrast pictures were taken from *quadruple* mutants (*hls1/+hlh1hlh2hlh3*). **(D)** Tri-cotyledon phenotypes in apical hook are dependent on the absence of HLS1. **(E)** Adult three-week-old plants displayed dwarfed phenotypes in long (Top panel) and short day conditions (Bottom panel).


A
HLS1
HLH2
ERF1
EBF2
Ein3 Binding
log2ratio(reference at 0hrs)
Expression
log2ratio(reference at 0hrs)
Ethylene treatment (hrs)
B
Wildtype (Col-0)
ein3-1
HLS1
HLH2
RPKM
Distance to EIN3 binding site
C
Col-0
quadruple
D
Col-0
hls1
hls1 hlh1hlh2
hls1 hlh1hlh3
quadruple
95.8%
71.2%
69.4%
79.0%
4.2%
28.8%
30.6%
21.0%
E
Col-0
hls1
quadruple
Col-0
hls1hlh1hlh2
hls1hlh1hlh3
quadruple

SUPPORTING ONLINE MATERIAL

MATERIALS AND METHODS

Plant material

The *Arabidopsis thaliana* ecotype Columbia (Col-0) was the parent strain for these experiments. Genotypes used for this study include wildtype Col-0, and mutants *ein3-1 (49)*, *ein3-1/eil1-1(10)*, *hls1-1 (hls1)(116), hlh1, hlh2, hlh3* (Fig. S7).

Growth of Arabidopsis seedlings

Three-day-old etiolated seedling tissue was used for these experiments unless otherwise noted. Seeds were sterilized and sown on Murashige and Skoog (cat#LSP03, Caisson) media pH 5.7, containing 1% sucrose and 1.8% agar. After stratification for three days in the dark at 4°C, exposure to light for 2-4 hours to induce germination, seeds were dark-grown in hydrocarbon free air at 24°C for three days. Etiolated seedlings were subsequently treated with ethylene gas at 10 µL L^{-1} for 0, 0.25, 0.5, 1, 4, 12, and 24 hours.

Chromatin preparation and immunoprecipitation

Etiolated seedlings were collected in the dark, immersed in 1% formaldehyde solution, and cross-linked under vacuum for 15 minutes. A final concentration of 125 mM glycine was used to quench the formaldehyde for 5 minutes under vacuum. Cross-linking under vacuum resulted in translucent etiolated seedling tissue. Tissue was liquid nitrogen ground and extraction of chromatin was performed as described in *(121)*.

Chromatin immunoprecipitation (ChIP) was performed as described in *(121)* with modifications, including the use of the Bioruptor sonicator (Diagenode). Bioruptor settings used were: H, 25 cycles of 0.5 min on, 0.5 min off, with 5 minute rests between every 5 cycles. Sonication was performed in a cooling water bath at 4°C. A small amount of chromatin (10 µl) was evaluated for shearing; the size range of chromatin was 150-700 bp, the majority of fragments at 300-400 bp.

Affinity-purified rabbit polyclonal antibodies capable of detecting the C-terminus of EIN3 were used in immunoprecipitation reactions. Details regarding the generation of EIN3 antibodies were previously described *(50)*. Prior to the experiments in this study, the amount of purified EIN3 antisera per immunoprecipitation reaction was optimized and 8ul of purified EIN3 antisera was determined to yield the optimal enrichment of the ERF1 promoter, the known target of EIN3 (data not shown). We then substituted Dynabeads Protein A (Invitrogen, cat#100-1D) and Dynabeads M-280 Sheep anti-Rabbit IgG (Invitrogen, cat#112-04D) for the salmon sperm DNA blocked Protein A agarose beads recommended in the protocol *(121)(121)(121)(4)*, as to avoid sequencing of salmon sperm DNA. Immunoprecipitation and washing of Dynabeads were performed using the buffers in 17*(121)*, otherwise Dynabeads were used as per the manufacturer's instructions. Multiple pipetting steps were performed while washing the beads to reduce non-specific binding carryover. Resulting ChIP DNA was purified as in *(121)*.

Quantitative PCR revealed that relative ChIP enrichment for the promoter of ERF1 performed with the Dynabeads M-280 Sheep anti-Rabbit IgG was higher in comparison to Dynabeads Protein A (Fig. S1A). Thus, Dynabeads M-280 Sheep anti-Rabbit IgG was used in all subsequent experiments. Primers for the ERF1 promoter

encompassing the EIN3 binding site, are as follows: F-GGGGGCATGTATCTTGAATC, R-TGCTGGATCAACTCAACAAAA. Actin primers were as in Mathieu et al. *(122)*. Enrichment was calculated using the Delta-Delta-Ct method with normalization to the reference Actin; fold change was calculated relative to the control for non-specific binding (EIN3 ChIP performed in *ein3-1* mutant).

ChIP was performed in chromatin derived from wildtype Col-0 three-day-old etiolated seedlings treated with 0, 0.25, 0.5, 1, 4, 12, and 24 hours of ethylene. Two independent biological replicates were used in two replicates experiments for timepoints, 0, 0.5, 1, 4 hours ethylene gas treatment. Single replicates exist for 0.25, 12, 24 hours of ethylene gas treatment.

Total RNA extraction

Total RNA was extracted from liquid nitrogen ground etiolated seedlings using the Qiagen RNeasy Plant Mini Kit with Qiashredder columns (cat#74904), with DNaseI (Qiagen, cat#79254) treatment prior to RNA precipitation in sodium acetate and ethanol. Concentrations of RNA were determined using the ND-1000 spectrometer (Nanodrop). Experiments were performed in three biological replicates for timepoints, 0, 0.25, 0.5, 1, 4, 12, 24 hours ethylene gas treatment.

ChIP-Seq library generation and sequencing

Resulting ChIP DNA from two pooled ChIP reactions above was used to generate a sequencing library as per the Illumina ChIP-Seq manufacturer's instructions. The Illumina Genome Analyzer II was used to sequence the single-read ChIP-Seq libraries as per manufacturer's instructions, for 36-43 bps (Table S1). Raw

sequencing data was analyzed using the Genome Analyzer Pipeline v.1.4.0. Reproducibility of the data is shown in Fig. S1. Although the general reproducibility of the data is lower than what was previously reported *(102, 103)*, it is clear that the reproducibility between biological replicates is much higher than that with respect to the control 0hr ethylene gas treatment timepoint. We did not extend raw reads for calculation of reproducibility but instead determined the reproducibility of RPKM values between replicates.

PolyA selection and mRNA-Seq library generation

At least 80 µg total RNA was subject to polyA selection using the Poly(A)Purist MAG Kit (Ambion, cat#AM1922). PolyA RNA was subsequently concentrated by ammonium acetate ethanol precipitation and concentrations were determined using the Qubit fluorometer (Invitrogen) and the Quant-iT RNA Assay Kit (Invitrogen, cat#Q33140). 50-100 ng of polyA RNA was used in a strand-specific library preparation as per the SOLiD Total RNA-Seq Kit protocol (Invitrogen, cat#4445374) and AMPure XP beads (Agencourt, cat#A63881) were used for purification of cDNA and amplified DNA. Samples were barcoded for multiplexing using the SOLiD RNA Barcoding Kit (Invitrogen, Module 1-16 cat#4427046, Module 17-32 cat#4453189, Module 33-48 cat#4453191) as per manufacturer's instructions; final size selection was performed using AMPure XP beads instead of the PAGE purification recommended in the protocol. Size selected libraries were then purified using the MinElute Gel Extraction Kit (Qiagen, cat#28604). Resulting concentrations of libraries were detecting using the Qubit fluorometer and Quant-iT dsDNA High-

Sensitivity Assay Kit (Invitrogen, cat #Q33120). RNA libraries were sequenced for 50 bps on the SOLiD4 platform (Life Technologies) (Table 2).

ChIP-Seq data analysis

The Illumina GERALD module was used to align the sequenced reads to the Col-0 reference genome, version TAIR10 (ftp://ftp.arabidopsis.org/). The analysis variable for the ELAND alignment program was set to eland_extended, as read length was greater than 32 bases (e.g. 36-43). Resulting aligned unique single copy reads were used in ChIP-Seq peak analysis (Table S1).

Saturation analysis of the ChIP libraries was conducted using the spp software *(123)*revealed that all samples were at least within 15% of saturation. Peak analysis was performed individually on each timepoint in each biological replicate using the corresponding 0 hour ethylene treated wildtype Col-0 EIN3 ChIP sample as a control. Two additional ethylene treated (4 hour) wildtype EIN3 ChIP biological replicates were included in the analysis, with corresponding mutant *ein3-1* ethylene treated (4 hour) EIN3 ChIP samples as controls. Three software packages: spp *(123)*, MACS *(124)*, PeakSeq *(125)* were originally used to identify peaks/regions of binding (Fig. S2). Parameters for each software was as follows: MACS (*p*-value = 0.01), spp (FDR = 0.1), PeakSeq (FDR = 0.1, mingap = 200, minhit = 20, minratio = 3.5). Binding regions were merged when the maximum gap between two peaks was less than 200 bp determined by separate software packages. Subsequent analysis was performed in R. Overlapping peaks in one biological replicate in one timepoint by more than one software package were retained as binding regions. Because of the variation of the number of called peaks in each software and each timepoint, we used

a majority vote to call peaks to identify all high stringent EIN3 targets (Fig. S2). PeakSeq results differed significantly from spp and MACS (12 - 76%), therefore only spp and MACS were ultimately used.

Using this method, 1460 EIN3 binding regions were identified. For each EIN3 binding region, the reads per kbp of binding site per million sample reads (RPKM) were calculated. Median normalization of the RPKM values between timecourse biological replicates was performed in R. Resulting RPKMs were log2 transformed with respect to the 0 hour ethylene treatment wildtype Col-0 EIN3 ChIP. Normalization with respect an input genomic control did not produce distinctively different EIN3 binding pattern profiles (data not shown). EIN3 binding regions were then associated to a gene if located within 5 kbp. The nearest expressed gene (RPKM>1) was assigned if there were more than one gene within 5 kbp. If both genes were not expressed, the nearest gene was selected. Distance was determined from the binding region center to the gene feature using the TAIR10 annotation (ftp://ftp.arabidopsis.org) (Fig. S1).

EIN3 binding profiles of previously determined targets are shown in Fig. S3. Data from biological replicate 1 is shown; biological replicate 2 results were similar. Four of seven previously determined EIN3 targets were identified as EIN3 targets in our dataset. Browser images of data were generated using AnnoJ *(126)*. ChIP browser images display read tracks normalized per library, the lowest number of reads for all ChIP samples was used as a minimum. This minimum number of reads was randomly selected from all other libraries for display, to effectively visualize enrichment among different samples. The trends in the data were reproducible

statistically (Fig. S1), and also evident in the visualization of data (see example of EIN3 binding for both biological replicates in EBF2 promoter depicted in Fig. S4).

Motif identification was performed with the matrix screening software Patser *(127)* and the known EIN3 consensus motif (TEIL) from TRANSFAC previously determined using SELEX*(128)*. ClustalW2 was used to align motifs *(www.ebi.ac.uk/Tools/msa/clustalw2/)*. Consensus motif representation of the three EIN3 binding sites in the promoter of EBF2 is shown in Fig. S4.

Gene ontology over-representation of ethylene-regulated EIN3 targets (EIN3-R)

Gene ontology over-representation of selected groups of genes were visualized and determined using the Cytoscape v.2.8.1 *(129)* plugin BiNGO v.2.44 *(107)*. The hypergeometric test was used with Benjamini and Hochberg multiple testing correction (FDR = 0.05). The GOSlim_Plants Ontology was used for *Arabidopsis thaliana* (Fig. S5).

Motif analysis of EIN3 binding regions

EIN3 binding sites were ranked using the R package timecourse, which has been previously used to analyze microarray timecourse data. We used this R package because no available software to analyze timecourse data for ChIP-Seq data exists. The top 50 EIN3 binding regions were determined and the repeatmasked. *De novo* motif analysis of these top 50 EIN3 binding regions was performed using SOMBRERO *(130)*, and alignment to known Arabidopsis motifs (AGRIS, http://arabidopsis.med.ohio-state.edu/) was performed using STAMP *(131)* (Fig. S3).

mRNA-Seq analysis

The SOLiD Bioscope v.1.3 software was used to align the reads to the Col-0 reference genome TAIR10 (ftp://ftp.arabidopsis.org/). Two perfect matches per location were allowed. Exonic expression was determined (RPKM) using mRNA-Seq reads mapping in exons in the direction of transcription. Genes were denoted as expressed if they contained RPKM values greater than one for at least one biological replicate in one timepoint. Differentially expressed genes were then called (t-test *p*-value = 0.05, 50% difference from prior timepoint of ethylene gas treatment), and log2 normalized with respect to the 0 hour ethylene gas treatment control (Fig. S6). Overlap of up- and down-regulated genes was ~1%.

Correlation of EIN3 binding and changes in mRNA steady-state levels

EIN3 ChIP targets were then classified as ethylene regulated (EIN3-R), non-ethylene-regulated (EIN3-NR), and transcription not detected in etiolated seedlings (EIN3-ND). The heatmap (Fig. S6) reveals that there is a singular binding pattern but various transcription profiles, as displayed in Fig. 2.

Although the majority of EIN3 targets were up-regulated by ethylene, consistent with the previously determined role of EIN3 as an activator, a subset of EIN3 targets was repressed upon ethylene treatment; one instance of EIN3 as a repressor has been previously reported *(71)*.

Generation of hormone crosstalk network

The most current protein-protein interaction network for *Arabidopsis (112)* containing high throughput yeast two hybrid and literature curated data was used as

the foundation for the hormone crosstalk network. The protein-DNA interaction network AtRegNet from AGRIS (http://arabidopsis.med.ohio-state.edu/; 7918 nodes, 10640 edges) 17included high throughput data (ChIP-chip and ChIP-Seq) for several transcription factors including AGL15, HY5, GL3, AtbHLH15, WRKY53, GL1, E2F, and SEP3 as well as literature curated data *(113)*. Transcription factor-DNA binding interactions from six additional studies were added, including TGA2 *(132)*, AP1*(102)*, BES1*(98)*, BZR1*(97)*, FLC *(100)* in addition to our data. This generated a protein-DNA interaction network of 8531 nodes and 11953 edges, which was then merged with the protein-protein interaction network. Protein-protein interaction and protein-DNA interaction edges were indicated by black and blue lines, respectively.

To identify genes associated with a hormone signal or response, we used the annotation in the Arabidopsis Hormone Database *(110)* (http://ahd.cbi.pku.edu.cn/) in addition to other datasets including relevant ethylene microrarray studies in etiolated seedlings *(78, 111)*. Hormone annotation attributes were imported into Cytoscape *(129)* and colored according to hormone. The amount of genes involved in hormone responses in the genome was 21% (5729/27416), where as the amount of genes involved in our EIN3 target group was 46%.

HLS1 gene family phenotypes in different stages of plant growth and development

Identification of loss-of-function mutants for the HLS1 homologs

We identified loss-of-function mutants and performed thorough genetic analyses of *HLS1* and its homologs to characterize the effect, if any, these genes have on the ethylene response. Three HLS1 homologs (HLHs) exist in *Arabidopsis*

genome. The protein sequences of the HLHs are homologous to the full-length protein (Fig. S8). Like *HLS1*, these homologs contain acetyltransferase domains at the N-terminal portion of the protein. Phylogenetic analysis of *HLS1*-like genes with acetyltransferase domain containing proteins from various organisms revealed that the HLS1 family of acetyltransferases form a unique plant-specific class (Fig. S7). We isolated the *bona fide* loss-of-function mutants in the coding regions of the genes for all the HLH genes using the Salk T-DNA mutant collection (Fig. S7) *(111)*. The single knockout mutants of the HLHs exhibited normal apical hook development and had no obvious developmental defects compared to wildtype (data not shown), indicating functional redundancy among HLS1 family members.

Embryo patterning phenotypes

Loss of *HLS1* and *HLH1* affected both cell growth and patterning and all *hls1hlh1* containing mutants were completely sterile. We therefore dissected the embryos of the individual *hlh* mutants, as well as the *hlh1hlh2, hlh1hlh3* and *hlhlh2hlh3* mutant combinations with the heterozygous *hls1-1* mutant. Abnormal embryos were observed in about a quarter of the total embryos examined starting from early globular stages of the embryo development (Fig. S9).

In wildtype embryos, both cell division rate and patterns are highly regulated during plant embryogenesis. However, all embryos from mutants containing *hls1hlh1* mutations had extra cells in comparison to wild type embryos. In normal embryogenesis, cell division rates and patterning are tightly regulated, however, the *hls1hlh1* containing mutants were accelerated in cell division rate and exhibited irregular cell division patterns (Fig. S9). At the globular stage, the mutant suspensor

contained additional longitudinal cell divisions resulting in multiple cell files in the suspensor in comparison to the single cell file that existed in wildtype (Fig. S9). Additional cell divisions also occurred in the apical part of the mutant embryos (indicated by arrows in Fig. S9), resulting in extra layers of cells during early heart stage (referred to as 'triangular stage'). The extra cells continued to divide and expanded irregularly during heart stage of the embryogenesis, giving rise to an extra organ at the apical part of the embryo (Fig. 4C), possibly accounting for the tri-cotyledon phenotype of the mutant seedlings.

Cell elongation phenotypes in light grown seedlings

Mutants with *hls1hlh1* combinations, *hls1hlh1, hls1hlh1hlh2, hls1hlh1hlh3* and *hls1hlh1hlh2hlh3*, displayed severe cell elongation and growth defects throughout all developmental stages. Hypocotyl and root measurements of 7 day old light grown seedlings revealed that knocking out *HLS1* and *HLH1*, the two closest members within the family, was sufficient to cause a severe cell elongation defect phenotype (Fig. S10).

Adult plant phenotypes

HLS1 and its homologs affect many plant development processes, including cell elongation, growth, apical dominance, embryo patterning, and floral development. While most mutant combinations did not result in obvious defects in adult plant growth and development, *hls1hlh1* containing double, triple and quadruple mutants displayed severely impaired growth and development (Fig. 4E). First, the overall statures of the *hls1hlh1* containing mutants were much smaller than that of wildtype (Fig. 4E).

hls1hlh1 and the *hls1hlh1hlh2* are similar in size and are slightly larger than the *hls1hlh1hlh3* and the *quadruple* mutants (Fig. 4E). Rosette leaves of these mutants were much smaller and circular-shaped with decreased petiole length (Fig. S11), consistent with cell elongation defects of these mutants at the seedling stage. Second, although rosettes of these mutant plants were dramatically reduced in comparison to wildtype, there was an increase in the number rosette leaves in these mutants (Fig. S11). At three weeks of growth, the *hls1hlh1hlh2* mutant contained the highest number of rosette leaves, followed by the *quadruple* mutant, and wildtype. However, the *hls1hlh1hlh3* mutant leaves are the most reduced in size. Third, these mutants flowered within two weeks (long day conditions) or three weeks (short day conditions), earlier in comparison to wildtype flowering times, four weeks (long day conditions), eight weeks (short day conditions) (data not shown). The increased number of rosette leaves (Fig. S11) and the early flowering phenotype (Fig. 4E) indicated that these mutants are possibly accelerated in development. Finally, these mutants completely lack apical dominance, producing numerous inflorescences, and are completely sterile.

Floral phenotypes

hls1hlh1 containing mutant flowers are reduced in size compared to wildtype, are completely sterile, and also have abnormal floral development (Fig. S12). Although all four types of floral organs (e.g. sepals, petals, stamens and carpels) are present in these mutant flowers, the anthers of these mutants are poorly developed and produce no pollen. In addition, sterility cannot be rescued by manual pollination with wildtype pollen, indicating both male and female gametophyte development was

severely affected in these mutants. Floral organ fusion also occurred in these mutants, e.g. stamen-petal fusion, stamen-stamen fusion, flower-flower fusion (Fig. S12).

All flowers of the *hls1hlh1* containing mutants contained two stigmas on a normal appearing gynoecium (Fig. S12), a phenotype very similar to that of the *arf3/ettin* mutant. Silique elongation did not occur in the *hls1hlh1* mutants; for example, *hlh1hlh3* and *hlh1hlh2hlh3* mutants, who showed no obvious defect during early stages of plant development, had shorter siliques compared to wildtype even after manual pollination with wildtype pollen (Fig S10). Cell files and guard cell patterning at the surface of the two fused carpels of the *hls1hlh1* containing mutants are abnormal in comparison to wildtype (Fig. S12).

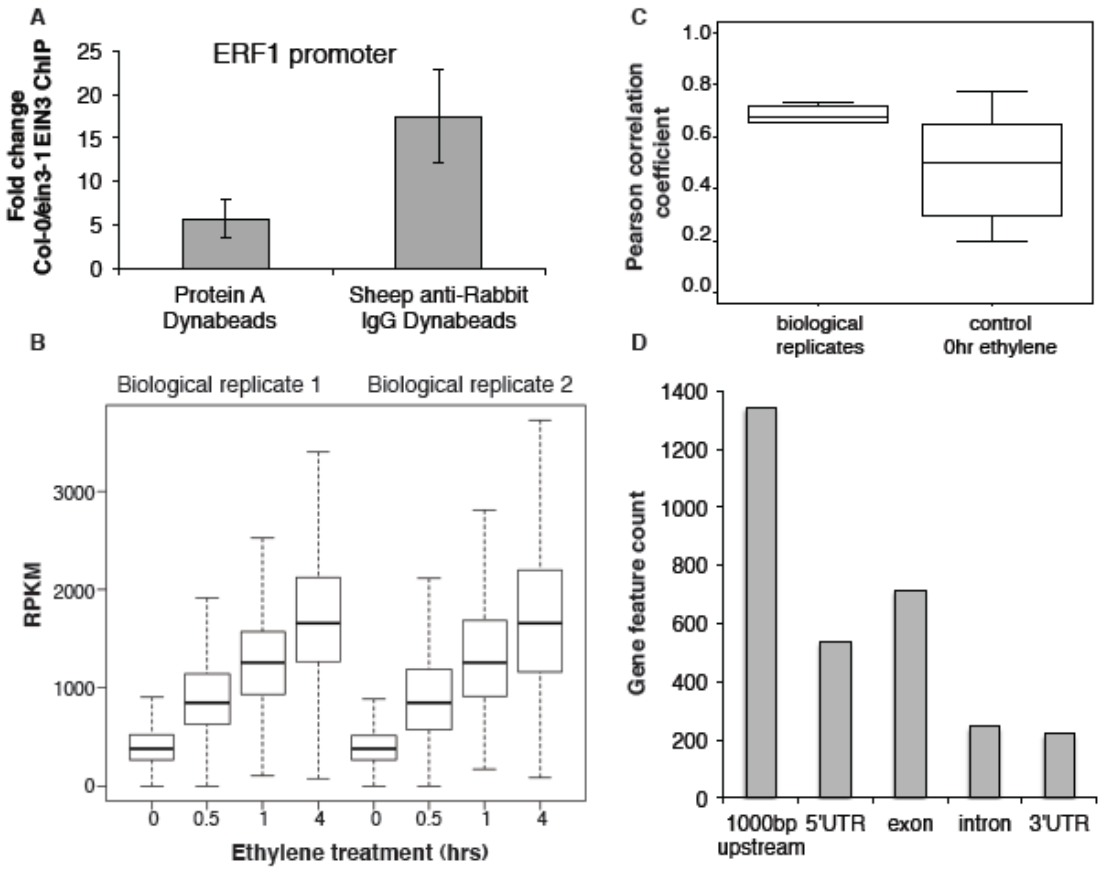


Fig. S1. EIN3 antibody reproducibly enriches DNA in chromatin immunoprecipitation. **(A)** Enrichment of the known target of EIN3, the promoter of ERF1, using Dynabeads Protein A and Dynabeads Sheep anti-Rabbit IgG to collect protein-DNA complexes. The average fold change for two technical ChIP replicates with three QPCR technical replicates each is shown. **(B)** Reproducibility in the two biological replicates for EIN3 ChIP performed upon treatment of ethylene gas for 0, 0.5, 1, and 4 hours. **(C)** Average RPKM of EIN3 binding sites 0, 0.5, 1, and 4 hours of ethylene gas treatment. **(D)** EIN3 binding preferentially occurs in the promoter regions of genes (1000 bp upstream of the TSS).

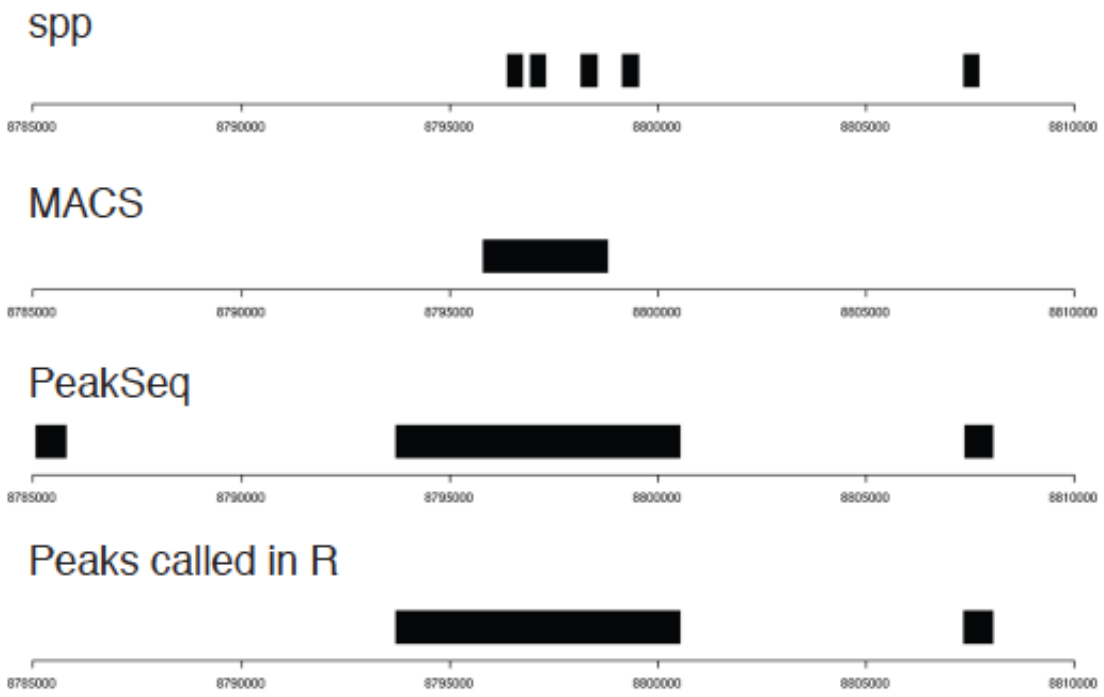


Fig. S2. Example of majority vote schema used to designate binding regions. Peaks called by spp, MACS, and PeakSeq are represented by black blocks, and are shown for a loci for one biological replicate for one timepoint. Peaks called in one biological replicate in one timepoint by more than one software package were retained as binding regions, “Peaks called in R”. Note the characteristics of the PeakSeq peaks are different from the spp and MACS peaks and were ultimately not used in the analysis.

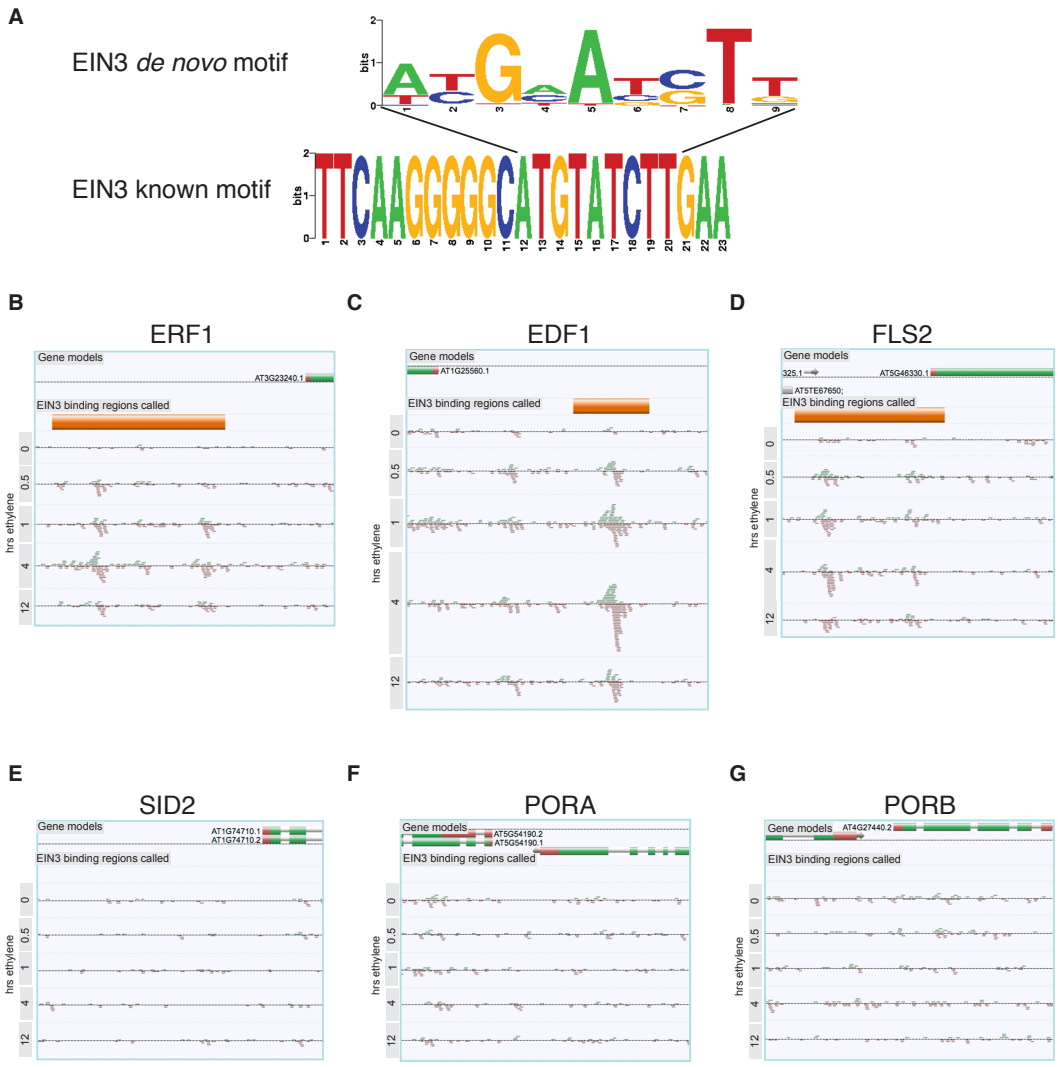


Fig. S3. Binding of EIN3 to previously known targets. **(A)** *De novo* motif from the top 50 EIN3 binding sites with the best match to the known EIN3 motif (E-value = 1.12 x 10^{-5}). EIN3 binding of the promoters of **(B)** ERF1, **(C)** EDF1, **(D)** FLS2, **(E)** SID2, **(F)** PORA, **(G)** PORB.

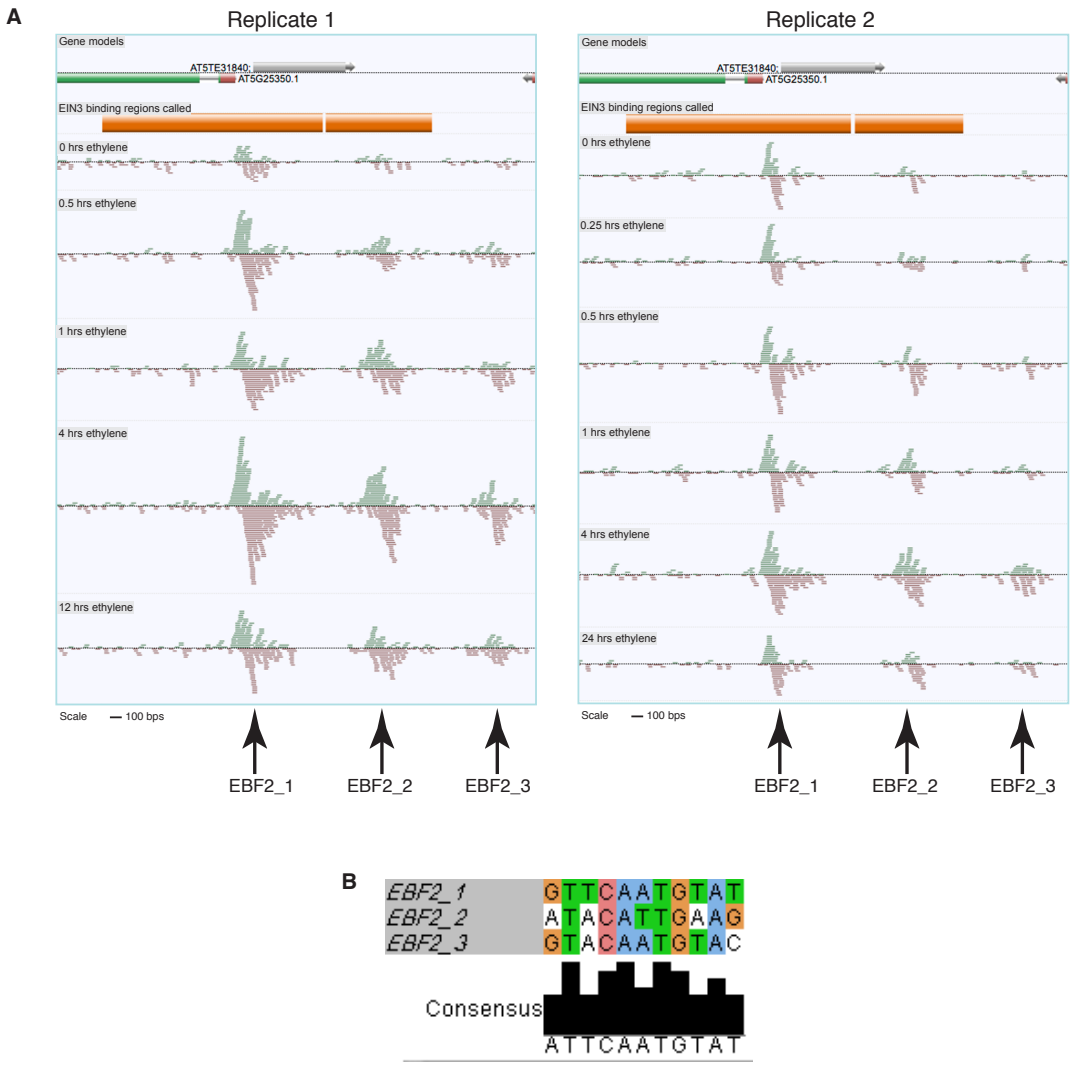


Fig. S4. EIN3 ChIP-Seq identified two additional binding sites in the EBF2 promoter. **(A)** Binding of EIN3 to all three sites in the EBF2 promoter increases upon ethylene gas treatment. Note that EIN3 binding is strongest in the most proximal site to the TSS, and weakest in the most distal site to the TSS. **(B)** Alignment of motifs of the three binding sites in the EBF2 promoter.

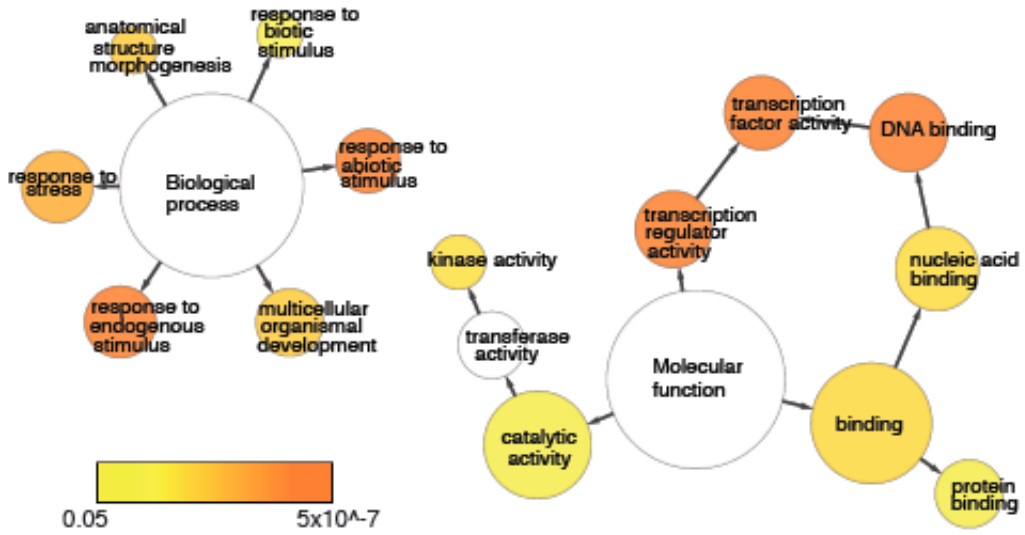


Fig. S5. Functional categories are over-represented for EIN3 targets that are ethylene-regulated (EIN3-R). Network was generated using BiNGO (v.2.44) using the GOSlim_Plants ontology, Benjamini and Hochberg *p*-value legend is indicated below.

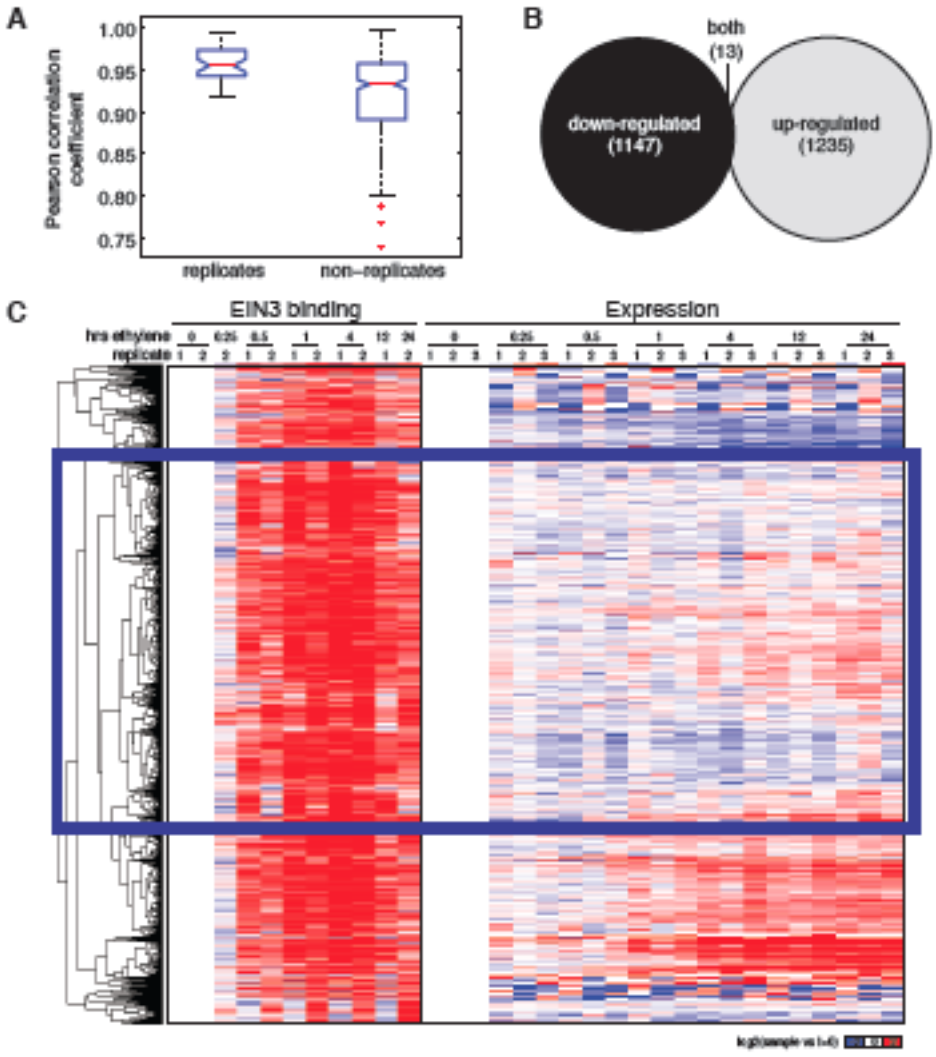


Fig. S6. Ethylene-regulated genes are induced and repressed. **(A)** Reproducibility of RNA-Seq experiments. Ethylene-regulated have a higher reproducibility amongst replicates than non-replicates. **(B)** Genes that are both up- and down-regulated occur at different timepoints of ethylene treatment. **(C)** The majority of EIN3 targets exhibit increased binding upon ethylene treatment, however, changes in steady-state levels of mRNA do not occur for the majority of these targets. EIN3-NR genes are indicated by the blue rectangle.

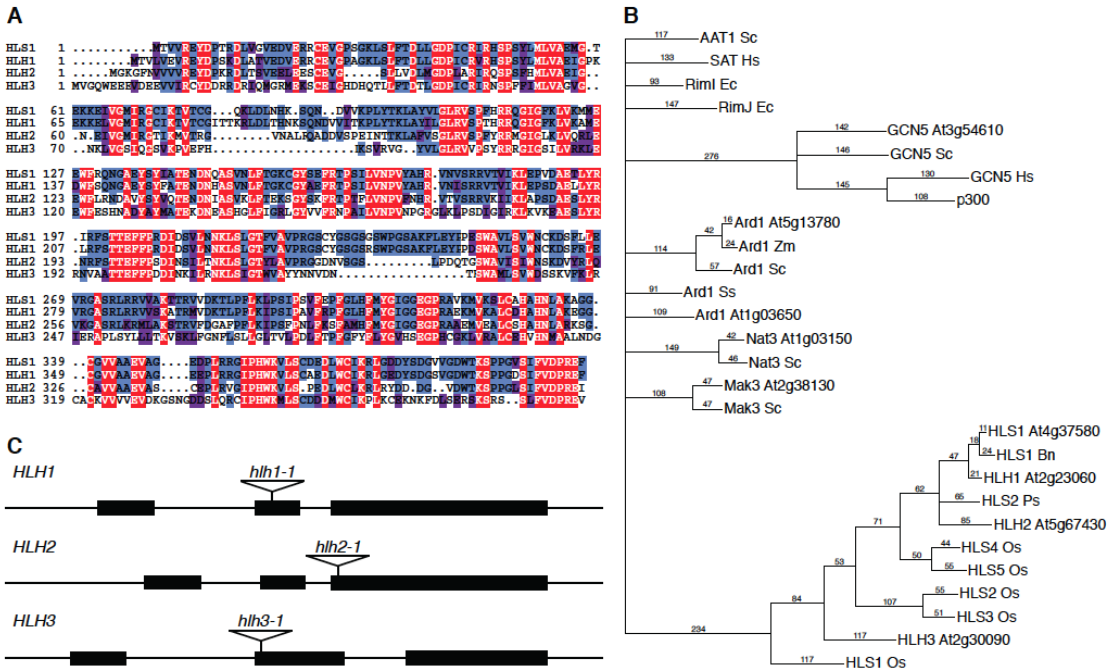


Fig. S7. HLS1-like homologs (HLHs) are similar to HLS1 in protein sequence and domain structure. **(A)** Conservation of HLS1 and HLHs proteins. Amino acid sequence alignment of HLS1 and its three homologs are shown. Gaps are represented as ".". Shading indicates identical sequences (black), conserved changes (gray), similar residues (light gray). **(B)** Phylogeny of HLS1 and HLHs and proteins from other organisms containing acetyltransferase domains. Amino acid sequences were aligned using Clustal, then a bootstrap 50% majority-rule consensus tree was constructed using PAUP. Abbreviations for species are as follows: Hs, Homo sapiens; Sc, Saccharomyces ceravisiae; Mm, Mus musculus; At, Arabidopsis thaliana; Bn, Brassica napus; Zm, Zea mays; Os, Oryza sativa; Ps, Pisum sativum; Ec, E. coli. **(C)** Location of T-DNA insertions in *HLH* genes. Boxes represent the exons of each *HLH* gene. Triangles represent the T-DNA alleles that are characterized in detail. Not all T-DNA insertion alleles in the *HLH* genes are shown.

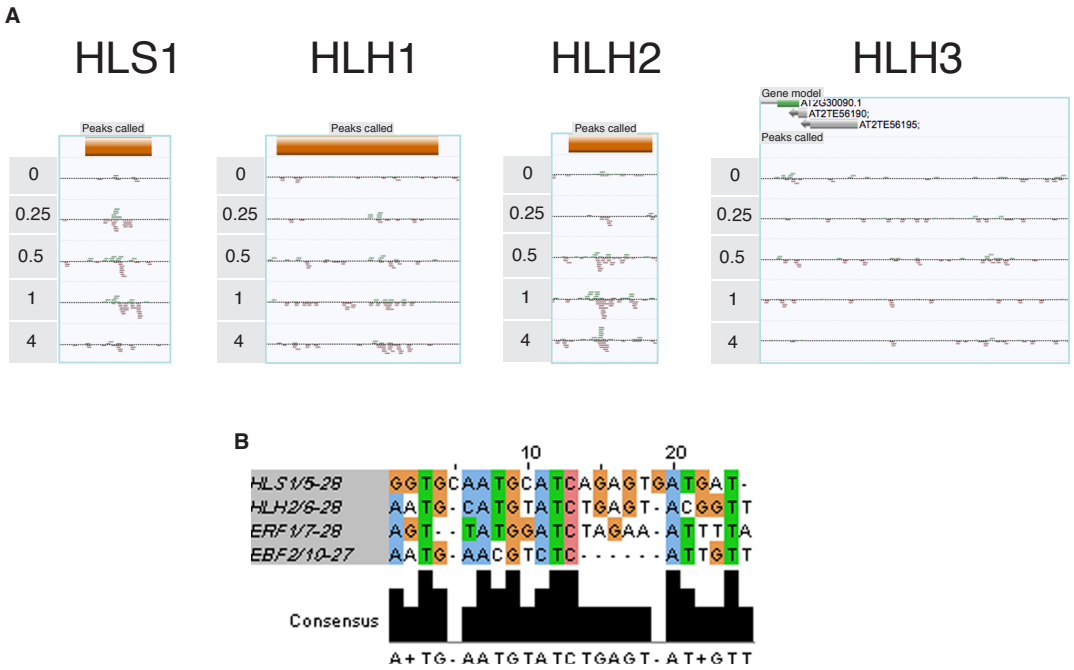


Fig. S8. *HLS1, HLH1,* and *HLH2* are targets of EIN3. **(A)** EIN3 binding in *HLS1, HLH1, HLH2* promoters. **(B)** EIN3 binding motifs in *HLS1, HLH2* reveal a consensus. *ERF1* and *EBF2* motifs are shown as a reference.

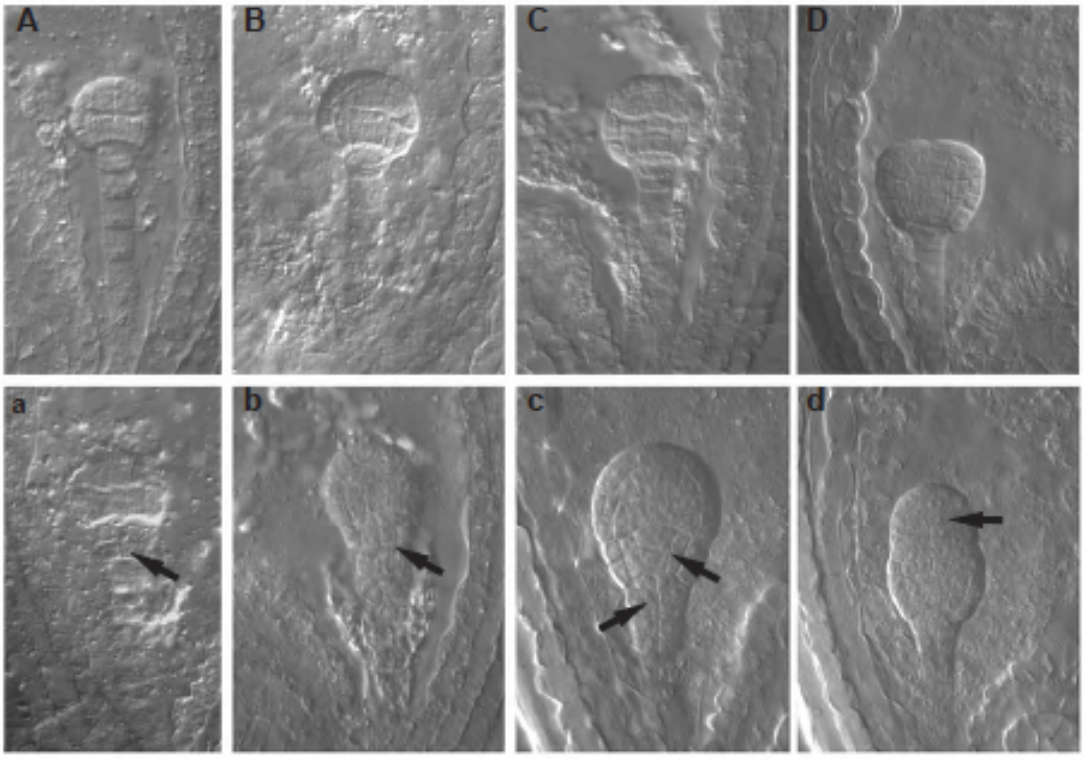


Fig. S9. *HLS1* gene family is involved in embryo patterning. Phase contrast pictures were taken for embryos from the *quadruple mutants* homozygous for *hlh1, hlh2, hlh3*, but heterozygous for *hls1* at various development stages. Lower case letters label the embryos with abnormal patterns, which are from the *quadruple* mutants. Upper case letters label the wild type looking embryos at similar stages corresponding to the mutant embryo stages. Arrows indicate additional cell divisions that occur in the mutant and not in the wildtype.

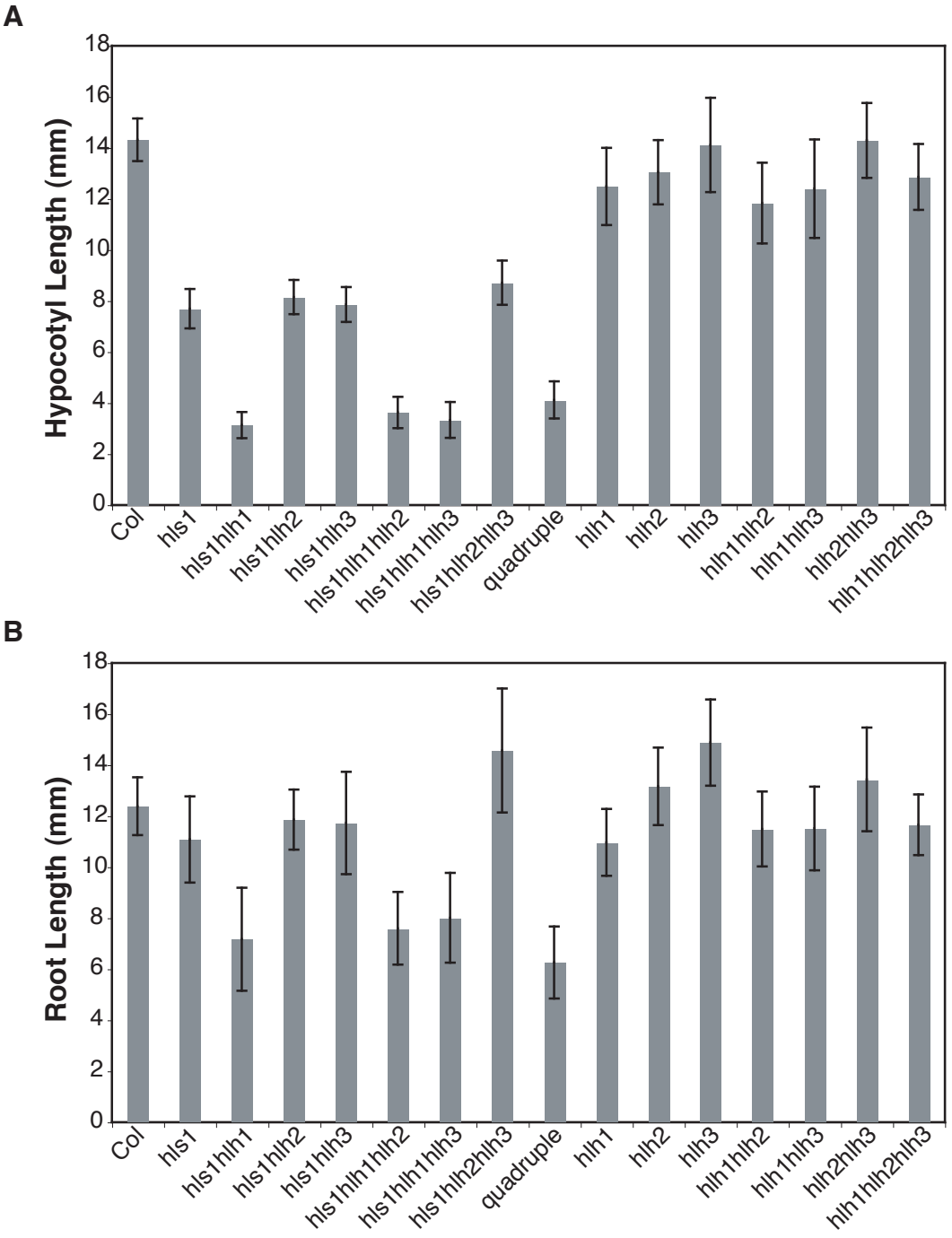


Fig. S10. *HLS1* gene family is involved in cell elongation in seedlings. **(A)** Hypocotyl and **(B)** Root elongation in 7-day-old light grown seedlings are severely impaired in *hls1hlh1* mutants. Error bars indicate standard deviation (N = 30).

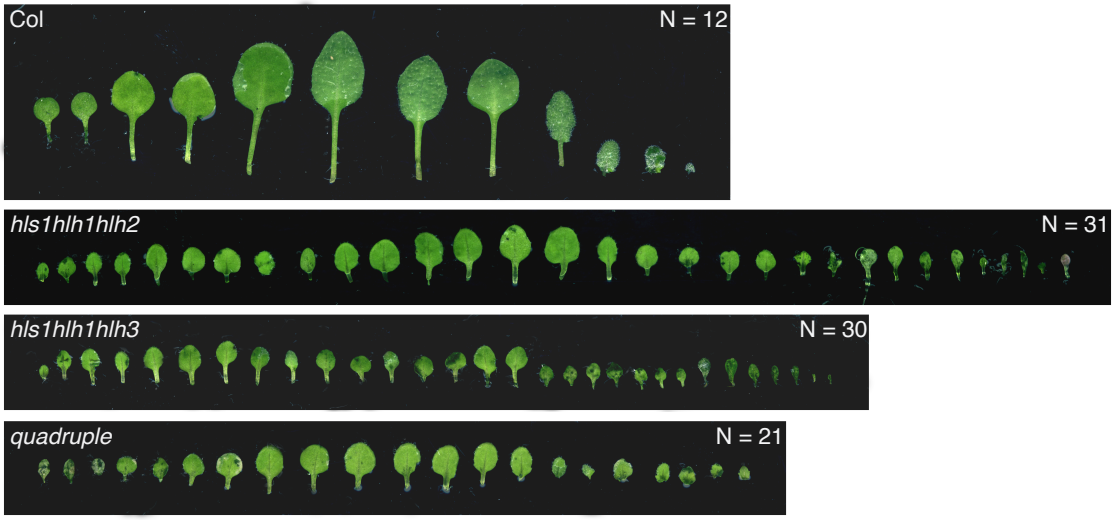


Fig. S11. Numbers of rosette leaves are affected in *hls1hlh1* mutants. Rosette leaf number and shape of the *hls1h1h1* containing mutants in three-week-old plants grown in short day conditions. Numbers listed include the cotyledons.

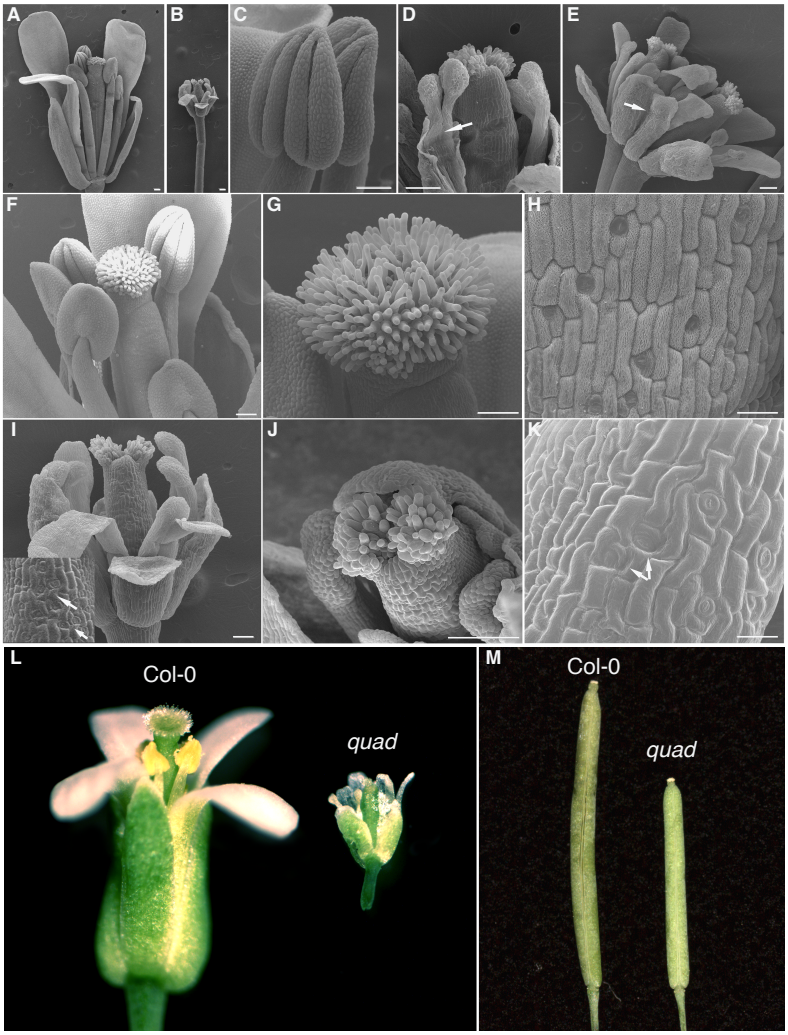


Fig. S12. *HLS1* gene family has a role in floral organ patterning and development. Scanning electronic microscopy images of wildtype (**A), (C), (F), (G),** (**H)** and *quadruple* mutant (**B), (D), (E), (I), (J),** (**K)** flowers reveal severe defects in mutant organs. Scale bars = 100 um. *Quadruple* mutants were not only affected in organ patterning, but overall size; mature flowers (**L)** and siliques (**M)** from wild type and *quadruple* mutants shown at the same magnification; SEM images of whole flowers of wild type (**A)** and the *quadruple* mutant (**B)** taken at the same magnification. Anthers from wild type flowers (**C)** are shown as a control; two fused stamens were observed in the *quadruple* mutant (**D)** (indicated by arrow). (**E)** Fusion of sepals between two flowers from the *quadruple* mutant (indicated by arrow). Wildtype (**F)** (**G)** and *quadruple* mutant (**I)** and (**J)** stigmas. Abnormal stomata patterning on carpel of the *quadruple* mutant (**K)** compared to wildtype (**H)**. Arrows point to two closely spaced stomata cells. Abnormal stomata patterning is also pictured in inset section of (**I)**.

Table S1. Summary of sequencing reads from EIN3 ChIP-Seq experiments.

	Ethylene treatment (hrs)						
Replicate	0	0.25	0.5	1	4	12	24
1	2937213		2129743	2921207	1552324	3021150	
2	3114917	4128762	2438606	3049491	2477851		3518255

Table S2. Summary of sequencing reads for mRNA-Seq experiments.

	Ethylene treatment (hrs)						
Replicate	0	0.25	0.5	1	4	12	24
1	22507787	32894861	37353263	22927060	27923583	27248282	30388612
2	30791072	26832664	25644083	21539330	34405113	29578538	33793078
3	26111636	33153195	27683922	28167560	25196268	40426314	28022905

Table S3. EIN3 targets.

EIN3-R (EIN3 targets, ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G01380	ENHANCER OF TRY AND CPC 1 (ETC1)
AT1G01650	SIGNAL PEPTIDE PEPTIDASE-LIKE 4 (SPPL4)
AT1G02205	Expression of the CER1 gene associated with production of stem epicuticular w
AT1G04220	3-KETOACYL-COA SYNTHASE 2 (KCS2)
AT1G04310	ETHYLENE RESPONSE SENSOR 2 (ERS2)
AT1G05010	ETHYLENE-FORMING ENZYME (EFE)
AT1G06080	DELTA 9 DESATURASE 1 (ADS1)
AT1G06180	MYB DOMAIN PROTEIN 13 (MYB13)
AT1G07150	MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 13 (MAPKKK13)
AT1G07440	NAD(P)-binding Rossmann-fold superfamily protein
AT1G08930	EARLY RESPONSE TO DEHYDRATION 6 (ERD6)
AT1G09530	PHYTOCHROME INTERACTING FACTOR 3 (PIF3)
AT1G09960	SUCROSE TRANSPORTER 4 (SUT4)
AT1G10060	BRANCHED-CHAIN AMINO ACID TRANSAMINASE 1 (BCAT-1)
AT1G10140	Uncharacterised conserved protein UCP031279
AT1G10480	ZINC FINGER PROTEIN 5 (ZFP5)
AT1G13930	Involved in response to salt stress. Knockout mutants are hypersensitive to sal
AT1G16370	ORGANIC CATION/CARNITINE TRANSPORTER 6 (OCT6)
AT1G17310	MADS-box transcription factor family protein
AT1G17345	SAUR-like auxin-responsive protein family
AT1G17810	BETA-TONOPLAST INTRINSIC PROTEIN (BETA-TIP)
AT1G18100	(E12A11)
AT1G18290	unknown protein
AT1G18570	MYB DOMAIN PROTEIN 51 (MYB51)
AT1G18740	FUNCTIONS IN: molecular_function unknown
AT1G19180	JASMONATE-ZIM-DOMAIN PROTEIN 1 (JAZ1)
AT1G19530	unknown protein
AT1G19770	PURINE PERMEASE 14 (PUP14)
AT1G19900	glyoxal oxidase-related protein
AT1G20900	ESCAROLA (ESC)
AT1G21050	Protein of unknown function, DUF617
AT1G21080	DNAJ heat shock N-terminal domain-containing protein
AT1G21130	O-methyltransferase family protein
AT1G21326	VQ motif-containing protein
AT1G21400	Thiamin diphosphate-binding fold (THDP-binding) superfamily protein
AT1G21910	encodes a member of the DREB subfamily A-5 of ERF/AP2 transcription factor
AT1G22990	Heavy metal transport/detoxification superfamily protein
AT1G23090	SULFATE TRANSPORTER 91 (AST91)
AT1G23730	BETA CARBONIC ANHYDRASE 4 (BCA3)
AT1G24530	Transducin/WD40 repeat-like superfamily protein
AT1G25560	TEMPRANILLO 1 (TEM1)
AT1G26270	Phosphatidylinositol 3- and 4-kinase family protein
AT1G26770	EXPANSIN A10 (EXPA10)
AT1G26930	Galactose oxidase/kelch repeat superfamily protein
AT1G27740	ROOT HAIR DEFECTIVE 6-LIKE 4 (RSL4)
AT1G28370	ERF DOMAIN PROTEIN 11 (ERF11)
AT1G28520	VASCULAR PLANT ONE ZINC FINGER PROTEIN (VOZ1)
AT1G29160	Dof-type zinc finger DNA-binding family protein

Table S3, con't. EIN3 targets.

EIN3-R (EIN3 targets, ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G30650	WRKY DNA-BINDING PROTEIN 14 (WRKY14)
AT1G31350	KAR-UP F-BOX 1 (KUF1)
AT1G32690	unknown protein
AT1G33260	Protein kinase superfamily protein
AT1G36060	encodes a member of the DREB subfamily A-6 of ERF/AP2 transcription factor
AT1G43160	RELATED TO AP2 6 (RAP2.6)
AT1G48690	Auxin-responsive GH3 family protein
AT1G48930	GLYCOSYL HYDROLASE 9C1 (GH9C1)
AT1G49200	RING/U-box superfamily protein
AT1G49660	CARBOXYESTERASE 5 (CXE5)
AT1G50040	Protein of unknown function (DUF1005)
AT1G50110	D-aminoacid aminotransferase-like PLP-dependent enzymes superfamily prote
AT1G50420	SCARECROW-LIKE 3 (SCL3)
AT1G51800	Leucine-rich repeat protein kinase family protein
AT1G52890	NAC DOMAIN CONTAINING PROTEIN 19 (NAC019)
AT1G53340	Cysteine/Histidine-rich C1 domain family protein
AT1G53690	Protein of unknown function that is homologous to At5g41010, which encodes a
AT1G53830	PECTIN METHYLESTERASE 2 (PME2)
AT1G55920	SERINE ACETYLTRANSFERASE 2
AT1G56220	Dormancy/auxin associated family protein
AT1G57560	MYB DOMAIN PROTEIN 50 (MYB50)
AT1G58170	Disease resistance-responsive (dirigent-like protein) family protein
AT1G59740	Major facilitator superfamily protein
AT1G59940	RESPONSE REGULATOR 3 (ARR3)
AT1G61370	S-locus lectin protein kinase family protein
AT1G62380	ACC OXIDASE 2 (ACO2)
AT1G62510	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily
AT1G63410	Protein of unknown function (DUF567)
AT1G64470	Ubiquitin-like superfamily protein
AT1G64940	CYTOCHROME P450, FAMILY 87, SUBFAMILY A, POLYPEPTIDE 6 (CYP89A
AT1G66180	Eukaryotic aspartyl protease family protein
AT1G67330	Protein of unknown function (DUF579)
AT1G67360	Rubber elongation factor protein (REF)
AT1G68840	REGULATOR OF THE ATPASE OF THE VACUOLAR MEMBRANE (RAV2)
AT1G69530	EXPANSIN A1 (EXPA1)
AT1G69870	NITRATE TRANSPORTER 1.7 (NRT1.7)
AT1G70230	TRICHOME BIREFRINGENCE-LIKE 27 (TBL27)
AT1G70710	GLYCOSYL HYDROLASE 9B1 (GH9B1)
AT1G71970	unknown protein
AT1G72450	JASMONATE-ZIM-DOMAIN PROTEIN 6 (JAZ6)
AT1G72510	Protein of unknown function (DUF1677)
AT1G73500	MAP KINASE KINASE 9 (MKK9)
AT1G73830	BR ENHANCED EXPRESSION 3 (BEE3)
AT1G74650	MYB DOMAIN PROTEIN 31 (MYB31)
AT1G75080	BRASSINAZOLE-RESISTANT 1 (BZR1)
AT1G75490	encodes a member of the DREB subfamily A-2 of ERF/AP2 transcription factor
AT1G75590	SAUR-like auxin-responsive protein family
AT1G76070	unknown protein

Table S3, con't. EIN3 targets.

EIN3-R (EIN3 targets, ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G76410	ATL8
AT1G76490	HYDROXY METHYLGLUTARYL COA REDUCTASE 1 (HMG1)
AT1G76500	SUPPRESSOR OF PHYB-4#3 (SOB3)
AT1G77145	Protein of unknown function (DUF506)
AT1G77330	similar to 1-aminocyclopropane-1-carboxylate oxidase GI:3386565 from (Sorgh
AT1G77640	encodes a member of the DREB subfamily A-5 of ERF/AP2 transcription factor
AT1G78480	Prenyltransferase family protein
AT1G79320	METACASPASE 6 (MC6)
AT1G79700	Integrase-type DNA-binding superfamily protein
AT1G79750	NADP-MALIC ENZYME 4 (NADP-ME4)
AT2G01530	MLP-LIKE PROTEIN 329 (MLP329)
AT2G01580	unknown protein
AT2G02680	Cysteine/Histidine-rich C1 domain family protein
AT2G04795	unknown protein
AT2G05510	Glycine-rich protein family
AT2G14170	ALDEHYDE DEHYDROGENASE 6B2 (ALDH6B2)
AT2G15760	Protein of unknown function (DUF1645)
AT2G15830	unknown protein
AT2G15960	unknown protein
AT2G16630	Pollen Ole e 1 allergen and extensin family protein
AT2G16740	UBIQUITIN-CONJUGATING ENZYME 29 (UBC29)
AT2G17880	Chaperone DnaJ-domain superfamily protein
AT2G18350	HOMEOBOX PROTEIN 24 (HB24)
AT2G18690	unknown protein
AT2G20670	Protein of unknown function (DUF506)
AT2G21185	unknown protein
AT2G22680	Zinc finger (C3HC4-type RING finger) family protein
AT2G22880	VQ motif-containing protein
AT2G22970	SERINE CARBOXYPEPTIDASE-LIKE 11 (SCPL11)
AT2G23180	CYTOCHROME P450, FAMILY 96, SUBFAMILY A, POLYPEPTIDE 1 (CYP96A
AT2G23290	MYB DOMAIN PROTEIN 70 (MYB70)
AT2G25490	EIN3-BINDING F BOX PROTEIN 1 (EBF1)
AT2G25940	ALPHA-VACUOLAR PROCESSING ENZYME (ALPHA-VPE)
AT2G26070	REVERSION-TO-ETHYLENE SENSITIVITY1 (RTE1)
AT2G26290	ROOT-SPECIFIC KINASE 1 (ARSK1)
AT2G26740	SOLUBLE EPOXIDE HYDROLASE (SEH)
AT2G27690	CYTOCHROME P450, FAMILY 94, SUBFAMILY C, POLYPEPTIDE 1 (CYP94C
AT2G27860	UDP-D-APIOSE/UDP-D-XYLOSE SYNTHASE 1 (AXS1)
AT2G28570	unknown protein
AT2G29470	GLUTATHIONE S-TRANSFERASE TAU 3 (GSTU3)
AT2G30150	UDP-Glycosyltransferase superfamily protein
AT2G31180	MYB DOMAIN PROTEIN 14 (MYB14)
AT2G31730	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT2G32150	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
AT2G32560	F-box family protein
AT2G34080	Cysteine proteinases superfamily protein
AT2G34650	PINOID (PID)
AT2G35260	unknown protein

Table S3, con't. EIN3 targets.

EIN3-R (EIN3 targets, ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT2G35930	PLANT U-BOX 23 (PUB23)
AT2G36080	Encodes a plant-specific B3 DNA-binding domain transcription factor. Has tran
AT2G36090	F-box family protein
AT2G36320	A20/AN1-like zinc finger family protein
AT2G36890	REGULATOR OF AXILLARY MERISTEMS 2 (RAX2)
AT2G37025	TRF-LIKE 8 (TRFL8)
AT2G39000	Acyl-CoA N-acyltransferases (NAT) superfamily protein
AT2G39180	ARABIDOPSIS THALIANA CRINKLY4 RELATED 2 (CCR2)
AT2G39210	Major facilitator superfamily protein
AT2G39400	alpha/beta-Hydrolases superfamily protein
AT2G39705	ROTUNDIFOLIA LIKE 8 (RTFL8)
AT2G39980	HXXXD-type acyl-transferase family protein
AT2G40540	POTASSIUM TRANSPORTER 2 (KT2)
AT2G40940	ETHYLENE RESPONSE SENSOR 1 (ERS1)
AT2G41300	STRICTOSIDINE SYNTHASE-LIKE 1 (SSL1)
AT2G42620	MORE AXILLARY BRANCHES 2 (MAX2)
AT2G43060	ILI1 BINDING BHLH 1 (IBH1)
AT2G43140	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT2G44070	NagB/RpiA/CoA transferase-like superfamily protein
AT2G44080	ARGOS-LIKE (ARL)
AT2G44570	GLYCOSYL HYDROLASE 9B12 (GH9B12)
AT2G46690	SAUR-like auxin-responsive protein family
AT2G46970	PHYTOCHROME INTERACTING FACTOR 3-LIKE 1 (PIL1)
AT2G47160	REQUIRES HIGH BORON 1 (BOR1)
AT2G47180	GALACTINOL SYNTHASE 1 (GolS1)
AT3G01290	SPFH/Band 7/PHB domain-containing membrane-associated protein family
AT3G01970	WRKY DNA-BINDING PROTEIN 45 (WRKY45)
AT3G03660	WUSCHEL RELATED HOMEOBOX 11 (WOX11)
AT3G05890	RARE-COLD-INDUCIBLE 2B (RCI2B)
AT3G07350	Protein of unknown function (DUF506)
AT3G07870	F-box and associated interaction domains-containing protein
AT3G09270	GLUTATHIONE S-TRANSFERASE TAU 8 (GSTU8)
AT3G10120	unknown protein
AT3G11410	ARABIDOPSIS THALIANA PROTEIN PHOSPHATASE 2CA (PP2CA)
AT3G11690	unknown protein
AT3G12920	SBP (S-ribonuclease binding protein) family protein
AT3G13520	ARABINOGALACTAN PROTEIN 12 (AGP12)
AT3G14230	RELATED TO AP2 2 (RAP2.2)
AT3G15450	Aluminium induced protein with YGL and LRDR motifs
AT3G15500	NAC DOMAIN CONTAINING PROTEIN 3 (NAC3)
AT3G16400	NITRILE SPECIFIER PROTEIN 1 (NSP1)
AT3G16770	ETHYLENE-RESPONSIVE ELEMENT BINDING PROTEIN (EBP)
AT3G16857	RESPONSE REGULATOR 1 (RR1)
AT3G18715	INFLORESCENCE DEFICIENT IN ABSCISSION (IDA)-LIKE 4 (IDL4)
AT3G19200	unknown protein
AT3G20300	Protein of unknown function (DUF3537)
AT3G20395	RING/U-box superfamily protein
AT3G20830	AGC (cAMP-dependent, cGMP-dependent and protein kinase C) kinase family

Table S3, con't. EIN3 targets.

EIN3-R (EIN3 targets, ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT3G20900	unknown protein
AT3G21510	HISTIDINE-CONTAINING PHOSPHOTRANSMITTER 1 (AHP1)
AT3G22120	CELL WALL-PLASMA MEMBRANE LINKER PROTEIN (CWLP)
AT3G22370	ALTERNATIVE OXIDASE 1A (AOX1A)
AT3G22750	Protein kinase superfamily protein
AT3G22800	Leucine-rich repeat (LRR) family protein
AT3G23150	ETHYLENE RESPONSE 2 (ETR2)
AT3G23170	unknown protein
AT3G23240	ETHYLENE RESPONSE FACTOR 1 (ERF1)
AT3G23810	S-ADENOSYL-L-HOMOCYSTEINE (SAH) HYDROLASE 2 (SAHH2)
AT3G23880	F-box and associated interaction domains-containing protein
AT3G24120	Homeodomain-like superfamily protein
AT3G24450	Heavy metal transport/detoxification superfamily protein
AT3G25730	ETHYLENE RESPONSE DNA BINDING FACTOR 3 (EDF3)
AT3G28340	GALACTURONOSYLTRANSFERASE-LIKE 10 (GATL10)
AT3G29035	NAC DOMAIN CONTAINING PROTEIN 3 (NAC3)
AT3G30775	EARLY RESPONSIVE TO DEHYDRATION 5 (ERD5)
AT3G44260	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
AT3G45280	SYNTAXIN OF PLANTS 72 (SYP72)
AT3G47510	unknown protein
AT3G48990	AMP-dependent synthetase and ligase family protein
AT3G49360	6-PHOSPHOGLUCONOLACTONASE 2 (PGL2)
AT3G50310	MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 20 (MAPKKK20)
AT3G50330	HECATE 2 (HEC2)
AT3G52480	unknown protein
AT3G53250	SAUR-like auxin-responsive protein family
AT3G54040	PAR1 protein
AT3G57550	GUANYLATE KINASE (AGK2)
AT3G58060	Cation efflux family protein
AT3G59900	AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE (ARGOS)
AT3G60200	unknown protein
AT3G60390	HOMEOBOX-LEUCINE ZIPPER PROTEIN 3 (HAT3)
AT3G60520	unknown protein
AT3G60530	GATA TRANSCRIPTION FACTOR 4 (GATA4)
AT3G60550	CYCLIN P3
AT3G60690	SAUR-like auxin-responsive protein family
AT3G61440	CYSTEINE SYNTHASE C1 (CYSC1)
AT3G63010	GA INSENSITIVE DWARF1B (GID1B)
AT4G00730	ANTHOCYANINLESS 2 (ANL2)
AT4G01335	FUNCTIONS IN: molecular_function unknown
AT4G01720	(WRKY47)
AT4G01870	tolB protein-related
AT4G02715	unknown protein
AT4G03140	NAD(P)-binding Rossmann-fold superfamily protein
AT4G03510	RING MEMBRANE-ANCHOR 1 (RMA1)
AT4G03540	Uncharacterised protein family (UPF0497)
AT4G05150	Octicosapeptide/Phox/Bem1p family protein
AT4G06744	Leucine-rich repeat (LRR) family protein

Table S3, con't. EIN3 targets.

EIN3-R (EIN3 targets, ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT4G08620	SULPHATE TRANSPORTER 1
AT4G10380	NODULIN26-LIKE INTRINSIC PROTEIN 5
AT4G11330	MAP KINASE 5 (MPK5)
AT4G11370	RING-H2 FINGER A1A (RHA1A)
AT4G12410	SAUR-like auxin-responsive protein family
AT4G12470	AZELAIC ACID INDUCED 1 (AZI1)
AT4G13395	ROTUNDIFOLIA LIKE 12 (RTFL12)
AT4G13660	PINORESINOL REDUCTASE 2 (PRR2)
AT4G14030	SELENIUM-BINDING PROTEIN 1 (SBP1)
AT4G14130	XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE 15 (XTH15)
AT4G14930	Survival protein SurE-like phosphatase/nucleotidase
AT4G15545	unknown protein
AT4G15550	INDOLE-3-ACETATE BETA-D-GLUCOSYLTRANSFERASE (IAGLU)
AT4G16442	Uncharacterised protein family (UPF0497)
AT4G17500	ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 1 (ERF-1)
AT4G19680	IRON REGULATED TRANSPORTER 2 (IRT2)
AT4G21410	CYSTEINE-RICH RLK (RECEPTOR-LIKE PROTEIN KINASE) 29 (CRK29)
AT4G21910	MATE efflux family protein
AT4G22470	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
AT4G22840	Sodium Bile acid symporter family
AT4G23510	Disease resistance protein (TIR-NBS-LRR class) family
AT4G26080	ABA INSENSITIVE 1 (ABI1)
AT4G27310	B-box type zinc finger family protein
AT4G27450	Aluminium induced protein with YGL and LRDR motifs
AT4G28040	nodulin MtN21 /EamA-like transporter family protein
AT4G28350	Concanavalin A-like lectin protein kinase family protein
AT4G30420	nodulin MtN21 /EamA-like transporter family protein
AT4G32285	ENTH/ANTH/VHS superfamily protein
AT4G32480	Protein of unknown function (DUF506)
AT4G34380	Transducin/WD40 repeat-like superfamily protein
AT4G34419	unknown protein
AT4G34710	ARGININE DECARBOXYLASE 2 (ADC2)
AT4G35060	Heavy metal transport/detoxification superfamily protein
AT4G35160	O-methyltransferase family protein
AT4G35190	Putative lysine decarboxylase family protein
AT4G35750	SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein
AT4G36780	BES1/BZR1 HOMOLOG 2 (BEH2)
AT4G37240	unknown protein
AT4G37260	MYB DOMAIN PROTEIN 73 (MYB73)
AT4G37410	CYTOCHROME P450, FAMILY 81, SUBFAMILY F, POLYPEPTIDE 4 (CYP81F
AT4G37580	HOOKLESS 1 (HLS1)
AT4G37890	EMBRYO SAC DEVELOPMENT ARREST 40 (EDA40)
AT4G38470	ACT-like protein tyrosine kinase family protein
AT5G01180	PEPTIDE TRANSPORTER 5 (PTR5)
AT5G01210	HXXXD-type acyl-transferase family protein
AT5G02230	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
AT5G02480	HSP20-like chaperones superfamily protein
AT5G02760	Protein phosphatase 2C family protein

Table S3, con't. EIN3 targets.

EIN3-R (EIN3 targets, ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT5G03280	ETHYLENE INSENSITIVE 2 (EIN2)
AT5G03310	SAUR-like auxin-responsive protein family
AT5G03700	D-mannose binding lectin protein with Apple-like carbohydrate-binding domain
AT5G03730	CONSTITUTIVE TRIPLE RESPONSE 1 (CTR1)
AT5G03890	unknown protein
AT5G04230	Member of Phenylalanine ammonialyase (PAL) gene family. Differs significantly
AT5G05440	PYRABACTIN RESISTANCE 1-LIKE 5 (PYL5)
AT5G05880	UDP-Glycosyltransferase superfamily protein
AT5G06570	alpha/beta-Hydrolases superfamily protein
AT5G06990	Protein of unknown function, DUF617
AT5G07010	SULFOTRANSFERASE 2A (ST2A)
AT5G07100	WRKY DNA-BINDING PROTEIN 26 (WRKY26)
AT5G08260	SERINE CARBOXYPEPTIDASE-LIKE 35 (scpl35)
AT5G12050	unknown protein
AT5G12170	CRT (CHLOROQUINE-RESISTANCE TRANSPORTER)-LIKE TRANSPORTEF
AT5G12330	A member of SHI gene family. Arabidopsis thaliana has ten members that enco
AT5G13080	WRKY DNA-BINDING PROTEIN 75 (WRKY75)
AT5G13330	RELATED TO AP2 6L (Rap2.6L)
AT5G13500	unknown protein
AT5G13910	LEAFY PETIOLE (LEP)
AT5G14780	FORMATE DEHYDROGENASE (FDH)
AT5G14920	Gibberellin-regulated family protein
AT5G17700	MATE efflux family protein
AT5G17780	alpha/beta-Hydrolases superfamily protein
AT5G17810	WUSCHEL RELATED HOMEOBOX 12 (WOX12)
AT5G18260	RING/U-box superfamily protein
AT5G18470	Curculin-like (mannose-binding) lectin family protein
AT5G18670	BETA-AMYLASE 3 (BMY3)
AT5G19040	ISOPENTENYLTRANSFERASE 5 (IPT5)
AT5G19120	Eukaryotic aspartyl protease family protein
AT5G20710	BETA-GALACTOSIDASE 7 (BGAL7)
AT5G20820	SAUR-like auxin-responsive protein family
AT5G21090	Leucine-rich repeat (LRR) family protein
AT5G21120	ETHYLENE-INSENSITIVE3-LIKE 2 (EIL2)
AT5G22270	unknown protein
AT5G22460	alpha/beta-Hydrolases superfamily protein
AT5G22890	C2H2 and C2HC zinc fingers superfamily protein
AT5G22940	FRA8 HOMOLOG (F8H)
AT5G24030	SLAC1 HOMOLOGUE 3 (SLAH3)
AT5G24570	unknown protein
AT5G25190	encodes a member of the ERF (ethylene response factor) subfamily B-6 of ERF
AT5G25350	EIN3-BINDING F BOX PROTEIN 2 (EBF2)
AT5G26220	ChaC-like family protein
AT5G27320	GA INSENSITIVE DWARF1C (GID1C)
AT5G38700	unknown protein
AT5G39610	NAC DOMAIN CONTAINING PROTEIN 6 (NAC6)
AT5G40590	Cysteine/Histidine-rich C1 domain family protein
AT5G42530	unknown protein

Table S3, con't. EIN3 targets.

EIN3-R (EIN3 targets, ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT5G44260	Zinc finger C-x8-C-x5-C-x3-H type family protein
AT5G44440	FAD-binding Berberine family protein
AT5G44580	unknown protein
AT5G46330	FLAGELLIN-SENSITIVE 2 (FLS2)
AT5G46710	PLATZ transcription factor family protein
AT5G47220	ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 2 (ERF2)
AT5G47230	ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 5 (ERF5)
AT5G48000	CYTOCHROME P450, FAMILY 708, SUBFAMILY A, POLYPEPTIDE 2 (CYP70
AT5G48175	FUNCTIONS IN: molecular_function unknown
AT5G48800	Phototropic-responsive NPH3 family protein
AT5G48890	C2H2-like zinc finger protein
AT5G49665	Zinc finger (C3HC4-type RING finger) family protein
AT5G49690	UDP-Glycosyltransferase superfamily protein
AT5G50335	unknown protein
AT5G51780	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT5G53830	VQ motif-containing protein
AT5G53980	HOMEOBOX PROTEIN 52 (HB52)
AT5G54490	PINOID-BINDING PROTEIN 1 (PBP1)
AT5G55620	unknown protein
AT5G57123	unknown protein
AT5G57400	unknown protein
AT5G57710	Double Clp-N motif-containing P-loop nucleoside triphosphate hydrolases supe
AT5G57760	unknown protein
AT5G57920	EARLY NODULIN-LIKE PROTEIN 10 (ENODL10)
AT5G59040	COPPER TRANSPORTER 3 (COPT3)
AT5G59220	HIGHLY ABA-INDUCED PP2C GENE 1 (HAI1)
AT5G59540	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT5G60520	Late embryogenesis abundant (LEA) protein-related
AT5G61290	Flavin-binding monooxygenase family protein
AT5G62280	Protein of unknown function (DUF1442)
AT5G63650	SNF1-RELATED PROTEIN KINASE 2.5 (SNRK2.5)
AT5G64410	OLIGOPEPTIDE TRANSPORTER 4 (OPT4)
AT5G64440	FATTY ACID AMIDE HYDROLASE (FAAH)
AT5G66170	SULFURTRANSFERASE 18 (STR18)
AT5G66440	unknown protein
AT5G66460	ENDO-BETA-MANNASE 7 (MAN7)
AT5G67020	unknown protein
AT5G67060	HECATE 1 (HEC1)
AT5G67080	MITOGEN-ACTIVATED PROTEIN KINASE KINASE KINASE 19 (MAPKKK19)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G01040	Encodes a Dicer homolog. Dicer is a RNA helicase involved in microRNA proce
AT1G01490	Heavy metal transport/detoxification superfamily protein
AT1G01800	NAD(P)-binding Rossmann-fold superfamily protein
AT1G02070	unknown protein
AT1G02400	GIBBERELLIN 2-OXIDASE 6 (GA2OX6)
AT1G02810	Plant invertase/pectin methylesterase inhibitor superfamily
AT1G02860	NITROGEN LIMITATION ADAPTATION (NLA)
AT1G03730	unknown protein
AT1G03780	Homolog of vertebrate TPX2. Protein has three domains involved in nuclear tar
AT1G04171	unknown protein
AT1G04690	POTASSIUM CHANNEL BETA SUBUNIT (KAB1)
AT1G04770	Tetratricopeptide repeat (TPR)-like superfamily protein
AT1G05140	Peptidase M50 family protein
AT1G06840	Leucine-rich repeat protein kinase family protein
AT1G06990	GDSL-like Lipase/Acylhydrolase superfamily protein
AT1G08300	NO VEIN-LIKE (NVL)
AT1G08500	EARLY NODULIN-LIKE PROTEIN 18 (ENODL18)
AT1G08510	FATTY ACYL-ACP THIOESTERASES B (FATB)
AT1G08920	ERD (EARLY RESPONSE TO DEHYDRATION) SIX-LIKE 1 (ESL1)
AT1G09080	(BIP3)
AT1G09250	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT1G09560	GERMIN-LIKE PROTEIN 5 (GLP5)
AT1G09570	PHYTOCHROME A (PHYA)
AT1G09970	(LRR XI-23)
AT1G10310	encodes a NADPH-dependent pterin aldehyde reductase that accepts pterin al
AT1G10460	GERMIN-LIKE PROTEIN 7 (GLP7)
AT1G10620	Protein kinase superfamily protein
AT1G10690	unknown protein
AT1G10740	alpha/beta-Hydrolases superfamily protein
AT1G11180	Secretory carrier membrane protein (SCAMP) family protein
AT1G11600	CYTOCHROME P450, FAMILY 77, SUBFAMILY B, POLYPEPTIDE 1 (CYP77B
AT1G11670	MATE efflux family protein
AT1G12740	encodes a protein with cytochrome P450 domain
AT1G13260	RELATED TO ABI3/VP1 1 (RAV1)
AT1G13370	Histone superfamily protein
AT1G13590	PHYTOSULFOKINE 1 PRECURSOR (PSK1)
AT1G13960	WRKY DNA-BINDING PROTEIN 4 (WRKY4)
AT1G14185	Glucose-methanol-choline (GMC) oxidoreductase family protein
AT1G14270	CAAX amino terminal protease family protein
AT1G14600	Homeodomain-like superfamily protein
AT1G14720	XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE 28 (XTH28)
AT1G14860	NUDIX HYDROLASE HOMOLOG 18 (NUDT18)
AT1G14910	ENTH/ANTH/VHS superfamily protein
AT1G15390	PEPTIDE DEFORMYLASE 1A (PDF1A)
AT1G15670	Galactose oxidase/kelch repeat superfamily protein
AT1G17230	Leucine-rich receptor-like protein kinase family protein
AT1G17370	OLIGOURIDYLATE BINDING PROTEIN 1B (UBP1B)
AT1G18400	BR ENHANCED EXPRESSION 1 (BEE1)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G18710	MYB DOMAIN PROTEIN 47 (MYB47)
AT1G18720	Protein of unknown function (DUF962)
AT1G19350	BRI1-EMS-SUPPRESSOR 1 (BES1)
AT1G19440	3-KETOACYL-COA SYNTHASE 4 (KCS4)
AT1G20640	Plant regulator RWP-RK family protein
AT1G20980	SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 14 (SPL14)
AT1G21000	PLATZ transcription factor family protein
AT1G21900	emp24/gp25L/p24 family/GOLD family protein
AT1G21920	Histone H3 K4-specific methyltransferase SET7/9 family protein
AT1G21940	unknown protein
AT1G21975	unknown protein
AT1G22180	Sec14p-like phosphatidylinositol transfer family protein
AT1G22767	unknown protein
AT1G23020	Encodes a ferric chelate reductase whose transcription is regulated by FIT1. E
AT1G23060	BEST Arabidopsis thaliana protein match is: TPX2 (targeting protein for Xklp2)
AT1G23080	PIN-FORMED 7 (PIN7)
AT1G24625	ZINC FINGER PROTEIN 7 (ZFP7)
AT1G25540	PHYTOCHROME AND FLOWERING TIME 1 (PFT1)
AT1G25550	myb-like transcription factor family protein
AT1G26290	unknown protein
AT1G26440	UREIDE PERMEASE 5 (UPS5)
AT1G26780	MYB DOMAIN PROTEIN 117 (MYB117)
AT1G27200	CONTAINS InterPro DOMAIN/s: Protein of unknown function DUF23 (InterPro:
AT1G27213	unknown protein
AT1G27470	transducin family protein / WD-40 repeat family protein
AT1G27980	DIHYDROSPHINGOSINE PHOSPHATE LYASE (DPL1)
AT1G28240	Protein of unknown function (DUF616)
AT1G28450	AGAMOUS-LIKE 58 (AGL58)
AT1G28580	GDSL-like Lipase/Acylhydrolase superfamily protein
AT1G29120	Hydrolase-like protein family
AT1G29400	MEI2-LIKE PROTEIN 5 (ML5)
AT1G29600	Zinc finger C-x8-C-x5-C-x3-H type family protein
AT1G29670	GDSL-like Lipase/Acylhydrolase superfamily protein
AT1G29760	Putative adipose-regulatory protein (Seipin)
AT1G30620	MURUS 4 (MUR4)
AT1G30960	GTP-binding family protein
AT1G31040	PLATZ transcription factor family protein
AT1G31310	hydroxyproline-rich glycoprotein family protein
AT1G31360	ARABIDOPSIS RECQ HELICASE L2 (RECQL2)
AT1G31810	FORMIN HOMOLOGY 14 (AFH14)
AT1G31880	BREVIS RADIX (BRX)
AT1G32200	(ATS1)
AT1G32530	RING/U-box superfamily protein
AT1G32700	PLATZ transcription factor family protein
AT1G32763	Encodes a defensin-like (DEFL) family protein.
AT1G32780	GroES-like zinc-binding dehydrogenase family protein
AT1G32930	Galactosyltransferase family protein
AT1G33020	F-box and associated interaction domains-containing protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G33390	FASCIATED STEM 4 (FAS4)
AT1G33790	jacalin lectin family protein
AT1G34041	unknown protein
AT1G34245	EPIDERMAL PATTERNING FACTOR 2 (EPF2)
AT1G34320	Protein of unknown function (DUF668)
AT1G34370	SENSITIVE TO PROTON RHIZOTOXICITY 1 (STOP1)
AT1G34400	unknown protein
AT1G35710	Protein kinase family protein with leucine-rich repeat domain
AT1G45249	Leucine zipper transcription factor that binds to the abscisic acid (ABA)–respon
AT1G45332	Translation elongation factor EFG/EF2 protein
AT1G47128	RESPONSIVE TO DEHYDRATION 21 (RD21)
AT1G47655	Dof-type zinc finger DNA-binding family protein
AT1G47870	(ATE2F2)
AT1G48000	MYB DOMAIN PROTEIN 112 (MYB112)
AT1G48320	Thioesterase superfamily protein
AT1G49435	LOW-MOLECULAR-WEIGHT CYSTEINE-RICH 16 (LCR16)
AT1G49500	unknown protein
AT1G49780	PLANT U-BOX 26 (PUB26)
AT1G50730	unknown protein
AT1G51090	Heavy metal transport/detoxification superfamily protein
AT1G51840	protein kinase-related
AT1G51940	protein kinase family protein / peptidoglycan-binding LysM domain-containing p
AT1G52565	unknown protein
AT1G53170	ETHYLENE RESPONSE FACTOR 8 (ERF8)
AT1G53180	unknown protein
AT1G53300	TETRATRICOPETIDE-REPEAT THIOREDOXIN-LIKE 1 (TTL1)
AT1G53310	PHOSPHOENOLPYRUVATE CARBOXYLASE 1 (PPC1)
AT1G53920	GDSL-MOTIF LIPASE 5 (GLIP5)
AT1G54130	RELA/SPOT HOMOLOG 3 (RSH3)
AT1G55020	LIPOXYGENASE 1 (LOX1)
AT1G55580	LATERAL SUPPRESSOR (LAS)
AT1G55880	Pyridoxal-5'-phosphate-dependent enzyme family protein
AT1G56145	Leucine-rich repeat transmembrane protein kinase
AT1G56200	EMBRYO DEFECTIVE 1303 (emb1303)
AT1G58350	(ZW18)
AT1G58400	Disease resistance protein (CC-NBS-LRR class) family
AT1G59820	AMINOPHOSPHOLIPID ATPASE3 (ALA3)
AT1G60010	unknown protein
AT1G60140	TREHALOSE PHOSPHATE SYNTHASE (TPS10)
AT1G60900	U2 snRNP auxilliary factor, large subunit, splicing factor
AT1G61460	S-locus protein kinase, putative
AT1G61870	PENTATRICOPEPTIDE REPEAT 336 (PPR336)
AT1G62181	unknown protein
AT1G62240	unknown protein
AT1G62370	RING/U-box superfamily protein
AT1G62420	Protein of unknown function (DUF506)
AT1G63500	Protein kinase protein with tetratricopeptide repeat domain
AT1G63810	CONTAINS InterPro DOMAIN/s: Nrap protein (InterPro:IPR005554)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G63840	RING/U-box superfamily protein
AT1G64080	unknown protein
AT1G65470	FASCIATA 1 (FAS1)
AT1G65920	Regulator of chromosome condensation (RCC1) family with FYVE zinc finger d
AT1G65930	CYTOSOLIC NADP+-DEPENDENT ISOCITRATE DEHYDROGENASE (cICDH
AT1G66080	unknown protein
AT1G66210	Subtilisin-like serine endopeptidase family protein
AT1G66270	(BGLU21)
AT1G66340	ETHYLENE RESPONSE 1 (ETR1)
AT1G66390	MYB DOMAIN PROTEIN 90 (MYB90)
AT1G66475	unknown protein
AT1G66790	unknown protein
AT1G67030	ZINC FINGER PROTEIN 6 (ZFP6)
AT1G67050	unknown protein
AT1G67090	RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN 1A (RBCS1A)
AT1G67620	Lojap-related protein
AT1G68050	FLAVIN-BINDING, KELCH REPEAT, F BOX 1 (FKF1)
AT1G68130	INDETERMINATE(ID)-DOMAIN 14 (IDD14)
AT1G68320	MYB DOMAIN PROTEIN 62 (MYB62)
AT1G68340	Protein of unknown function (DUF1639)
AT1G68550	encodes a member of the ERF (ethylene response factor) subfamily B-6 of ERF
AT1G68630	PLAC8 family protein
AT1G68670	myb-like transcription factor family protein
AT1G69270	RECEPTOR-LIKE PROTEIN KINASE 1 (RPK1)
AT1G69310	WRKY DNA-BINDING PROTEIN 57 (WRKY57)
AT1G69430	unknown protein
AT1G69500	CYTOCHROME P450, FAMILY 704, SUBFAMILY B, POLYPEPTIDE 1 (CYP70
AT1G69810	WRKY DNA-BINDING PROTEIN 36 (WRKY36)
AT1G69960	SERINE/THREONINE PROTEIN PHOSPHATASE 2A (PP2A)
AT1G70100	unknown protein
AT1G70420	Protein of unknown function (DUF1645)
AT1G70590	F-box family protein
AT1G70720	Plant invertase/pectin methylesterase inhibitor superfamily protein
AT1G70750	Protein of unknown function, DUF593
AT1G71080	RNA polymerase II transcription elongation factor
AT1G71692	AGAMOUS-LIKE 12 (AGL12)
AT1G71980	Protease-associated (PA) RING/U-box zinc finger family protein
AT1G72150	PATELLIN 1 (PATL1)
AT1G72416	Chaperone DnaJ-domain superfamily protein
AT1G72750	TRANSLOCASE INNER MEMBRANE SUBUNIT 23-2 (TIM23-2)
AT1G72880	Survival protein SurE-like phosphatase/nucleotidase
AT1G73390	Endosomal targeting BRO1-like domain-containing protein
AT1G73880	UDP-GLUCOSYL TRANSFERASE 89B1 (UGT89B1)
AT1G74080	MYB DOMAIN PROTEIN 122 (MYB122)
AT1G74430	MYB DOMAIN PROTEIN 95 (MYB95)
AT1G74640	alpha/beta-Hydrolases superfamily protein
AT1G75180	Erythronate-4-phosphate dehydrogenase family protein
AT1G75380	BIFUNCTIONAL NUCLEASE IN BASAL DEFENSE RESPONSE 1 (BBD1)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G75388	CONSERVED PEPTIDE UPSTREAM OPEN READING FRAME 5 (CPuORF5)
AT1G75440	UBIQUITIN-CONJUGATING ENZYME 16 (UBC16)
AT1G75540	SALT TOLERANCE HOMOLOG2 (STH2)
AT1G75710	C2H2-like zinc finger protein
AT1G75750	GAST1 PROTEIN HOMOLOG 1 (GASA1)
AT1G76160	SKU5 SIMILAR 5 (sks5)
AT1G76180	EARLY RESPONSE TO DEHYDRATION 14 (ERD14)
AT1G76380	DNA-binding bromodomain-containing protein
AT1G76405	unknown protein
AT1G76870	BEST Arabidopsis thaliana protein match is: sequence-specific DNA binding tra
AT1G77660	Histone H3 K4-specific methyltransferase SET7/9 family protein
AT1G77670	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein
AT1G77740	PHOSPHATIDYLINOSITOL-4-PHOSPHATE 5-KINASE 2 (PIP5K2)
AT1G77890	DNA-directed RNA polymerase II protein
AT1G78000	SULFATE TRANSPORTER 1
AT1G78080	RELATED TO AP2 4 (RAP2.4)
AT1G78100	F-box family protein
AT1G78340	GLUTATHIONE S-TRANSFERASE TAU 22 (GSTU22)
AT1G78600	light-regulated zinc finger protein 1 (LZF1)
AT1G79070	SNARE-associated protein-related
AT1G79110	zinc ion binding
AT1G79120	Ubiquitin carboxyl-terminal hydrolase family protein
AT1G79160	unknown protein
AT1G80500	SNARE-like superfamily protein
AT2G01023	unknown protein
AT2G01050	zinc ion binding
AT2G01140	Aldolase superfamily protein
AT2G01410	NHL domain-containing protein
AT2G01505	CLAVATA3/ESR-RELATED 16 (CLE16)
AT2G01670	NUDIX HYDROLASE HOMOLOG 17 (NUDT17)
AT2G01830	WOODEN LEG (WOL)
AT2G02100	LOW-MOLECULAR-WEIGHT CYSTEINE-RICH 69 (LCR69)
AT2G02220	PHYTOSULFOKIN RECEPTOR 1 (PSKR1)
AT2G02770	4'-phosphopantetheinyl transferase superfamily
AT2G03750	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT2G03890	PHOSPHOINOSITIDE 4-KINASE GAMMA 7 (PI4K GAMMA 7)
AT2G04041	Protein of unknown function (DUF784)
AT2G04100	MATE efflux family protein
AT2G05580	Glycine-rich protein family
AT2G06541	TTF-type zinc finger protein with HAT dimerisation domain
AT2G07050	CYCLOARTENOL SYNTHASE 1 (CAS1)
AT2G07180	Protein kinase superfamily protein
AT2G07640	NAD(P)-binding Rossmann-fold superfamily protein
AT2G07673	unknown protein
AT2G07687	Cytochrome c oxidase, subunit III
AT2G07692	unknown protein
AT2G07705	unknown protein
AT2G07721	unknown protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT2G07732	Ribulose bisphosphate carboxylase large chain, catalytic domain
AT2G07734	Alpha-L RNA-binding motif/Ribosomal protein S4 family protein
AT2G07738	unknown protein
AT2G07739	Ycf1 protein
AT2G07775	unknown protein
AT2G07798	unknown protein
AT2G07820	unknown protein
AT2G07835	unknown protein
AT2G11773	unknown protein
AT2G13950	Cysteine/Histidine-rich C1 domain family protein
AT2G14260	PROLINE IMINOPEPTIDASE (PIP)
AT2G14390	unknown protein
AT2G15390	FUCOSYLTRANSFERASE 4 (FUT4)
AT2G16430	PURPLE ACID PHOSPHATASE 10 (PAP10)
AT2G16720	MYB DOMAIN PROTEIN 7 (MYB7)
AT2G18160	BASIC LEUCINE-ZIPPER 2 (bZIP2)
AT2G18170	MAP KINASE 7 (MPK7)
AT2G18876	Afadin/alpha-actinin-binding protein
AT2G18960	H(+)-ATPASE 1 (HA1)
AT2G19190	FLG22-INDUCED RECEPTOR-LIKE KINASE 1 (FRK1)
AT2G19560	ENHANCED ETHYLENE RESPONSE 5 (EER5)
AT2G19580	TETRASPANIN2 (TET2)
AT2G19820	LOB DOMAIN-CONTAINING PROTEIN 9 (LBD9)
AT2G19880	Nucleotide-diphospho-sugar transferases superfamily protein
AT2G20210	RNI-like superfamily protein
AT2G20900	DIACYLGLYCEROL KINASE 5 (DGK5)
AT2G21110	Disease resistance-responsive (dirigent-like protein) family protein
AT2G21300	ATP binding microtubule motor family protein
AT2G21520	Sec14p-like phosphatidylinositol transfer family protein
AT2G22030	Galactose oxidase/kelch repeat superfamily protein
AT2G22201	unknown protein
AT2G22470	ARABINOGALACTAN PROTEIN 2 (AGP2)
AT2G22650	FAD-dependent oxidoreductase family protein
AT2G22760	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT2G22800	(HAT9)
AT2G22860	PHYTOSULFOKINE 2 PRECURSOR (PSK2)
AT2G23220	CYTOCHROME P450, FAMILY 81, SUBFAMILY D, POLYPEPTIDE 6 (CYP81D
AT2G23300	Leucine-rich repeat protein kinase family protein
AT2G23310	(ATRER1C1)
AT2G23450	Protein kinase superfamily protein
AT2G23680	Cold acclimation protein WCOR413 family
AT2G23800	GERANYLGERANYL PYROPHOSPHATE SYNTHASE 2 (GGPS2)
AT2G24000	SERINE CARBOXYPEPTIDASE-LIKE 22 (scpl22)
AT2G24765	ADP-RIBOSYLATION FACTOR 3 (ARF3)
AT2G25190	PPPDE putative thiol peptidase family protein
AT2G25310	FUNCTIONS IN: carbohydrate binding
AT2G25737	Sulfite exporter TauE/SafE family protein
AT2G26660	SPX DOMAIN GENE 2 (SPX2)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT2G27060	Leucine-rich repeat protein kinase family protein
AT2G27310	F-box family protein
AT2G27830	unknown protein
AT2G28050	Pentatricopeptide repeat (PPR) superfamily protein
AT2G28490	RmlC-like cupins superfamily protein
AT2G29080	FTSH PROTEASE 3 (ftsh3)
AT2G29440	GLUTATHIONE S-TRANSFERASE TAU 6 (GSTU6)
AT2G30020	Encodes AP2C1. Belongs to the clade B of the PP2C-superfamily. Acts as a M
AT2G30060	Pleckstrin homology (PH) domain superfamily protein
AT2G30230	unknown protein
AT2G30660	ATP-dependent caseinolytic (Clp) protease/crotonase family protein
AT2G30700	unknown protein
AT2G31160	LIGHT SENSITIVE HYPOCOTYLS 3 (LSH3)
AT2G31240	Tetratricopeptide repeat (TPR)-like superfamily protein
AT2G31750	UDP-GLUCOSYL TRANSFERASE 74D1 (UGT74D1)
AT2G32240	FUNCTIONS IN: molecular_function unknown
AT2G32415	Polynucleotidyl transferase, ribonuclease H fold protein with HRDC domain
AT2G33170	Leucine-rich repeat receptor-like protein kinase family protein
AT2G33205	Serinc-domain containing serine and sphingolipid biosynthesis protein
AT2G33700	Protein phosphatase 2C family protein
AT2G33710	encodes a member of the ERF (ethylene response factor) subfamily B-4 of ERF
AT2G33810	SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 3 (SPL3)
AT2G34480	Ribosomal protein L18ae/LX family protein
AT2G34610	unknown protein
AT2G34730	myosin heavy chain-related
AT2G36350	Protein kinase superfamily protein
AT2G36690	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT2G36950	Heavy metal transport/detoxification superfamily protein
AT2G37170	PLASMA MEMBRANE INTRINSIC PROTEIN 2 (PIP2B)
AT2G37220	Encodes a chloroplast RNA binding protein. A substrate of the type III effector I
AT2G37760	NAD(P)-linked oxidoreductase superfamily protein
AT2G38120	AUXIN RESISTANT 1 (AUX1)
AT2G38310	PYR1-LIKE 4 (PYL4)
AT2G39370	unknown protein
AT2G39480	P-GLYCOPROTEIN 6 (PGP6)
AT2G39500	unknown protein
AT2G39990	EUKARYOTIC TRANSLATION INITIATION FACTOR 2 (EIF2)
AT2G40004	unknown protein
AT2G40230	HXXXD-type acyl-transferase family protein
AT2G40260	Homeodomain-like superfamily protein
AT2G40340	(DREB2C)
AT2G40830	RING-H2 FINGER C1A (RHC1A)
AT2G41220	GLUTAMATE SYNTHASE 2 (GLU2)
AT2G41233	unknown protein
AT2G41380	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
AT2G41540	(GPDHC1)
AT2G41640	Glycosyltransferase family 61 protein
AT2G41690	HEAT SHOCK TRANSCRIPTION FACTOR B3 (HSFB3)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT2G41870	Remorin family protein
AT2G41940	ZINC FINGER PROTEIN 8 (ZFP8)
AT2G42280	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT2G42500	PROTEIN PHOSPHATASE 2A-3 (PP2A-3)
AT2G42520	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT2G42610	LIGHT SENSITIVE HYPOCOTYLS 10 (LSH10)
AT2G43240	Nucleotide-sugar transporter family protein
AT2G43320	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
AT2G43590	Chitinase family protein
AT2G44500	O-fucosyltransferase family protein
AT2G44660	ALG6, ALG8 glycosyltransferase family
AT2G44670	Protein of unknown function (DUF581)
AT2G45600	alpha/beta-Hydrolases superfamily protein
AT2G46200	unknown protein
AT2G46420	Plant protein 1589 of unknown function
AT2G46580	Pyridoxamine 5'-phosphate oxidase family protein
AT2G46680	HOMEOBOX 7 (HB-7)
AT2G46870	NGATHA1 (NGA1)
AT2G46920	POLTERGEIST (POL)
AT2G47010	unknown protein
AT2G47490	NAD+ TRANSPORTER 1 (NDT1)
AT2G47510	FUMARASE 1 (FUM1)
AT2G47660	unknown protein
AT2G47844	unknown protein
AT2G48010	RECEPTOR-LIKE KINASE IN IN FLOWERS 3 (RKF3)
AT3G01470	HOMEOBOX 1 (HB-1)
AT3G02120	hydroxyproline-rich glycoprotein family protein
AT3G02180	SPIRAL 1-LIKE3 (SP1L3)
AT3G02910	AIG2-like (avirulence induced gene) family protein
AT3G03270	Adenine nucleotide alpha hydrolases-like superfamily protein
AT3G04010	O-Glycosyl hydrolases family 17 protein
AT3G04070	NAC DOMAIN CONTAINING PROTEIN 47 (NAC047)
AT3G04680	CLP-SIMILAR PROTEIN 3 (CLPS3)
AT3G04721	unknown protein
AT3G05160	Major facilitator superfamily protein
AT3G05165	Major facilitator superfamily protein
AT3G05400	Major facilitator superfamily protein
AT3G05490	RALF-LIKE 22 (RALFL22)
AT3G05840	(ATSK12)
AT3G05936	unknown protein
AT3G05937	unknown protein
AT3G06019	unknown protein
AT3G06500	Plant neutral invertase family protein
AT3G07360	PLANT U-BOX 9 (PUB9)
AT3G07390	AUXIN-INDUCED IN ROOT CULTURES 12 (AIR12)
AT3G07760	Sterile alpha motif (SAM) domain-containing protein
AT3G07840	Pectin lyase-like superfamily protein
AT3G09510	Ribonuclease H-like superfamily protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT3G09920	PHOSPHATIDYL INOSITOL MONOPHOSPHATE 5 KINASE (PIP5K9)
AT3G10320	Glycosyltransferase family 61 protein
AT3G10410	SERINE CARBOXYPEPTIDASE-LIKE 49 (scpl49)
AT3G10760	Homeodomain-like superfamily protein
AT3G10985	SENESCENCE ASSOCIATED GENE 20 (SAG20)
AT3G11780	MD-2-related lipid recognition domain-containing protein / ML domain-containin
AT3G12600	NUDIX HYDROLASE HOMOLOG 16 (NUDT16)
AT3G12960	unknown protein
AT3G13310	Chaperone DnaJ-domain superfamily protein
AT3G13480	unknown protein
AT3G13686	unknown protein
AT3G13710	PRENYLATED RAB ACCEPTOR 1.F4 (PRA1.F4)
AT3G13882	Ribosomal protein L34
AT3G14020	NUCLEAR FACTOR Y, SUBUNIT A6 (NF-YA6)
AT3G14060	unknown protein
AT3G14225	GDSL-MOTIF LIPASE 4 (GLIP4)
AT3G14380	Uncharacterised protein family (UPF0497)
AT3G14450	CTC-INTERACTING DOMAIN 9 (CID9)
AT3G14560	unknown protein
AT3G14590	(NTMC2T6.2)
AT3G14595	Ribosomal protein L18ae family
AT3G14680	CYTOCHROME P450, FAMILY 72, SUBFAMILY A, POLYPEPTIDE 14 (CYP72
AT3G14840	Leucine-rich repeat transmembrane protein kinase
AT3G15200	Tetratricopeptide repeat (TPR)-like superfamily protein
AT3G15300	VQ motif-containing protein
AT3G15790	METHYL-CPG-BINDING DOMAIN 11 (MBD11)
AT3G16830	TOPLESS-RELATED 2 (TPR2)
AT3G17860	JASMONATE-ZIM-DOMAIN PROTEIN 3 (JAZ3)
AT3G18040	MAP KINASE 9 (MPK9)
AT3G18550	BRANCHED 1 (BRC1)
AT3G18710	PLANT U-BOX 29 (PUB29)
AT3G18773	RING/U-box superfamily protein
AT3G18940	clast3-related
AT3G19000	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT3G19390	Granulin repeat cysteine protease family protein
AT3G19520	Protein of unknown function (DUF626)
AT3G19550	unknown protein
AT3G19680	Protein of unknown function (DUF1005)
AT3G20270	lipid-binding serum glycoprotein family protein
AT3G20510	Transmembrane proteins 14C
AT3G20820	Leucine-rich repeat (LRR) family protein
AT3G21320	BEST Arabidopsis thaliana protein match is: hydroxyproline-rich glycoprotein fa
AT3G21700	(SGP2)
AT3G21710	unknown protein
AT3G22160	VQ motif-containing protein
AT3G22380	TIME FOR COFFEE (TIC)
AT3G22890	ATP SULFURYLASE 1 (APS1)
AT3G23030	INDOLE-3-ACETIC ACID INDUCIBLE 2 (IAA2)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT3G23040	unknown protein
AT3G23310	AGC (cAMP-dependent, cGMP-dependent and protein kinase C) kinase family
AT3G23440	EMBRYO SAC DEVELOPMENT ARREST 6 (EDA6)
AT3G23620	Ribosomal RNA processing Brix domain protein
AT3G23727	SCR-LIKE 12 (SCRL12)
AT3G24100	Uncharacterised protein family SERF
AT3G24160	PUTATIVE TYPE 1 MEMBRANE PROTEIN (PMP)
AT3G24420	alpha/beta-Hydrolases superfamily protein
AT3G24480	Leucine-rich repeat (LRR) family protein
AT3G24503	ALDEHYDE DEHYDROGENASE 2C4 (ALDH2C4)
AT3G25250	(AGC2-1)
AT3G25410	Sodium Bile acid symporter family
AT3G25600	Calcium-binding EF-hand family protein
AT3G25610	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase famil
AT3G25720	RNA-directed DNA polymerase (reverse transcriptase)-related family protein
AT3G25780	ALLENE OXIDE CYCLASE 3 (AOC3)
AT3G26870	Plant self-incompatibility protein S1 family
AT3G27010	TEOSINTE BRANCHED 1, CYCLOIDEA, PCF (TCP)-DOMAIN FAMILY PROTI
AT3G27590	unknown protein
AT3G27809	unknown protein
AT3G28850	Glutaredoxin family protein
AT3G28940	AIG2-like (avirulence induced gene) family protein
AT3G29100	VESICLE TRANSPORT V-SNARE 13 (VTI13)
AT3G29370	unknown protein
AT3G29575	ABI FIVE BINDING PROTEIN 3 (AFP3)
AT3G29810	COBRA-LIKE PROTEIN 2 PRECURSOR (COBL2)
AT3G30580	unknown protein
AT3G42060	myosin heavy chain-related
AT3G42770	F-box/RNI-like/FBD-like domains-containing protein
AT3G43120	SAUR-like auxin-responsive protein family
AT3G43148	FUNCTIONS IN: molecular_function unknown
AT3G43153	cAMP-dependent protein kinase inhibitor-related
AT3G43160	MATERNAL EFFECT EMBRYO ARREST 38 (MEE38)
AT3G43280	unknown protein
AT3G43290	unknown protein
AT3G44990	XYLOGLUCAN ENDO-TRANSGLYCOSYLASE-RELATED 8 (XTR8)
AT3G45050	unknown protein
AT3G45443	unknown protein
AT3G45730	unknown protein
AT3G45830	unknown protein
AT3G46280	protein kinase-related
AT3G46620	zinc finger (C3HC4-type RING finger) family protein
AT3G46830	RAB GTPASE HOMOLOG A2C (RABA2c)
AT3G47420	PHOSPHATE STARVATION-INDUCED GENE 3 (PS3)
AT3G47600	MYB DOMAIN PROTEIN 94 (MYB94)
AT3G48240	Octicosapeptide/Phox/Bem1p family protein
AT3G48360	BTB AND TAZ DOMAIN PROTEIN 2 (bt2)
AT3G48510	unknown protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT3G48920	MYB DOMAIN PROTEIN 45 (MYB45)
AT3G49115	unknown protein
AT3G49530	NAC DOMAIN CONTAINING PROTEIN 62 (NAC062)
AT3G49570	RESPONSE TO LOW SULFUR 3 (LSU3)
AT3G49580	RESPONSE TO LOW SULFUR 1 (LSU1)
AT3G49590	Autophagy-related protein 13
AT3G50530	CDPK-related kinase
AT3G50770	CALMODULIN-LIKE 41 (CML41)
AT3G50800	unknown protein
AT3G51890	Clathrin light chain protein
AT3G52290	IQ-DOMAIN 3 (IQD3)
AT3G52340	SUCROSE-6F-PHOSPHATE PHOSPHOHYDROLASE 2 (SPP2)
AT3G52510	F-box associated ubiquitination effector family protein
AT3G52870	IQ calmodulin-binding motif family protein
AT3G52890	KCBP-INTERACTING PROTEIN KINASE (KIPK)
AT3G53200	MYB DOMAIN PROTEIN 27 (MYB27)
AT3G53480	PLEIOTROPIC DRUG RESISTANCE 9 (PDR9)
AT3G54650	(FBL17)
AT3G55250	unknown protein
AT3G55560	AT-HOOK PROTEIN OF GA FEEDBACK 2 (AGF2)
AT3G55730	MYB DOMAIN PROTEIN 109 (MYB109)
AT3G56200	Encodes a putative amino acid transporter.
AT3G56230	BTB/POZ domain-containing protein
AT3G57030	Calcium-dependent phosphotriesterase superfamily protein
AT3G57180	BRASSINAZOLE(BRZ) INSENSITIVE PALE GREEN 2 (BPG2)
AT3G57410	VILLIN 3 (VLN3)
AT3G57450	unknown protein
AT3G59050	POLYAMINE OXIDASE 3 (PAO3)
AT3G59080	Eukaryotic aspartyl protease family protein
AT3G59220	PIRIN (PRN)
AT3G59710	NAD(P)-binding Rossmann-fold superfamily protein
AT3G60140	DARK INDUCIBLE 2 (DIN2)
AT3G61420	BSD domain (BTF2-like transcription factors, Synapse-associated proteins and
AT3G61930	unknown protein
AT3G62090	PHYTOCHROME INTERACTING FACTOR 3-LIKE 2 (PIL2)
AT3G62630	Protein of unknown function (DUF1645)
AT3G62830	(AUD1)
AT3G63460	transducin family protein / WD-40 repeat family protein
AT4G01245	FUNCTIONS IN: molecular_function unknown
AT4G01340	CHP-rich zinc finger protein-related
AT4G01516	unknown protein
AT4G01535	unknown protein
AT4G01671	unknown protein
AT4G01690	(PPOX)
AT4G02330	(ATPMEPCRB)
AT4G02380	SENESCENCE-ASSOCIATED GENE 21 (SAG21)
AT4G03292	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
AT4G03390	STRUBBELIG-RECEPTOR FAMILY 3 (SRF3)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT4G03460	Ankyrin repeat family protein
AT4G04610	APS REDUCTASE 1 (APR1)
AT4G04690	F-box and associated interaction domains-containing protein
AT4G04745	unknown protein
AT4G04960	Concanavalin A-like lectin protein kinase family protein
AT4G05020	NAD(P)H dehydrogenase B2 (NDB2)
AT4G05620	Galactose oxidase/kelch repeat superfamily protein
AT4G06479	zinc ion binding
AT4G07526	unknown protein
AT4G08073	unknown protein
AT4G08097	BEST Arabidopsis thaliana protein match is: myosin heavy chain-related (TAIR:
AT4G08580	microfibrillar-associated protein-related
AT4G08878	Major facilitator superfamily protein
AT4G08949	unknown protein
AT4G09150	T-complex protein 11
AT4G09570	CALCIUM-DEPENDENT PROTEIN KINASE 4 (CPK4)
AT4G09880	unknown protein
AT4G09890	Protein of unknown function (DUF3511)
AT4G10600	RING/FYVE/PHD zinc finger superfamily protein
AT4G10700	BEST Arabidopsis thaliana protein match is: CDC68-related (TAIR:AT4G10660
AT4G10860	unknown protein
AT4G11521	Receptor-like protein kinase-related family protein
AT4G11640	SERINE RACEMASE (SR)
AT4G12090	Cornichon family protein
AT4G12110	STEROL-4ALPHA-METHYL OXIDASE 1-1 (SMO1-1)
AT4G12250	UDP-D-GLUCURONATE 4-EPIMERASE 5 (GAE5)
AT4G12420	(SKU5)
AT4G12730	FASCICLIN-LIKE ARABINOGALACTAN 2 (FLA2)
AT4G13040	Integrase-type DNA-binding superfamily protein
AT4G14350	AGC (cAMP-dependent, cGMP-dependent and protein kinase C) kinase family
AT4G14580	CBL-INTERACTING PROTEIN KINASE 4 (CIPK4)
AT4G14940	AMINE OXIDASE 1 (AO1)
AT4G15056	FUNCTIONS IN: molecular_function unknown
AT4G15236	ABC-2 and Plant PDR ABC-type transporter family protein
AT4G15800	RALF-LIKE 33 (RALFL33)
AT4G16146	cAMP-regulated phosphoprotein 19-related protein
AT4G16150	calmodulin binding
AT4G16380	Heavy metal transport/detoxification superfamily protein
AT4G16490	ARM repeat superfamily protein
AT4G16500	Cystatin/monellin superfamily protein
AT4G16990	RESISTANCE TO LEPTOSPHAERIA MACULANS 3 (RLM3)
AT4G17230	SCARECROW-LIKE 13 (SCL13)
AT4G17615	CALCINEURIN B-LIKE PROTEIN 1 (CBL1)
AT4G17830	Peptidase M20/M25/M40 family protein
AT4G18310	unknown protein
AT4G18450	encodes a member of the ERF (ethylene response factor) subfamily B-3 of ERF
AT4G18640	MORPHOGENESIS OF ROOT HAIR 1 (MRH1)
AT4G18660	unknown protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT4G18700	CBL-INTERACTING PROTEIN KINASE 12 (CIPK12)
AT4G18950	Integrin-linked protein kinase family
AT4G19110	Protein kinase superfamily protein
AT4G19150	Ankyrin repeat family protein
AT4G19220	Tetratricopeptide repeat (TPR)-like superfamily protein
AT4G19420	Pectinacetylesterase family protein
AT4G19960	K+ UPTAKE PERMEASE 9 (KUP9)
AT4G20250	unknown protein
AT4G20320	CTP synthase family protein
AT4G21990	APS REDUCTASE 3 (APR3)
AT4G22820	A20/AN1-like zinc finger family protein
AT4G23010	UDP-galactose transporter 2 (UTR2)
AT4G23060	IQ-DOMAIN 22 (IQD22)
AT4G23190	CYSTEINE-RICH RLK11 (CRK11)
AT4G23810	(WRKY53)
AT4G23940	FtsH extracellular protease family
AT4G24060	Dof-type zinc finger DNA-binding family protein
AT4G24480	Protein kinase superfamily protein
AT4G24570	DICARBOXYLATE CARRIER 2 (DIC2)
AT4G24940	SUMO-ACTIVATING ENZYME 1A (SAE1A)
AT4G24960	ARABIDOPSIS THALIANA HVA22 HOMOLOGUE D (HVA22D)
AT4G25630	FIBRILLARIN 2 (FIB2)
AT4G26950	Protein of unknown function, DUF584
AT4G27180	ARABIDOPSIS THALIANA KINESIN 2 (ATK2)
AT4G27260	(WES1)
AT4G27270	Quinone reductase family protein
AT4G27300	S-locus lectin protein kinase family protein
AT4G27410	RESPONSIVE TO DESICCATION 26 (RD26)
AT4G27657	unknown protein
AT4G28240	Wound-responsive family protein
AT4G28530	NAC DOMAIN CONTAINING PROTEIN 74 (NAC074)
AT4G28700	AMMONIUM TRANSPORTER 1
AT4G28840	unknown protein
AT4G28910	NOVEL INTERACTOR OF JAZ (NINJA)
AT4G29140	MATE efflux family protein
AT4G29190	Zinc finger C-x8-C-x5-C-x3-H type family protein
AT4G29700	Alkaline-phosphatase-like family protein
AT4G29880	PLANT INTRACELLULAR RAS GROUP-RELATED LRR 7 (PIRL7)
AT4G29905	unknown protein
AT4G30060	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
AT4G30190	belongs to the P-type ATPase superfamily of cation-transporting ATPases, pum
AT4G30360	CYCLIC NUCLEOTIDE-GATED CHANNEL 17 (CNGC17)
AT4G30370	RING/U-box superfamily protein
AT4G30400	RING/U-box superfamily protein
AT4G30470	NAD(P)-binding Rossmann-fold superfamily protein
AT4G30490	AFG1-like ATPase family protein
AT4G30610	BRI1 SUPPRESSOR 1 (BRS1)
AT4G30710	QWRF DOMAIN CONTAINING 8 (QWRF8)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT4G31500	CYTOCHROME P450 MONOOXYGENASE 83B1 (CYP83B1)
AT4G31875	unknown protein
AT4G32020	unknown protein
AT4G32070	Octicosapeptide/Phox/Bem1p (PB1) domain-containing protein / tetratricopepti
AT4G32280	INDOLE-3-ACETIC ACID INDUCIBLE 29 (IAA29)
AT4G32300	S-DOMAIN-2 5 (SD2-5)
AT4G32460	FUNCTIONS IN: molecular_function unknown
AT4G33040	Thioredoxin superfamily protein
AT4G33170	Tetratricopeptide repeat (TPR)-like superfamily protein
AT4G33310	unknown protein
AT4G33580	BETA CARBONIC ANHYDRASE 5 (BCA5)
AT4G33960	unknown protein
AT4G33980	FUNCTIONS IN: molecular_function unknown
AT4G34200	EMBRYO SAC DEVELOPMENT ARREST 9 (EDA9)
AT4G34530	CRYPTOCHROME-INTERACTING BASIC-HELIX-LOOP-HELIX 1 (CIB1)
AT4G34740	GLN PHOSPHORIBOSYL PYROPHOSPHATE AMIDOTRANSFERASE 2 (ASE
AT4G34770	SAUR-like auxin-responsive protein family
AT4G34980	SUBTILISIN-LIKE SERINE PROTEASE 2 (SLP2)
AT4G35070	SBP (S-ribonuclease binding protein) family protein
AT4G35720	Arabidopsis protein of unknown function (DUF241)
AT4G35790	PHOSPHOLIPASE D DELTA (PLDDELTA)
AT4G35850	Pentatricopeptide repeat (PPR) superfamily protein
AT4G36400	Encodes a (D)-2-hydroxyglutarate dehydrogenase.
AT4G36650	PLANT-SPECIFIC TFIIB-RELATED PROTEIN (PBRP)
AT4G36925	unknown protein
AT4G36970	Remorin family protein
AT4G37235	Uncharacterised protein family (UPF0497)
AT4G37380	Tetratricopeptide repeat (TPR)-like superfamily protein
AT4G37550	Acetamidase/Formamidase family protein
AT4G37730	BASIC LEUCINE-ZIPPER 7 (bZIP7)
AT4G37740	GROWTH-REGULATING FACTOR 2 (GRF2)
AT4G37790	(HAT22)
AT4G38330	Integral membrane protein hemolysin-III homolog
AT4G38620	MYB DOMAIN PROTEIN 4 (MYB4)
AT4G38850	SMALL AUXIN UPREGULATED 15 (SAUR15)
AT4G39390	NUCLEOTIDE SUGAR TRANSPORTER-KT 1 (NST-K1)
AT4G39890	RAB GTPASE HOMOLOG H1C (RABH1c)
AT4G39930	unknown protein
AT4G40010	SNF1-RELATED PROTEIN KINASE 2.7 (SNRK2.7)
AT5G01530	LIGHT HARVESTING COMPLEX PHOTOSYSTEM II (LHCB4.1)
AT5G01720	RNI-like superfamily protein
AT5G01734	unknown protein
AT5G01820	SERINE/THREONINE PROTEIN KINASE 1 (SR1)
AT5G01830	ARM repeat superfamily protein
AT5G02021	unknown protein
AT5G02530	RNA-binding (RRM/RBD/RNP motifs) family protein
AT5G03370	acylphosphatase family
AT5G03530	RAB GTPASE HOMOLOG C2A (RABC2A)

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT5G03660	Family of unknown function (DUF662)
AT5G03720	HEAT SHOCK TRANSCRIPTION FACTOR A3 (HSFA3)
AT5G04430	BINDING TO TOMV RNA 1L (LONG FORM) (BTR1L)
AT5G04470	SIAMESE (SIM)
AT5G04530	3-KETOACYL-COA SYNTHASE 19 (KCS19)
AT5G04590	SULFITE REDUCTASE (SIR)
AT5G04750	F1F0-ATPase inhibitor protein, putative
AT5G04820	ARABIDOPSIS THALIANA OVATE FAMILY PROTEIN 13 (OFP13)
AT5G04930	AMINOPHOSPHOLIPID ATPASE1 (ALA1)
AT5G05090	Homeodomain-like superfamily protein
AT5G05140	Transcription elongation factor (TFIIS) family protein
AT5G05590	PHOSPHORIBOSYLANTHRANILATE ISOMERASE 2 (PAI2)
AT5G05598	Encodes a Defensin-like (DEFL) family protein
AT5G05730	ASA1 encodes the alpha subunit of anthranilate synthase, which catalyzes the
AT5G05965	unknown protein
AT5G06270	unknown protein
AT5G06300	Putative lysine decarboxylase family protein
AT5G07440	GLUTAMATE DEHYDROGENASE 2 (GDH2)
AT5G07680	NAC DOMAIN CONTAINING PROTEIN 80 (NAC080)
AT5G07870	HXXXD-type acyl-transferase family protein
AT5G07980	dentin sialophosphoprotein-related
AT5G08240	unknown protein
AT5G08650	Small GTP-binding protein
AT5G08790	(ATAF2)
AT5G09760	Plant invertase/pectin methylesterase inhibitor superfamily
AT5G09850	Transcription elongation factor (TFIIS) family protein
AT5G09976	FUNCTIONS IN: molecular_function unknown
AT5G10720	HISTIDINE KINASE 5 (HK5)
AT5G10960	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
AT5G11090	serine-rich protein-related
AT5G11230	Nucleotide-sugar transporter family protein
AT5G11650	alpha/beta-Hydrolases superfamily protein
AT5G11950	Encodes a protein of unknown function. It has been crystallized and shown to b
AT5G12340	unknown protein
AT5G12890	UDP-Glycosyltransferase superfamily protein
AT5G12940	Leucine-rich repeat (LRR) family protein
AT5G13170	SENESCENCE-ASSOCIATED PROTEIN 29 (SAG29)
AT5G13180	NAC DOMAIN CONTAINING PROTEIN 83 (NAC083)
AT5G13360	Auxin-responsive GH3 family protein
AT5G13390	NO EXINE FORMATION 1 (NEF1)
AT5G13740	ZINC INDUCED FACILITATOR 1 (ZIF1)
AT5G13760	Plasma-membrane choline transporter family protein
AT5G14180	MYZUS PERSICAE-INDUCED LIPASE 1 (MPL1)
AT5G14350	Pentatricopeptide repeat (PPR) superfamily protein
AT5G14500	aldose 1-epimerase family protein
AT5G14760	L-ASPARTATE OXIDASE (AO)
AT5G15210	HOMEOBOX PROTEIN 30 (HB30)
AT5G15220	Ribosomal protein L27 family protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT5G15260	Ribosomal protein L34e superfamily protein
AT5G15580	LONGIFOLIA1 (LNG1)
AT5G15950	Adenosylmethionine decarboxylase family protein
AT5G16480	Phosphotyrosine protein phosphatases superfamily protein
AT5G16560	KANADI (KAN)
AT5G16567	unknown protein
AT5G16640	Pentatricopeptide repeat (PPR) superfamily protein
AT5G16880	Target of Myb protein 1
AT5G17330	GLUTAMATE DECARBOXYLASE (GAD)
AT5G17640	Protein of unknown function (DUF1005)
AT5G17795	unknown protein
AT5G18560	(PUCHI)
AT5G18630	alpha/beta-Hydrolases superfamily protein
AT5G18650	CHY-type/CTCHY-type/RING-type Zinc finger protein
AT5G18760	RING/U-box superfamily protein
AT5G19140	(AILP1)
AT5G19160	TRICHOME BIREFRINGENCE-LIKE 11 (TBL11)
AT5G19257	unknown protein
AT5G19650	ARABIDOPSIS THALIANA OVATE FAMILY PROTEIN 8 (OFP8)
AT5G20110	Dynein light chain type 1 family protein
AT5G20181	unknown protein
AT5G20510	ALFIN-LIKE 5 (AL5)
AT5G22250	Polynucleotidyl transferase, ribonuclease H-like superfamily protein
AT5G22300	NITRILASE 4 (NIT4)
AT5G22490	O-acyltransferase (WSD1-like) family protein
AT5G22570	WRKY DNA-BINDING PROTEIN 38 (WRKY38)
AT5G23000	MYB DOMAIN PROTEIN 37 (MYB37)
AT5G23340	RNI-like superfamily protein
AT5G23870	Pectinacetylesterase family protein
AT5G23930	Mitochondrial transcription termination factor family protein
AT5G24210	alpha/beta-Hydrolases superfamily protein
AT5G24620	Pathogenesis-related thaumatin superfamily protein
AT5G24660	RESPONSE TO LOW SULFUR 2 (LSU2)
AT5G24760	GroES-like zinc-binding dehydrogenase family protein
AT5G25130	CYTOCHROME P450, FAMILY 71, SUBFAMILY B, POLYPEPTIDE 12 (CYP71
AT5G25180	CYTOCHROME P450, FAMILY 71, SUBFAMILY B, POLYPEPTIDE 14 (CYP71
AT5G25560	CHY-type/CTCHY-type/RING-type Zinc finger protein
AT5G25820	Exostosin family protein
AT5G25840	Protein of unknown function (DUF1677)
AT5G26010	Protein phosphatase 2C family protein
AT5G27000	ARABIDOPSIS THALIANA KINESIN 4 (ATK4)
AT5G27340	unknown protein
AT5G27350	(SFP1)
AT5G28050	Cytidine/deoxycytidylate deaminase family protein
AT5G28300	Duplicated homeodomain-like superfamily protein
AT5G28960	unknown protein
AT5G29560	caleosin-related family protein
AT5G31412	hAT transposon superfamily protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT5G33370	GDSL-like Lipase/Acylhydrolase superfamily protein
AT5G33406	hAT dimerisation domain-containing protein / transposase-related
AT5G35180	FUNCTIONS IN: phosphoinositide binding
AT5G37260	REVEILLE 2 (RVE2)
AT5G37890	Protein with RING/U-box and TRAF-like domains
AT5G38020	encodes a protein whose sequence is similar to SAM:salicylic acid carboxyl me
AT5G38040	UDP-Glycosyltransferase superfamily protein
AT5G38200	Class I glutamine amidotransferase-like superfamily protein
AT5G38210	Protein kinase family protein
AT5G38240	Protein kinase family protein
AT5G38410	Ribulose bisphosphate carboxylase (small chain) family protein
AT5G38480	GENERAL REGULATORY FACTOR 3 (GRF3)
AT5G38980	unknown protein
AT5G39050	HXXXD-type acyl-transferase family protein
AT5G39320	UDP-glucose 6-dehydrogenase family protein
AT5G39581	unknown protein
AT5G39950	THIOREDOXIN 2 (TRX2)
AT5G40330	MYB DOMAIN PROTEIN 23 (MYB23)
AT5G40470	RNI-like superfamily protein
AT5G40690	CONTAINS InterPro DOMAIN/s: EF-Hand 1, calcium-binding site (InterPro:IPR
AT5G40800	unknown protein
AT5G40840	(SYN2)
AT5G42650	ALLENE OXIDE SYNTHASE (AOS)
AT5G43150	unknown protein
AT5G43700	AUXIN INDUCIBLE 2-11 (ATAUX2-11)
AT5G43810	ZWILLE (ZLL)
AT5G44220	F-box family protein
AT5G44390	FAD-binding Berberine family protein
AT5G44790	RESPONSIVE-TO-ANTAGONIST 1 (RAN1)
AT5G45100	SBP (S-ribonuclease binding protein) family protein
AT5G45110	NPR1-LIKE PROTEIN 3 (NPR3)
AT5G45280	Pectinacetylesterase family protein
AT5G45720	AAA-type ATPase family protein
AT5G45820	CBL-INTERACTING PROTEIN KINASE 20 (CIPK20)
AT5G45940	NUDIX HYDROLASE HOMOLOG 11 (NUDT11)
AT5G46080	Protein kinase superfamily protein
AT5G46700	TORONADO 2 (TRN2)
AT5G47040	LON PROTEASE 2 (LON2)
AT5G47100	CALCINEURIN B-LIKE PROTEIN 9 (CBL9)
AT5G47740	Adenine nucleotide alpha hydrolases-like superfamily protein
AT5G48150	PHYTOCHROME A SIGNAL TRANSDUCTION 1 (PAT1)
AT5G48500	unknown protein
AT5G48550	F-box associated ubiquitination effector family protein
AT5G48860	unknown protein
AT5G49120	Protein of unknown function (DUF581)
AT5G49610	F-box family protein
AT5G50120	Transducin/WD40 repeat-like superfamily protein
AT5G50460	secE/sec61-gamma protein transport protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT5G50760	SAUR-like auxin-responsive protein family
AT5G50900	ARM repeat superfamily protein
AT5G50920	CLPC HOMOLOGUE 1 (CLPC1)
AT5G51090	unknown protein
AT5G51180	alpha/beta-Hydrolases superfamily protein
AT5G51990	C- REPEAT-BINDING FACTOR 4 (CBF4)
AT5G52780	Protein of unknown function (DUF3464)
AT5G53200	TRIPTYCHON (TRY)
AT5G53220	unknown protein
AT5G53280	PLASTID DIVISION1 (PDV1)
AT5G53410	unknown protein
AT5G53451	unknown protein
AT5G53500	Transducin/WD40 repeat-like superfamily protein
AT5G53540	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT5G53905	unknown protein
AT5G54145	unknown protein
AT5G54160	O-METHYLTRANSFERASE 1 (OMT1)
AT5G54165	unknown protein
AT5G54225	LOW-MOLECULAR-WEIGHT CYSTEINE-RICH 83 (LCR83)
AT5G54540	Uncharacterised conserved protein (UCP012943)
AT5G54650	FORMIN HOMOLOGY5 (Fh5)
AT5G55080	RAS-RELATED NUCLEAR PROTEIN 4 (RAN4)
AT5G56050	FUNCTIONS IN: molecular_function unknown
AT5G56230	PRENYLATED RAB ACCEPTOR 1.G2 (PRA1.G2)
AT5G56865	unknown protein
AT5G56870	BETA-GALACTOSIDASE 4 (BGAL4)
AT5G56980	unknown protein
AT5G57040	Lactoylglutathione lyase / glyoxalase I family protein
AT5G57330	Galactose mutarotase-like superfamily protein
AT5G57391	unknown protein
AT5G57567	unknown protein
AT5G57700	BNR/Asp-box repeat family protein
AT5G57720	AP2/B3-like transcriptional factor family protein
AT5G57800	ECERIFERUM 3 (CER3)
AT5G58110	chaperone binding
AT5G58350	WITH NO K (=LYSINE) 4 (WNK4)
AT5G58900	Homeodomain-like transcriptional regulator
AT5G59450	GRAS family transcription factor
AT5G59520	ZRT/IRT-LIKE PROTEIN 2 (ZIP2)
AT5G59820	RESPONSIVE TO HIGH LIGHT 41 (RHL41)
AT5G59960	unknown protein
AT5G60660	PLASMA MEMBRANE INTRINSIC PROTEIN 2
AT5G60680	Protein of unknown function, DUF584
AT5G60690	REVOLUTA (REV)
AT5G60880	BREAKING OF ASYMMETRY IN THE STOMATAL LINEAGE (BASL)
AT5G61440	ATYPICAL CYS HIS RICH THIOREDOXIN 5 (ACHT5)
AT5G61780	TUDOR-SN PROTEIN 2 (TUDOR2)
AT5G61990	Pentatricopeptide repeat (PPR) superfamily protein

Table S3, con't. EIN3 targets.

EIN3-NR (EIN3 targets, not ethylene-regulated)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT5G62470	MYB DOMAIN PROTEIN 96 (MYB96)
AT5G63087	Encodes a Plant thionin family protein
AT5G63420	EMBRYO DEFECTIVE 2746 (emb2746)
AT5G63595	FLAVONOL SYNTHASE 4 (FLS4)
AT5G63790	NAC DOMAIN CONTAINING PROTEIN 102 (NAC102)
AT5G64260	EXORDIUM LIKE 2 (EXL2)
AT5G64530	XYLEM NAC DOMAIN 1 (XND1)
AT5G64570	BETA-D-XYLOSIDASE 4 (XYL4)
AT5G64650	Ribosomal protein L17 family protein
AT5G64660	CYS, MET, PRO, AND GLY PROTEIN 2 (CMPG2)
AT5G64910	unknown protein
AT5G65310	HOMEOBOX PROTEIN 5 (HB5)
AT5G65470	O-fucosyltransferase family protein
AT5G65640	BETA HLH PROTEIN 93 (bHLH093)
AT5G65670	INDOLE-3-ACETIC ACID INDUCIBLE 9 (IAA9)
AT5G65910	BSD domain-containing protein
AT5G66650	Protein of unknown function (DUF607)
AT5G66790	Protein kinase superfamily protein
AT5G66880	SUCROSE NONFERMENTING 1(SNF1)-RELATED PROTEIN KINASE 2.3 (SI
AT5G67290	FAD-dependent oxidoreductase family protein
AT5G67300	MYB DOMAIN PROTEIN R1 (MYBR1)
AT5G67420	LOB DOMAIN-CONTAINING PROTEIN 37 (LBD37)

Table S3, con't. EIN3 targets.

EIN3-ND (EIN3 targets, expression not detected)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G06148	Unknown gene
AT1G07180	ALTERNATIVE NAD(P)H DEHYDROGENASE 1 (NDA1)
AT1G11250	SYNTAXIN OF PLANTS 125 (SYP125)
AT1G11330	S-locus lectin protein kinase family protein
AT1G17235	ROTUNDIFOLIA LIKE 11 (RTFL11)
AT1G20130	GDSL-like Lipase/Acylhydrolase superfamily protein
AT1G20310	unknown protein
AT1G21970	LEAFY COTYLEDON 1 (LEC1)
AT1G23965	unknown protein
AT1G26210	SOB FIVE-LIKE 1 (SOFL1)
AT1G29020	Calcium-binding EF-hand family protein
AT1G31290	ARGONAUTE 3 (AGO3)
AT1G52343	unknown protein
AT1G57613	unknown protein
AT1G60095	Mannose-binding lectin superfamily protein
AT1G69990	Leucine-rich repeat protein kinase family protein
AT1G76230	unknown protein
AT2G25460	CONTAINS InterPro DOMAIN/s: C2 calcium-dependent membrane targeting (I
AT2G27250	CLAVATA3 (CLV3)
AT2G28560	(RAD51B)
AT2G28690	Protein of unknown function (DUF1635)
AT2G31230	ETHYLENE-RESPONSIVE ELEMENT BINDING FACTOR 15 (ERF15)
AT2G32487	unknown protein
AT2G37980	O-fucosyltransferase family protein
AT2G38920	SPX (SYG1/Pho81/XPR1) domain-containing protein / zinc finger (C3HC4-type
AT2G41225	unknown protein
AT2G44800	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT2G46640	unknown protein
AT3G11180	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT3G15700	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT3G19090	RNA-binding protein
AT3G19700	HAIKU2 (IKU2)
AT3G21780	UDP-GLUCOSYL TRANSFERASE 71B6 (UGT71B6)
AT3G23120	RECEPTOR LIKE PROTEIN 38 (RLP38)
AT3G26200	CYTOCHROME P450, FAMILY 71, SUBFAMILY B, POLYPEPTIDE 22 (CYP71
AT3G26790	FUSCA 3 (FUS3)
AT3G58780	One of two genes (SHP1 and SHP2) that are required for fruit dehiscence. The
AT4G18350	NINE-CIS-EPOXYCAROTENOID DIOXYGENASE 2 (NCED2)
AT4G27290	S-locus lectin protein kinase family protein
AT4G27850	Glycine-rich protein family
AT4G37310	CYTOCHROME P450, FAMILY 81, SUBFAMILY H, POLYPEPTIDE 1 (CYP81H
AT4G38530	PHOSPHOLIPASE C1 (PLC1)
AT4G38560	Arabidopsis phospholipase-like protein (PEARLI 4) family
AT5G03680	PETAL LOSS (PTL)
AT5G25290	CONTAINS InterPro DOMAIN/s: F-box domain, cyclin-like (InterPro:IPR001810
AT5G35770	STERILE APETALA (SAP)
AT5G40900	Nucleotide-diphospho-sugar transferase family protein
AT5G48100	TRANSPARENT TESTA 10 (TT10)

Table S3, con't. EIN3 targets.

EIN3-ND (EIN3 targets, expression not detected)

AtID	TAIR10 Primary Gene Symbol/Gene Model Description
AT1G06148	Unknown gene
AT5G49620	Member of the R2R3 factor gene family.
AT5G54720	Ankyrin repeat family protein
AT5G61740	ABC2 HOMOLOG 14 (ATH14)
AT5G65320	basic helix-loop-helix (bHLH) DNA-binding superfamily protein

Table S4. EIN3 Target Gene Distribution of Gene Ontology Terms.

Term	Observed Frequency	Expected Frequency	p-value
response to hormone stimulus	107 out of 1306 genes, 8.2%	702 out of 28924 genes, 2.4%	3.92E-26
response to endogenous stimulus	115 out of 1306 genes, 8.8%	801 out of 28924 genes, 2.8%	7.21E-26
response to chemical stimulus	143 out of 1306 genes, 10.9%	1271 out of 28924 genes, 4.4%	1.35E-21
response to ethylene stimulus	37 out of 1306 genes, 2.8%	138 out of 28924 genes, 0.5%	2.65E-16
hormone-mediated signaling	44 out of 1306 genes, 3.4%	225 out of 28924 genes, 0.8%	6.12E-14
regulation of transcription	147 out of 1306 genes, 11.3%	1617 out of 28924 genes, 5.6%	1.15E-13
regulation of nucleobase, nucleoside, nucleotid	147 out of 1306 genes, 11.3%	1636 out of 28924 genes, 5.7%	3.09E-13
regulation of cellular metabolic process	150 out of 1306 genes, 11.5%	1703 out of 28924 genes, 5.9%	8.19E-13
regulation of cellular process	170 out of 1306 genes, 13.0%	2033 out of 28924 genes, 7.0%	1.05E-12
transcription	150 out of 1306 genes, 11.5%	1713 out of 28924 genes, 5.9%	1.33E-12
regulation of metabolic process	150 out of 1306 genes, 11.5%	1719 out of 28924 genes, 5.9%	1.79E-12
regulation of biological process	176 out of 1306 genes, 13.5%	2178 out of 28924 genes, 7.5%	7.42E-12
&regulation of gene expression	150 out of 1306 genes, 11.5%	1763 out of 28924 genes, 6.1%	1.42E-11
regulation of gene expression	150 out of 1306 genes, 11.5%	1763 out of 28924 genes, 6.1%	1.42E-11
ethylene mediated signaling pathway	20 out of 1306 genes, 1.5%	55 out of 28924 genes, 0.2%	4.83E-11
biological regulation	192 out of 1306 genes, 14.7%	2521 out of 28924 genes, 8.7%	1.12E-10
two-component signal transduction system (pho	22 out of 1306 genes, 1.7%	73 out of 28924 genes, 0.3%	2.41E-10
response to stimulus	205 out of 1306 genes, 15.7%	2823 out of 28924 genes, 9.8%	1.41E-09
response to abscisic acid stimulus	37 out of 1306 genes, 2.8%	227 out of 28924 genes, 0.8%	4.94E-09
response to auxin stimulus	40 out of 1306 genes, 3.1%	273 out of 28924 genes, 0.9%	2.23E-08
response to jasmonic acid stimulus	26 out of 1306 genes, 2.0%	139 out of 28924 genes, 0.5%	2.61E-07
regulation of transcription, DNA-dependent	85 out of 1306 genes, 6.5%	934 out of 28924 genes, 3.2%	2.96E-07
regulation of RNA metabolic process	85 out of 1306 genes, 6.5%	940 out of 28924 genes, 3.2%	4.02E-07
regulation of ethylene mediated signaling pathw	8 out of 1306 genes, 0.6%	11 out of 28924 genes, 0.0%	9.96E-07
negative regulation of ethylene mediated signal	8 out of 1306 genes, 0.6%	11 out of 28924 genes, 0.0%	9.96E-07
cell communication	94 out of 1306 genes, 7.2%	1109 out of 28924 genes, 3.8%	1.41E-06
transcription, DNA-dependent	85 out of 1306 genes, 6.5%	967 out of 28924 genes, 3.3%	1.54E-06
RNA biosynthetic process	85 out of 1306 genes, 6.5%	968 out of 28924 genes, 3.3%	1.61E-06
intracellular signaling cascade	59 out of 1306 genes, 4.5%	581 out of 28924 genes, 2.0%	2.59E-06
signal transduction	83 out of 1306 genes, 6.4%	974 out of 28924 genes, 3.4%	9.50E-06
nucleobase, nucleoside, nucleotide and nucleic	176 out of 1306 genes, 13.5%	2657 out of 28924 genes, 9.2%	5.14E-05
response to gibberellin stimulus	21 out of 1306 genes, 1.6%	126 out of 28924 genes, 0.4%	8.90E-05
response to salt stress	26 out of 1306 genes, 2.0%	183 out of 28924 genes, 0.6%	9.30E-05
response to abiotic stimulus	73 out of 1306 genes, 5.6%	883 out of 28924 genes, 3.1%	0.00021
&regulation of response to stimulus	18 out of 1306 genes, 1.4%	107 out of 28924 genes, 0.4%	0.00055
regulation of response to stimulus	18 out of 1306 genes, 1.4%	107 out of 28924 genes, 0.4%	0.00055
negative regulation of signal transduction	10 out of 1306 genes, 0.8%	36 out of 28924 genes, 0.1%	0.00118
response to osmotic stress	26 out of 1306 genes, 2.0%	210 out of 28924 genes, 0.7%	0.00133
&regulation of signal transduction	16 out of 1306 genes, 1.2%	93 out of 28924 genes, 0.3%	0.00156
regulation of signal transduction	16 out of 1306 genes, 1.2%	93 out of 28924 genes, 0.3%	0.00156
response to salicylic acid stimulus	19 out of 1306 genes, 1.5%	131 out of 28924 genes, 0.5%	0.00279

Chapter Two, in full, consists of the following manuscript in preparation in the format of the journal *Science* as a report.

Chang KN, Li H, Hon G, Pelizzola M, Schmitz RJ, Urich M, Kuo P. Dwight, Nery J, Qiao H, Ideker T, Ecker JR. Dynamic transcription factor binding reveals role of EIN3 in plant hormone crosstalk.

I was the primary researcher and author for this manuscript and Joseph R. Ecker directed and supervised the research that formed the basis of this chapter. Hai Li contributed the genetic and phenotypic analysis in Chapter Two.

CHAPTER 3:

DISCUSSION

SUMMARY

The sessile nature of plants necessitates a phenotypic plasticity that enables plants to respond to changes in the environment throughout growth and development. However, little is known about how signals are integrated to produce a specific response. To further understand the transcriptional regulation of the ethylene response, we identified targets of the master regulator of the ethylene transcriptional response, EIN3. Overexpression of the transcription factor EIN3 in the absence of ethylene is sufficient to produce an ethylene response in etiolated seedlings, and characterization of the targets of EIN3 may help us understand how ethylene mediates many morphological responses. We studied the dynamic ethylene transcriptional response in etiolated seedlings by using a combination ChIP-Seq and mRNA-Seq. The results reveal EIN3 targets that are downstream effectors of the ethylene response, negative regulators of the ethylene signaling pathway, and genes involved in hormone crosstalk.

In a minority of cases ethylene induced EIN3 targets seem to behave in a manner consistent with a canonical view of transcription, in which a bound activator elicits transcription upon physical recruitment to DNA. Interestingly, ~70% of EIN3 physical binding targets do not behave in this manner suggesting that possibly other transcription factor coregulators and/or chromatin modifiers are likely to strongly influence transcriptional outcomes and a complete understanding of the ethylene transcriptional response will require an understanding of additional modulatory factors at the chromatin level. Also, it is possible EIN3 may be playing a major role in biological processes not relating to the ethylene response.

The major contribution of this study to the field of plant biology is that it

establishes a network view of the transcriptional response mediated by ethylene, allowing for the generation of a theoretical framework for the understanding of the crosstalk between ethylene and other plant hormones, the temporal progression of the ethylene response, and confirmation of genetically identified ethylene targets in a mechanistic manner.

SIGNIFICANCE OF RESULTS AND FUTURE IMPLICATIONS

The implications of this study for the plant hormone are two-fold: 1) we identified many interconnections between ethylene and other hormone pathways via the mechanism of EIN3 DNA binding which can now be studied on an individual basis, 2) the role of ethylene or EIN3 in growth and developmental processes can be further examined mechanistically. Those studying the ethylene response, or other hormones that share crosstalk elements with ethylene, can now easily identify whether EIN3 protein-DNA binding is involved in a specific biological phenomenon. Conversely, those interested in studying a specific biological phenomenon (e.g. plant immune response) can determine whether EIN3 may be involved in the transcriptional regulation of that specific process. Homozygous T-DNA mutants exist for ~70% of the EIN3 targets (923 of 1300+ targets, http://signal.salk.edu). One example of a follow-up study of an EIN3 target was performed in collaboration with Boutrot et al. *(73)*. Boutrot and colleagues performed a screen to isolate mutants insensitive to the bacterial flagellin peptide (flg22), and discovered that *ein2-5* and *ein3-1 eil1-1* had deficient pathogen responses that were dependent on the leucine-rich receptor kinase, FLAGELLIN SENSITIVE2 (FLS2). Coincidentally, we found that EIN3 targets the FLS2 gene at two sites in its promoter, and its induction by ethylene made it a

likely direct target of EIN3. This protein-DNA interaction was then confirmed. Several gene families and gene ontology terms were over-represented in our dataset, and it is likely that this study may further serve as a discovery tool for mechanistic explanations of transcriptional regulation by ethylene.

Because hormone regulation is prevalent in growth and development, the role of hormones in plant growth and development are apparent in our data. In our case study of HLS1 and its HLH homologs as targets of EIN3, we observed pleiotropic phenotypes at many stages of plant development. This is a complementary finding to previous studies examining targets of the floral transcription factors. Kaufmann and colleagues found an abundance of hormone crosstalk genes when querying targets of floral transcription factors AP1 and SEP3. Conversely, we found many genes involved in plant growth and development, and redundancy was revealed in the binding of EIN3 to targets within the same gene family. Other examples of redundancy in targets of EIN3 exist in our data, for example, HECATE1 and HECATE2, transcription factors required for floral development, are both targets of EIN3. Furthermore, as previously mentioned, gene families of transcription factors such as BZR1, bHLH, TIFY, and RAVs are significantly represented in our data.

Plant growth and development is intricately tied into hormone regulation. Identification of the connections between these elements by mapping all protein-protein interactions or protein-DNA interactions will yield the possible solution set of network connections possible. For a given time, specific condition, and spatial location in a plant, a subset of these connections will most likely be relevant. Therefore, this study, in addition to the previous high throughput protein-DNA

interaction studies and protein-protein interaction studies, provide a framework to build the hormone crosstalk network.

This work is a technical advance both in the plant biology field and in the area of transcriptional regulation mediated by protein-DNA interactions. This represents a method development milestone because of the limited number of plant chromatin immunoprecipitation studies. This work is one of three ChIP-Seq studies performed in *Arabidopsis thaliana*, the model system for dicots. Only two successful plant ChIP-Seq studies have been previously reported in the world, both were performed by the same group in floral tissue *(102, 103)*. To our knowledge, this is the first report of ChIP-Seq in etiolated seedling tissue, for the master regulator of a hormone signaling pathway. There are two recent reports of ChIP-chip studies investigating protein-DNA interactions of transcription factors involved in the brassinosteroid (BR) response *(97, 98)*. Most of the recent successful genome-wide protein DNA studies have used transgenic strains containing epitope-tagged proteins, thus negating the need of the generation of many native antibodies to study transcription factor binding. In addition cloning strategies using inducible transgenic lines or native promoter constructs may help create a system that is most similar to the native system, thus discovering *bona fide* transcription factor targets.

This work is also a technical advance in the general protein-DNA interaction field because of the temporal resolution. Few temporal protein-DNA binding studies have been undertaken *(133-135)*. Temporal protein-DNA interactions are very difficult to reconcile with temporal gene expression profiles. The complexity of regulation that occurs transcriptionally is very challenging to characterize and interpret biologically.

None of the previous studies found a single protein-DNA binding profile for a transcription factor over time, and because of this result, this study is unique.

COMPARISON TO BRASSINOSTEROID TRANSCRIPTIONAL REGULATION

As mentioned previously, this study is one of three genome-wide ChIP-Seq analyses of hormone transcription factor binding in *Arabidopsis thaliana*. ChIP-chip studies have been performed for two other brassinosteroid transcription factors, BRI1 EMS SUPPRESSOR1 (BES1) and BRASSINAZOLE RESISTANT1 (BZR1). Transcription factor binding of BES1, BZR1, and EIN3 reveal the complexity of hormone crosstalk that occurs in plant growth and development as evidenced by the number and annotation of targets. BZR1 has been shown to negatively regulate BR signaling while BES1 positively regulates BR signaling. Similarly, EIN3 is positively regulates the ethylene response. BES1 and BZR1 are also involved in feedback regulation of the BR pathway; both target the BRI1 receptor, and BZR1 additionally targets BRI1 SUPPORESSOR 1 (BSU1) and BRASSINOSTEROID INSENSITIVE2 (BIN2) signaling components as well as BR biosynthesis genes. In our study, EIN3 was found to target ethylene signaling components, including negative regulators of the ethylene response (CTR1, ETR2, ERS1, ERS2, RTE1, EBF1, EBF2) as well as positive regulators of the ethylene response (ACO2, EIN2, EIL2).

Similarly to EIN3, BES1 and BZR1 targets genes involved in cell elongation, however most of the selected targets of BES1 exhibited negligible differences to sensitivity of BR. BZR1 binding regulation extends to the biological processes such as cell expansion, root development, and development genes, as well as other signaling pathways, including ethylene. BZR1 targets ethylene biosynthesis (ACO2, ACS11),

signaling components (CTR1, EBF2, EIL1), and downstream effectors of the ethylene signaling pathway (ERF1/2/8/11/12) *(98)*. Other hormone responses that are linked to BR transcriptional regulation through the targeting of BES1 and BZR1 include abscisic acid, auxin, cytokinin, gibberellin, methyl jasmonate, and light. Both BES1 and BZR1 target IAA and ARF proteins involved in the regulation of the auxin response; EIN3 also targets IAAs. Because our study was conducted in etiolated seedlings, where the ethylene triple response is observed, we did not observe many light-related targets of EIN3 (e.g. PORA/B), whereas BES1 and BZR1 targeted more light related genes.

Each of these transcription factors regulates a large set of shared targets, as represented as the large circle in the center of the protein-DNA interaction network. Also, each transcription factor regulates a smaller, unique set of targets (Fig. 1). It is likely the unique set of targets represent those that are relevant to the transcriptional response mediated specifically by EIN3, BZR1, and BES1. This is in agreement with the previous conjecture that hormones regulate the same processes through nonoverlapping transcriptional modules *(78)*. The unique set of targets of each transcription factor may also indicate the specific complement of genes required for a morphological response mediated by a specific transcription factor.

UNANSWERED QUESTIONS

Although this study represents a significant advance for plant biology and hormone crosstalk, there are several questions that remain unanswered. The most significant lingering questions are discussed below. These include: 1) Why does EIN3 binding occur in a singular increasing pattern upon ethylene gas treatment? 2) Why

are there an excess of EIN3 targets that are not transcriptionally regulated? What is the significance of these, and are these morphologically functional? 3) Can other transcription factors such as homologs of EIN3 regulate the same or similar sets of targets? 4) What is the functional relevance of EIN3 binding in hormone crosstalk?

Singularity of EIN3 binding pattern

The simple binding pattern of EIN3 for all targets was surprising but not illogical based on previous observations of the ethylene transcriptional response and the protein stability of EIN3. This singular EIN3 binding pattern was surprising because the ethylene response occurs in two phases, an initial rapid growth inhibition phase, followed by a lag, and a prolonged growth inhibition *(106)*. The first growth inhibition response to ethylene is EIN3 and EIL1 independent. However, we do know that EIN3 accumulates to a maximum at one hour, and is rapidly degraded in thirty minutes following ethylene gas removal. Although the EIN3 protein is short-lived, the degradation of EIN3 is was shown not to be required for the ethylene response *(50)*.

The increasing pattern of EIN3 binding correlates to the amount of EIN3 protein present upon ethylene gas treatment. There are two simple explanations for this: 1) an increased number of cells or cell types experience the ethylene response upon a longer duration of ethylene gas treatment, thus the signal of EIN3 binding increases over time, or 2) the proportion of EIN3 binding increases with the accumulation of EIN3, based on the ratio of EIN3 protein to accessible DNA binding sites. It is also possible that this pattern is a result of the combination of these occurrences. For example, we know that different cells display a different amount of nuclear localization of EIN3 protein, even when in close vicinity of each other *(45)*, it

is possible that only a fraction of cells are experiencing an ethylene response at any given time. With an increasing duration of ethylene treatment, more cells may exhibit an ethylene response. In previous EIN3-GFP localization studies, an excessive amount of ACC, the precursor of ethylene was used (50 μM, 100-fold more than is required to induce an ethylene response). Localization of EIN3 to the nucleus appeared vary in different cells, although a majority displayed the EIN3-GFP signal in the nucleus. A timecourse of ACC treatment revealed an accumulation of EIN3-GFP in the nucleus, for the same cells, so it is unknown whether there was cell-to-cell variation over the duration of ethylene treatment. A caveat to the aforementioned study is that the effects of ACC and ethylene gas are not identical *(136)*. Therefore, we are unable to make conclusions as to why we observed this specific EIN3 binding pattern over a timecourse of ethylene gas treatment.

Disconnect between ethylene transcription steady-state levels and EIN3 DNA-binding

Another striking feature of our study was that in general EIN3 binding, although induced by ethylene over time, did not correlate to changes in steady-state levels of transcription (Fig. 2). The lack of correlation between these datasets was not due to issues in the mRNA-Seq data. Our mRNA-Seq was highly reproducible with at least 21 million reads per sample. Approximately 60% of all genes were expressed in etiolated seedlings (19,500 of 32,678 with RPKM > 1). We were also able to detect ethylene-induced root-specific genes (ACO2) in our expression data, suggesting that the sensitivity of the mRNA-Seq experiments was sufficient. Furthermore, other genes that have been known to be up-regulated by ethylene, including ERF1, were detected

in our data.

Many recent studies examining the protein-DNA binding of transcription factors have found that a large number of target sites do not exhibit detectable changes in levels of transcripts or steady-state levels of transcripts, suggesting a disconnect between protein-DNA binding and actual transcriptional regulation. This is common for master regulators, which most likely bind a small subset of downstream effectors, which then bind specific sets of genes thus leading to a morphological response or phenotype. However, why is there such an excess amount of protein-DNA interactions that may or may not be functional? Macquarrie and colleagues hypothesize that the lack of correlation between protein-DNA interactions and transcriptional regulation may be due to site accessibility, chromosome looping, global chromatin changes, and selective advantage *(137)*.

The site accessibility model posits that transcription factor binding sites in intergenic or repetitive elements helps maintain optimum levels of the transcription factor in the nucleus. In this model, gene expression is regulated by transcription factor concentrations. The binding of transcription factors to these intergenic or repeat elements can serve to decreasing the available amount of transcription factors, thus inhibiting binding at other sites which require cofactors and/or cooperative binding. The chromosome looping model is most applicable for genes that are regulated at a distance as in enhancer-mediated gene regulation. The structure of DNA in terms of looping would enable long-range protein-DNA interactions. Indeed, many cases of enhancer-mediated gene regulation occur in mammalian systems, however the contribution of enhancers to the general transcriptional regulatory network in *Arabidopsis* is unknown. The global chromatin structure model suggests that

transcription factors function to change chromatin state, in addition to gene regulation. In this model, transcription factors bind DNA and may also bind or recruit proteins involved in chromatin modification. There are several examples of general regulators affecting chromatin structure in the literature. Finally, in the selective advantage model, transcription factors target many genes because there is an evolutionary advantage. Transcriptional regulatory networks may have been driven to their current state by a specific selective advantage and are robust, but require a high probability to generate a new network. It is possible that more than one of these models may apply to a specific transcription factor.

Given that we find a large number of DNA binding sites for EIN3 for which we do not observe transcriptional differences, the application of these models becomes relevant. Out of these four hypotheses, the first three can be discounted due to the observation that in these cases the distribution of binding would be more uniform across the genome. We do not observe a random distribution of binding; for the majority of EIN3 binding regions (> 60%), binding occurs in the promoter or 5' UTR of the associated gene. In addition, the distribution of the EIN3 binding sites is concentrated near the TSS.

We can test whether there is a selective advantage to maintain many EIN3 binding sites in the *Arabidopsis* genome. If the selective advantage model is true, all EIN3 targets regardless of their transcriptional regulation by ethylene have the potential to be functional in some context, perhaps in a specific condition, tissue or cell type, or developmental stage. We can test this two ways. The first approach is to evaluate the EIN3-mediated ethylene transcriptional response in different conditions or developmental stages. The EBS:GUS reporter plants *(81)* which contain a

synthetic EIN3-responsive promoter (5 copies of the EIN3 binding site) driving the beta-glucuronidase (GUS) gene, may be used to determine which tissues are important at specific developmental stages, and then these can be further dissected to generate a specific transcriptional profile. However, because plants contain more than 200 different cell types, this may prove difficult. Also, separation of different groups of cell types or tissue types is tedious and may require laser dissection or generation of transgenics for sorting using flow cytometry after protoplasting. Plants are not easily amenable to grow in liquid culture, as hormones auxin and cytokinin may trigger transcriptional programs, yielding artifacts.

Being that a laboratory exploration to dissect whether EIN3 binding sites are functional in other cell types, we can turn to other information available to test whether the selective advantage model is true in our case. Are the EIN3 binding sites functional or have the potential to be functional? We can determine this by sequence conservation. If the selective advantage model were true, we would expect the sequence conservation of both the transcriptionally regulated targets vs. the non-transcriptionally regulated targets to be roughly the same. Current databases (VISTA, http://pipeline.lbl.gov/cgi-bin/gateway2?bg=ara2mask&selector=vista, PHYTOZOME, http://www.phytozome.net/) contain sequence conservation data between species; however conservation scores of sequence across the genome do not exist (e.g. phastcons, http://compgen.bscb.cornell.edu/phast). In some cases, the conservation of between plant species is not sufficient to calculate nucleotide resolution conservation scores. For example, VISTA contains the percent sequence conservation for regions of a genome that has at least 50% conservation with *Arabidopsis thaliana*. We are currently working on generating sequence conservation

scores of EIN3 binding regions in *Arabidopsis lyrata*, Poplar, Moss, Rice, Selaginella, and Sorghum with respect to *Arabidopsis thaliana*. In addition, sequences of 200 Arabidopsis ecotypes will soon be publicly available. We are currently working on determining the amount of SNPs that occur in EIN3 binding sites with respect to the SNP frequency in these 200 genomes.

Another way of addressing the conversation of EIN3 binding sites is to determine whether EIN3 binding sites are associated with genes with intragenomic conserved non-coding sequences (CNSs) *(138)*. Based on analysis of gene pairs in *Arabidopsis thaliana*, almost 15,000 conserved non-coding sequences were found in the Arabidopsis genome. These 15,000 CNSs were associated with 4,337 genes. Of the 1314 EIN3 targets, 110 were in a gene with at least one CNS (Hypergeometric p-value = 10^{-17}). However, the frequency of CNSs in EIN3 target genes (8%) was less than the occurrence of CNSs in the genome (16%, 4337/26751, TAIR6). More analysis on the relative amount of conservation of the target sequences needs to be performed to understand the function, if any, of the non-transcriptionally regulated targets.

If the selective advantage model is pertinent to explain the disconnect between EIN3 binding and ethylene-mediated transcriptional regulation, we may observe other evidence of functionality, e.g. footprints of coregulator binding that differ among sets of EIN3 targets. However, when we clustered EIN3 binding and ethylene-induced transcriptional profiles (kmeans, hierarchical, STEM *(139)*), we did not observe co-regulated modules of genes, corresponding to significant motif signatures. Perhaps this calls for a more thorough motif analysis of the EIN3 binding sites. It is also possible that coregulators involved in biological processes other than that

examined in our study gate the transcriptional activity of the non-ethylene regulated targets. Correlation of interactors of EIN3 with the over-representation of motifs in the EIN3 target sequences can help us identify possible coregulators. We did find that there are a few AP2/EREBP transcription factors that may interact with EIN3. In addition, *de novo* motif analysis revealed that the EIN3 binding regions contained GCC motifs. AP2/EREBP transcription factors, such as ERF1, have been demonstrated to bind the GCC box motif *(56)*. The putative EIN3 and AP2/EREBP protein-protein interactions as well as the presence of GCC motifs represented in the EIN3 target binding sites suggest that EIN3 and AP2/EREBPs may act in a feed forward regulation of the transcriptional ethylene response. However, as both EIN3 binding profiles and the expression of ERF1 was detected at 0.5 hrs of ethylene gas treatment, it is possible that our timecourse resolution was too course to detect the feedforward transcriptional ethylene response.

Thus far, these observations support a selective advantage model, but further work must be performed to confirm the relevance of this model. Permutations in feed-forward network motifs may enable the plasticity of transcriptional regulatory networks. Regulation of complex biological pathways often employ feedforward regulation, transcriptional regulation can occur in a temporal manner, but cessation of target expression is easily attained. Internal or external changes in transcriptional regulation can thus be switched on and off rapidly. Macquarrie et al. mention the prevalence of feed-forward circuitry in yeast transcriptional regulatory networks, a convergent evolution based on widespread utility. Redundancy of EIN3 and ERF1 suggest a flexible transcriptional response network, as discussed later.

If a transcription factor binds numerous places in the genome, it increases the

possibilities of regulation in different conditions, cell types, etc. A recent study examining transcriptional initiation and elongation in real-time using fluorescently labeled RNA *(140)* found that memory of transcription initiation does not occur. Therefore internal and external cues must be called constantly, for example, EIN3 accumulation occurs while binding increases in the presence of exogenous ethylene. Larson and colleagues also found that transcription occurs in bursts, whereas elongation could vary based on cell cycle progression. The rapidity of transcription observed in Larson et al. may explain the prevalence of the feedforward regulatory circuitry of transcriptional networks. Another interesting finding by Larson et al. was that initiation kinetics of transcription was determined by trans-activating factors. Therefore, coregulators or other cooperative factors may be required. Perhaps the diversity of factors required for transcription translates simplicity of a rapid response into complexity. Again, permutations of transcription factors, and an excess number of binding sites can lead to a wide variety of transcriptional profiles, which is later discussed.

Finally, under the selective advantage of widespread protein-DNA binding, we could have the highest probability of generating a new network by mutating a master regulator, which would change the expression of large numbers of genes. Simultaneously, the current transcriptional regulatory network would be buffered. If this was true, master regulators' downstream effectors could be mutated with little phenotypic consequence. The selective advantage model to explain the excess binding events that are not transcriptionally regulated argues for a transcription regulatory network homeostasis. This buffering capacity of the transcriptional regulatory network would be useful as the transcription network might need to change

drastically on rare occasion for survival, but must maintain the ability to respond rapidly, and not allow spontaneous mutations to be lethal. Transcriptional regulatory networks should thus contain hubs that are master regulators, resulting in many connections, a scale-free network that can be altered or under evolution.

Possible combinations of protein-DNA interactions with EIN3 homologs

In order for the transcriptional regulatory network to be robust, redundancy in the transcriptional regulatory network may occur at many different levels. As mentioned previously, under the selective advantage model, duplication and sequence divergence of a transcription factor may yield gene family members that have overlapping but distinct DNA binding. EIN3 has five family members, ETHYLENE INSENSITIVE-LIKE1-5 (EIL1-5). *EIL1* and *EIL2* are capable of rescuing the *ein3* mutant *(49)*, and are thus functionally redundant to EIN3. Studies also show that a null mutation of EIN3 has a milder phenotype than is expected, and the *ein3 eil1* mutant is much more insensitive to ethylene. However, overexpression of EIN3 can induce the ethylene response in the absence of hormone, indicating that EIN3 protein accumulation alone is sufficient to induce the ethylene response.

As discussed in the previous chapter, a target of EIN3, HLS1 contains three homologs, two of which are targeted by EIN3. It is possible that the homologs of EIN3 target the HLHs with more affinity than EIN3, however this has yet to be tested. Although it has been shown that EIN3 and EIL1 differentially regulate growth *(55)*, we did not address the binding of EIL1 to DNA in this study. Future studies will determine the similarities and differences of protein-DNA binding of EIN3 and its homologs.

Dependency of ethylene hormone crosstalk on EIN3 protein-DNA interactions

One of the unanswered questions of this study is how significant is the binding of EIN3 for hormone crosstalk, specifically between ethylene and auxin? Many auxin mutants have been identified in screens for ethylene insensitive mutants *(7, 81)*, however, ethylene mutants do not display an aberrant auxin response phenotype in seedlings. *ein2, ein3, ein3 eil1* mutants exhibit a normal auxin response measured by root length. What does this mean? Auxin biosynthesis, perception, signaling and response do not require the ethylene response, or EIN3 protein-DNA binding. When conducting phenotypic analyses, it is possible that general phenotypes which only require a basal level of auxin are scored (e.g. gravitropism, vasculature patterning, seedling morphology, floral morphology, organ shape and size, plant architecture, shape, size) *(81, 141)*. Ethylene has been shown to sensitize cells to the auxin response, increasing the basal auxin response. The analysis of more refined phenotypes (e.g. root hair length) may be necessary to uncover the significance of EIN3 binding in hormone crosstalk.

As one example of the mechanism of ethylene-auxin hormone crosstalk, we showed that EIN3 targets HLS1, an N-acetyltransferase that regulates the protein levels of ARF2. ARF2 is part of a gene family involved in regulating the auxin response. However, because the auxin response is not dependent on ethylene, as discussed above, it is difficult to determine what role, if any, EIN3 protein-DNA binding has on the auxin response. Previous analysis of two *HLS1* alleles has revealed that mutations in the *HLS1* promoter were sufficient to yield a weak *hookless* phenotype *(116)*. To further determine the significance of EIN3 binding in ethylene-auxin hormone crosstalk, we decided to evaluate whether an auxin molecular

phenotype/signature existed for EIN3 targets. Using a previous microarray that determined the auxin response genes in light grown seedlings *(117)*, we found that of the 2882 auxin response genes (two-sided t-test *p*-value = 0.01), 233 were EIN3 targets. Approximately 18% of EIN3 targets have been implicated in auxin response ($P < 10^{-15}$). Although these microarray studies were performed in light grown seedlings, in comparison to dark grown three-day-old seedlings used in our study, this suggests that ethylene and auxin crosstalk may occur at a gene regulatory network level. Taken together, the genetic and molecular phenotypes suggest that EIN3 does have a role in ethylene-auxin hormone crosstalk, even if ethylene is not required for the basal auxin response. The significance and relevance of EIN3 binding in HLS1-mediated ethylene-auxin crosstalk must be further investigated.

Most studies have found that hormone crosstalk does not occur at the biosynthesis, perception, signaling levels *(77)*. Therefore, two hormones do not truly share the same components in their biosynthesis, perception, signaling pathways. Signal integration is thought to occur at the regulatory gene network level. When conducting a study to find a core growth regulating transcriptional network, Nemhauser and colleagues found that hormones regulate the same processes with nonoverlapping transcriptional responses. It is possible that the publicly available microarray data did not have the spatial or temporal resolution suitable for detecting coregulated modules. This study does not necessarily disclaim the binding of transcription factors functioning in different hormone pathways to the same targets, especially when only a fraction of targets are transcriptionally regulated. Comparing EIN3 targets to those of brassinosteroid transcription factors BES1 and BZR1, we found that EIN3 shares targets with BES1 and BZR1, however each transcription

factor has a unique set of targets. Perhaps coregulator combinations are complex, and each combination of transcription factors yields a specific transcriptional response.

CONCLUSIONS

Although we have made a significant contribution to mapping the protein-DNA interactions of plant hormone crosstalk, studies are required to determine the importance of these interactions in the context of development and evolution. Because of the robustness of plant growth and developmental processes, mechanistic instances of hormone crosstalk will most likely be observed in specific cell types or developmental stages of the plant. The addition of more studies examining the targets of transcription factors important in hormone signaling will help map the possible interactions between hormone pathways and lead to the understanding of the significance of hormone crosstalk.

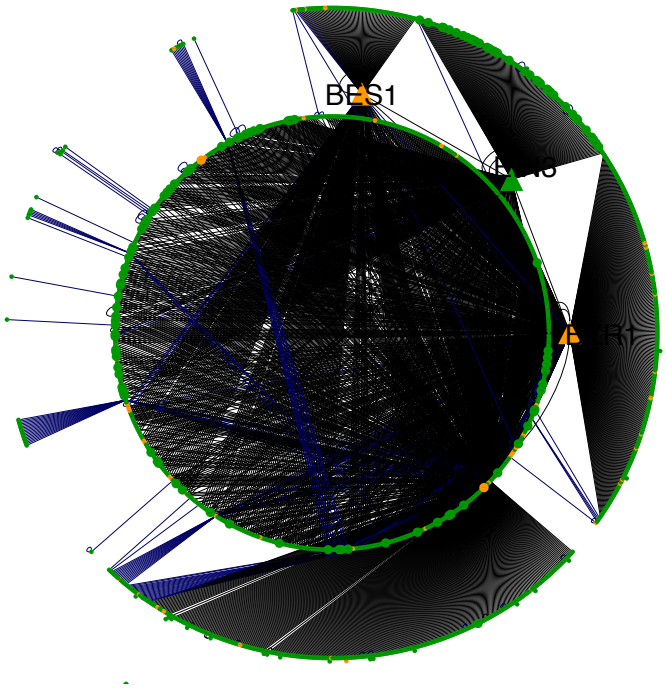


Fig. 1. Ethylene and brassinosteroid protein-DNA interaction network. Black edges represent protein-DNA interactions. Large circular nodes are EIN3 targets. Green nodes represent those annotated as ethylene-related, while orange nodes represent those that are annotated as brassinosteroid related.

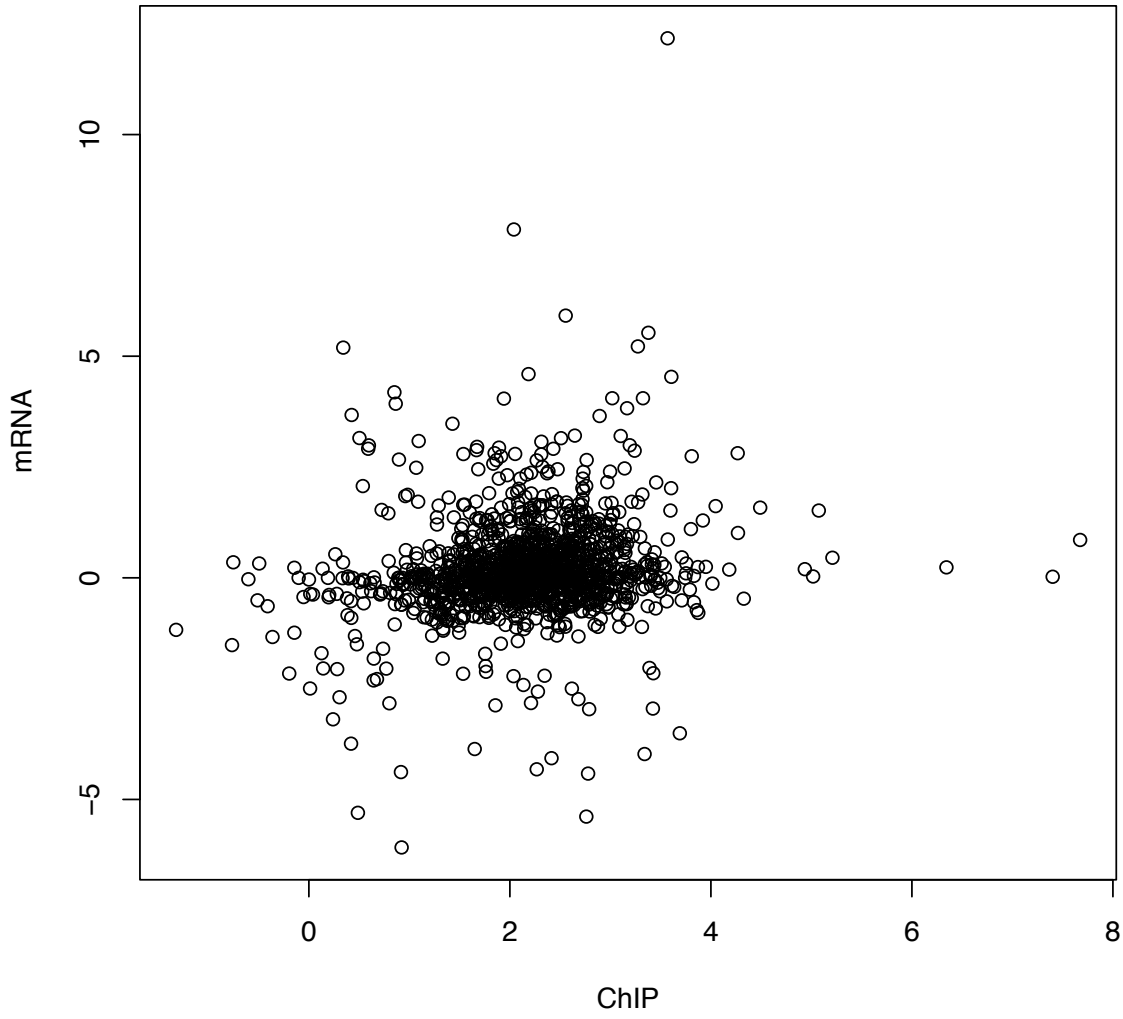


Fig. 2. Correlation between EIN3 binding and transcription of associated genes. Values shown are log2 ratios of reads per kilobasepair per million reads in sample (RPKM) with respect to the 0 hour ethylene treatment control.

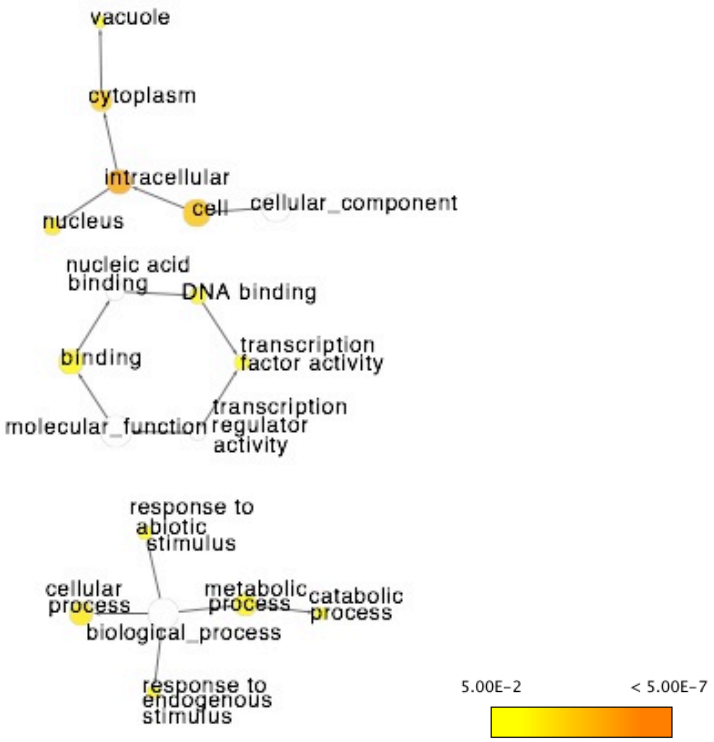


Fig. 3. Over-representation of gene ontology terms for EIN3 interactors. A significant amount of interactors have DNA binding/transcription factor activity annotation.

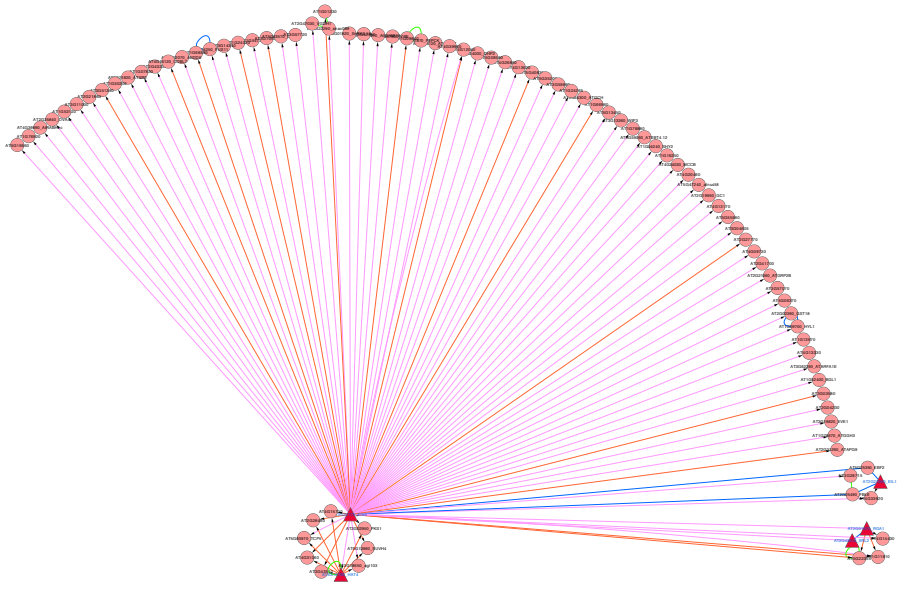


Fig. 4. EIN3 protein interaction network reveals possible coregulators. EIN3 is the triangular-shaped node with the most interactors. EIN3 interacts with a few AP2/EREBP transcription factor family members.

REFERENCES

1. F. Abeles, P. Morgan, M. J. Saltveit, *Ethylene in Plant Biology* (Academic Press, New York, ed. 2).

2. R. Gane, *Nature* (1934).

3. J. Ecker, R. Davis, *Proc Natl Acad Sci U S A* **84**, 5202-5206 (1987).

4. M. Arshad, W. T. Frankenberger, *Ethylene: agricultural sources and applications* (Kluwer Academic/Plenum Publishers, New York, 2002), p. 342.

5. A. B. Bleecker, H. Kende, *Annu Rev Cell Dev Biol* **16**, 1-18 (2000).

6. H. Klee, *Proc Natl Acad Sci USA* **103**, 7537-7538 (2006).

7. P. Guzmán, J. R. Ecker, *Plant Cell* **2**, 513-523 (1990).

8. M. Kendrick, C. Chang, *Curr Opin Plant Biol* **11**, 479-485 (2008).

9. Y. Cho, S. Yoo, *J. Plant Biol.* **52**, 283-288 (2009).

10. J. M. Alonso et al., *Proc Natl Acad Sci USA* **100**, 2992-2997 (2003).

11. P. B. Larsen, C. Chang, *Plant Physiol* **125**, 1061-1073 (2001).

12. M. J. Christians, L. M. Robles, S. M. Zeller, P. B. Larsen, *The Plant Journal* **55**, 467-477 (2008).

13. S. Yang, *Annual Review of Plant Physiology* (1984).

14. K. L. Wang, H. Li, J. R. Ecker, *Plant Cell* **14 Suppl**, S131-51 (2002).

15. K. Woeste, J. Vogel, *Physiologia Plantarum* (1999).

16. K. Skottke, G. Yoon, J. Kieber, *PLoS Genet* (2011).

17. K. L. Wang, H. Yoshida, C. Lurin, J. R. Ecker, *Nature* **428**, 945-950 (2004).

18. A. B. Bleecker, M. A. Estelle C. Somerville, H. Kende, *Science* **241**, 1086-1089 (1988).

19. C. Chang, S. F. Kwok, A. B. Bleecker, E. M. Meyerowitz, *Science* **262**, 539-544 (1993).

20. J. Hua, C. Chang, Q. Sun, E. M. Meyerowitz, *Science* **269**, 1712-1714 (1995).

21. J. Hua, E. M. Meyerowitz, *Cell* **94**, 261-271 (1998).

22. H. Sakai et al., *Proc Natl Acad Sci USA* **95**, 5812-5817 (1998).

23. F. I. Rodríguez et al., *Science* **283**, 996-998 (1999).

24. G. E. Schaller, J. J. Kieber, *The Arabidopsis Book* **12**, 1 (2002).

25. T. Hirayama et al., *Cell* **97**, 383-393 (1999).

26. B. M. Binder, F. I. Rodríguez, A. B. Bleecker, *J Biol Chem* **285**, 37263-37270 (2010).

27. R. C. O'Malley et al., *Plant J* **41**, 651-659 (2005).

28. B. M. Binder et al., *Plant Physiol* **136**, 2913-2920 (2004).

29. P. Moussatche, H. J. Klee, *J Biol Chem* **279**, 48734-48741 (2004).

30. Y. Chen et al., *PLoS ONE* **5**, e8640 (2010).

31. Z. Gao et al., *Journal of Biological Chemistry* **283**, 23801-23810 (2008).

32. X. Qu, G. E. Schaller, *Plant Physiol* **136**, 2961-2970 (2004).

33. J. Voet-Van-Vormizeele, G. Groth, *Molecular Plant* **1**, 380-387 (2008).

34. W. Wang, A. E. Hall, R. O'Malley, A. B. Bleecker, *Proc Natl Acad Sci USA* **100**, 352-357 (2003).

35. H. Kim et al., *Plant Physiol* (2011), doi:10.1104/pp.110.170621.

36. J. S. Resnick, C. Wen, J. A. Shockey, C. Chang, *Proc Natl Acad Sci USA* **103**, 7917-7922 (2006).

37. X. Zhou, Q. Liu, F. Xie, C. Wen, *Plant Physiol* **145**, 75-86 (2007).

38. C. Dong, M. Rivarola, J. S. Resnick, B. D. Maggin, C. Chang, *Plant J* **53**, 275-286 (2008).

39. J. S. Resnick, M. Rivarola, C. Chang, *The Plant Journal*, ???-??? (2008).

40. M. Rivarola, C. A. McClellan, J. S. Resnick, C. Chang, *Plant Physiol* **150**, 547-551 (2009).

41. C. Dong et al., *Journal of Biological Chemistry* **285**, 40706-40713 (2010).

42. J. J. Kieber, M. Rothenberg, G. Roman, K. A. Feldmann, J. R. Ecker, *Cell* **72**, 427-441 (1993).

43. K. L. Clark, P. B. Larsen, X. Wang, C. Chang, *Proc Natl Acad Sci USA* **95**, 5401-

5406 (1998).

44. Z. Gao et al., *J Biol Chem* **278**, 34725-34732 (2003).

45. Q. Hong, J. Ecker.

46. M. M. Bisson, A. Bleckmann, S. Allekotte, G. Groth, *Biochemical Journal* **424**, 1-6 (2010).

47. J. M. Alonso, T. Hirayama, G. Roman, S. Nourizadeh, J. R. Ecker, *Science* **284**, 2148-2152 (1999).

48. H. Qiao, K. N. Chang, J. Yazaki, J. R. Ecker, *Genes Dev* **23**, 512-521 (2009).

49. Q. Chao et al., *Cell* **89**, 1133-1144 (1997).

50. H. Guo, J. R. Ecker, *Cell* **115**, 667-677 (2003).

51. T. Potuschak et al., *Cell* **115**, 679-689 (2003).

52. J. M. Gagne et al., *Proc Natl Acad Sci USA* **101**, 6803-6808 (2004).

53. M. Konishi, S. Yanagisawa, *The Plant Journal* **55**, 821-831 (2008).

54. G. Olmedo et al., *Proc Natl Acad Sci USA* **103**, 13286-13293 (2006).

55. F. An et al., *Plant Cell* **22**, 2384-2401 (2010).

56. R. Solano, A. Stepanova, Q. Chao, J. R. Ecker, *Genes Dev* **12**, 3703-3714 (1998).

57. A. N. Stepanova, J. R. Ecker, *Curr Opin Plant Biol* **3**, 353-360 (2000).

58. B. Noh et al., *Plant Cell* **16**, 2601-2613 (2004).

59. A. de Paepe, L. de Grauwe, S. Bertrand, J. Smalle, D. van der Straeten, *J Exp Bot* **56**, 2409-2420 (2005).

60. L. M. Robles, J. S. Wampole, M. J. Christians, P. B. Larsen, *J Exp Bot* **58**, 2627-2639 (2007).

61. S. Yoo, Y. Cho, G. Tena, Y. Xiong, J. Sheen, *Nature* **451**, 789-795 (2008).

62. J. Xu et al., *J Biol Chem* **283**, 26996-27006 (2008).

63. H. Guo, J. R. Ecker, *Curr Opin Plant Biol* **7**, 40-49 (2004).

64. B. M. Binder, L. A. Mortimore, A. N. Stepanova, J. R. Ecker, A. B. Bleecker, *Plant Physiol* **136**, 2921-2927 (2004).

65. B. M. Binder et al., *Plant Cell* **19**, 509-523 (2007).

66. R. Solano, J. R. Ecker, *Curr Opin Plant Biol* **1**, 393-398 (1998).

67. A. N. Stepanova, thesis, University of Pennsylvania (2001).

68. O. Lorenzo, R. Piqueras, J. J. Sánchez-Serrano, R. Solano, *Plant Cell* **15**, 165-178 (2003).

69. G. Y. Zhong, G. V. Zhong, J. K. Burns, *Plant Mol Biol* **53**, 117-131 (2003).

70. A. de Paepe, M. Vuylsteke, P. van Hummelen, M. Zabeau, D. van der Straeten, *Plant J* **39**, 537-559 (2004).

71. H. Chen et al., *Plant Cell* **21**, 2527-2540 (2009).

72. S. Zhong et al., *Proc Natl Acad Sci USA* **106**, 21431-21436 (2009).

73. F. Boutrot et al., *Proc Natl Acad Sci* **107**, 14502-14507 (2010).

74. K. Yamasaki et al., *J Mol Biol* **348**, 253-264 (2005).

75. T. T. Paull, D. Cortez, B. Bowers, S. J. Elledge, M. Gellert, *Proc Natl Acad Sci USA* **98**, 6086-6091 (2001).

76. H. Li et al., *Proteomics* **9**, 1646-1661 (2009).

77. Y. Jaillais, J. Chory, *Nat Struct Mol Biol* **17**, 642-645 (2010).

78. J. L. Nemhauser, F. Hong, J. Chory, *Cell* **126**, 467-475 (2006).

79. T. Sato, A. Theologis, *… of the National Academy of Sciences …* (1989).

80. H. Li, P. Johnson, A. Stepanova, J. M. Alonso, J. R. Ecker, *Dev Cell* **7**, 193-204 (2004).

81. A. N. Stepanova, J. Yun, A. V. Likhacheva, J. M. Alonso, *Plant Cell* **19**, 2169-2185 (2007).

82. R. Swarup et al., *Plant Cell* **19**, 2186-2196 (2007).

83. K. Růzicka et al., *Plant Cell* **19**, 2197-2212 (2007).

84. A. N. Stepanova et al., *Cell* **133**, 177-191 (2008).

85. A. Zarei et al., *Plant Mol Biol* **75**, 321-331 (2011).

86. C. Zhou, L. Zhang, J. Duan, B. Miki, K. Wu, *Plant Cell* **17**, 1196-1204 (2005).

87. J. Yazaki, A. Kim, M. Galli, J. R. Ecker.

88. V. Orlando, R. Paro, *Cell* **75**, 1187-1198 (1993).

89. V. Orlando, H. Strutt, R. Paro, *METHODS* **11**, 205-214 (1997).

90. M. H. Kuo, C. D. Allis, *METHODS* **19**, 425-433 (1999).

91. B. Ren et al., *Science* **290**, 2306-2309 (2000).

92. D. S. Johnson, A. Mortazavi, R. M. Myers, B. Wold, *Science* **316**, 1497-1502 (2007).

93. A. Valouev et al., *Nat Methods* **5**, 829-834 (2008).

94. T. S. Mikkelsen et al., *Nature* **448**, 553-560 (2007).

95. G. Robertson et al., *Nat Methods* **4**, 651-657 (2007).

96. A. Barski, K. Zhao, *J. Cell. Biochem.* **107**, 11-18 (2009).

97. Y. Sun et al., *Dev Cell* **19**, 765-777 (2010).

98. X. Yu et al., *The Plant Journal* **65**, 634-646 (2011).

99. C. M. Winter et al., *Dev Cell* **20**, 430-443 (2011).

100. W. Deng et al., *Proc Natl Acad Sci U S A* **108**, 6680-6685 (2011).

101. J. Lee et al., *Plant Cell* **19**, 731-749 (2007).

102. K. Kaufmann et al., *Science* **328**, 85-89 (2010).

103. K. Kaufmann et al., *PLoS Biol* **7**, e90 (2009).

104. K. Kaufmann et al., *Nature protocols* **5**, 457-472 (2010).

105. M. Konishi, S. Yanagisawa, *Plant J* **5**, 821-831 (2008).

106. B. M. Binder, L. A. Mortimore, A. N. Stepanova, J. R. Ecker, A. B. Bleecker, *Plant Physiol* **136**, 2921-2927 (2004).

107. S. Maere, *Bioinformatics* **21**, 3448-3449 (2005).

108. K. L. Macquarrie, A. P. Fong, R. H. Morse, S. J. Tapscott, *Trends Genet* **27**, 141-148 (2011).

109. Y. Chen et al., *J Biol Chem* **282**, 24752-24758 (2007).

110. Z. Peng et al., *Nucleic Acids Res*, 8 (2008).

111. J. M. Alonso et al., *Science* **301**, 653-657 (2003).

112. Arabdopsis Interactome Mapping Consortium. In review.

113. A. Yilmaz et al., *Nucleic Acids Res* **39**, D1118-D1122 (2010).

114. Y. Assenov, F. Ramirez, S. E. Schelhorn, T. Lengauer, M. Albrecht, *Bioinformatics* **24**, 282-284 (2008).

115. M. Ohto, S. Hayashi, S. Sawa, A. Hashimoto-Ohta, K. Nakamura, *Plant Cell Physiol* **47**, 1603-1611 (2006).

116. A. Lehman, R. Black, J. R. Ecker, *Cell* **85**, 183-194 (1996).

117. J. L. Nemhauser, T. C. Mockler, J. Chory, *PLoS Biol* **2**, E258 (2004).

118. G. Roman, B. Lubarsky, J. J. Kieber, M. Rothenberg, J. R. Ecker, *Genetics* **139**, 1393-1409 (1995).

119. H. Leyser, C. Lincoln, C. Timpte, D. Lammer, *Nature* (1993).

120. R. A. Sessions, P. C. Zambryski, *Development* **121**, 1519-1532 (1995).

121. Z. Lippman, A. Gendrel, V. Colot, *Nat Methods* (2005).

122. O. Mathieu et al., *Plant Cell* **15**, 2929-2939 (2003).

123. P. V. Kharchenko, M. Y. Tolstorukov, P. J. Park, *Nat Biotechnol* **26**, 1351-1359 (2008).

124. Y. Zhang et al., *Genome Biol* **9**, R137 (2008).

125. J. Rozowsky et al., *Nat Biotechnol* **27**, 66-75 (2009).

126. R. Lister et al., *Cell* **133**, 523-536 (2008).

127. G. Z. Hertz, G. D. Stormo, *Bioinformatics* **15**, 563-577 (1999).

128. S. Kosugi, Y. Ohashi, *Nucleic Acids Res* **28**, 960-967 (2000).

129. P. Shannon et al., *Genome Res* **13**, 2498-2504 (2003).

130. S. Mahony, A. Golden, T. J. Smith, P. V. Benos, *Bioinformatics* **21 Suppl 1**, i283-91 (2005).

131. S. Mahony, P. V. Benos, *Nucleic Acids Res* **35**, W253-W258 (2007).

132. F. Thibaud-Nissen et al., *Plant J* **47**, 152-162 (2006).

133. L. Ni et al., *Genes Dev* **23**, 1351-1363 (2009).

134. H. Hiroi, *Molecular Endocrinology* **18**, 791-806 (2004).

135. R. P. Zinzen, C. Girardot, J. Gagneur, M. Braun, E. E. M. Furlong, *Nature* **461**, 65-70 (2009).

136. A. Tsuchisaka et al., *Genetics* **183**, 979-1003 (2009).

137. K. L. Macquarrie, A. P. Fong, R. H. Morse, S. J. Tapscott, *Trends Genet* **27**, 141-148 (2011).

138. B. C. Thomas, L. Rapaka, E. Lyons, B. Pedersen, M. Freeling, *Proc Natl Acad Sci USA* **104**, 3348-3353 (2007).

139. Z. Bar-Joseph, G. Gerber, I. Simon, D. K. Gifford, T. S. Jaakkola, *Proc Natl Acad Sci USA* **100**, 10146-10151 (2003).

140. D. R. Larson, D. Zenklusen, B. Wu, J. A. Chao, R. H. Singer, *Science* **332**, 475-478 (2011).

141. A. N. Stepanova, *Personal communication* (2011).