



Agrobacterium VirE2 Protein Modulates Plant Gene Expression and Mediates Transformation From Its Location Outside the Nucleus

Rachelle A. Lapham^{1†}, Lan-Ying Lee¹, Eder Xhako^{1†}, Esteban Gañán Gómez^{1,2}, V. M. Nivya^{1,3} and Stanton B. Gelvin^{1*}

¹ Department of Biological Sciences, Purdue University, West Lafayette, IN, United States, ² Departamento de Ciencias Biológicas, Universidad EAFIT, Medellín, Colombia, ³ Department of Plant Science, School of Biological Science, Central University of Kerala, Kasaragod, India

OPEN ACCESS

Edited by:

Manuel Martinez,
Polytechnic University of Madrid,
Spain

Reviewed by:

Satoko Nonaka,
University of Tsukuba, Japan
Paul Van Heusden,
Leiden University, Netherlands

*Correspondence:

Stanton B. Gelvin
gelvin@purdue.edu

† Present address:

Rachelle A. Lapham,
Calyxt, Inc., Roseville, MN,
United States
Eder Xhako,
Program in Genetics and Genomics,
Department of Molecular and Human
Genetics, Baylor College of Medicine,
Houston, TX, United States Esteban

Specialty section:

This article was submitted to
Plant Pathogen Interactions,
a section of the journal
Frontiers in Plant Science

Received: 22 March 2021

Accepted: 10 May 2021

Published: 04 June 2021

Citation:

Lapham RA, Lee L-Y, Xhako E,
Gómez EG, Nivya VM and Gelvin SB
(2021) Agrobacterium VirE2 Protein
Modulates Plant Gene Expression
and Mediates Transformation From Its
Location Outside the Nucleus.
Front. Plant Sci. 12:684192.
doi: 10.3389/fpls.2021.684192

Agrobacterium effector protein VirE2 is important for plant transformation. VirE2 likely coats transferred DNA (T-DNA) in the plant cell and protects it from degradation. VirE2 localizes to the plant cytoplasm and interacts with several host proteins. Plant-expressed VirE2 can complement a *virE2* mutant *Agrobacterium* strain to support transformation. We investigated whether VirE2 could facilitate transformation from a nuclear location by affixing to it a strong nuclear localization signal (NLS) sequence. Only cytoplasmic-, but not nuclear-localized, VirE2 could stimulate transformation. To investigate the ways VirE2 supports transformation, we generated transgenic *Arabidopsis* plants containing a *virE2* gene under the control of an inducible promoter and performed RNA-seq and proteomic analyses before and after induction. Some differentially expressed plant genes were previously known to facilitate transformation. Knockout mutant lines of some other VirE2 differentially expressed genes showed altered transformation phenotypes. Levels of some proteins known to be important for transformation increased in response to VirE2 induction, but prior to or without induction of their corresponding mRNAs. Overexpression of some other genes whose proteins increased after VirE2 induction resulted in increased transformation susceptibility. We conclude that cytoplasmically localized VirE2 modulates both plant RNA and protein levels to facilitate transformation.

Keywords: *Agrobacterium*, *Arabidopsis*, plant transformation, VirE2, protein subcellular localization, transcriptome, proteome

INTRODUCTION

Agrobacterium tumefaciens, the causative agent of crown gall disease, transfers virulence effector proteins to infected host plants to facilitate the transfer of T-(transfer) DNA into and trafficking through plant cells. Once in the nucleus, transferred DNA (T-DNA) uses the host's machinery to express transgenes, and may integrate into the host genome. Scientists have used this process to insert beneficial genes into plants by replacing native T-DNA genes with other genes of interest, making *Agrobacterium*-mediated transformation the preferred method for plant genetic

engineering (Gelvin, 2003, 2012; Pitzschke and Hirt, 2010; Lacroix and Citovsky, 2013, 2019; Hiei et al., 2014; Nester, 2015; Van Eck, 2018).

VirE2 is one of the *A. tumefaciens* effector proteins that is important for plant transformation (Gelvin, 2003, 2012). *A. tumefaciens* mutant strains lacking a functional *virE2* gene are severely attenuated in virulence (Stachel and Nester, 1986), and integrated T-DNAs delivered from such strains often exhibit large deletions (Rossi et al., 1996). VirE2 can coat single-stranded DNA molecules *in vitro* (Gietl et al., 1987; Christie et al., 1988; Citovsky et al., 1988, 1989; Das, 1988; Sen et al., 1989) and has been proposed to coat single-stranded T-DNA molecules (T-strands) and protect them from nucleases as they traffic through the plant cell (Gietl et al., 1987; Citovsky et al., 1988; Tinland et al., 1994; Yusibov et al., 1994). Expression of VirE2 in the plant can complement a *virE2* mutant *Agrobacterium* strain to full virulence (Citovsky et al., 1992; Simone et al., 2001), suggesting that one of VirE2's functions in transformation occurs in the plant and involves the maintenance of T-DNA integrity (Gietl et al., 1987; Citovsky et al., 1988).

VirE2 has been proposed to assist with nuclear import of T-strands through its interaction with the transcription factor VIP1 (VirE2-interacting protein 1; Tzfira et al., 2001). This observation led to the model that T-DNA-bound VirE2 binds VIP1 and uses VIP1 nuclear localization to deliver T-DNA into the nucleus (the "Trojan Horse" model; Djamei et al., 2007). However, conflicting reports of VirE2 subcellular localization exist in the literature (Citovsky et al., 1992, 1994, 2004; Tzfira and Citovsky, 2001; Tzfira et al., 2001; Li et al., 2005; Bhattacharjee et al., 2008; Grange et al., 2008; Lee et al., 2008; Shi et al., 2014; Lapham et al., 2018). In contrast to the Trojan Horse model, our laboratory showed that VirE2 holds at least a portion of the VIP1 pool outside the nucleus (Shi et al., 2014), and that VIP1 and its homologs are not required for *Agrobacterium*-mediated transformation (Shi et al., 2014; Lapham et al., 2018).

In addition to its proposed structural role in T-strand binding, we investigated other possible functions of VirE2 in transformation. VirE2 interacts with numerous plant proteins (Lee et al., 2008, 2012) including the transcription factors VIP1 and VIP2 (Tzfira et al., 2001; Anand et al., 2007; Pitzschke et al., 2009). We hypothesized that these interactions could lead to changes in plant gene expression, perhaps facilitating transformation. To determine which subcellular site of localization is required for VirE2 to facilitate transformation, we generated plants expressing cytoplasmic localized VirE2-Venus or nuclear localized VirE2-Venus-NLS under the control of a β -estradiol inducible promoter (Zuo et al., 2000). Following induction, these plants were assayed for transformation using a *virE2* mutant *Agrobacterium* strain. Only cytoplasmic localized VirE2 could support transformation, indicating that VirE2's major function in transformation occurs in the cytoplasm. We also performed RNA-seq and proteomic analyses on transgenic *Arabidopsis thaliana* roots before and after VirE2 expression. Genes previously shown to be important for transformation were differentially expressed in the presence of VirE2, and proteins known to be important for transformation were more prevalent after VirE2 induction, possibly facilitating transformation.

Knockout mutant lines of some of the differentially expressed genes (DEGs) exhibited altered transformation phenotypes. Transgenic plants overexpressing cDNAs encoding some of the proteins shown to be more prevalent in the presence of VirE2 had enhanced transformation susceptibility. Taken together, our results suggest that VirE2 alters expression of specific plant genes and proteins to facilitate transformation, and that VirE2's major role in transformation occurs from its position in the plant cytoplasm.

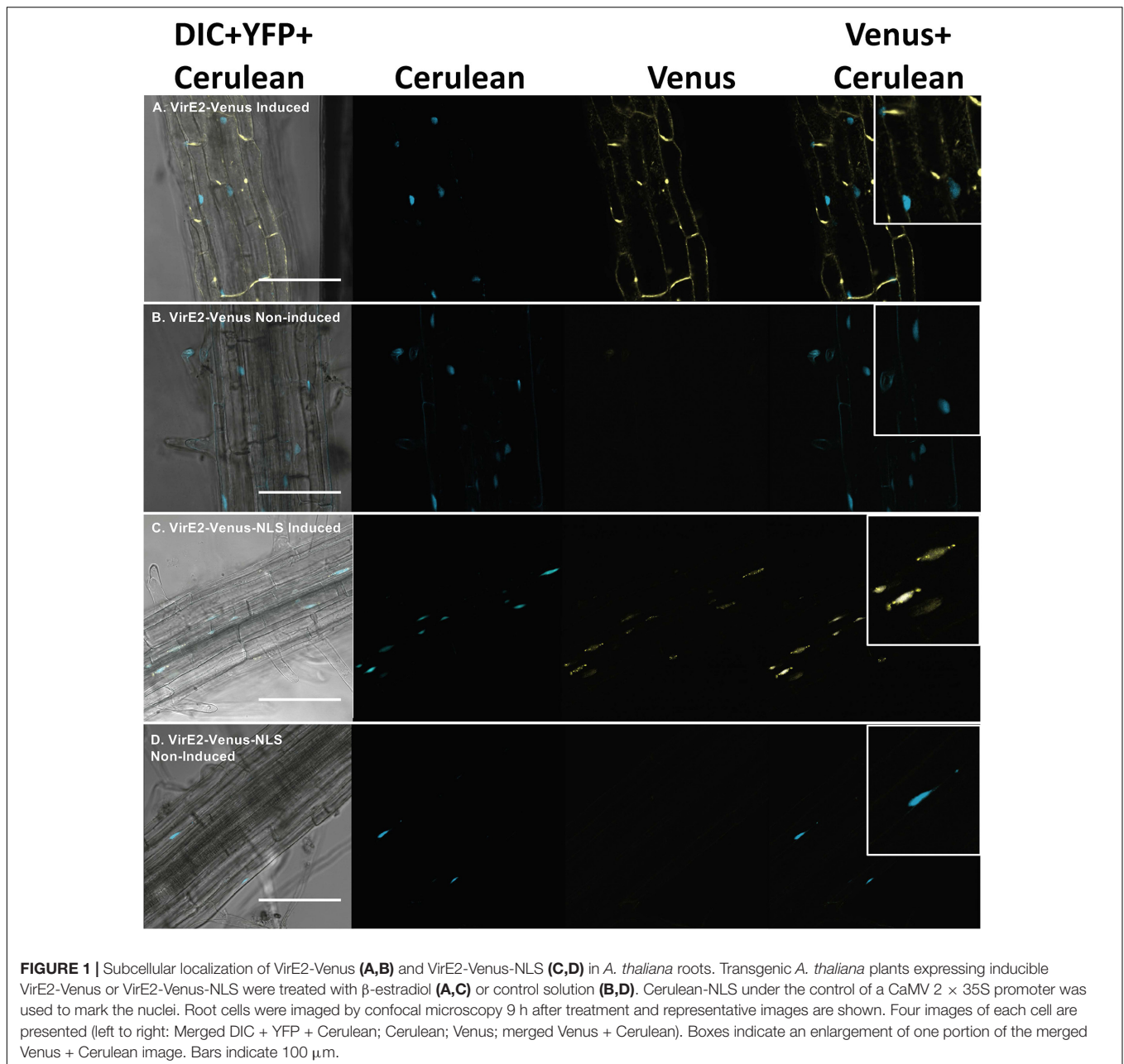
RESULTS

Cytoplasmic but Not Nuclear Localized VirE2 Can Support Transformation

To determine the subcellular localization of VirE2 that is required to facilitate transformation, we first constructed plasmids to express the recombinant proteins VirE2-Venus or VirE2-Venus-NLS [containing a nuclear localization signal (NLS)] constitutively. Tobacco BY-2 protoplasts were individually co-transfected with DNA from each of these constructs and a plasmid containing a red fluorescence protein (RFP) nuclear marker. The protoplasts were imaged 16 h later using confocal microscopy (**Supplementary Figure 1**). Consistent with data from previous studies (Bhattacharjee et al., 2008; Grange et al., 2008; Lee et al., 2008; Li et al., 2014, 2020; Shi et al., 2014; Li and Pan, 2017; Lapham et al., 2018; Roushan et al., 2018), VirE2-Venus localized to the cytoplasm (**Supplementary Figure 1A**); however, VirE2-Venus-NLS localized to the nucleus (**Supplementary Figure 1B**). Although we have not precisely identified where in the cytoplasm VirE2-Venus localizes, it does not localize to the nucleus, and for convenience we shall hereafter refer to the subcellular localization of VirE2-Venus as "cytoplasmic."

We generated transgenic *A. thaliana* plant lines expressing either VirE2-Venus or VirE2-Venus-NLS under the control of a β -estradiol inducible promoter (Zuo et al., 2000), and a Cerulean-NLS nuclear marker under the control of a constitutive Cauliflower Mosaic Virus (CaMV) double 35S promoter. After incubating the plants in either control (non-induced) or β -estradiol (induced) solution for 9 h, we imaged the roots using confocal microscopy. Only induced, but not non-induced, roots showed a yellow fluorescence signal (**Figures 1A,C**), whereas the Cerulean marked nuclei were evident in both non-induced and induced roots (**Figures 1A–D**). VirE2-Venus localized outside of the nucleus and throughout the cytoplasm (**Figure 1A**), whereas VirE2-Venus-NLS co-localized with the Cerulean nuclear marker (**Figure 1C**) in transgenic *Arabidopsis* roots.

We performed transient *Agrobacterium*-mediated transformation assays on wild-type (Col-0), and three independent lines each of inducible VirE2-Venus, and inducible VirE2-Venus-NLS transgenic plants. These lines were chosen based upon equivalent levels of expression of the fluorescently tagged VirE2 protein. Plant roots were treated with either control or β -estradiol solution for 24 h before cutting the roots into small segments and infecting them with a *virE2* mutant



Agrobacterium strain containing the T-DNA binary vector pBISN2 or a *virE2*⁺ control strain containing pBISN1. The T-DNAs of pBISN1 and pBISN2 are identical and contain a plant-active *gusA*-intron gene (Narasimhulu et al., 1996). A low level of transformation was observed in all non-induced samples infected with the *virE2* mutant *Agrobacterium* strain (**Figure 2A**). Such low-level *virE2*-independent transformation has been observed previously (Stachel and Nester, 1986; Rossi et al., 1996; Dombek and Ream, 1997). Induction of only transgenic plants encoding cytoplasmic-localized VirE2-Venus, but not nuclear-localized VirE2-Venus-NLS, increased transient transformation efficiency compared to that of non-induced levels. The inability of nuclear-localized VirE2-Venus-NLS

to complement the *virE2* mutant strain to full virulence was not due to a toxic effect of the protein because both inducible VirE2-Venus and inducible VirE2-Venus-NLS plants showed comparable transformation rates when infected with a *virE2*⁺*Agrobacterium* strain (**Figure 2B**). In addition, the similar transformation efficiencies of induced VirE2-Venus and VirE2-Venus-NLS plants by a *virE2*⁺*Agrobacterium* strain indicates that nuclear-localized VirE2 does not prevent T-strand uncoating in the nucleus, thus preventing *gusA* transgene expression. Thus, these results indicate that in order for VirE2 to complement the transformation deficiency of a *virE2* mutant *Agrobacterium* strain, it must be localized in the cytoplasm.

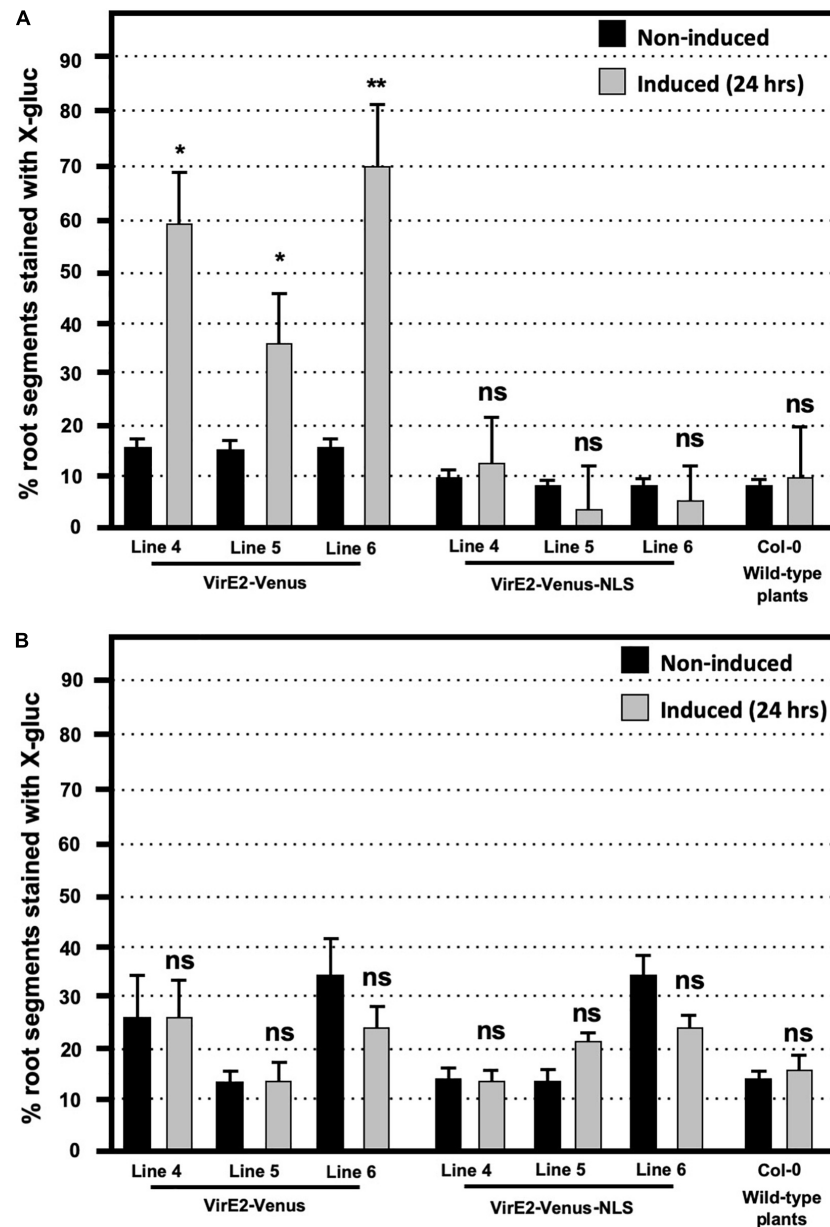


FIGURE 2 | Transformation susceptibility of *Arabidopsis* wild-type (Col-0) and β -estradiol inducible transgenic VirE2-Venus and VirE2-Venus-NLS plants.

Agrobacterium-mediated transient transformation assays were conducted on roots of three transgenic lines of inducible VirE2-Venus, three transgenic lines of inducible VirE2-Venus-NLS, and wild-type Col-0 plants. Following treatment for 24 h with β -estradiol or control solutions, root segments were inoculated with (A) 10^8 cfu/mL of the *virE2* mutant strain *A. tumefaciens* At1879 containing pBISN2 or (B) 10^5 cfu/mL of the wild-type *VirE2* strain EHA105:pBISN1 (At1529). Root segments were stained with X-gluc 6 days after infection. Bars represent an average of three biological replicates (each replicate containing >60 root segments) + SE. ANOVA test * p -value < 0.05, ** p -value < 0.01, ns, not significant.

Cytoplasmic-Localized VirE2 Alters Expression of Numerous *Arabidopsis* Genes, Including Those Involved in Defense Response and Transformation Susceptibility

We generated multiple transgenic *A. thaliana* lines expressing untagged *VirE2* under the control of a β -estradiol inducible

promoter and tested them for *VirE2* induction by RT-PCR. Root tissue pooled from ~30 plants of one representative line was harvested after treating either with β -estradiol or control (non-induced) solution for 3 or 12 h. Both the control and β -estradiol solutions contained the avirulent strain *A. tumefaciens* A136 that lacks a Ti-plasmid (Sciaky et al., 1978) at a concentration of 10^8 cfu/mL. The inclusion of this bacterial strain was done to mimic more closely natural infection conditions because

a plant cell will only be exposed to VirE2 in the presence of *Agrobacterium*. RNA was extracted from each sample and induced expression of *VirE2* was confirmed using RT-PCR. VirE2 transcripts were detectable within 1 h of induction (**Supplementary Figure 2A**). RT-qPCR was performed on samples collected from 3 to 12 h after induction (**Supplementary Figure 2B**) before submitting the samples for RNA-seq analysis. This analysis was initially performed on one biological replicate composed of roots from ~30 plants as a pilot study to identify potential target genes to test for transformation phenotypes. DEG were determined using Cufflinks (Trapnell et al., 2012). For this pilot study and considering all time points, a total of 443 *A. thaliana* genes (~1.5% of the annotated protein coding genome) were differentially expressed in VirE2-induced versus non-induced samples (**Supplementary Data Sheet 1**).

We later conducted RNA-seq analysis on two additional biological replicates, each composed of roots from ~30 plants, of the same inducible *VirE2* line. Using more stringent criteria than used in our pilot study, we identified in total 145 unique up-regulated genes and 25 unique down-regulated genes in induced versus non-induced samples by at least two computational methods with an adjusted *P*-value cut-off of 0.1 across all analyses (**Supplementary Data Sheet 2**). Of these unique 170 DEGs, 61 were identified in the pilot study (**Table 1** and **Supplementary Data Sheet 2**, yellow highlighted rows). DEGs identified in both studies were displayed according to their annotated Gene Ontology (GO) biological process (**Figure 3**; Ashburner et al., 2000). Some genes which showed significant changes in expression were tested using RT-qPCR to validate the RNA-seq results. All genes tested by RT-qPCR showed changes in expression consistent with the RNA-seq data (**Supplementary Data Sheet 1** and **Supplementary Figure 3**). Genes involved in response to stress (both biotic and abiotic), regulation of gene expression, biological regulation, and various other developmental, biosynthetic, and metabolic processes were differentially expressed in VirE2-induced plants (**Figure 3A**). A subset of genes differentially expressed following VirE2 induction are involved in defense responses, particularly those involved in responding to bacteria (**Figure 3B**).

A GO enrichment analysis was performed to determine which categories of genes were over-represented 2-fold or more in the RNA-seq dataset for those 61 DEGs common to both RNA-seq experiments (**Figure 4** and **Table 1**). Genes involved in cellular response to hypoxia, abiotic and chemical stimuli, and stress were enriched. Some stress associated genes, such as those encoding protein phosphatase 2C (down-regulated) and *HEAT SHOCK PROTEIN 90.1* (up-regulated), have previously been shown to be important for transformation (Tao et al., 2004; Park et al., 2014; **Supplementary Data Sheets 1, 2**). These VirE2-induced changes may facilitate transformation.

It is possible that the differential expression of *Arabidopsis* genes following VirE2 induction is merely a stress response to overexpression of a protein in the plant cytoplasm, as indicated by induction of the *HSP90.1* gene. To control for this possibility, we generated *Arabidopsis* lines that inducibly overexpress *VIP1* (Lapham et al., 2018). *VIP1* encodes a protein that localizes both to the nucleus and the cytoplasm, depending upon the time

after osmotic or thigmostimulation (Tsugama et al., 2012, 2014, 2016). Unlike VirE2, *VIP1* is not important for *Agrobacterium*-mediated transformation (Shi et al., 2014; Lapham et al., 2018). We conducted RNA-seq analysis of RNA from *Arabidopsis* roots extracted at various times after *VIP1* induction. **Supplementary Data Sheet 2** compares VirE2-DEGs with *VIP1*-DEGs. Under the same conditions (in the presence of *Agrobacterium*, and at the same time point), only two DEG (At3G13437 and At4G26200) overlap between the VirE2 and *VIP1* overexpression analyses, and neither of these encode stress-response or heat shock/chaperonin proteins. Thus, overexpression of VirE2 elicits a specific DEG response that differs from that elicited by overexpression of another protein.

Arabidopsis Lines Harboring Mutations in Some Genes Differentially Expressed by VirE2 Exhibit Altered Transformation Phenotypes

Transferred DNA insertion mutant lines of a subset of the VirE2 DEGs identified in our pilot RNA-seq study were tested for transformation susceptibility (**Table 2**). Transformation results for mutants of VirE2 up-regulated and down-regulated genes are shown in **Supplementary Figures 4, 5**, respectively, and are summarized in **Table 2**. If a mutant showed no statistically significant difference in transformation efficiency at any of the tested bacterial concentrations, the results are reported as “No change” in **Table 2**. However, some of these mutations may still have a minor impact on transformation.

The *atpsk3*, *tst18*, and *miR163* mutant lines (**Table 2** and **Supplementary Figures 4B,C,G**) showed decreased transformation compared to that of wild-type plants. All three genes are up-regulated in the presence of VirE2 and may therefore facilitate transformation. The *pr5* mutant showed an increase in transient transformation (**Table 2** and **Supplementary Figure 4D**). *PR5* is up-regulated in the presence of VirE2, and because of its role in defense response and effector-triggered immunity (ETI) (Wu et al., 2014) one would predict that the *pr5* mutant would be more susceptible to *Agrobacterium*-mediated infection. At least for transient transformation, this prediction is consistent with our results (**Supplementary Figure 4D**).

Several of the mutants for genes down-regulated in the presence of VirE2 showed increased transient or stable transformation efficiency compared to that of wild-type plants (**Supplementary Figures 5B–F**). These genes may act to inhibit transformation, and their VirE2-dependent down-regulation may facilitate transformation, as reflected by the increased susceptibility of their respective knockout mutant lines to *Agrobacterium*-mediated transformation. A *PROTEIN PHOSPHATASE 2C* (**Supplementary Figure 5F**) was previously identified as a transformation inhibitor (Tao et al., 2004). Conversely, the *exl1*, *oep6*, and *rld17* mutants showed decreased transformation (**Table 2** and **Supplementary Figures 5A,D,E**) even though they are down-regulated in the presence of VirE2. These genes may be important for transformation, but their mechanism of action and regulation during transformation remain unknown.

TABLE 1 | VirE2 differentially expressed genes in both RNA-seq studies.

Gene ID	Encoded Protein	Up/Down-regulated Second study (Fold-change)	Up/Down-regulated Pilot study (Fold-change)
	VirE2	Up (194.0)-3 h; Up (2342.3)-12 h	Up (188)-3 h; Up (1961.7)-12 h
At1g01580	Ferric reduction oxidase 2	Down (2.5)-3 h; Down (2.4)-12 h	Down (1.4)-3 h; Down (2.5)-12 h
At1g09932	Phosphoglycerate mutase family protein	Up (1.8)	Up (3.6)
At1g14200	RING/U-box superfamily protein SNIPER1	Up (3.4)	Up (1.9)
At1g23730	Beta carbonic anhydrase 3	Up (5.2)	Up (2.8)
At1g26800	E3 ubiquitin-protein ligase MPSR1	Up (3.0)	Up (1.4)
At1g32350	Alternative oxidase 1D	Up (6.9)	Up (2.5)
At1g61560	MILDEW RESISTANCE LOCUS O 6	Up (1.9)	Up (1.7)
At1g61820	Beta-glucosidase 46	Up (1.9)	Up (1.8)
At1g62370	RING/U-box superfamily protein	Up (2.5)	Up (1.9)
At1g63530	Hypothetical protein	Up (3.0)	Up (1.5)
At1g66090	Disease resistance protein (TIR-NBS class)	Up (4.6)	Up (6.3)
At1g73120	F-box/RNI superfamily protein	Down (5.5)	Down (5.7)
At2g16660	Major facilitator superfamily protein	Down (2.4)	Down (3.8)
At2g17040	NAC domain containing protein 36	Up (2.6)	Up (5.0)
At2g23270	Transmembrane protein	Up (2.9)	Up (4.0)
At2g26150	Heat stress transcription factor A-2	Up (11.6)	Up (1.7)
At2g28160	Transcription factor FER-LIKE IRON DEFICIENCY-INDUCED TRANSCRIPTION FACTOR	Up (1.8)	Up (1.4)
At2g29450	Glutathione S-transferase tau 5	Up (2.4)	Up (2.7)
At2g42850	Cytochrome P450, family 718	Up (2.1)	Up (1.4)
At2g44010	Hypothetical protein	Up (1.6)	Up (3.2)
At2g44578	RING/U-box superfamily protein	Up (2.9)	Up (1.6)
At2g45920	U-box domain-containing protein	Up (1.9)	Up (1.4)
At3g07090	PPPDE putative thiol peptidase family protein	Up (2.0)	Up (1.4)
At3g09290	Telomerase activator1 (TAC1)	Up (3.2)	Up (2.2)
At3g09350	Fes1A	Up (3.2)	Up (1.3)
At3g13437	Enhancer of vascular Wilt Resistance 1; EWR1	Up (2.2)	Up (3.6)
At3g14362	DEVIL 19; DVL19; ROTUNDIFOLIA like 10	Up (2.6)	Up (1.7)
At3g15340	Proton pump interactor 2 (PPI2)	Up (2.8)	Up (1.3)
At3g29000	Calcium-binding EF-hand family protein	Up (2.8)	Up (2.6)
At3g48920	Myb domain protein 45	Up (2.1)	Up (3.4)
At3g46810	Cysteine/Histidine-rich C1 domain family protein	Down (2.7)	Down (2.7)
At3g53150	UDP-glucosyl transferase 73D1	Up (2.3)	Up (1.5)
At3g54150	Embryonic abundant protein-like	Up (2.4)	Up (1.8)
At3g61400	1-aminocyclopropane-1-carboxylate oxidase homolog 8	Down (9.7)	Down (2.7)
At4g04990	Serine/arginine repetitive matrix-like protein (DUF761)	Up (2.1)	Up (2.4)
At4g19690	Fe(2+) transport protein 1	Down (2.5)-3 h; Down (2.8)-12 h	Down (2.3)-3 h
At4g26200	1-aminocyclopropane-1-carboxylate synthase 7	Up (4.8)	Up (1.7)
At4g30230	Uncharacterized protein At4g30230	Up (26.2)	Up (2.1)
At4g30960	CBL-interacting serine/threonine-protein kinase 6	Up (1.6)	Up (1.4)
At4g33050	Calmodulin-binding family protein	Up (1.9)	Up (1.3)
At4g34950	Major facilitator superfamily protein	Down (2.3)	Down (3.7)
At4g37290	Transmembrane protein	Up (5.1)	Up (5.7)
At4g39670	ACD11 homolog protein	Up (2.1)	Up (1.5)
At5g02490	Probable mediator of RNA polymerase II transcription subunit 37c	Up (2.3)	Up (2.8)
At5g03545	Expressed in response to phosphate starvation protein	Down (1.6)	Down (2.3)
At5g06760	LEA4-5	Up (8.3)-3 h; Up (4.4)-12 h	Up (6.6)-12 h
At5g13320	Auxin-responsive GH3 family protein	Up (8.3)	Up (27.3)
At5g25450	Cytochrome b-c1 complex subunit 7	Up (2.7)	Up (3.3)
At5g39050	Phenolic glucoside malonyltransferase 1	Up (2.4)	Up (1.4)

(Continued)

TABLE 1 | Continued

Gene ID	Encoded Protein	Up/Down-regulated Second study (Fold-change)	Up/Down-regulated Pilot study (Fold-change)
At5g39360	EID1-like 2	Up (1.7)	Up (1.5)
At5g39670	Probable calcium-binding protein CML46	Up (2.5)	Up (1.9)
At5g40010	AAA-ATPase ASD, mitochondrial	Up (2.6)	Up (1.3)
At5g43450	1-aminocyclopropane-1-carboxylate oxidase homolog 10	Up (3.9)	Up (3.4)
At5g45840	Phytosulfokin receptor 1	Up (2.5)	Up (1.3)
At5g51440	23.5 kDa heat shock protein, mitochondrial	Up (5.9)	Up (2.1)
At5g54165	Avr9/Cf-9 rapidly elicited protein	Up (4.7)	Up (21.4)
At5g57010	IQ domain-containing protein IQM5	Up (5.4)	Up (1.6)
At5g57510	Cotton fiber protein	Up (4.4)	Up (2.4)
At5g59820	Zinc finger protein ZAT12	Up (2.0)	Up (1.3)
At5g64810	Probable WRKY transcription factor 51	Up (3.0)	Up (1.7)

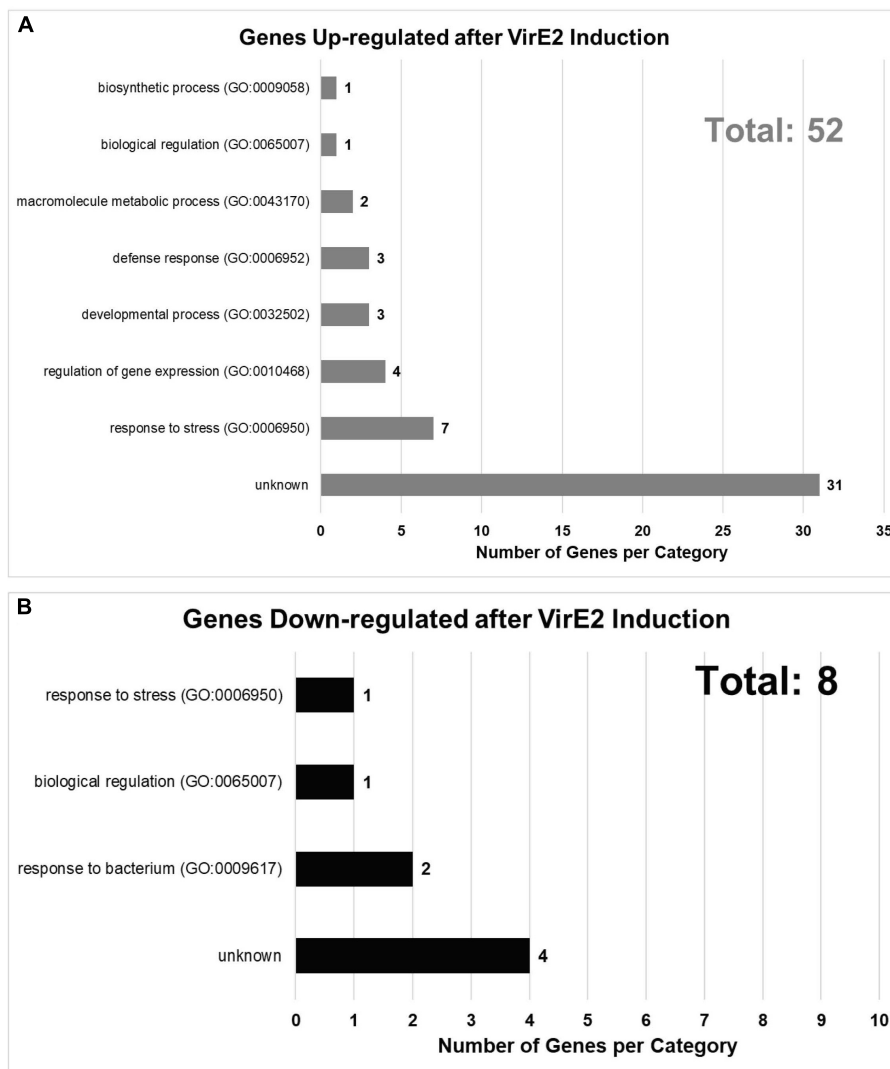
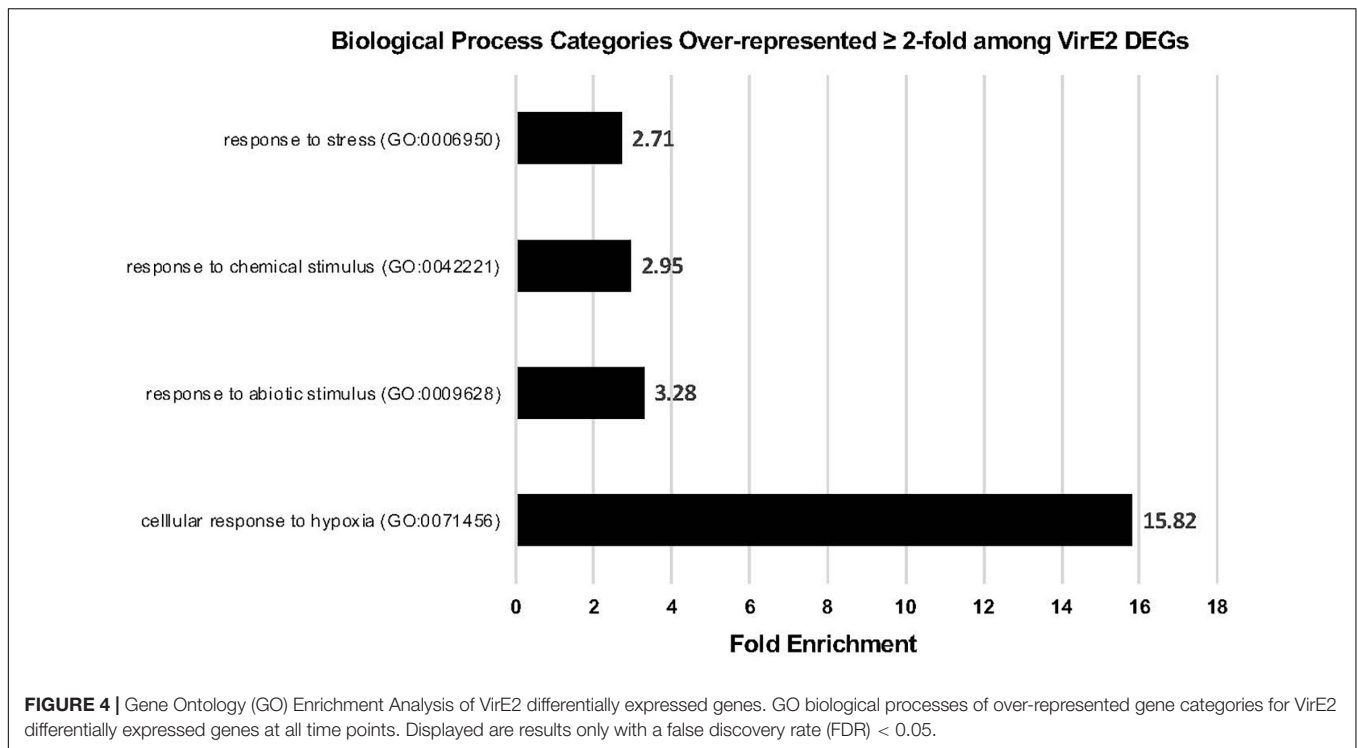


FIGURE 3 | Gene Ontology (GO) Biological Process Categories of up- (**A**) and down-regulated (**B**) genes in the presence of VirE2. Displayed are categories of genes with 1.3-fold or greater change in expression, considering all time points.



The Subcellular Location of VirE2 Results in Different *Arabidopsis* Root Transcriptome Patterns

Roots of transgenic inducible VirE2-Venus or VirE2-Venus-NLS plants were induced for 3 or 12 h with β -estradiol in the presence of *A. tumefaciens* A136. Total RNA was extracted from infected root samples. A subset of genes which exhibited significant changes in expression after the induction of an untagged VirE2 line (determined by RNA-seq) was tested by RT-qPCR for changes in expression in the inducible VirE2-Venus and inducible VirE2-Venus-NLS samples compared with non-induced samples (Table 3, Figure 5 and Supplementary Datasheets 1, 2). The inducible VirE2-Venus (cytoplasmic localized) samples showed either no change or a similar pattern of gene expression changes to those observed for untagged VirE2 (Table 1 and Figure 5). However, the genes *FERRIC REDUCTION OXIDASE2 (FRO2)*, *TRANSMEMBRANE PROTEIN (TMP)*, and *LATE EMBRYOGENESIS ABUNDANT 4-5 (LEA4-5)* showed the opposite pattern of gene expression changes in the VirE2-Venus-NLS (nuclear localized) line compared to that of the VirE2-Venus (cytoplasmic localized) samples (Figures 5A,B,D). These results suggest that the changes in expression of these genes resulted from the cytoplasmic localization of VirE2. Interestingly, both cytoplasmic (VirE2-Venus) and nuclear-localized (VirE2-Venus-NLS) VirE2 caused up-regulation of *HEAT SHOCK PROTEIN 90-1 (HSP90)* and *CALMODULIN-BINDING FAMILY PROTEIN (CBFP)*; Figures 5C,E), but to different extents. Up-regulation of these genes occurred after 12 h of induction and could have resulted from downstream effects caused by the presence of VirE2 in the

plant regardless of localization. It is also possible that a small amount of VirE2 or VirE2-Venus could have entered the nucleus and was sufficient to induce expression of these two genes.

VirE2 Alters the *Arabidopsis* Proteome to Facilitate Transformation

We investigated the effect of VirE2 on the *Arabidopsis* root proteome using the same transgenic inducible VirE2 *Arabidopsis* line that we employed for transcriptome analysis. A total of 135 unique *A. thaliana* proteins (~0.6% of the detectable proteins) showed a statistically significant change in abundance of at least 20% in all three biological replicates of VirE2-induced samples (Supplementary Datasheet 3). These proteins were graphed according to their annotated GO biological process (Figure 6; Ashburner et al., 2000). Proteins previously shown to be important for transformation, such as histones and histone modifying proteins, arabinogalactan proteins (AGPs), and cyclophilins, showed increased abundance in the presence of VirE2 (Table 4; Supplementary Datasheet 3; Deng et al., 1998; Nam et al., 1999; Gaspar et al., 2004; Crane and Gelvin, 2007; Tenea et al., 2009). These VirE2-induced elevations in protein level likely facilitate transformation. Proteins whose levels changed in the presence of VirE2 did not show changes in their RNA levels (Supplementary Datasheet 3), suggesting that VirE2-induced changes to protein levels occur at the translational or post-translational level.

We generated transgenic lines of *A. thaliana* that constitutively overexpressed selected genes whose proteins showed increased abundance in response to VirE2-induction

TABLE 2 | Transformation phenotypes of mutants of VirE2 differentially expressed genes.

Gene Name	Gene_ID	Encoded Protein	Up/Down-regulated (Fold-change)	ABRC Stock ID	Transformation Result
<i>lncRNA</i>	At3g12965	Long non-coding RNA	Up (5.8)	SALK_086573	No change
<i>atpsk3</i>	At3g44735	Phytosulfokine 3 precursor	Up (5)	SALK_044781	*Decreased transient
<i>acs6</i>	At4g11280	1-aminocyclopropane-1-carboxylate synthase 6	Up (3)	SALK_054467	No change
<i>tst18</i>	At5g66170	Thiosulfate sulfurtransferase 18	Up (3.7)	CS867285	*Decreased transient and stable
<i>pr5</i>	At1g75040	Pathogenesis-related protein 5	Up (14)	SALK_055063C	*Increased transient
<i>agp14</i>	At5g56540	Arabinogalactan protein 14	Up (4.9)	SALK_096806	No change
<i>tasi4</i>	At3g25795	Trans-acting siRNA 4	Up (15.1)	SALK_066997	No change
<i>miR163</i>	At1g66725	microRNA 163	Up (3.3)	CS879797	**Decreased stable
<i>samp</i>	At2g41380	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein	Up (10.1)	SALK_209995C	No change
<i>tasi3</i>	At3g17185	Trans-acting siRNA 3	Up (3)	GABI-Kat Stock N432182 (N2051875)	No change
<i>ex1</i>	At1g23720	Proline-rich extensin-like family protein 1	Down (3.3)	SALK_010243C	**Decreased stable
<i>mee39</i>	At3g46330	Maternal effect embryo arrest 39 (putative LRR receptor-like serine/threonine-protein kinase)	Down (4.7)	SALK_065070C	No change
<i>rbc3b</i>	At5g38410	Ribulose biphosphate carboxylase small chain 3B	Down (7.4)	SALK_117835	No change
<i>abah3</i>	At5g45340	Abscisic acid 8'-hydroxylase 3	Down (3.4)	SALK_078170	*Increased transient
<i>ntr2.6</i>	At3g45060	High affinity nitrate transporter 2.6	Down (28)	SALK_204101C	*Increased transient
<i>cup</i>	At3g60270	Cupredoxin superfamily protein	Down (31.3)	SALK_201444C	**Increased transient
<i>ntr2:1</i>	At1g08090	Nitrate transporter 2:1	Down (35.7)	SALK_035429C	*Increased transient
<i>oep6</i>	At3g63160	Outer envelope protein 6 (chloroplast)	Down (5.6)	CS862774	*Decreased stable
<i>esm1</i>	At3g14210	Epithiospecifier modifier 1	Down (10)	SALK_150833C	**Increased stable
<i>rd17</i>	At2g17850	Rhodanese-like domain-containing protein 17	Down (4.7)	SALK_115776C	***Decreased transient and stable
<i>pp2c25</i>	At2g30020	Putative protein phosphatase 2C 25	Down (3.5)	SALK_104445	**Increased transient
<i>adh1</i>	At1g77120	Alcohol dehydrogenase 1	Down (23.2)	SALK_052699	***Increased transient

ANOVA test **p*-value < 0.05, ***p*-value < 0.01, ****p*-value < 0.001.

TABLE 3 | VirE2 subcellular localization impacts changes in plant gene expression.

Gene Name	Gene_ID	Encoded Protein	Up/Down-Regulated in the presence of VirE2 (untagged)	Up/Down-regulated in the presence of VirE2-Venus (cytoplasmic)	Up/Down-regulated in the presence of VirE2-Venus-NLS (nuclear)
<i>FRO2</i>	At1g01580	FERRIC REDUCTION OXIDASE 2	Down 2-fold	Down 2.9-fold	Up 6.3-fold
<i>TMP</i>	At4g37290	TRANSMEMBRANE PROTEIN	Up 5-fold	Up 3.9-fold	Down 6.9-fold
<i>HSP90</i>	At5g52640	HEAT SHOCK PROTEIN 90-1	Up 6-fold	Up 2.9-fold	Up 7.1-fold
<i>LEA4-5</i>	At5g06760	LATE EMBRYOGENESIS ABUNDANT 4-5	Up ≥3-fold	Up 1.6-fold	Down 2.5-fold
<i>CBFP</i>	At5g57010	CALMODULIN-BINDING FAMILY PROTEIN	Up 5-fold	Up 6.1-fold	Up 4.0-fold

(Table 5). Although statistical analysis of the iBAC (intensity-Based Absolute Quantitation) scores showed that the increased ARABINOGALACTAN PROTEIN 31 (AGP31) abundance was statistically significant at only $p = 0.27$, we included this gene in our overexpression analysis because our previous study indicated that AGPs were important for transformation (Gaspar et al., 2004). Roots from multiple T2 generation transgenic lines were assayed for transient and stable transformation susceptibility (Table 5 and Supplementary Figure 6). Some transgenic

lines containing a PEROXIDASE 34 (PERX34) overexpressing construct showed decreased stable transformation (Table 5 and Supplementary Figure 6A), whereas most of the ROTAMASE CYCLOPHILIN 2 (ROC2) overexpression lines showed increased transient transformation (Table 5; Supplementary Figure 6B). Transgenic lines containing an overexpression construct for HISTONE DEACTYLASE 3 (HDA3; Supplementary Figure 6C) showed increased transient transformation. Some plant lines containing overexpression constructs for

TABLE 4 | Proteins previously identified as important for transformation show increased abundance in the presence of VirE2.

Gene ID	Gene Name	Encoded Protein	% Change in Protein Level (Time Post-VirE2 Induction) ^a
At2g28740	<i>HIS4</i> (formerly <i>HFO4</i>)	Histone H4	+37% (3 h); $p = 0.006$
At4g27230	<i>HTA2</i>	Histone H2A2	+140% (3 h); $p = ns$
At5g03740	<i>HD2C</i> (formerly <i>HDT3</i>)	Histone deacetylase 2C	+35% (3 h); $p = 0.007$
At3g44750	<i>HDA3</i> (formerly <i>HDT1</i>)	Histone deacetylase 3	+50% (12 h); $p = 0.002$
At2g16600	<i>ROC3</i>	Rotamase cyclophilin 3	+20% (12 h); $p = 0.061$
At3g56070	<i>ROC2</i>	Rotamase cyclophilin 2	+85% (12 h); $p = 0.046$
At1g03870	<i>FLA9</i>	FASCICLIN-like arabinogalactan 9	+25% (3 h); $p = 0.035$
At1g28290	<i>AGP31</i>	Arabinogalactan protein 31	+50% (3 h); $p = ns$

^a p -value according to IBAQ analysis; ns, not significant at $p < 0.1$.

both *HISTONE DEACTYLASE 2C* (*HD2C*; **Supplementary Figure 6D**) and *AGP31* (**Supplementary Figure 6E**) also showed increased transformation efficiency. These data suggest that VirE2-induced changes to levels of specific proteins may facilitate transformation.

Taken together, our results suggest that VirE2 impacts the plant cell on both the RNA and protein levels to facilitate transformation, and that the effect of VirE2 occurs from its position in the plant cytoplasm.

DISCUSSION

VirE2 Must Localize to the Plant Cytoplasm to Facilitate Transformation

We have shown that only cytoplasmic localized VirE2-Venus, but not nuclear localized VirE2-Venus-NLS, could complement the loss of virulence of a *virE2*⁻ *Agrobacterium* mutant. These results suggest that the major role of VirE2 in transformation occurs in the cytoplasm.

The reported subcellular localization of VirE2 is controversial. When tagged on its N-terminus, VirE2 was reported to localize to the nucleus (Citovsky et al., 1992, 1994, 2004; Tzfira and Citovsky, 2001; Tzfira et al., 2001; Li et al., 2005). Other studies showed that both N- and C-terminally tagged VirE2 localized to the cytoplasm (Bhattacharjee et al., 2008; Grange et al., 2008; Lee et al., 2008; Shi et al., 2014; Lapham et al., 2018). However, only the C-terminally tagged fusion protein, when expressed in a plant, could complement a *virE2* mutant strain and restore efficient transformation (Bhattacharjee et al., 2008).

More recently, Li et al. (2014) showed that an *Agrobacterium* strain expressing VirE2 with an internal small GFP fragment (GFP11) is virulent. Using this strain and a split-GFP approach, VirE2-GFP11 delivered from *Agrobacterium* could refold with GFP1-10 expressed *in planta* to restore GFP fluorescence.

TABLE 5 | Transformation phenotypes of *A. thaliana* lines containing overexpression constructs of genes whose proteins show increased abundance post-VirE2 induction.

Gene ID	Gene Name	Encoded Protein	% Change in Protein Level (Time Post-VirE2 Induction)	Transformation Result
At3g49120	<i>PERX34</i>	Peroxidase 34	+38% (3 h)	*Decreased stable
At5g03740	<i>HD2C</i> (formerly <i>HDT3</i>)	Histone deacetylase 2C	+35% (3 h)	**Increased transient and *Increased stable
At3g44750	<i>HDA3</i> (formerly <i>HDT1</i>)	Histone deacetylase 3	+50% (12 h)	*Increased transient
At3g56070	<i>ROC2</i>	Rotamase cyclophilin 2	+85% (12 h)	**Increased transient
At1g28290	<i>AGP31</i>	Arabinogalactan protein 31	+50% (3 h)	*Increased transient

ANOVA test * p -value < 0.05, ** p -value < 0.01.

In the plant cell, VirE2-GFP complexes formed filamentous structures mainly in the cytoplasm and with a few that appeared within the nucleus. Roushan et al. (2018) used phiLOV2.1 to tag VirE2 internally and showed that, when transferred from *Agrobacterium*, the protein localized to the cytoplasm of *Arabidopsis* roots and *Nicotiana tabacum* leaves. Li et al. (2020) further demonstrated that only very small amounts of VirE2 could be detected in the nucleus in the presence of VirD2 and T-strands, solving the conundrum of conflicting results from different laboratories. Our results indicate that, regardless of its site of synthesis, only when VirE2-Venus protein localizes to the cytoplasm can it complement a *virE2* mutant *Agrobacterium* strain. An inducible nuclear-localized VirE2-Venus-NLS protein could not complement the *virE2* mutant strain. These results confirm our previous observations (Bhattacharjee et al., 2008) and indicate that VirE2 must localize to the cytoplasm to perform its functions in *Agrobacterium*-mediated transformation.

VirE2 interacts with several *Arabidopsis* importin α (Imp α) isoforms in a yeast two hybrid system and in plant cells when overexpressed (Bhattacharjee et al., 2008; Lee et al., 2008). VirE2 interacts with many Imp α isoforms in the cytoplasm, but only VirE2-Imp α -4 interaction localizes to the nucleus of BY-2 protoplasts. Although VirE2 protein contains two putative bipartite NLS sequences (Citovsky et al., 1992, 1994), structural analyses indicated that the interactions between rice Imp α 1a and the VirE2 NLS sequences are weak (Chang et al., 2014). Ziemienowicz et al. (2001) observed that VirE2 bound to ssDNA was not imported into isolated tobacco nuclei, but they did observe the import of free VirE2 molecules into the nucleus. On the other hand, VirE2, in addition to the effector protein VirD2, was required for nuclear import of large ssDNA molecules in this *in vitro* system (Ziemienowicz et al., 2001). It is possible that a small amount of VirE2 localizes to the nucleus during transformation. However, based on our results, exclusive nuclear localization of VirE2 does not support transformation.

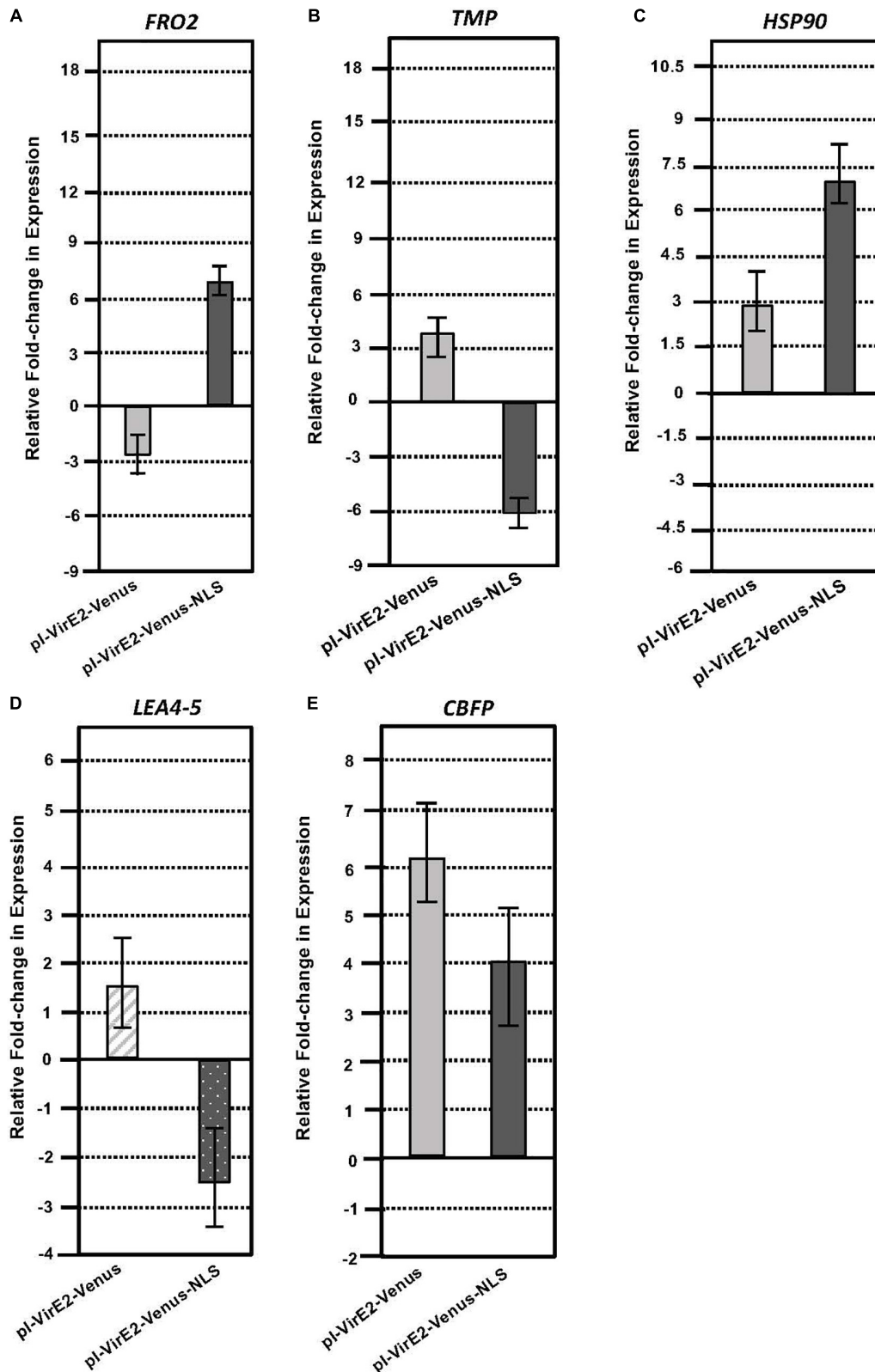
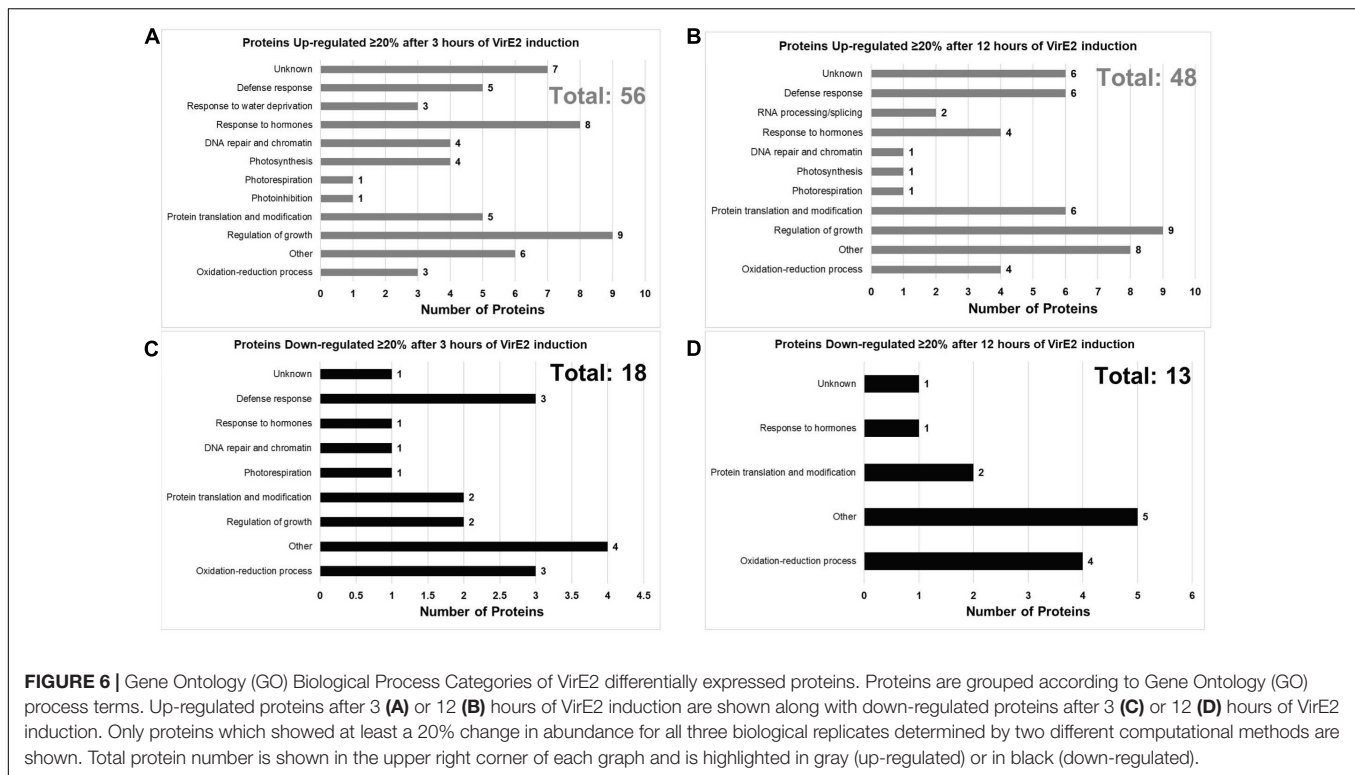


FIGURE 5 | Quantitative RT-PCR analysis of selected VirE2 differentially expressed genes in inducible VirE2-Venus (cytoplasmic) versus inducible VirE2-Venus-NLS (nuclear) plants. VirE2-Venus (left) and VirE2-Venus-NLS (right) results of (A) *FRO2*, (B) *TMP*, (C) *HSP90*, (D) *LEA4-5*, and (E) *CBFP* gene expression in induced relative to non-induced roots. Bars represent an average of three technical replicates \pm SE for one representative biological replicate of one transgenic line. Relative expression is shown after 3 (*LEA4-5* only) or 12 h after induction in the presence of *A. tumefaciens* A136.



VirE2 Alters the *Arabidopsis* Root Transcriptome and Proteome to Facilitate Transformation

We cannot rule out that VirE2 has no function within the nucleus. We therefore investigated possible functions of VirE2 in transformation other than its proposed structural roles in protecting T-strands (Howard and Citovsky, 1990) and/or shaping T-strands to traverse the nuclear pores (Ziemienowicz et al., 2001; Li et al., 2020). VirE2 interacts with the *Arabidopsis* transcription factors VIP1 and VIP2 (Tzfira et al., 2001; Anand et al., 2007; Pitzschke et al., 2009) and various other plant proteins (Lee et al., 2008, 2012). Although VIP1 and its orthologs do not play a role in *Agrobacterium*-mediated transformation (Shi et al., 2014; Lapham et al., 2018), interactions with VIP2 or other proteins could lead to changes in plant gene expression, perhaps facilitating transformation. RNA-seq analysis of transgenic *A. thaliana* roots expressing VirE2 revealed that most transcript abundance changes occurred 12 h post-VirE2 induction (Supplementary Data Sheets 1, 2). Conversely, proteomics analysis indicated that numerous proteins changed abundance 3 h after VirE2 induction, but none of the transcripts for these proteins changed abundance at that early time (Supplementary Data Sheet 3). These results suggest that alterations in mRNA and protein abundance in response to VirE2 expression occur post-transcriptionally, most likely at the translational or post-translational level. This hypothesis is consistent with cytoplasmic- rather than nuclear-localized VirE2. It is also supported by our data showing that proteins involved in translation also exhibited rapid changes in their steady-state

levels in response to VirE2 induction (Figure 6). The mechanism by which VirE2 increases the abundance of particular proteins requires further investigation.

Genes involved in plant defense were differentially expressed in response to VirE2 induction (Figure 3 and Supplementary Data Sheets 1, 2). Duan et al. (2018) noted that expression of several defense genes was up-regulated in *A. thaliana* constitutively expressing VirE2 24 h after the plants were treated with the avirulent *Agrobacterium* strain A136. They also found that plants constitutively expressing VirE2 had reduced transformation efficiency compared to wild-type plants. They proposed that this inhibition was caused by enhanced defense responses in the VirE2-expressing plants. We also observed up-regulation of genes involved in innate immune responses 12 h after VirE2 induction in the presence of the avirulent *Agrobacterium* strain A136 (Figure 3 and Supplementary Data Sheet 1, 2), but the genes we identified differed from those identified previously by Duan et al. (2018; Supplementary Data Sheet 1, 2). Ditt et al. (2006) found that genes involved in response to biotic stimulus, abiotic stimulus, and stress were enriched for transcripts up-regulated 48 h after infection of *Arabidopsis* cell cultures (ecotype Ler) by the tumorigenic *Agrobacterium* strain A348. We also observed up-regulation of these same gene categories 12 h after VirE2 induction in the presence of the avirulent *Agrobacterium* strain A136. Veena et al. (2003) observed an increase in defense response gene transcripts early (3–6 h) after *Agrobacterium* infection of *N. tabacum* BY-2 suspension cells, but expression of these genes was suppressed at later infection times (30–36 h) in the presence of *Agrobacterium* strains that could transfer

virulence proteins. However, suppression of this delayed defense response did not occur when the plants were infected with the transfer-deficient *Agrobacterium* strain A136 (Veena et al., 2003).

The stress-response associated *ALCOHOL DEHYDROGENASE 1 (ADH1)* gene was strongly down-regulated in the presence of VirE2 (Table 2; Supplementary Figure 3A and Supplementary Data Sheet 1) and a knockout mutant line of this gene showed increased transformation (Supplementary Figure 5F). Veena et al. (2003) also found that a tobacco alcohol dehydrogenase gene was down-regulated in the presence of a virulent *Agrobacterium* strain at later infection time points. In addition, our RNA-seq experiments revealed that the transcription factor *WRKY33* was up-regulated 12 h after VirE2 induction (Supplementary Data Sheet 2). Zheng et al. (2006) showed that ectopic over-expression of *WRKY33* resulted in increased susceptibility to the bacterial pathogen *Pseudomonas syringae*, and that *WRKY33* could act as a negative regulator of bacterial defense responses.

Genes known to be important for transformation, including those encoding a protein phosphatase 2C (Tao et al., 2004), AGPs (Nam et al., 1999; Gaspar et al., 2004), and heat shock proteins (Park et al., 2014), showed changes in expression in response to VirE2 (Supplementary Data Sheets 1, 2). *PROTEIN PHOSPHATASE 2C 25 (PP2C25)* was down-regulated by VirE2 (Table 2) and its knockout mutant line exhibited increased transformation (Supplementary Figure 5F). A tomato protein phosphatase 2C (*DIG3*) was previously shown to act as a negative regulator of transformation by dephosphorylating a serine residue in VirD2 that is critical for VirD2 nuclear import (Tao et al., 2004). VirE2-mediated down-regulation of *PP2C25* may therefore facilitate more efficient nuclear import of VirD2/T-strand complexes.

Induction of VirE2 increased transcript and protein levels of some AGP genes (Supplementary Data Sheets 1, 3). *ARABINO GALACTAN PROTEIN 17 (AGP17)* was previously shown to be important for transformation by enhancing attachment of *Agrobacterium* to plant cells (Nam et al., 1999; Gaspar et al., 2004). We assayed a knockout mutant of the *AGP14* gene for transformation susceptibility but did not observe any significant difference in transformation compared to wild-type plants (Supplementary Figure 4E). Schultz et al. (2002) identified 50 *Arabidopsis* genes encoding AGPs, and it is plausible that many have redundant functions in the plant cell. *AGP31* showed increased protein levels (although at a p -value = 0.27 by iBAQ analysis) in the presence of VirE2 (Table 3 and Supplementary Data Sheet 3) and plants overexpressing *AGP31* exhibited increased transient transformation susceptibility (Table 5 and Supplementary Figure 6E). Therefore, VirE2 may modulate both the transcript and protein levels of some AGPs to facilitate transformation.

Some heat shock protein transcript and protein levels increased in response to VirE2 induction (Supplementary Data Sheet 2), including the transcript encoding *HEAT SHOCK*

PROTEIN 90 (HSP90). Park et al. (2014) demonstrated that over-expression of *HSP90* increased *Arabidopsis* root transformation susceptibility and proposed that *HSP90* could act as a molecular chaperone to stabilize VirE2 and other proteins important for transformation. Up-regulation of *HSP90* by VirE2 could also facilitate transformation.

Histones, histone modifying enzymes, and cyclophilins showed increased protein levels in response to VirE2 (Table 4 and Supplementary Data Sheet 3) and have previously been shown, or proposed, to play important roles in transformation (Deng et al., 1998; Nam et al., 1999; Bakó et al., 2003; Crane and Gelvin, 2007; Tenea et al., 2009). Histone H2A2 (*HTA2*) and histone H4 (*HIS4*; formerly *HFO4*) protein levels increased 3 h after induced VirE2 expression (Table 4 and Supplementary Data Sheet 3). Over-expression of *HIS4*, *HTA2*, and some other histone H2A variants increased transformation susceptibility of *Arabidopsis* (Tenea et al., 2009). The histone deacetylases *HD2C* (formerly *HDT3*) and *HDA3* (formerly *HDT1*) also showed increased protein levels in response to VirE2 (Table 4 and Supplementary Data Sheet 3). Crane and Gelvin (2007) showed that RNAi-mediated silencing of *HDA3* and other chromatin-related genes resulted in reduced transformation and T-DNA integration. Plants overexpressing *HDA3* had enhanced transient transformation susceptibility (Table 5 and Supplementary Figure 6C), whereas *HD2C* overexpressing plants had increased transient and stable transformation rates compared to wild-type plants (Table 5 and Supplementary Figure 6D). Increased levels of these histones and histone modifying proteins in response to VirE2 may also facilitate transformation.

VirD2 interacts with various cyclophilin proteins, and this interaction is important for efficient transformation (Deng et al., 1998; Bakó et al., 2003). Two cyclophilin proteins, *ROC2* and *ROC3*, showed increased protein levels post-VirE2 induction (Table 4 and Supplementary Data Sheet 3). Our results also show that plants overexpressing *ROC2* have increased transformation susceptibility (Table 5 and Supplementary Figure 6B). Taken together, these data suggest that VirE2 increases the levels of some cyclophilin proteins, facilitating transformation.

CONCLUSION

VirE2 alters the steady-state levels of specific plant RNAs and proteins which are known to be important for transformation. VirE2 likely mediates these changes post-transcriptionally. This model is supported by the rapid changes in levels of certain proteins and more delayed changes in levels of specific RNAs we observed in response to VirE2 induction. Coupled with our observation that cytoplasmic localization of VirE2 is required for it to function in transformation, our results are consistent with a post-transcriptional role in modulating mRNA and protein levels. We conclude that VirE2, from its location in the plant cytoplasm, modulates specific plant steady-state RNA and protein levels post-transcriptionally to facilitate transformation.

MATERIALS AND METHODS

Plasmid and Strain Constructions

Supplementary Table 2 lists the plasmids and strains used in this study. To make a cloning vector with an inducible promoter (Pi), a blunted *SphI-XhoI* fragment containing the *LexA* operator and a minimal CaMV 35S promoter from pER8 (Zuo et al., 2000) was ligated to the blunted *AgeI-XhoI* plasmid pE3542 to make pE4224.

To make the pPi-VirE2-Venus construction, a *SwaI-NotI* fragment containing the VirE2-Venus fragment from pE3759 was cloned into the *SwaI-NotI* sites of pE4224 to make Pi-VirE2-Venus (pE4282). The *AscI* fragment from pE4282 containing the expression cassette pPi-VirE2-Venus and an *I-SceI* fragment containing P_{nos}-Cerulean-NLS from pE4373 were cloned into the *AscI* and *I-SceI* sites, respectively, of a binary vector derived from pE4215 containing an XVE expression cassette to make pE4438 (pPZP-Pi-VirE2-Venus-P_{nos}-Cerulean-NLS).

To make the pPi-VirE2-Venus-NLS construction, pSAT1-P_{35S}-Venus-VirD2 (pE3561) was digested with *HindIII* before self-ligating the backbone fragment to create pSAT1-P_{35S}-Venus-NLS (pE4433). A *PstI-NotI* fragment from pE4433 was used to replace the *PstI-NotI* fragment of pE3759 to make pE4434 (pSAT1-P_{35S}-VirE2-Venus-NLS). A *SmaI-NotI* fragment from pE4434 was cloned into the *SmaI-NotI* sites of pE4224 to make pE4436 (pSAT1-P_i-VirE2-Venus-NLS). An *AscI* fragment containing the Pi-VirE2-Venus-NLS expression cassette from pE4436 was cloned into the *AscI* site (to replace the Pi-VirE2-Venus expression cassette) of pE4389 to make pE4435. pE4435 was digested with *I-CeuI* and self-ligated to make pE4439 (pPZP-Pi-VirE2-Venus-NLS-P_{nos}-Cerulean-NLS). pE4438 and pE4439 were separately introduced into *A. tumefaciens* GV3101 (Van Larebeke et al., 1974) by electroporation to make *A. tumefaciens* At2155 and At2156, respectively.

To generate a binary vector carrying the Pi-VirE2 expression cassette, a *SwaI-NotI* fragment containing the *VirE2* gene from pE4229 was cloned into the *SmaI-NotI* sites of pE4224 to create pE4276. The *AscI* fragment containing pPi-VirE2 was cloned into the *AscI* sites of pE4215 to generate pE4289. pE4289 was electroporated into *A. tumefaciens* GV3101 to make *A. tumefaciens* At2091.

To generate the constitutive overexpression constructs for proteins whose levels are increased in the presence of VirE2, cDNA clones were ordered from the Arabidopsis Biological Resource Center (ABRC)¹ for each selected gene (**Supplementary Table 2**). Each gene was amplified from the cDNA clone using PCR and primers with flanking sequences containing restriction enzyme sites (**Supplementary Table 3**). Either Phusion High-Fidelity DNA Polymerase (New England Biolabs) or Platinum SuperFi DNA Polymerase (Invitrogen) was used and the reactions were conducted according to the manufacturers' protocols. The PCR fragments containing *PERX34* (At3g49120) were digested with restriction enzymes which recognized their flanking sequences (**Supplementary Table 3**) before cloning those fragments into the same sites

on pE4297 to create pE4622 (**Supplementary Table 2**). The blunt-end PCR fragments containing *AGP31* (At1g28290), *HDA3* (At3g44750), *HD2C* (At5g03740), and *ROC2* (At3g56070) were cloned into pBluescript KS+ cut with *EcoRV* to make pE4626, pE4629, pE4633, and pE4637, respectively (**Supplementary Tables 2, 3**). These plasmids were also sequenced. The *EcoRI-BamHI* fragments from pE4629 (*HDA3*) and pE4637 (*ROC2*) were cloned into the same sites of pE4515 to make pE4630 and pE4638, respectively. The *Sall-BamHI* fragment from pE4626 (*AGP31*) and the *BglII-BamHI* fragment from pE4633 (*HD2C*) were cloned into the same sites of pE4297 to make pE4627 and pE4634, respectively. The *AscI* fragments containing the overexpression cassettes from pE4622 (*PERX34*), pE4627 (*AGP31*), pE4630 (*HDA3*), pE4634 (*HD2C*), and pE4638 (*ROC2*) were cloned into the *AscI* site of the binary vector pE4145 to make pE4623, pE4628, pE4631, pE4635, and pE4639, respectively. Each binary vector was electroporated into *A. tumefaciens* GV3101 to make *A. tumefaciens* strains At2259, At2264, At2265, At2267, and At2268, respectively.

Isolation and Transfection of Tobacco BY-2 Protoplasts

Protoplasts were isolated from tobacco BY-2 cells and transfected as described by Lee et al. (2012). A plasmid encoding a nuclear mRFP marker (pE3170) was co-transfected with the appropriate clones into the protoplasts. Imaging was performed 16 h post-transfection using a Nikon A1R Confocal Laser Microscope System as described in Shi et al. (2014).

Generation and Selection of Inducible VirE2, VirE2-Venus, VirE2-Venus-NLS, VIP1, and Transgenic *A. thaliana* Plants Constitutively Overexpressing Selected Genes

Wild-type *A. thaliana* plants (ecotype Col-0) were individually transformed by *A. tumefaciens* At2155, At2156, At2091, At2259, At2264, At2265, At2267, or At2268 using a flower dip protocol (Clough and Bent, 1998). T0 generation seeds from the transformed plants were surface sterilized for 15–20 min in a 50% commercial bleach and 0.1% sodium dodecylsulfate (SDS) solution before washing five times with sterile water. After overnight incubation in water at 4°C, the seeds were plated on solidified Gamborg's B5 medium (Caisson Labs) containing 100 mg/mL Timentin and 20 mg/mL hygromycin. The seeds were placed at 23°C under a 16/8-h light/dark cycle. T1 generation hygromycin-resistant seedlings for the inducible lines were transplanted to soil and grown under the same temperature and light conditions. For inducible VirE2 plants, hygromycin was used to select for homozygous plants. Homozygous T2 plants containing the inducible VirE2-Venus and VirE2-Venus-NLS constructions were used for future experiments. T1 generation hygromycin-resistant seedlings for each of the constitutive overexpression lines were transferred to baby food jars containing solidified B5 medium for 10–14 days. Roots of each plant were cut into 3–5 mm segments and assayed as described in Tenea et al. (2009). Root segments were infected with *A. tumefaciens* At849

¹www.arabidopsis.org

[GV3101:pMP90 (Koncz and Schell, 1986) containing pBISN1 (Narasimhulu et al., 1996)] to measure transient transformation at a concentration of 10^6 cfu/mL. Shoots were re-rooted in solidified B5 medium in the jars for 7–10 days before transferring plantlets to soil.

Transgenic plants overexpressing VIP1 were generated using *A. tumefaciens* At2082 as previously described (Lapham et al., 2018).

Imaging of VirE2-Venus and VirE2-Venus-NLS Transgenic *A. thaliana* Roots

Inducible VirE2-Venus and VirE2-Venus-NLS seedlings (T2 generation) were germinated on B5 medium containing 100 mg/mL Timentin and 20 mg/mL hygromycin. The seedlings were transferred after 2 weeks to plates containing B5 medium lacking antibiotics. These plates were placed vertically in racks to promote root growth on the surface of the medium. After 10 days, the plates were placed horizontally and B5 liquid medium containing 10 μ M β -estradiol dissolved in DMSO (β -estradiol solution) or B5 plus DMSO only (control solution) was pipetted onto the surface until a thin layer covered the root tissue (4–5 mL). The roots were incubated in the solution for 9 h before imaging using a Nikon A1R Confocal Laser Microscope System as described in Shi et al. (2014).

Assaying Inducible VirE2-Venus and VirE2-Venus-NLS Transgenic *A. thaliana* Roots for Complementation of *virE2*⁻ Mutant *Agrobacterium*

Three transgenic lines of Inducible VirE2-Venus (Lines #4-6) and VirE2-Venus-NLS (Lines #4-6) seedlings (T2 generation) were grown and treated with either 10 μ M β -estradiol induction or control solution for 24 h as described above. Root segments were infected as described in Tenea et al. (2009) using either *A. tumefaciens* At1529 or the *virE2*⁻ mutant strain At1879 at a concentration of 10^6 or 10^8 cfu/mL, respectively (Supplementary Table 2). Three replicates were assayed for each line with root segments pooled from 10 to 30 plants for each replicate. A total of 80 or more root segments were scored for each data point and statistical analysis was performed using ANOVA.

VirE2, VirE2-Venus, VirE2-Venus-NLS, and VIP1 Induction in the Presence of *Agrobacterium*

Inducible VirE2 (line #10) or inducible VIP1 (line #12) T3 generation plants were grown and assayed as described above, except that *A. tumefaciens* A136 (lacking a Ti plasmid) were added either to induction (1 μ M β -estradiol) or control solution at a concentration of 10^8 cfu/mL. Roots from 30 plants were cut after 0, 3, or 12 h treatment, rinsed with sterile water, dried on a paper towel, and frozen in liquid nitrogen before RNA extraction.

Inducible VirE2-Venus Line #4 and inducible VirE2-Venus-NLS Line #4 T2 generation plants were also grown, treated, and harvested in the same manner as the inducible VirE2

plants before isolating RNA for quantitative RT-PCR (RT-qPCR) analysis.

Preparation of Samples for RNA-seq Analysis and Quantitative RT-PCR

For both RNA-seq and RT-qPCR analyses, RNA was isolated from non-induced and induced roots in the presence of *Agrobacterium* after 0, 3, and 12 h of treatment using TriZol reagent². Three biological replicates of inducible VirE2 *A. thaliana* transgenic line #10 were analyzed by both RNA-seq and RT-qPCR. The inducible VIP1 *A. thaliana* transgenic line #12 was analyzed by RNA-seq and two biological replicates were analyzed by RT-qPCR (Lapham et al., 2018). Two biological replicates of inducible VirE2 Venus transgenic line #4 and inducible VirE2-Venus-NLS transgenic line #4 were analyzed by RT-qPCR. cDNA was made from polyA⁺ RNA using an Illumina TruSeq Stranded mRNA kit without rRNA depletion. One biological replicate was sequenced at the Purdue Genomics Core Facility on an Illumina HiSeq 2500 DNA sequencer using single-end, 100 cycle rapid run chemistry for the initial VirE2 pilot study and the VIP1 study. RNA from two additional VirE2 biological replicates was similarly sequenced by the Cornell University Institute of Biotechnology Genomics Facility, using an Illumina TruSeq-3' RNA-seq kit to make cDNA.

A total of 2 μ g of total RNA was treated with Ambion DNase I (Thermo Fisher Scientific) before submitting the RNA for sequencing. For RT-qPCR, cDNA was synthesized from 1.45 μ g of total RNA treated with Ambion DNase I using SuperScript III reverse transcriptase (Thermo Fisher Scientific) following the manufacturer's protocols. RT-qPCR was performed using FastStart Essential Green Master reagents (Roche) on a Roche LightCycler 96. Primer sequences for gene amplification are listed in Supplementary Table 3. RT-qPCR data were analyzed using the LightCycler 96 software and Microsoft Excel.

RNA-seq Bioinformatic Analysis: Pilot Study

RNA was submitted to the Purdue Genomics Core Facility for sequencing after treatment with DNase I to remove any contaminating genomic DNA. Ribosomal RNA was depleted and cDNA libraries (stranded) were prepared from each of the samples before sequencing. Between 15 and 23 million reads were obtained for each sample (100 nucleotides per read) which were quality trimmed and mapped to the *A. thaliana* genome using TopHat (Trapnell et al., 2010). DEGs were determined from the mapped (bam) files using Cuffdiff from the Cufflinks suite of programs (Trapnell et al., 2010). Custom perl scripts were used to extract genes for which fold-changes of 3 or greater occurred between the induced and non-induced control samples at their respective time points. The resulting genes were annotated by hand and separated into categories based on their GO functions which were found in the National Center for Biotechnology Information (NCBI) database³.

²<http://www.thermofischer.com>

³<https://www.ncbi.nlm.nih.gov/>

RNA-seq Bioinformatic Analysis by Purdue Bioinformatics Core: Second Study

Sequence quality was assessed using FastQC (v 0.11.7)⁴ for all samples and quality and adapter trimming was done using TrimGalore (0.4.4) (Krueger, 2017) to remove the sequencing adapter sequences and bases with Phred33 scores less than 30. The resulting reads of length >25 bases were retained (original read length = 50 and lib type = unstranded) respectively. The quality trimmed reads were mapped against the reference genome using STAR (Dobin et al., 2013) (v 2.5.4b). STAR derived mapping results and annotation (GTF/GFF) file for reference genome were used as input for HTSeq (Anders et al., 2015) package (v 0.7.0) to obtain the read counts for each gene feature for each replicate. Counts from all replicates were merged using custom Perl scripts to generate a read count matrix for all samples.

The merged counts matrix was used for downstream differential gene expression analysis. Genes that did not have counts in all samples were removed from the count matrix and genes that had counts in some samples but not in others were changed from 0 to 1 in order to avoid having infinite values calculated for the fold change. Differential gene expression (DEG) analysis between treatment and control was carried out using “R” (v 3.5.1)⁵ with two different methods (DESeq2 and edgeR). Basic exploration of the read count data file such as accessing data range, library sizes, etc. was performed to ensure data quality. An edgeR object was created by combining the count’s matrix, library sizes, and experimental design using the edgeR (Robinson et al., 2010) (v 3.24.3) package. Normalization factors were calculated for the count’s matrix, followed by estimation of common dispersion of counts. An exact test for differences between the negative binomial distribution of counts for the two experimental conditions resulted in finding differential expression, which was then adjusted for multiple hypothesis testing. DESeq2 (Love et al., 2014) (v 1.22.2) was also used to find DEGs. Both use an estimate variance-mean test based on a model using the negative binomial distribution. The significant genes were identified by examining the adjusted *p*-value.

Additionally, STAR mapping (bam) files were used for analysis by the Cuffdiff from Cufflinks (v 2.2.1) (Trapnell et al., 2010) suite of programs which perform DE analysis based on FPKM values. Cuffdiff uses bam files to calculate Fragments per Kilobase of exon per Million fragments mapped (FPKM) values, from which differential gene expression between the pairwise comparisons can be ascertained. DEG lists detected by at least two or more methods (DESeq2, edgeR, and Cufflinks) were generated using custom Perl scripts.

Gene annotations were retrieved from BioMart databases using biomart package in “R.” The “transcript_biotype,” “description” attributes were extracted using `mart = “plants_mart”` and `dataset = “athaliana_eg_gene.”`

GO enrichment analysis was also performed using DEGs from two or more methods while using two replicates. Singular Enrichment Analysis (SEA) from agriGO (Du et al., 2010) was used to perform GO enrichment analysis (count = 5 with Fisher exact *t*-test with multiple testing). A GO enrichment analysis was performed using the PANTHER Classification system and online tools⁶.

Genotyping and *Agrobacterium*-Mediated Transient and Stable Transformation Assays of T-DNA Insertion Lines

Arabidopsis thaliana T-DNA insertion lines tested in this study are listed in **Table 2**. Seeds for these lines were obtained from the ABRC (see text footnote 1). For genotyping, DNA was isolated from leaves sampled from 10 to 15 individual plants after freezing the tissue in liquid nitrogen and grinding it into a fine powder using a sterile tube pestle. A total of 0.5 mL of extraction buffer (100 mM Tris pH 8.0, 50 mM EDTA, 500 mM NaCl) was added to the ground tissue before mixing thoroughly. A total of 26 μ L of 20% SDS solution was added to each sample before mixing by inverting the tubes. The samples were incubated in a 65°C water bath for 20 min and were mixed by inverting every 5 min during the incubation. After removing the samples from the water bath, 125 μ L of potassium acetate buffer was added to each sample before mixing. The potassium acetate buffer is made by mixing 60 mL of 5 M KOAc from crystals, 11.5 mL glacial acetic acid, and 28.5 mL of filtered H₂O to make 100 mL (3 M of potassium and 5 M of acetate in the final solution). The tubes were placed on ice for up to 20 min before centrifugation at top speed for 10 min in a microcentrifuge at 4°C. The supernatant solution was transferred to a fresh tube (~600 μ L). The samples were centrifuged a second time if cellular debris were still evident within the supernatant solution. A 0.7 volume (420 μ L) of isopropanol was added to the supernatant fluid before mixing the samples and placing them at –20°C for at least 1 h to precipitate the DNA. The samples were centrifuged at top speed for 10 min in a microcentrifuge at 4°C to pellet the DNA. The DNA pellets were washed with 500 μ L of 70% ethanol by flicking the tube until the pellets released from the bottom of the tube. The samples were centrifuged again for 5 min before carefully removing the ethanol. The pellets were then allowed to air-dry for 5–10 min before resuspending the pellets in 30 μ L of 1 xTE buffer (10 mM Tris-Cl, 1 mM EDTA [pH 8.0]) plus 20 μ g/mL RNase A.

Lines homozygous for the annotated T-DNA insertions were confirmed by PCR (primer sequences are listed in **Supplementary Table 3**). PCR reaction mixes were made using ExTaq Buffer (TaKaRa), dNTPs (0.2 mM), the appropriate forward and reverse primers (0.2 μ M each), homemade Taq polymerase, and water with a tenth volume of sample added to act as a template. The reactions were incubated at 95°C for 3 min before performing 35 cycles of a 30 s, 95°C denaturation step, followed by a 30 s annealing step (temperature was ~5°C lower

⁴<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

⁵<http://www.r-project.org/>

⁶<http://geneontology.org/docs/go-enrichment-analysis/>

than the average melting temperature for each primer set), and a 1 min, 72°C extension step (1 min). A final 10 min extension step at 72°C followed the last cycle before PCR products were visualized using gel electrophoresis.

Arabidopsis thaliana plants homozygous for their annotated T-DNA insertion were grown for 20 days in baby food jars containing sterile Gamborg's B5 medium before cutting their roots into 3–5 mm segments. The segments were assayed as described in Tenea et al. (2009). *A. tumefaciens* At849 (GV3101:pMP90 containing pBISN1) was used to measure transient transformation, whereas *A. tumefaciens* A208 (Sciaky et al., 1978) was used for stable transformation (Supplementary Table 2). Three replicates were assayed for each experiment with root segments from 10 plants pooled for each replicate. A minimum of 80 root segments were scored for each data point and statistical analysis was performed using ANOVA.

Protein Isolation and Proteomics Analysis

Roots were homogenized in 8 M urea using a Percellys® 24 homogenizer (Bertin) and incubated at room temperature for 1 h with continuous vortexing before centrifugation at 14,000 rpm for 15 min at 4°C. The supernatant solution was transferred to a new tube and the protein concentration was determined using a Pierce™ BCA assay (Thermo Fisher Scientific). A total of 100 µg protein from each sample (equivalent volume) was taken for digestion. Proteins were first precipitated using four volumes of cold acetone (−20°C) overnight before centrifugation at 14,000 rpm for 15 min at 4°C to collect the precipitated proteins. Protein pellets were washed twice with 80% cold (−20°C) acetone, dried in a speed-vac for 5 min, and then solubilized in 8 M urea. Samples were reduced using 10 mM dithiothreitol and cysteine alkylated using 20 mM iodoacetamide. This was followed by digestion using sequence grade Lys-C/Trypsin (Promega) mix at a 1:25 (enzyme:substrate) ratio to enzymatically digest the proteins. All digestions were carried out at 37°C overnight. The samples were cleaned over C18 MicroSpin columns (Nest Group), dried, and resuspended in 97% purified H₂O/3% acetonitrile (ACN)/0.1% formic acid (FA). After BCA at the peptide level, 1 µg of each sample was loaded onto the column.

Digested samples were analyzed using a Dionex UltiMate 3000 RSLC Nano System coupled with a Q Exactive™ HF Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific, Waltham, MA, United States). Peptides were first loaded onto a 300 µm × 5 mm C18 PepMap™ 100 trap column and washed with 98% purified water/2% acetonitrile (ACN)/0.01% formic acid (FA) using a flow rate of 5 µL/min. After 5 min, the trap column was switched in-line with a 75 µm × 50 cm reverse phase Acclaim™ PepMap™ RSLC C18 analytical column heated to 50°C. Peptides were separated over the analytical column using a 120 min method at a flow rate of 300 nL min^{−1}. Mobile phase A contained 0.01% FA in purified water while mobile phase B consisted of 0.01% FA/80% ACN in

purified water. The linear gradient began at 2% B and reached 10% B in 5 min, 30% B in 80 min, 45% B in 91 min, and 100% B in 93 min. The column was held at 100% B for the next 5 min before returning to 5% B where it was equilibrated for 20 min. Samples were injected into the QE HF through the Nanospray Flex™ Ion Source fitted with an emitter tip from New Objective. MS spectra were collected from 400 to 1600 m/z at 120,000 resolution, a maximum injection time of 100 ms, and a dynamic exclusion of 15 s. The top 20 precursors were fragmented using higher-energy C-trap dissociation (HCD) at a normalized collision energy of 27%. MS/MS spectra were acquired in the Orbitrap at a resolution of 15,000 with a maximum injection time of 20 ms. The raw data were analyzed using MaxQuant software (v. 1.5.3.28) against a TAIR 10 protein database combined with VirE2 proteins (Cox and Mann, 2008; Cox et al., 2011, 2014). The search was performed with the precursor mass tolerance set to 10 ppm and MS/MS fragment ions tolerance was set to 20 ppm. The enzyme was set to trypsin and LysC, allowing up to two missed cleavages. Oxidation of methionine was defined as a variable modification, and carbamidomethylation of cysteine was defined as a fixed modification. The “unique plus razor peptides” (razor peptides are the non-unique peptides assigned to the protein group with the most other peptides) were used for peptide quantitation. The false discovery rate (FDR) of peptides and proteins identification was set at 0.01. iBAQ scores and MS/MS counts for each identified protein were compared between the non-induced and induced samples. Proteins which showed a 0.2-fold (20%) increase or decrease in abundance in the induced versus non-induced samples for at least two biological replicates by comparing both iBAQ scores and MS/MS counts were considered to have levels which changed in response to VirE2 induction.

DATA AVAILABILITY STATEMENT

The RNA-seq data presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NCBI GEO with the accession number GSE172314.

AUTHOR CONTRIBUTIONS

SG and RL designed the research and wrote the manuscript with input from all authors. RL, L-YL, and SG analyzed data. All authors performed the experiments.

FUNDING

RL was supported by fellowships from Purdue University (Frederick N. Andrews Graduate Assistantship, Purdue Research Foundation Graduate Award, and Bilsland Dissertation Fellowship). This work was supported by grants from the

National Science Foundation to SG and a P30 grant (CA023168) to the Purdue Cancer Center.

ACKNOWLEDGMENTS

We thank Fang-Yu Hsu for plasmid construction and generating *Arabidopsis* lines containing the inducible VirE2-Venus and VirE2-Venus-NLS genes, and Walt Ream for critical reading of an early version of the manuscript. We acknowledge the help of the Purdue University Bindley Imaging Facility, the Purdue University Genomics Core Facility, The Cornell University Institute of Biotechnology Genomics Facility, the

Purdue University Bioinformatics Core Facility, and the Purdue University Proteomics Core Facility. We especially thank Uma Aryal for assistance with the analysis and interpretation of the proteomics data, and Andy Schaber for valuable advice for confocal imaging.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.684192/full#supplementary-material>

REFERENCES

- Anand, A., Krichevsky, A., Schornack, S., Lahaye, T., Tzfira, T., Tang, Y., et al. (2007). *Arabidopsis* VirE2 INTERACTING PROTEIN 2 is required for *Agrobacterium* T-DNA integration in plants. *Plant Cell* 19, 1695–1708. doi: 10.1105/tpc.106.042903
- Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: Tool for the unification of biology. *Nat. Genet.* 25, 25–29.
- Bakó, L., Umeda, M., Tiburcio, A. F., Schell, J., and Koncz, C. (2003). The VirD2 pilot protein of *Agrobacterium*-transferred DNA interacts with the TATA box-binding protein and a nuclear protein kinase in plants. *Proc. Natl. Acad. Sci. USA* 100, 10108–10113. doi: 10.1073/pnas.1733208100
- Bhattacharjee, S., Lee, L.-Y., Oltmanns, H., Cao, H., Veena, Cuperus, J., et al. (2008). AtImpa-4, an *Arabidopsis* importin α isoform, is preferentially involved in *Agrobacterium*-mediated plant transformation. *Plant Cell* 20, 2661–2680. doi: 10.1105/tpc.108.060467
- Chang, C. W., Counago, R. L., Williams, S. J., Boden, M., and Kobe, B. (2014). Crystal structure of rice importin- α and structural basis of its interaction with plant-specific nuclear localization signals. *Plant Cell* 24, 5074–5088. doi: 10.1105/tpc.112.104422
- Christie, P. J., Ward, J. E., Winans, S. C., and Nester, E. W. (1988). The *Agrobacterium tumefaciens* virE2 gene product is a single-stranded-DNA-binding protein that associates with T-DNA. *J. Bacteriol.* 170, 2659–2667.
- Citovsky, V., De Vos, G., and Zambryski, P. (1988). Single-stranded DNA binding protein encoded by the *virE* locus of *Agrobacterium tumefaciens*. *Science* 240, 501–504. doi: 10.1126/science.240.4851.501
- Citovsky, V., Wong, M. L., and Zambryski, P. (1989). Cooperative interaction of *Agrobacterium* VirE2 protein with single-stranded DNA: implications for the T-DNA transfer process. *Proc. Natl. Acad. Sci. USA* 86, 1193–1197. doi: 10.1073/pnas.86.4.1193
- Citovsky, V., Zupan, J., Warnick, D., and Zambryski, P. (1992). Nuclear localization of *Agrobacterium* VirE2 protein in plant cells. *Science* 256, 1802–1805. doi: 10.1126/science.1615325
- Citovsky, V., Warnick, D., and Zambryski, P. (1994). Nuclear import of *Agrobacterium* VirD2 and VirE2 proteins in maize and tobacco. *Proc. Natl. Acad. Sci. USA* 91, 3210–3214. doi: 10.1073/pnas.91.8.3210
- Citovsky, V., Kapelnikov, A., Oliel, S., Zakai, N., Rojas, M. R., and Gilbertson, R. L. (2004). Protein interactions involved in nuclear import of the *Agrobacterium* VirE2 protein in vivo and in vitro. *J. Biol. Chem.* 279, 29528–29533. doi: 10.1074/jbc.m403159200
- Clough, S. J., and Bent, A. F. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743. doi: 10.1046/j.1365-3113x.1998.00343.x
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372. doi: 10.1038/nbt.1511
- Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011). Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* 10, 1794–1805. doi: 10.1021/pr101065j
- Cox, J., Hein, M. Y., Lubber, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteomics* 13, 2513–2526. doi: 10.1074/mcp.m113.031591
- Crane, Y. M., and Gelvin, S. B. (2007). RNAi-mediated gene silencing reveals involvement of *Arabidopsis* chromatin-related genes in *Agrobacterium*-mediated root transformation. *Proc. Natl. Acad. Sci. USA* 104, 15156–15161. doi: 10.1073/pnas.0706986104
- Das, A. (1988). *Agrobacterium tumefaciens* virE operon encodes a single-stranded DNA-binding protein. *Proc. Natl. Acad. Sci. USA* 85, 2909–2913. doi: 10.1073/pnas.85.9.2909
- Deng, W., Chen, L., Wood, D. W., Metcalfe, T., Liang, X., Gordon, M. P., et al. (1998). *Agrobacterium* VirD2 protein interacts with plant host cyclophilins. *Proc. Natl. Acad. Sci. USA* 95, 7040–7045. doi: 10.1073/pnas.95.12.7040
- Ditt, R. F., Kerr, K. F., de Figueiredo, P., Delrow, J., Comai, L., and Nester, E. W. (2006). The *Arabidopsis thaliana* transcriptome in response to *Agrobacterium tumefaciens*. *Mol. Plant-Microbe Interact.* 19, 665–681. doi: 10.1094/mpmi-19-0665
- Djamei, A., Pitzschke, A., Nakagami, H., Rajh, I., and Hirt, H. (2007). Trojan horse strategy in *Agrobacterium* transformation: Abusing MAPK defense signaling. *Science* 318, 453–456. doi: 10.1126/science.1148110
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- Dombek, P., and Ream, W. (1997). Functional domains of *Agrobacterium tumefaciens* single-stranded DNA-binding protein VirE2. *J. Bacteriol.* 179, 1165–1173. doi: 10.1128/jb.179.4.1165-1173.1997
- Du, Z., Zhou, X., Ling, Y., Zhang, Z., and Su, Z. (2010). agriGO: A GO analysis toolkit for the agricultural community. *Nucl. Acids Res.* 2010:310. doi: 10.1093/nar/gkq310
- Duan, K., Willig, C. J., De Tar, J. R., Spollen, W. G., and Zhang, Z. J. (2018). Transcriptomic analysis of *Arabidopsis* seedlings in response to *Agrobacterium*-mediated transformation process. *Mol. Plant-Microbe Interact.* 31, 445–459. doi: 10.1094/mpmi-10-17-0249-r
- Gaspar, Y. M., Nam, J., Schultz, C. J., Lee, L.-Y., Gilson, P. R., Gelvin, S. B., et al. (2004). Characterization of the *Arabidopsis* lysine-rich arabinogalactan-protein AtAGP17 mutant (*rat1*) that results in a decreased efficiency of *Agrobacterium* transformation. *Plant Physiol.* 135, 2162–2171. doi: 10.1104/pp.104.045542
- Gelvin, S. B. (2003). *Agrobacterium*-mediated plant transformation: The biology behind the “gene-jockeying” tool. *Microbiol. Mol. Biol. Rev.* 67, 16–37. doi: 10.1128/mmbr.67.1.16-37.2003
- Gelvin, S. B. (2012). Traversing the cell: *Agrobacterium* T-DNA’s journey to the host genome. *Front. Plant Sci.* 3, 1–11. doi: 10.3389/fpls.2012.00052
- Gietl, C., Koukolikova-Nicola, Z., and Hohn, B. (1987). Mobilization of T-DNA from *Agrobacterium* to plant cells involves a protein that binds single-stranded DNA. *Proc. Natl. Acad. Sci. USA* 84, 9006–9010. doi: 10.1073/pnas.84.24.9006

- Grange, W., Duckely, M., Husale, S., Jacob, S., Engel, A., and Hegner, M. (2008). VirE2: A unique ssDNA-compacting molecular machine. *PLoS Biol.* 6:e44. doi: 10.1371/journal.pbio.0060044
- Hiei, Y., Ishida, Y., and Komari, T. (2014). Progress of cereal transformation technology mediated by *Agrobacterium tumefaciens*. *Front. Plant Sci.* 5:628. doi: 10.3389/fpls.2014.00628
- Howard, E., and Citovsky, V. (1990). The emerging structure of the *Agrobacterium* T-DNA transfer complex. *BioEssays* 12, 103–108. doi: 10.1002/bies.950120302
- Koncz, C., and Schell, J. (1986). The promoter of T_L-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol. Gen. Genet.* 204, 383–396. doi: 10.1007/bf00331014
- Krueger, F. (2017). Babraham Bioinformatics. *Trim Galore!* 2017, 0153. doi: 10.1007/s10439010-0153-9.Impact
- Lacroix, B., and Citovsky, V. (2013). The roles of bacterial and host plant factors in *Agrobacterium*-mediated genetic transformation. *Int. J. Dev. Biol.* 57, 467–481. doi: 10.1387/ijdb.130199bl
- Lacroix, B., and Citovsky, V. (2019). Pathways of DNA transfer to plants from *Agrobacterium tumefaciens* and related bacterial species. *Annu. Rev. Phytopathol.* 57, 231–251.
- Lapham, R., Lee, L.-Y., Tsugama, D., Lee, S., Mengiste, T., and Gelvin, S. B. (2018). VIP1 and its homologs are not required for *Agrobacterium*-mediated transformation, but play a role in *Botrytis* and salt stress responses. *Front. Plant Sci.* 9:1–15. doi: 10.3389/fpls.2018.00749
- Lee, L.-Y., Fang, M.-J., Kuang, L.-Y., and Gelvin, S. B. (2008). Vectors for multi-color bimolecular fluorescence complementation to investigate protein-protein interactions in living plant cells. *Plant Methods* 4:24. doi: 10.1186/1746-4811-4-24
- Lee, L.-Y., Wu, F.-H., Hsu, C.-T., Shen, S.-C., Yeh, H.-Y., Liao, D.-C., et al. (2012). Screening of a cDNA library for protein-protein interactions directly in planta. *Plant Cell* 24, 1746–1759. doi: 10.1105/tpc.112.097998
- Li, J., Krichevsky, A., Vaidya, M., Tzifira, T., and Citovsky, V. (2005). Uncoupling of the functions of the *Arabidopsis* VIP1 protein in transient and stable plant genetic transformation by *Agrobacterium*. *Proc. Natl. Acad. Sci. USA* 102, 5733–5738. doi: 10.1073/pnas.0404118102
- Li, X., and Pan, S. Q. (2017). *Agrobacterium* delivers VirE2 protein into host cells via clathrin-mediated endocytosis. *Sci. Adv.* 3, e1601528. doi: 10.1126/sciadv.1601528
- Li, X., Yang, Q., Tu, H., Lim, Z., and Pan, S. Q. (2014). Direct visualization of *Agrobacterium*-delivered VirE2 in recipient cells. *Plant J.* 77, 487–495. doi: 10.1111/tpj.12397
- Li, X., Yang, Q., Peng, L., Tu, H., Lee, L.-Y., Gelvin, S. B., et al. (2020). *Agrobacterium*-delivered VirE2 interacts with host nucleoporin CG1 to facilitate the nuclear import of VirE2-coated T-complex. *Proc. Natl. Acad. Sci. USA* 117, 26389–26397. doi: 10.1073/pnas.2009645117
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014, 0550. doi: 10.1186/s13059-014-0550-8
- Nam, J., Mysore, K. S., Zheng, C., Knue, M. K., Matthysee, A. G., and Gelvin, S. B. (1999). Identification of T-DNA tagged *Arabidopsis* mutants that are resistant to transformation by *Agrobacterium*. *Mol. Gen. Genet.* 261, 429–438. doi: 10.1007/s004380050985
- Narasimhulu, S. B., Deng, X.-B., Sarria, R., and Gelvin, S. B. (1996). Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *Plant Cell* 8, 873–886. doi: 10.2307/3870289
- Nester, E. W. (2015). *Agrobacterium*: Nature's genetic engineer. *Front. Plant Sci.* 5:730. doi: 10.3389/fpls.2014.00730
- Park, S.-Y., Yin, X., Duan, K., Gelvin, S. B., and Zhang, Z. J. (2014). Heat shock protein 90.1 plays a role in *Agrobacterium*-mediated plant transformation. *Mol. Plant J.* 1793–1796. doi: 10.1093/mp/ssu091
- Pitzschke, A., Djamei, A., Teige, M., and Hirt, H. (2009). VIP1 response elements mediate mitogen-activated protein kinase 3-induced stress gene expression. *Proc. Natl. Acad. Sci. USA* 106, 18414–18419. doi: 10.1073/pnas.0905599106
- Pitzschke, A., and Hirt, H. (2010). New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. *EMBO J.* 29, 1021–1032. doi: 10.1038/emboj.2010.8
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Rossi, L., Hohn, B., and Tinland, B. (1996). Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. USA* 93, 126–130. doi: 10.1073/pnas.93.1.126
- Roushan, M. R., de Zeeuw, M. A. M., Hooykaas, P. J. J., and van Heusden, G. P. H. (2018). Application of phiLOV2.1 as a fluorescent marker for visualization of *Agrobacterium* effector protein translocation. *Plant J.* 96, 685–699. doi: 10.1111/tpj.14060
- Schultz, C. J., Rumsewicz, M. P., Johnson, K. L., Jones, B. J., Gaspar, Y. M., and Bacic, A. (2002). Using genomic resources to guide research directions. The arabinogalactan-protein gene family as a test case. *Plant Physiol.* 129, 1448–1463. doi: 10.1104/pp.003459
- Sciaky, D., Montoya, A. L., and Chilton, M.-D. (1978). Fingerprints of *Agrobacterium* Ti plasmids. *Plasmid* 1, 238–253. doi: 10.1016/0147-619x(78)90042-2
- Sen, P., Pazour, G. J., Anderson, D., and Das, A. (1989). Cooperative binding of *Agrobacterium tumefaciens* VirE2 protein to single-stranded DNA. *J. Bacteriol.* 171, 2573–2580. doi: 10.1128/jb.171.5.2573-2580.1989
- Shi, Y., Lee, L.-Y., and Gelvin, S. B. (2014). Is VIP1 important for *Agrobacterium*-mediated transformation? *Plant J.* 79, 848–860. doi: 10.1111/tpj.12596
- Simone, M., McCullen, C. A., Stahl, L. E., and Binns, A. N. (2001). The carboxy-terminus of VirE2 from *Agrobacterium tumefaciens* is required for its transport to host cells by the virB- encoded type IV transport system. *Mol. Microbiol.* 41, 1283–1293. doi: 10.1046/j.1365-2958.2001.02582.x
- Stachel, S. E., and Nester, E. W. (1986). The genetic and transcriptional organization of the vir region of the A6 Ti plasmid of *Agrobacterium tumefaciens*. *EMBO J.* 5, 1445–1454. doi: 10.1002/j.1460-2075.1986.tb04381.x
- Tao, Y., Rao, P. K., Bhattacharjee, S., and Gelvin, S. B. (2004). Expression of plant protein phosphatase 2C interferes with nuclear import of the *Agrobacterium* T-complex protein VirD2. *Proc. Natl. Acad. Sci. USA* 101, 5164–5169. doi: 10.1073/pnas.0300084101
- Tenea, G. N., Spantzel, J., Lee, L.-Y., Zhu, Y., Lin, K., Johnson, S. J., et al. (2009). Over-expression of several *Arabidopsis* histone genes increases *Agrobacterium*-mediated transformation and transgene expression in plants. *Plant Cell* 21, 3350–3367. doi: 10.1105/tpc.109.070607
- Tinland, B., Hohn, B., and Puchta, H. (1994). *Agrobacterium tumefaciens* transfers single-stranded transferred DNA (T-DNA) into the plant cell nucleus. *Proc. Natl. Acad. Sci. USA* 91, 8000–8004. doi: 10.1073/pnas.91.17.8000
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016
- Tsugama, D., Liu, S., and Takano, T. (2012). A bZIP protein, VIP1, is a regulator of osmosensory signaling in *Arabidopsis*. *Plant Physiol.* 159, 144–155. doi: 10.1104/pp.112.197020
- Tsugama, D., Liu, S., and Takano, T. (2014). Analysis of functions of VIP1 and its close homologs in osmosensory responses of *Arabidopsis thaliana*. *PLoS One* 9:e103930. doi: 10.1371/journal.pone.0103930
- Tsugama, D., Liu, S., and Takano, T. (2016). The bZIP protein VIP1 is involved in touch responses in *Arabidopsis* roots. *Plant Physiol.* 171, 1355–1365.
- Tzifira, T., Vaidya, M., and Citovsky, V. (2001). VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. *EMBO J.* 20, 3596–3607. doi: 10.1093/emboj/20.13.3596
- Tzifira, T., and Citovsky, V. (2001). Comparison between nuclear localization of nopaline- and octopine-specific *Agrobacterium* VirE2 proteins in plant, yeast and mammalian cells. *Mol. Plant Pathol.* 2, 171–176. doi: 10.1046/j.1364-3703.2001.00065.x
- Van Eck, J. (2018). Genome editing and plant transformation of solanaceous food crops. *Curr. Opin. Biotechnol.* 49, 35–41. doi: 10.1016/j.copbio.2017.07.012
- Van Larebeke, N., Engler, G., Holsters, M., Van Den Elzacker, S., Zaenen, I., Schilperoort, R. A., et al. (1974). Large plasmid in *Agrobacterium tumefaciens*

- essential for crown gall-inducing ability. *Nature* 252, 169–170. doi: 10.1038/252169a0
- Veena, Jiang, H., Doerge, R. W., and Gelvin, S. B. (2003). Transfer of T-DNA and Vir proteins to plant cells by *Agrobacterium tumefaciens* induces expression of host genes involved in mediating transformation and suppresses host defense gene expression. *Plant J.* 35, 219–236. doi: 10.1046/j.1365-313x.2003.01796.x
- Wu, L., Chen, H., Curtis, C., and Fu, Z. Q. (2014). Go in for the kill: How plants deploy effector-triggered immunity to combat pathogens. *Virulence* 5, 710–721. doi: 10.4161/viru.29755
- Yusibov, V. M., Steck, T. R., Gupta, V., and Gelvin, S. B. (1994). Association of single-stranded transferred DNA from *Agrobacterium tumefaciens* with tobacco cells. *Proc. Natl. Acad. Sci. USA* 91, 2994–2998. doi: 10.1073/pnas.91.8.2994
- Zheng, Z., Qamar, S. A., Chen, Z., and Mengiste, T. (2006). *Arabidopsis* WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J.* 48, 592–605. doi: 10.1111/j.1365-313x.2006.02901.x
- Ziemienowicz, A., Merkle, T., Schoumacher, F., Hohn, B., and Rossi, L. (2001). Import of *Agrobacterium* T-DNA into plant nuclei: two distinct functions of VirD2 and VirE2 proteins. *Plant Cell* 13, 369–383. doi: 10.2307/3871282
- Zuo, J., Niu, Q.-W., and Chua, N.-H. (2000). An estrogen receptor-based system transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J.* 24, 265–227. doi: 10.1046/j.1365-313x.2000.00868.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Lapham, Lee, Xhako, Gómez, Nivya and Gelvin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.