Transcriptome Profiling of Plant Genes in Response to *Agrobacterium tumefaciens*-Mediated Transformation

Christopher J. Willig, Kaixuan Duan and Zhanyuan J. Zhang

Contents

Abstract *Agrobacterium tumefaciens* is a plant pathogen that causes crown gall disease. During infection of the host plant, *Agrobacterium* transfers T-DNA from its Ti plasmid into the host cell, which can then be integrated into the host genome. This unique genetic transformation capability has been employed as the dominant technology for producing genetically modified plants for both basic research and biotechnological applications. *Agrobacterium* has been well studied as a disease-causing agent. The *Agrobacterium*-mediated transformation process involves early attachment of the bacterium to the host's surface, followed by transfer of T-DNA and virulence proteins into the plant cell. Throughout this process, the host plants exhibit dynamic gene expression patterns at each infection

C. J. Willig · K. Duan · Z. J. Zhang (✉)
Division of Plant Sciences, University of Missouri, Columbia, MO, USA
e-mail: zhangzh@missouri.edu

Current Topics in Microbiology and Immunology (2018) 418:319–348
DOI 10.1007/82_2018_115

Published Online: 31 July 2018

stage or in response to *Agrobacterium* strains with varying pathogenic capabilities. Shifting host gene expression patterns throughout the transformation process have effects on transformation frequency, host morphology, and metabolism. Thus, gene expression profiling during the *Agrobacterium* infection process can be an important approach to help elucidate the interaction between *Agrobacterium* and plants. This review highlights recent findings on host plant differential gene expression patterns in response to *A. tumefaciens* or related elicitor molecules.

1 Introduction

Agrobacterium tumefaciens is a gram-negative soil-borne bacterium that is the causative agent of crown gall disease, which affects a wide range of host species (DeCleene and DeLay 1976). Through the course of its infection of a host plant, *Agrobacterium* mobilizes a single-strand DNA segment originating from a sequence located on its tumor-inducing (Ti) plasmid, referred to as transfer DNA (or T-DNA), into host plant cells (Chilton et al. 1977). Integration of T-DNA into the genome of the host and the subsequent expression of the gene products it codes for leads to the formation of tumorous growths that are characteristic of crown gall disease (Escobar and Dandekar 2003). This uncommon ability to transmit DNA sequences into hosts and modify their gene expression as part of its infection strategy has made *Agrobacterium* an important tool in the development of transgenic plants for crop breeding and basic research through the utilization of a method known as *Agrobacterium*-mediated transformation (Azpiroz-Leehan and Feldmann 1997). The close contact between pathogen and host throughout the different stages of infection gives rise to heavily intertwined and multi-layered molecular interactions among them. The successful genomic integration and expression of T-DNA in a host is dependent on a few different steps occurring during pathogenesis including signal perception, expression of pathogen virulence genes, attachment of *Agrobacterium* to the host cell, and T-DNA processing (the presence of a Ti plasmid is required for transfer of T-DNA), transport of T-DNA and accessory proteins into the host cell, nuclear import, and finally uncoating of the T-DNA strand and recombination with the host genome (Gelvin 2000). Molecular interactions at any one of these stages can affect the success of *Agrobacterium* at infecting the plant. To attempt to understand these interactions, we need to examine each stage in greater detail.

In nature, the infection process begins with the production of phenolic compounds by a potential host. This may be caused by environmental stress or damage due to wounding. Perception of plant-derived phenolic molecules can induce the expression of a suite of genes known as the *vir* regulon, which is housed on the Ti plasmid (Bhattacharya et al. 2010). The *Agrobacterium* cell attaches to the surface of the host plant, mediated by pathogen-secreted polysaccharide and protein adhesin compounds. Following attachment, *Agrobacterium* cells become sessile. Although the Ti plasmid is required for T-DNA transfer and tumorigenesis, it is not strictly required for this polar attachment process (Tomlinson and Fuqua 2009). It is at this stage that the plant has its first opportunity for defense against *Agrobacterium* infection. Some

bacteria-derived compounds known as pathogen-associated molecular patterns (PAMPs) may be perceived by the host, thereby invoking a quick, but relatively weak immune response termed PAMP-triggered immunity (PTI) (Zipfel and Robatzek 2010). The most well-studied *Agrobacterium*-derived PAMP that elicits a PTI response is EF-Tu, which was found to be recognized by the host receptor kinase protein EFR (EF-Tu receptor). In addition to EF-Tu, other yet to be characterized PAMPs may play a role in the molecular interactions at this stage (Zipfel et al. 2006).

Provided that a virulent strain of *Agrobacterium* has undergone attachment to the plant surface and the expression of *vir* genes has been induced, the T-DNA segment on the Ti plasmid is nicked by VirD2 at right and left border sequences and becomes covalently attached to this VirD2 protein at the 5' end of nicked T-strand (Mysore et al. 1998). The T-strand is shuttled through a type IV secretion system into the cytoplasm of a host plant cell where it is thought to be coated by molecules of the VirE2 protein (Ziemienowicz et al. 2001). Mutually bound VirD2, single-stranded T-DNA, and VirE2 are believed to form what is termed the "T-complex" (Vergunst et al. 2000). In addition to VirD2 and VirE2, other virulence proteins are known to be delivered into the host cell including VirE3, VirD5, and VirF. (Thompson et al. 1988; Ward et al. 1988, 2002; Kuldau et al. 1990; Shirasu et al. 1990; Beijersbergen et al. 1994; Christie and Vogel 2000; Schrammeijer et al. 2003). VirE2 contains a plant-active nuclear localization signal and has been shown to interact with the plant protein VIP1, which was initially thought to be involved in T-DNA integration into the genome (Tzfira and Citovsky 2001; Djamei et al. 2007). However, its participation in this process has been questioned in more recent studies (Shi et al. 2014; Lapham et al. 2018). The T-DNA from wild-type strains of *Agrobacterium* encodes the plant-active genes *iaaH*, *iaaM*, and *ipt*, which induce the biosynthesis of auxin and cytokinin. The expression of these genes gives rise to the formation of crown galls (Morris 1986; Binns and Costantino 1998). For the purpose of adapting T-DNA transfer for biotechnological applications, *Agrobacterium* strains which have had their Ti plasmids "disarmed" (i.e., non-tumorigenic) are used (Barton and Brill 1983). This means that the oncogenes present in the T-DNA segment of the plasmid have been deleted, whereas the *vir* genes remain. Therefore, non-tumorigenic strains of *Agrobacterium* often used in the laboratory still transfer proteins into host cells via a type IV secretion system, but do not lead to the formation of crown galls.

Bacterial pathogens other than *Agrobacterium* are known to transfer proteins into the plant cell during infection, which may serve to dampen the PTI response. However, host plants are often able to perceive these foreign proteins (termed effectors), thereby initiating a strong defense response called effector-triggered immunity (ETI) (Cui et al. 2009). This series of defenses include the hypersensitive response (HR) consisting of rapid programmed cell death, which prevents the spread of the disease to uninfected tissue. However, *Agrobacterium* is able to avoid eliciting HR in most plants despite its transfer of foreign Vir proteins. This is a key advantage *Agrobacterium* has over other plant bacterial pathogens (Pu and Goodman 1993; Staskawicz et al. 1995; Wood et al. 2001). Although the secreted Vir proteins seem not to elicit an immune response, they have been shown in several cases to have *in planta* functions that could modulate host gene expression patterns (Tzfira et al. 2004; Lacroix et al. 2005; García-Rodríguez et al. 2006; Magori and Citovsky 2011; Wang et al. 2014; Niu et al. 2015;

Zhang et al. 2017). In order to examine the differences between the effects caused by *Agrobacterium* PAMPs or other non-transferred factors and those caused by the *in planta* activities of the Vir proteins, some of the studies reviewed here utilized "cured" strains of *Agrobacterium*, meaning that they no longer contain a Ti plasmid and thus cannot express or transfer Vir proteins (Watson et al. 1975). Strains lacking Ti plasmids are often referred to in the literature as "avirulent" in contrast to "virulent" strains that contain a disarmed Ti plasmid (Veena et al. 2003). Both of these are considered non-tumorigenic as they cannot induce crown gall formation. To date, many functions of the *Agrobacterium* genes involved in the pathogen-host interaction are known and most of the Vir proteins have been well-characterized (Tzfira and Citovsky 2006). By contrast, many of the host plant factors that are directly involved in the transformation process, including attachment, pathogen recognition, T-DNA transfer, trafficking through the cell cytoplasm and integration of T-DNA into the host genome remain unidentified (Gelvin 2010, 2017; Lacroix and Citovsky 2013).

One of the many approaches that have been implemented to understand the molecular players and large-scale processes involved in transformation is differential gene expression analysis. The basic premise behind this analysis is that by contrasting the transcript levels of genes between two different conditions or treatments, one can elucidate which genes are altered by the treatment and thus might be involved in facilitating or suppressing a given process. The technology for evaluating transcript expression has advanced exponentially over the past 20 years, resulting in drastically decreased cost and labor and increased capacity for data generation. Each of the techniques that have been developed generally involves the production of a cDNA pool generated by reverse transcription of total mRNA extracts from plant tissues. The experimental methodologies utilized in the studies reviewed here include cDNA-AFLP (amplified fragment length polymorphism), RT-PCR (reverse transcription), suppression subtractive hybridization, DNA macroarrays, microarrays, quantitative real-time RT-PCR, and RNA-Seq. For a detailed review on these methods and more, see Casassola et al. (2013). The widespread use of microarrays and RNA-Seq has led to the accumulation of massive amounts transcriptional data as there can be coverage over nearly all of the genome in one experiment yielding hundreds or thousands of differentially expressed genes. As a result, it has become popular to analyze the data generated by such experiments by classifying them according to the probable molecular or biochemical function, biological process, or cellular compartment of their predicted gene product using tools such as Gene Ontology (GO), MapMan, and the Kyoto Encyclopedia of Genes and Genomes (KEGG). This analysis serves to give an overall picture of systems in the organism that are altered, dependent on the experimental treatment.

Here, we review the findings on the gene expression dynamics of plants during *Agrobacterium*-mediated transformation upon heterologous expression of *Agrobacterium*derived virulence proteins in the host, and upon treatment with PAMPs, with an emphasis on the commonalities and differences between *Agrobacterium* and other plant bacterial pathogens. We will then discuss the relative benefits and drawbacks of using modern transcriptomic methodologies to examine *Agrobacterium*–plant interactions. For ease of reference, Tables 1, 2, 3, and 4 summarizing the reviewed transcriptomic studies are listed at the end of this chapter.

2 Differential Gene Expression in Response to Inoculation with Various *Agrobacterium Tumefaciens* Strains

2.1 Transcriptomic Analysis of Model Plants During Agrobacterium-Mediated Transformation

Several studies have been carried out over the last 20 years that utilized different methods for detection of differential gene expression in response to inoculation with *Agrobacterium*. In a few cases, they have attempted to untangle the unique responses relating to particular steps of *Agrobacterium* infection and T-DNA transfer by various means. These include contrasting *Agrobacterium* response with that of other bacterial organisms, contrasting the responses to different *Agrobacterium* strains that are necessarily arrested at some point in the infection process owing to their genotype (including using tumorigenic versus non-tumorigenic strains), and temporally analyzing different infection stages during a time course. Other important differences exist among these studies with respect to plant species/genotype, tissue type sampled, culture conditions, inoculation method, *Agrobacterium* strain(s) used, and experimental design, which could account for some of the variability in results. Therefore, these details will be thoroughly outlined for each study.

In one of the earliest studies on the plant response to *Agrobacterium* (Ditt et al. 2001), researchers used a cell suspension culture of the tropical plant *Ageratum conyzoides* to construct a cDNA library of 16,000 sequences and then implemented cDNA-AFLP to identify cDNA sequences that were differentially regulated in response to the non-tumorigenic *Agrobacterium* strain EHA105. A total of 179 unique gene fragments were upregulated in response to *Agrobacterium*, whereas 72 were downregulated. Twenty of the most strongly induced of these sequences were used as queries to search for sequence similarity to other species. Top gene candidates in these similarity searches were involved in cellular functions such as signal perception, signal transduction, and defense. Using RT-PCR as an independent gene expression method, it was shown that four of the genes predicted to function in defense response were also induced by treatment with non-pathogenic *E. coli* cells, whereas two genes (one encoding a protein similar to a nodulin from *Oryza sativa*) and one encoding a protein similar to a lectin-like protein kinase from *Populus nigra* showed an *Agrobacterium*-specific response. This same research group later used the same experimental system to evaluate the expression of three of the putative defense genes identified in this study in response to the attachment of *Agrobacterium* to the host (Ditt et al. 2005). Plant cell cultures were infected with four different non-tumorigenic strains of *Agrobacterium*: EHA105 harboring the T-DNA binary vector pBISN1 (T-DNA transfer competent), LBA4404 (carrying a disarmed Ti plasmid, but no T-DNA), A136 (avirulent; no Ti plasmid) and *chvB* (contains an oncogenic Ti plasmid, but is attachment-deficient mutant). RT-PCR revealed that the ability of the *Agrobacterium* to attach to host cells had a negative effect on the expression levels of the three defense genes being measured compared to the attachment-deficient strain, which induced expression levels similar to that of

the non-pathogenic bacterial control. The authors suggested that exopolysaccharide (a product of the mutated gene in *chvB*) may play a role in suppressing the plant defense response.

In yet another published study by this group, the authors used a microarray to analyze wide-scale gene expression in suspension cell cultures of the *Arabidopsis thaliana* Ler-0 ecotype in response to infection by the tumorigenic *Agrobacterium* strain A348 along a time course (Ditt et al. 2006). Although this study used a 26,000 oligonucleotide array, surprisingly, no statistically significant differentially expressed genes were found at the first three time-point comparisons with the mock control (4, 12, and 24 h post-inoculation). In the comparison for the samples collected 48 h after infection, 303 differentially regulated genes were identified that showed consistency among the four experimental replications. Of these, 115 were upregulated and 188 were downregulated. GO biological process category analysis was performed on these two gene sets. The authors reported that genes falling into the "cell organization and biogenesis" and "protein metabolism" terms were overrepresented among the downregulated genes and "electron transport or energy pathways," "response to abiotic or biotic stimulus," and "response to stress" were overrepresented among the upregulated genes. They also compared their microarray data with other publicly available data sets evaluating the transcriptomic response of *Arabidopsis* to various other plant pathogens and to treatment with auxin compounds. Of these comparisons, the one that had the most overlapping differentially expressed genes with the *Agrobacterium* data set (at 53 genes) was one from mature *Arabidopsis* leaves that had been infiltrated with *Pseudomonas syringae* (Tao et al. 2003).

Working in a tobacco BY-2 suspension cell culture system, researchers in a 2003 study used suppression subtractive hybridization and macroarrays to determine changes in gene expression in host cells inoculated with *Agrobacterium* (Veena et al. 2003). Cells were exposed to both virulent and avirulent non-tumorigenic *Agrobacterium* strains with or without the capability to transfer Vir proteins or to transfer T-DNA containing a GUS-intron reporter gene. They found that a suite of genes related to defense responses including glutathione-S-transferase genes and alcohol dehydrogenase were induced in the earlier time-points following infection by *Agrobacterium*, regardless of its ability to transfer T-DNA. However, during the later time-points the relative expression of these genes was higher among cells infected by the avirulent *Agrobacterium* strain. Another interesting finding was that genes associated with cell division and plant growth such as the core histone gene family members encoding H2A, H2B, H3, and H4 along with ribosomal proteins were increased in their expression levels by exposure to the virulent strain at the later time-points after infection significantly above the levels induced by the avirulent strain at the same stage of infection. This timing coincided with the earliest point that expression of T-DNA could be detected in protoplasts, suggesting that these genes may play some important role in T-DNA integration specifically.

A study using mature *A. thaliana* plants examined the changes in both gene expression and phytohormone accumulation in response to infection by the tumorigenic *Agrobacterium* strain C58 or by the non-tumorigenic strain GV3101

(Lee et al. 2009). The lower part of inflorescence stalks of plants was inoculated just above the basal leaves. This location was chosen in order to allow the formation of crown gall tumors. Using microarray analysis to sample transcript levels at three different time-points (3 h and 6 days for both strains, and 35 days post-inoculation for the tumorigenic strain only), they found that over four times as many *Arabidopsis* genes were significantly altered in their expression levels upon exposure to the tumorigenic *Agrobacterium* strain as were altered by the disarmed strain using wounded, but uninoculated, inflorescence stalks as a control. It was found that both strains induced genes falling into the MapMan functional category of "stress," while hormone-related genes were affected by both, but with a stronger response to strain C58. Genes of functional classes associated with changes in host morphology were activated only by exposure to C58, but not by GV3101.

A recently published study that was conducted in our laboratory used next-generation sequencing to characterize the changes to the whole transcriptome through time in *A. thaliana* seedlings infected with either of two different non-tumorigenic strains of *Agrobacterium*: strain At804 (virulent) or A136 (avirulent). Seedlings were sampled across a time course from 0 to 48 h after infection. As was found in the two studies previously discussed in this section, substantially more genes were significantly altered in their expression levels by exposure to the T-DNA transfer competent bacterial strain than by the avirulent strain relative to a mock treatment. Because a greater number of transcripts can be detected and there is a greater sensitivity of measurement using RNA-Seq compared to the older differential gene expression techniques, we were able to find many more genes with altered expression levels than had been reported in previous studies. Using Gene Ontology (GO) enrichment analysis, Duan et al. (2018) found that categories such as "cell wall organization or biogenesis," "DNA replication," and "external encapsulating structure organization" were all overrepresented among downregulated genes and that "defense response," "response to stress," and "response to reactive oxygen species" were overrepresented among upregulated genes for the treatment for both strains. This agrees with the previous findings that indicate *Agrobacterium* seems to generally repress normal plant growth and cell division and to activate defense response pathways (at least in the early stages of infection) irrespective of its ability to transfer T-DNA and Vir proteins to the host. We found that in the latter time-points following infection that certain functional categories were unique to treatment with one strain or the other among upregulated genes. For instance, "cellular response to stress" and "secondary metabolite biosynthetic process" were found only in upregulated genes from the avirulent treatment, indicating that these responses may be attenuated by transfer of Vir proteins and/or T-DNA. On the other hand, categories having to do with cell growth, transcription and RNA metabolism, as well as "heterocycle metabolic process" and "response to abscisic acid," were found in upregulated genes only in the virulent treatment condition. It was suggested in Veena et al. (2003) that genes associated with cell division and growth processes were induced by successful *Agrobacterium*-mediated transformation. The results of our recent study support this

conclusion along with suggesting other biological and metabolic processes that may be activated specifically by Vir genes or T-DNA.

2.2 *Transcriptomic Analysis of Crop Species During Agrobacterium-Mediated Transformation*

The majority of gene expression studies that have been used to uncover the interaction between *Agrobacterium* and host plants have been carried out using *A. thaliana* and, to some extent, tobacco suspension cell cultures. However, there have been three studies published within the last few years that have made use of transcriptomic data to examine the effect of *Agrobacterium* on economically important crop species. In each of these studies, transcriptome profiling by microarray or RNA-Seq was one experimental approach of several that were used to gain insight into overcoming a crop plant's recalcitrance to transformation.

In Tie et al. (2012), the researchers used microarrays to measure gene expression throughout infection by *Agrobacterium* strain EHA105 in the two rice cultivars "Nipponbare" (Nip) and "Zhenshan 97" (ZS), representing *O. sativa* ssp. *japonica* and ssp. *indica*, respectively. Representatives from both subspecies were selected for comparison because *indica* varieties of rice are overall much more resistant to transformation than are *japonica* varieties. Embryogenic calli from either variety were sampled at 0, 1, 6, 12, and 24 h after inoculation. The highest number of differentially regulated genes occurred at the 1 or 6 h time-points, and of all unique locus identifiers that were differentially expressed in either type of callus over time (11,105 sequences) only 35% were shared between the two callus cultivars. GO enrichment analysis found some differences in biological process terms between the callus types when contrasted at the same time-point. Genes involved in "defense response" and "response to biotic stimulus" were overrepresented among the set of upregulated genes in the *indica* callus variety. Meanwhile, categories relating to "cell cycle," "cell division," and "DNA repair" were overrepresented among the downregulated genes in ZS callus at the earliest time-points. Some genes that are involved in ubiquitin-proteasome degradation were repressed in the more recalcitrant variety, leading the authors to suggest that degradation of the proteins coating the T-complex may be inhibited in ZS, giving rise to a deficiency in T-DNA integration and lowering the overall transformation frequency.

Another study combined transcriptomic analysis (using RNA-Seq) and proteomics (2-DE and MS) to uncover *Agrobacterium*–host interactions in transformed wheat immature embryo tissue (Zhou et al. 2013). Embryos of the Chinese commercial wheat variety "Yangmai12" were extracted from immature seed and pre-cultured for 4 days before being transformed with the non-tumorigenic *Agrobacterium* strain C58C1. Tissue samples were collected 36 h following inoculation. In total, the researchers found 4889 genes that showed significant differential

expression compared with the mock treatment. GO biological process term analysis showed categories such as "chromatin assembly or disassembly," "signal transduction," "biosynthesis of secondary metabolites," and "phenylpropanoid biosynthesis" were overrepresented among these genes. The molecular function terms showed that a large portion (20.5%) of the functionally classified genes fell into the "nucleic acid binding" category. The proteomic analysis uncovered 90 differentially expressed proteins (DEPs) between the two conditions. Notably, only 24 of these DEPs corresponded to gene sequences found in the transcriptomic data. Of these overlapping DEPs, half of them were predicted to play a role in response to stress or immunity. The differences observed between the results of the transcriptomic and proteomic analyses in this study demonstrate that measured mRNA levels do not perfectly correspond to final expression of a gene product.

In a study aimed at uncovering the mechanism of the observed improvement of *Agrobacterium*-mediated transformation in soybean by employment of sonication and α-aminooxyacetic acid (AOA), RNA-Seq was used to monitor transcriptomic changes in response to *Agrobacterium* infection combined with sonication of the explant tissue and media amended with AOA or to *Agrobacterium* infection without these additions to the protocol (standard *Agrobacterium* transformation) (Zhang et al. 2016). Cotyledonary nodes of the soybean genotype "Jidou17" were infected with *Agrobacterium* EHA105 with or without sonication plus AOA during inoculation and samples were collected five hours after infection. As in the previous studies, plant defense and immune responses were upregulated dependent on *Agrobacterium* infection. A total of 2158 differentially expressed genes (55.1% of them upregulated) were responsive to standard *Agrobacterium* transformation in contrast to mock treatment, whereas 5062 genes showed a difference in expression between standard transformation and transformation plus sonication and AOA with 69.6% of these representing downregulated genes. Upregulated and downregulated genes were categorized by their pathway function using KEGG. Interestingly, this study showed that in addition to defense genes related to PTI (which *Agrobacterium* has been long understood to stimulate) "R" genes known for participating in ETI such as RPM1, RPS2, RPS5, RIN4, and PBS1 were also induced. This is notable as *Agrobacterium* had not previously been shown to induce ETI. Consistent with the role of AOA as an inhibitor of phenylpropanoid biosynthesis, the authors showed that genes involved in this metabolic process are upregulated by *Agrobacterium* treatment (in agreement with the study discussed above), but their expression is ameliorated by sonication and AOA supplementation in tissue culture following transformation.

2.3 Evaluation of GO Terms and Individual Genes Common to Gene Sets from Multiple Studies

In spite of the variety of host plant species, explant tissues, culture systems, *Agrobacterium* strains, and timing of sample collection used in the above studies, there are common patterns seen in many of the data sets. Using GO biological

process terms as a representation of wide-scale gene expression changes, processes such as defense response, plant cell wall modification, cell growth and development and transcription have been shown to be consistently altered by *Agrobacterium* infection. For GO categories overrepresented among upregulated genes in response to disarmed, virulent *Agrobacterium* strains in four studies (Tie et al. 2012; Zhou et al. 2013; Duan et al. 2018; and Niu et al. 2015, discussed below) that had such data readily available, ones that showed the most consistency (called in three out of four datasets) were the broad classes "DNA-dependent," "response to stress," and "transport." Other categories of note included "regulation of gene expression," "regulation of macromolecule metabolic process," "M phase of mitotic cell cycle," and "response to oxidative stress." GO terms for the downregulated gene sets were less consistent between studies with only seven terms shared between at least two studies. These categories included "catabolic process," "cell wall polysaccharide metabolic process," and "signal transduction." For a complete list of GO terms overlapping between datasets from these studies, see Table 3. In the same way, comparing sets of individual transcripts that were upregulated in response to virulent disarmed *Agrobacterium* (or in response to expression of *vir* genes) across studies yielded some consensus genes worth evaluating. A total of 97 genes were identified in at least two of the three datasets compared (Lee et al. 2009; Niu et al. 2015; Duan et al. 2018). The three genes that were present in all three datasets encoded an oxidoreductase (FOX1), caffeoyl CoA 3-O-methyltransferase (CCOAMT; a component of the phenylpropanoid biosynthesis pathway), and a cytochrome P450 protein (CYP71A12). Other notable genes shared by two of the datasets coded for the VirE2-binding F-box protein (VBF), a non-symbiotic hemoglobin (HB1; GLB1), several peroxidase proteins (including PER37, PER71, and PER4), and a regulator of the chromosome condensation family protein (AT4G14368). These concurring upregulated genes could serve as potential candidates for future functional validation of their role in the *Agrobacterium*–plant interaction. For a complete list of these genes, see Table 4.

3 Transcriptional Profiling in Response to PAMP Exposure

A plant's first line of defense against pathogen attack depends on its ability to perceive and respond to pathogen-associated molecular patterns (PAMPs). These usually consist of evolutionarily conserved structural molecules (or epitopes within these molecules), and thus hosts are able to recognize broad ranges of potential pathogens for each PAMP. PAMP perception induces a series of responses known as PAMP-triggered immunity (PTI), which includes reactive oxygen species (ROS) burst, calcium flux, MAP kinase activation, ethylene production, callose deposition, and transcriptional reprogramming (Zipfel and Robatzek 2010). One of

the most well-known bacteria-derived PAMPs is flg22, derived from flagellin, which is perceived by the protein receptor FLS2 in *Arabidopsis* (Gómez-Gómez and Boller 2000; Chinchilla et al. 2006). Recognition of flg22 confers resistance to several bacterial pathogens; however, *Agrobacterium* is able to evade the perception of its flagellin due to a mutation near the N-terminus of the protein (Felix et al. 1999; Kunze et al. 2004). Another bacteria-derived PAMP, Elongation Factor Tu (EF-Tu), is recognized by *Arabidopsis* plants upon exposure to *Agrobacterium*. This response is mediated through the recognition of receptor protein EFR, and there is some convergence between the signaling responses initiated by either flagellin or EF-Tu perception. Furthermore, *Arabidopsis efr* mutants show increased susceptibility to *Agrobacterium* (Zipfel et al. 2006). Clearly, PTI has an important role to play during the *Agrobacterium* transformation process. Several studies have used transcriptomic profiling to investigate the response to PAMP perception. Although some of these PAMPs were derived from bacteria other than *Agrobacterium*, these studies can yield insight into which genes form part of a common defense pathway and which, if any, are more of an *Agrobacterium*-specific response.

Zipfel et al. (2004) explored the gene expression response of *Arabidopsis* Ler-0 seedlings 30 min after treatment with flg22 using an Affymetrix whole-genome array chip. In total, they found that 966 genes were upregulated and 202 genes were downregulated. Treatment of *Arabidopsis fls2* mutants with flg22, however, only altered the expression patterns of six genes when compared with the control, none of which were contained in the set of those that showed a response to flg22 in the wild-type plant. This result showed that FLS2 is solely required for the flg22-dependent defense response to take place. Of the upregulated genes with known functions, a large fraction of them were predicted to be involved in signal perception, signal transduction, or transcriptional regulation. This included 155 receptor-like kinase (RLK) proteins, one of which was FLS2. Treatment of *fls2* mutant seedlings with crude extracts of pathogenic bacteria including *Agrobacterium* gave similar physiological responses to that of wild-type plants, indicating that there were PAMPs other than flg22 being perceived by the plant immune system. Another paper published by this group, in addition to identifying the *Arabidopsis* receptor-like kinase EFR as the receptor for EF-Tu and demonstrating that EF-Tu recognition occurs during *Agrobacterium*-mediated transformation, compared the transcriptomic responses to EF-Tu or flagellin perception (Zipfel et al. 2006). *Arabidopsis* seedlings were this time treated with the EF-Tu-derived peptides elf26 and elf18 and sampled at 30 and 60 min after induction. A total of 866 genes were called as upregulated and 83 were downregulated after 60 min. There was a high degree of overlap demonstrated between the genes affected by either EF-Tu or flagellin. For instance, over 100 of the 610 predicted RLK proteins in the *Arabidopsis* genome were induced by both PAMPs. Treatment with either PAMP also increased the abundance of receptor sites for the other PAMP after only a few hours, indicating a positive feedback in signaling shared between the two. Another noteworthy result is that *efr* mutant plants treated

with *Agrobacterium* crude extracts still showed a defense response, suggesting that there are yet other PAMPs from *Agrobacterium* that plants can perceive.

Other studies have explored, along similar lines, the transcriptomic effects of treatment with types of bacterial PAMPs other than flagellin and EF-Tu on *Arabidopsis* plants. One study looked at the response of mature *Arabidopsis* Col-0 leaves to infiltration with *Staphylococcus aureus*-derived peptidoglycan (PGN) after 4 h (Gust et al. 2007). When contrasted with a water-treated control, 236 genes were shown to be upregulated greater than twofold. Once again, many of these genes overlapped with those induced by flg22 at the same time after treatment. Some of the specific classes of these genes included chitinases, peroxidases, and the phenylpropanoid pathway enzyme phenylalanine ammonia lyase 1 (PAL1), along with WRKY and AP2 family transcription factors. Researchers in another study looked at differential gene expression induced by treatment with lipopolysaccharide (LPS), which is a cell membrane component of gram-negative bacteria including *Agrobacterium*, and harpin, a component of the type III secretion system which *Agrobacterium* does not possess (Livaja et al. 2008). They used *Arabidopsis* cell suspension cultures sampled at six time-points up to 24 h after treatment and analyzed gene expression via microarrays. A total of 1573 genes had altered expression levels greater than twofold by either PAMP with 313 responding to both, 309 unique to LPS, and 915 unique to harpin. Harpin showed a much quicker and stronger response than did LPS, with the highest amount of DE genes overlapping between the two PAMPs being observed with harpin 30 min after treatment and LPS 24 h after treatment. There was a substantial difference in the induction of several gene categories between treatment with harpin and LPS, including WRKY family transcription factors, genes associated with ROS burst, and MAP kinase signaling components.

These studies that explore the host transcriptomic response to treatment by bacterial PAMPs can give clues into the mechanism of plant defense against *Agrobacterium*-mediated transformation. Even though *Agrobacterium* does not induce PTI through flagellin or harpin perception, we can see from these studies that there is a high degree of commonality in the response to each of these PAMPs, and thus, *Agrobacterium* must overcome some of the same barriers to infection as do other bacterial plant pathogens. Comparing the sets of genes upregulated or downregulated by PAMP treatment (Zipfel et al. 2004, 2006; Gust et al. 2007; Livaja et al. 2008), treatment with avirulent *Agrobacterium* strain A136 (Duan et al. 2018) yielded 400 genes that were shared between at least three of the sets. Of these, 212 of them were not upregulated upon treatment with any virulent *Agrobacterium* strain or upon VirE3 expression. Some notable genes included in this set were the transcription factor *WRKY33*, *MILDEW RESISTANCE LOCUS O 12* (*MLO12*), a PBS1-like gene (*AT3G55450*), *ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR 4* (*ERF4*), and a paralog of the *Arabidopsis* defense signaling gene *NPR1* (*NPR3*). There were only 10 genes represented in three or more of the datasets for PAMP-downregulated genes. These included *PHOSPHOENOLPYRUVATE*

CARBOXYLASE KINASE (*PPCK*), *EXPANSIN-LIKE A1* (*EXLA1*), and a type A cytokinin signaling gene *Arabidopsis RESPONSE REGULATOR 6* (*ARR6*). Some of the genes that are upregulated by PAMP treatment, but not by virulent *Agrobacterium* or Vir proteins, could be those that *Agrobacterium* is able to suppress the expression of in order to ameliorate the plant defense response. See Table 2 for a complete list of these consensus PAMP-responding genes. It is plausible that the variations in the expression of the some of the key genes involved in this common defense response may account for differences seen in the amenability of different cell types, explant tissue, or plant genotypes/species to T-DNA transfer. Transcriptomic profiling, in combination with other molecular techniques, will help to shed light on the key molecules and pathways that could potentially be targeted for improving plant transformation capabilities.

4 Transcriptional Response to Heterologous Expression of *Vir* Genes in Host Plants

There have been two studies published so far that used transgenic *Arabidopsis thaliana* plants expressing *Agrobacterium*-derived *vir* genes in order to examine differences in gene expression compared with wild-type plants. In Niu et al. (2015), the investigators transformed *Arabidopsis* Col-0 plants with the *virE3* gene sequence under the control of a tamoxifen-inducible promoter. The VirE3 protein had been previously implicated as a possible plant-active transcription factor because it could induce transcription in yeast when fused with a DNA-binding domain (García-Rodríguez et al. 2006). Fourteen-day-old seedlings of these *virE3* transgenic plants were treated with tamoxifen to induce transgene expression and RNA-Seq was used to profile the differential gene expression compared to wild-type and mock-treated transgenic plants. They found 607 genes that were upregulated and 132 that were repressed specifically by *virE3* expression (using a fold-change cutoff of 3 or 0.33). Among the upregulated genes were those encoding the plant protein VBF (a functional homolog of the *Agrobacterium* VirF protein), which had been previously shown to be induced by *Agrobacterium* infection and to play a role in destabilizing VirE2 (Zaltsman et al. 2010). Another gene that they found to be highly induced by VirE3 was that coding for NIMIN1, which binds the salicylic acid signaling protein NPR1 and reduces expression of the pathogenesis-related (PR) genes. (Weigel et al. 2005). They showed through additional experiments that the promoters of both *VBF* and *NIMIN1* could be bound by VirE3. The authors performed GO term enrichment analysis on their datasets of differentially expressed genes. Biological process categories including "signal transduction," "response to stress," "DNA-dependent transcription," and "developmental processes" were overrepresented among upregulated genes in tamoxifen-induced VirE3 plants. Examples of the molecular

function categories that were highlighted are “kinase activity,”, “transporter activity,” and “Transcription factor activity.”

Duan et al. (2018) produced transgenic *virE3* as well as *virE2 Arabidopsis* Col-0 plants with their expression under the control of the CaMV 35S promoter. These plants were tested for expression of *virE3* and *virE2* mRNA, respectively, using qRT-PCR. As determined by qRT-PCR that for both transgenic plants, the expression levels of defense genes, including *PR3*, *CRK41*, and *CRK18*, were elevated compared to wild-type controls upon treatment with the avirulent (cured) *Agrobacterium* strain A136. Meanwhile, the defense genes *FRK1*, *PR2*, and *PR4* showed increased expression only in the *virE3*-overexpression plants. These limited gene expression data can be interpreted to indicate that overexpression of VirE2, and to a greater extent VirE3, enhances the plant defense response to *Agrobacterium*, though an alternative interpretation is that constitutive expression of the virulence transgenes has interrupted the normal functioning of the defense pathways in the plant.

Heterologous expression of *vir* genes in plants is an attractive method for investigating their specific functions as it allows the gene expression changes they induce in the host to be decoupled from expression changes caused by the presence of the *Agrobacterium* itself (i.e., defense responses due to induction of PTI). It would be far more difficult to evaluate gene expression from the same angle using modified *Agrobacterium* strains to induce gene expression changes. However, there are necessarily caveats when attempting to compare the expression of a virulence gene in the host cell to what actually occurs during the infection process. For example, delivery of the molecule from *Agrobacterium* to the host may be precisely controlled, and thus, expression levels of the transgene from the plant may not correlate with the level of protein that occurs in *Agrobacterium*-mediated transformation. It is also possible that differential expression of some genes may represent secondary transcriptional effects following induction of the virulence gene, which do not represent the normal function of the *vir* gene during T-DNA transfer. Taking this into consideration, it is obviously advantageous to design transgene cassettes using an inducible system such as that used by Niu et al. (2015) rather than constitutive expression. This way at least the timing of expression can be controlled, potentially allowing for temporal isolation of primary transcriptional effects from those of more indirect responses. The *Arabidopsis* genes *VBF* and *NIMIN1*, which Niu et al. (2015) demonstrated could be transcriptionally regulated by VirE3, were both included in a set of genes of upregulated in response to both tamoxifen-induced VirE3 expression and treatment with virulent *Agrobacterium* strain At804 (see Table 2). Surprisingly, however, both of these genes were almost equally upregulated in seedlings treated with either At804 or the avirulent strain A136 relative to the mock treatment (Duan et al. 2018). This result suggests either that some *Agrobacterium*-derived PAMP, or at least an extracellular signal, initiates a VirE3-independent route by which these genes can be induced, or that VirE3 is dispensable not only for overall T-DNA transfer efficiency, but also for its role in

transcriptional activation, and thus is functionally redundant with some host protein. This question could be resolved by evaluating the expression levels of these genes in plants treated with a *virE3* mutant *Agrobacterium* strain.

5 Advantages and Limitations of Using Transcriptomic Analysis as an Approach to Discover Plant Genes and Pathways Associated with *Agrobacterium*-Mediated Transformation

There are obvious advantages to employing a transcriptomic approach for the purpose of gene discovery and pathway elucidation. Transcriptome-wide analysis allows for a quick survey of nearly all expressed genes in a given organism under specific conditions at a relatively low cost. The assays used are very sensitive and can give an accurate reflection of gene expression levels at specific times. Relative to older techniques, these methods also have a high degree of reproducibility. The advent of next-generation sequencing technology has made transcriptomic analysis the preferred choice for gene discovery and, as demonstrated above, it has been instrumental for uncovering the molecular players in specific pathways such as those involved in the *Agrobacterium*-mediated transformation process. Some of the main reasons for the rapid adoption and wide use of RNA-Seq include the low background signal, its capability for detecting a wide dynamic range of expression, and its generation of novel sequence data at a single-base resolution. Unlike microarray technology, it is not reliant on preexisting sequence information and is not susceptible to false signals generated by potential cross-hybridization of similar transcript species to the same sequence probes. All of this allows a much greater amount of much higher quality data to be generated than was previously possible for gene expression studies.

As a tool for gene discovery, transcriptomic analysis has intrinsic limitations. The sequences detected by these methods, after all, reflect only mRNA transcripts and not the final products of the respective gene's expression (i.e., protein or metabolites). Thus, a survey of the transcriptome is blind to the effect of translational or posttranslational controls on gene expression. Another drawback to RNA-Seq is that the amount of data generated can be overwhelming and requires significant storage space and some specialty to manage and analyze the data. Too much data availability can be an obstacle to gene discovery as it is often difficult to decide which genes to study further out of a large pool of candidates. Furthermore, the degree of differential expression does not necessarily indicate which genes have the most critical functions in a given pathway because genes that are likely induced through multiple steps in a regulatory cascade often display the highest fold changes, whereas genes that function earlier in a pathway may show comparatively subtler effects. Additionally, some portion of the genes called as differentially

expressed may have had their expression patterns altered as secondary effects from the original treatment and may not, therefore, have any relationship to the process being studied. These drawbacks limit the conclusions that can be drawn about the genes involved in *Agrobacterium*-mediated transformation from transcriptomic data alone. However, by using other molecular biology techniques to evaluate characteristics of pathways such as protein–protein interactions and protein-DNA binding, as well as using transformation experiments to validate the importance of specific genes, we can compensate for the constraints inherent in transcriptomic analysis in order to develop a more complete picture of the molecular mechanism underlying this process.

6 Summary and Outlook

Taken together, the studies that have used differential gene expression to examine the host response to *Agrobacterium* or elicitor compounds have given us a fuller picture of which genes could be important in *Agrobacterium*–plant interactions. However, many questions still need to be addressed in order to apply our understanding of the process toward the development or improvement of *Agrobacterium*-mediated transformation techniques for the benefit of crop breeding and basic research (Altpeter et al. 2016). The interaction between *Agrobacterium* and plants is complex and multi-layered. More focused transcriptomic analyses of specific pathways known to be involved would be desirable to maximize the utility of this approach. Next-generation sequencing technology has allowed greater sensitivity when measuring gene expression and offers attractive novel possibilities for experimental designs, which could give a more refined view of molecular processes. For instance, in order to explore gene regulatory interactions that occur in response to *Agrobacterium*, RNA-Seq could be combined with ChIP-Seq to determine the direct induction of genes by key transcription factors.

With respect to biotechnological applications, one could make improvements to the transformation process in various ways informed by differential gene expression data in response to *Agrobacterium*. For example, an *Agrobacterium* strain was modified to elicit a weaker plant defense in a study on potato (Vences-Guzman et al. 2013). In other studies, the composition of tissue culture medium and/or growth conditions was modified in order to minimize plant defense responses (Zhang et al. 2013, 2016). Because of the similarity of the plant defense response between *Agrobacterium* and other bacterial pathogens, effector proteins originating from other bacteria have even been utilized to repress plant defense gene induction. In one instance, the AvrPto protein from *Pseudomonas syringae* was conditionally expressed in *Arabidopsis* plants leading to higher efficiency of *Agrobacterium*-mediated transformation due to its suppression of plant defenses (Tsuda et al. 2012). As more of the mechanisms involved in *Agrobacterium*–plant interactions are uncovered, plant transformation researchers will be able to apply some of these discoveries to make new improvements to transformation techniques.

Table 1 Summary of transcriptomic analysis of *Agrobacterium*-mediated plant transformation or PAMP exposure

Authors	Analysis method	*Agrobacterium* strain (tumorigenic or disarmed) or PAMP	Plant materials/treatment/time frame	Key discovery
Ditt et al. (2001)	cDNA-AFLP	EHA105 (pBISN1); disarmed	*Ageratum conyzoides* cell cultures or Tobacco BY-2 cell cultures 24 or 48 h post-inoculation	Plant cells rapidly responded to *Agrobacterium* infection. About 250 unique gene fragments differentially regulated
Veena et al. (2003)	Suppressive subtractive hybridization and DNA macroarrays	At739, At1221 (avirulent), At543, At804, At1222 (virulent); disarmed	Tobacco BY-2 cell suspension cultures at 0 to 36 h post-inoculation	Transfer of T-DNA/Vir proteins modulated the expression of host genes differently than transfer-deficient *Agrobacterium*
Zipfel et al. (2004)	Affymetrix ATH1 whole-genome microarray chip	None; flg22	*Arabidopsis* Ler-0 seedlings 30 min after flg22 treatment	1168 genes differentially regulated including 155 RLKs
Ditt et al. (2006)	*Arabidopsis* 26,000-gene oligonucleotide microarray	A348; tumorigenic	*Arabidopsis* cell cultures at 4 to 48 h post-inoculation	Differential expression only observed at 48 h time-point. Induced genes encode known defense proteins and repressed genes encode proteins involved in cell proliferation
Zipfel et al. (2006)	Affymetrix ATH1 whole-genome microarray chip	None; elf26	*Arabidopsis* Ler-0 seedlings 30 and 60 min after elf26 treatment	427 genes were upregulated and 7 were suppressed at 30 min following treatment with elf26. 866 were upregulated and 83 were downregulated at 60 min. A large portion overlap with those induced by flg22
Gust et al. (2007)	Affymetrix ATH1 whole-genome microarray chip	None; Peptidoglycan (PGN)	*Arabidopsis* Col-0 mature plants at 4 h following infiltration with PGN	236 genes were upregulated. A large involved in defense response including several *WRKY* transcription factors
Livaja et al. (2008)	Agilent *Arabidopsis* cDNA microarray	None; Lipopolysaccharide (LPS), harpin	*Arabidopsis* cell suspension culture at 30 min to 24 h following treatment with either LPS or harpin	1573 genes showed altered expression pattern in response in total with 313 responding to both PAMPs. LPS showed much weaker and later defense induction than harpin

(continued)

Table 1 (continued)

Authors	Analysis method	*Agrobacterium* strain (tumorigenic or disarmed) or PAMP	Plant materials/treatment/time frame	Key discovery
Lee et al. (2009)	Affymetrix ATH1 whole-genome microarray chip	C58 (tumorigenic), GV3101 (disarmed)	*Arabidopsis* inflorescence stalks at 3 h and 6 d post-inoculation, and 35-day-old tumors	Pathogen defense genes responded predominantly to oncogenic strain C58. Host responses were much stronger toward the oncogenic strain C58 than to the disarmed strain GV3101
Tie et al. (2012)	Microarray	EHA105; disarmed	*Oryza sativa* ssp. *japonica* genotype "Nipponbare" and ssp. *indica* genotype "Zhenshan 97" callus tissue at 0 to 24 h post-inoculation	Representative genotypes of japonica and indica rice showed divergent transcriptional responses to transformation with EHA105. Defense response was stronger in Zhenshan 97, cell division and growth was repressed
Zhou et al. (2013)	RNA-Seq	C58C1; disarmed	*Triticum aestivum* genotype "Yangmai12" 4-day pre-cultured immature embryo tissue at 36 h post-inoculation	4889 wheat genes showed significant differential expression. Major categories included stress or immunity response and secondary metabolite biosynthesis
Niu et al. (2015)	RNA-Seq	None; *virE3* tamoxifen-inducible transgenic line	*Arabidopsis* seedlings at 4 h after tamoxifen treatment	Expression of *virE3* upregulated 607 genes > three-fold and repressed 132 genes > three-fold
Zhang et al. (2016)	RNA-Seq	EHA105; disarmed	*Glycine max* genotype "Jidou17" cotyledonary nodes with or without sonication and AOA treatment at 5 h post-inoculation	AOA and sonication treatment, which leads to greater transformation efficiency, suppresses expression of some defense-related genes and isoflavone biosynthesis enzymes relative to *Agrobacterium* alone
Duan et al. (2018)	RNA-Seq	A136, At804; disarmed	*Arabidopsis* Col-0 seedlings at 0 to 48 h post-inoculation	*Agrobacterium* infection suppressed genes important for plant growth and development but induced defense response genes

Table 2 *Arabidopsis* genes upregulated in at least three studies by PAMPs or avirulent *Agrobacterium* strain, but not upregulated by any virulent *Agrobacterium* treatment, and genes downregulated by PAMPs

	Function	Gene names
Upregulated PAMP genes	Carbohydrate metabolism	UGT74B1 (AT1G24100), AT1G78820, AT2G15480, AT2G27500, PEN2 (AT2G44490), AT3G10720, CHIV (AT3G54420), AT3G62720, AT4G02330, AT4G30290, AT5G03700, AT5G18470
	Cell growth related	EXLA1 (AT3G45970), PSK4 (AT3G49780)
	Defense signaling	MLO2 (AT1G11310), PP2-A5 (AT1G65390), RLP12 (AT1G71400), PSKR1 (AT2G02220), AT2G38870, MLO12 (AT2G39200), NHL2 (AT3G11650), SYP122 (AT3G52400), EXO70H1 (AT3G55150), MPK5 (AT4G11330), RAF27 (AT4G18950), NHL3 (AT5G06320), NPR3 (AT5G45110), EXO70H7 (AT5G59730), BON1 (AT5G61900), PROPEP3 (AT5G64905), MAPKKK5 (AT5G66850)
	DNA modification	PEARLI4 (AT2G20960), AT4G35110
	Hormone metabolism	GA2OX6 (AT1G02400)
	Hormone signaling	JAZ1 (AT1G19180), AT5G35735
	Metabolic enzyme—phenylpropanoid biosynthesis	CCR2 (AT1G80820)
	Other metabolic enzyme	KCS1 (AT1G01120), AT1G17420, AT1G30370, AT1G55450, AT1G71697, AT2G39400, SDR5 (AT2G47140), AT3G15530, AGK2 (AT3G57550), APS3 (AT4G14680), AT4G24160, AT4G24380, GGT1 (AT4G39640), NADP-ME2 (AT5G11670)
	Protein folding	J20 (AT4G13830)
	Protein metabolism	PUB20 (AT1G66160), AT2G35930, PUB22 (AT3G52450), AT3G59080, SGT1A (AT4G23570), PUB51 (AT5G61560)
	Protein modification	PAPP2C (AT1G22280), AT3G59350, AT5G36250, AT5G46080
	Receptor-like kinase	LRR XI-23 (AT1G09970), LRK10L1.2 (AT1G18390), AT1G25390, AT1G51790, AT1G51800, AT1G51850, AT1G61370, SD1-29 (AT1G61380), AT1G70530, RIPK (AT2G05940), SOBIR1 (AT2G31880), AT2G33580, BIK1 (AT2G39660), PTI1-4 (AT2G47060), AT3G02880, CERK1 (AT3G21630), PBL1 (AT3G55450), AT4G08850, CRK10 (AT4G23180), CRK11 (AT4G23190), LecRKA4.1 (AT5G01540), AT5G20050, AT5G22690, AT5G35370, AT5G40170, AT5G41550, AT5G46330, AT5G58120, AT5G60270

(continued)

Table 2 (continued)

	Function	Gene names
	Stress response—cytochrome P450	CYP707A3 (AT5G45340)
	Stress response—glutathione-S-transferase	GSTF8 (AT2G47730)
	Other stress response	HEMA2 (AT1G09940), ACA.l (AT1G13210), PHO1;H3 (AT1G14040), CAD1 (AT1G29690), AT1G33590, AT1G33600, AT1G61360, AT1G72900, AT1G72920, NUDT21 (AT1G73540), UCP5 (AT2G22500), HSPRO2 (AT2G40000), ATL6 (AT3G05200), TH9 (AT3G08710), ATSIK (AT3G08760), ACA.k (AT3G22910), DTX18 (AT3G23550), ERD5 (AT3G30775), CAF1a (AT3G44260), ZAR1 (AT3G50950), CNGC13 (AT4G01010), AT4G02200, NUDT7 (AT4G12720), ATL17 (AT4G15975), CBL1 (AT4G17615), CML42 (AT4G20780), AT4G31000, AT4G39830, AT5G12940, CAF1b (AT5G22250), AT5G24430, FC1 (AT5G26030), AT5G39020, CAD1 (AT5G44070), AT5G44910, RDUF2 (AT5G59550), ACHT5 (AT5G61440)
	Transcription factor	MYB51 (AT1G18570), SCL1 (AT1G21450), AT1G21910, AT1G28370, AT1G35350, BZIP60 (AT1G42990), ERF8 (AT1G53170), WRKY6 (AT1G62300), RAV2 (AT1G68840), WRKY17 (AT2G24570), AT2G42360, MYB2 (AT2G47190), ATL2 (AT3G16720), AT3G49530, GATA-8 (AT3G54810), WRKY22 (AT4G01250), HAT1 (AT4G17460), ERF1 (AT4G17500), WRKY28 (AT4G18170), SHB1 (AT4G25350), WRKY11 (AT4G31550), RHA3B (AT4G35480), MYB73 (AT4G37260), ZAT6 (AT5G04340), WRKY72 (AT5G15130), ERF5 (AT5G47230), BZIP01 (AT5G49450), MYB96 (AT5G62470), AZF1 (AT5G67450)
	Uncategorized	AT1G03740, AT1G18380, AT1G19380, AT1G21010, AT1G22890, AT1G51920, AT1G52200, ZCF37 (AT1G59590), FH7 (AT1G59910), AT1G65400, AT1G70740, AT1G72060, AT1G80450, J8 (AT1G80920), AT2G16900, AT2G18680, VQ12 (AT2G22880), AT2G24600, AT2G25735, AT2G27660, AT2G27830, AT2G32140, PHT1;4 (AT2G38940), AT2G44500, AT3G13430, AT3G18560, AT3G19010, AT3G25600, AT3G46110, CaM9 (AT3G51920), AT3G55470, SIB1 (AT3G56710), AT3G56880, AT3G57450, AT3G60420, SAG21 (AT4G02380), AMT1;1 (AT4G13510), AT4G20000, TET9 (AT4G30430), LHT7 (AT4G35180), SEN1 (AT4G35770), AGP18 (AT4G37450), AT5G08350, EXL4 (AT5G09440), AT5G14700, AT5G18150, AT5G20400, AT5G28630, TCH2 (AT5G37770), AT5G42010, AT5G47710, AT5G53110, EXL2 (AT5G64260)

(continued)

Table 2 (continued)

	Function	Gene names
Downregulated PAMP genes	Carbohydrate metabolism	PPCK1 (AT1G08650), XTH12 (AT5G57530)
	Cell growth related	EXLA1 (AT3G45970)
	Hormone signaling	ARR6 (AT5G62920)
	Other stress response	GRXS8 (AT4G15660)
	Transcription factor	DEWAX (AT5G61590), LBD39 (AT4G37540)
	Transport	UMAMIT29 (AT4G01430), MFS1 (AT4G34950), G3Pp1 (AT3G47420)

Table 3 GO biological process categories differentially regulated by *Agrobacterium* infection or Vir proteins in more than one of the reviewed studies

	Studies represented	GO categories	Number of studies represented
Upregulated categories	Tie et al. (2012) (Nipponbare); Niu et al. (2015); Duan et al. (2018) (At804)	Response to stress	3
		Transport	
	Zhou et al. (2013); Niu et al. (2015); Duan et al. (2018) (At804)	DNA-dependent	3
	Tie et al. (2012) (Nipponbare); Duan et al. (2018) (At804)	Nitrogen compound metabolic process	2
		Biological regulation	
		Carboxylic acid metabolic process	
		Cellular metabolic process	
		Cellular nitrogen compound metabolic process	
		Cellular process	
		Cellular protein metabolic process	
		Cellular response to chemical stimulus	
		Cellular response to organic substance	
		Cellular response to stimulus	
		Establishment of localization	
		Hormone-mediated signaling pathway	
		Metabolic process	
		Monocarboxylic acid metabolic process	
		Nucleic acid metabolic process	
		Organic acid metabolic process	
		Oxoacid metabolic process	
		Regulation of RNA metabolic process	

(continued)

Table 3 (continued)

	Studies represented	GO categories	Number of studies represented
		Regulation of biological process	
		Regulation of biosynthetic process	
		Regulation of cellular biosynthetic process	
		Regulation of cellular metabolic process	
		Regulation of cellular process	
		Regulation of gene expression	
		Regulation of macromolecule biosynthetic process	
		Regulation of macromolecule metabolic process	
		Regulation of metabolic process	
		Regulation of nitrogen compound metabolic process	
		Regulation of primary metabolic process	
		Regulation of transcription	
		Response to endogenous stimulus	
		Response to organic substance	
		Response to oxidative stress	
		Response to stimulus	
	Zhou et al. (2013); Duan et al. (2018) (At804)	mRNA splicing, via spliceosome	2
		RNA biosynthetic process	
		RNA metabolic process	
		RNA processing	
		RNA splicing	

(continued)

Table 3 (continued)

	Studies represented	GO categories	Number of studies represented
	Niu et al. (2015); Duan et al. (2018) (At804)	Transcription	2
	Tie et al. (2012); Zhou et al. (2013)	DNA-dependent DNA replication	2
		L-serine metabolic process	
		M phase	
		M phase of meiotic cell cycle	
		M phase of mitotic cell cycle	
		DNA metabolic process	
		Nucleoside	
		Nucleotide and nucleic acid metabolic process	
		DNA repair	
		DNA replication	
Downregulated categories	Tie et al. (2012) (Nipponbare); Duan et al. (2018) (At804)	Catabolic process	2
		Cell wall macromolecule metabolic process	
		Cell wall polysaccharide metabolic process	
		Hemicellulose metabolic process	
		Polysaccharide metabolic process	
	Tie et al. (2012) (Nipponbare); Niu et al. (2015)	Response to stress	2
		Signal transduction	

Table 4 Genes upregulated in response to virulent *Agrobacterium* strain or Vir proteins in more than one of the reviewed studies

Studies represented in	Function	Gene names	Number of studies represented
Lee et al. (2009) (GV3101); Niu et al. (2015); Duan et al. (2018) (At804)	Metabolic enzyme—phenylpropanoid biosynthesis	CCoAOMT6 (AT1G67980)	3
	Other stress response	FOX1 (AT1G26380)	
	Stress response—cytochrome P450	CYP71A12 (AT2G30750)	
Lee et al. (2009) (GV3101), Duan et al. (2018) (At804)	Stress response—cytochrome P450	CYP71A13 (AT2G30770), CYP71B15 (AT3G26830), CYP82C2 (AT4G31970), CYP79B2 (AT4G39950)	2
	Stress response—glutathione S-transferase	AtGSTF7 (AT1G02920), AtGSTF6 (AT1G02930)	
	Stress response—peroxidase	PER37 (AT4G08770), PER71 (AT5G64120)	
	Other stress response	ATBBE10 (AT1G30720), HB1/GLB1 (AT2G16060), ARD3 (AT2G26400), AT2G43570, AT2G43620, AT4G33560	
	Other metabolic enzyme	AT1G43800, PXMT1 (AT1G66700), ALDH2C4 (AT3G24503)	
	Uncategorized	MBL1 (AT1G78850), AT3G29970, AT4G12500, AT4G13180, AT5G42830	
Niu et al. (2015); Duan et al. (2018) (At804)	Cell growth related	AT5G27420	2
	Defense signaling	MPK11 (AT1G01560), NIMIN1 (AT1G02450), FMO1 (AT1G19250), GRX480 (AT1G28480), NHL25 (AT5G36970)	

(continued)

Table 4 (continued)

Studies represented in	Function	Gene names	Number of studies represented
	DNA modification	EFD (AT3G54150)	
	Hormone metabolism	ACS7 (AT4G26200)	
	Metabolic enzyme—phenylpropanoid biosynthesis	COMT-like3 (AT1G21120)	
	Protein metabolism	AT3-MMP (AT1G24140), VBF (AT1G56250)	
	Protein modification	AtDSP13 (AT3G02800), PIMT2 (AT5G50240)	
	Receptor-like kinase	AT1G21240, AT1G21241, AT1G21242, AT1G21243, AT1G21244, RMG1 (AT1G21245), CRK31 (AT1G21246), AT1G21247, AT1G21248, AT1G21249, LECRK-VI.3 (AT1G21250), AT1G21251, AT1G21252	
	Stress response—cytochrome P450	CYP706A2 (AT4G22710), CYP81F2 (AT5G57220)	
	Stress response—glutathione S-transferase	AtGSTU3 (AT2G29470), AT5G44990	
	Stress response—peroxidase	PER4 (AT1G14540), PER5 (AT1G14550), PER62 (AT5G39580)	
	Other stress response	MC8 (AT1G16420), ATBBE4 (AT1G26390), ATBBE6 (AT1G26410), AT1G57630, AT1G66090, ATGLR2.8 (AT2G29110), PMAT2 (AT3G29670), AT3G60140, CBP60G (AT5G26920), ATGLR1.2 (AT5G48400)	

(continued)

Table 4 (continued)

Studies represented in	Function	Gene names	Number of studies represented
	Transcription factor	AT2G43000, AT3G23230, AT3G46080, ZAT72 (AT3G46090), BHLH041 (AT5G56960)	
	Other metabolic enzyme	ASB1 (AT1G25220), AT3G44830, AT4G16820, AT4G39670	
	Uncategorized	AT1G05675, ARI12 (AT1G05880), AT1G32350, AGP5 (AT1G35230), AT1G53620, AT1G53625, AT1G56060, AT1G76650, AT2G23270, AT2G38790, AT3G01830, AT3G02840, AT3G19615, AT3G25240, AT4G14368, AT4G37290, VQ29 (AT4G37710), AT5G22530, AT5G25260, AT5G42380	

Acknowledgements We want to thank members of the Zhanyuan Zhang's laboratory for their technical assistance during the course of the RNA-Seq study. This review work was supported by a University of Missouri Research Board grant and the Life Science Fellowship.

References

Altpeter F, Springer NM, Bartley LE et al (2016) Advancing crop transformation in the era of genome editing. Plant Cell 28:1510–1520

Azpiroz-Leehan R, Feldmann KA (1997) T-DNA insertion mutagenesis in *Arabidopsis*: going back and forth. Trends Genet 13:152–156

Barton KA, Brill WJ (1983) Prospects in plant genetic engineering. Science 219:671–682

Beijersbergen A, Smith SJ, Hooykaas PJ (1994) Localization and topology of VirB proteins of *Agrobacterium tumefaciens*. Plasmid 32:212–218

Bhattacharya A, Priyanka S, Citovsky V (2010) The roles of plant phenolics in defence and communication during *Agrobacterium* and *Rhizobium* infection. Mol Plant Pathol 11:705–719

Binns AN, Costantino P (1998) The *Agrobacterium* oncogenes. In: Spaink HP, Kondorosi A, Hooykaas PJJ (eds) The Rhizobiaceae. Springer, Netherlands, pp 251–266

Casassola A, Brammer SP, Chaves MS et al (2013) Gene expression: a review on methods for the study of defense-related gene differential expression in plants. Am J Plant Sci 4:64–73

Chilton MD, Drummond MH, Merio DJ et al (1977) Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. Cell 11:263–271

Chinchilla D, Bauer Z, Regenass M et al (2006) The *Arabidopsis* receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. Plant Cell 18:465–476

Christie PJ, Vogel JP (2000) Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol 8:354–360

Cui H, Xiang T, Zhou JM (2009) Plant immunity: a lesson from pathogenic bacterial effector proteins. Cell Microbiol 11:1453–1461

DeCleene M, DeLay J (1976) The host range of crown gall. Bot Rev 42:389–466

Ditt RF, Nester EW, Comai L (2001) Plant gene expression response to *Agrobacterium tumefaciens*. Proc Natl Acad Sci U S A 98:10954–10959

Ditt RF, Nester E, Comai L (2005) The plant cell defense and *Agrobacterium tumefaciens*. FEMS Microbiol Lett 247:207–213

Ditt F, Kerr KF, de Figueiredo P et al (2006) The *Arabidopsis thaliana* transcriptome in response to *Agrobacterium tumefaciens*. Mol Plant Microbe Interact 19:665–681

Djamei A, Pitzschke A, Nakagami H, Rajh I, Hirt H (2007) Trojan horse strategy in *Agrobacterium* transformation—abusing MAPK defense signaling. Science 318:453–456

Duan K, Willig C, de Tar JR et al (2018) Transcriptomic analysis of *Arabidopsis* seedlings in response to *Agrobacterium*-mediated transformation process. Mol Plant Microbe Interact 31:445–459

Escobar MA, Dandekar AM (2003) *Agrobacterium tumefaciens* as an agent of disease. Trends Plant Sci 8:380–386

Felix G, Duran JD, Volko S et al (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J 18:265–276

García-Rodríguez FM, Schrammeijer B, Hooykaas PJ (2006) The *Agrobacterium* VirE3 effector protein: a potential plant transcriptional activator. Nucl Acids Res 34:6496–6504

Gelvin SB (2000) *Agrobacterium* and plant genes involved in T-DNA transfer and integration. Annu Rev Plant Physiol Plant Mol Biol 51:223–256

Gelvin SB (2010) Plant proteins involved in *Agrobacterium*-mediated genetic transformation. Annu Rev Phytopathol 48:45–68

Gelvin SB (2017) Integration of *Agrobacterium* T-DNA into the plant genome. Annu Rev Genet 51:195–217

Gómez-Gómez L, Boller T (2000) FLS2: An LRR receptor–like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. Mol Cell 5:1003–1011

Gust AA, Biswas R, Lenz HD et al (2007) Bacteria-derived peptidoglycans constitute pathogen-associated molecular patterns triggering innate immunity in *Arabidopsis*. J Biol Chem 282:32338–32348

Kuldau GA, DeVos G, Owen J et al (1990) The *virB* operon of *Agrobacterium tumefaciens* pTiC58 encodes 11 open reading frames. Mol Gen Genet 221:256–266

Kunze G, Zipfel C, Robatzek S et al (2004) The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. Plant Cell 16:3496–3507

Lacroix B, Citovsky V (2013) The roles of bacterial and host plant factors in *Agrobacterium*-mediated genetic transformation. Int J Dev Biol 57:467–481

Lacroix B, Vaidya M, Tzfira T, Citovsky V (2005) The VirE3 protein of *Agrobacterium* mimics a host cell function required for plant genetic transformation. EMBO J 24:428–437

Lapham R, Lee LY, Tsugama D et al (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 9:749

Lee CW, Efetova M, Engelmann JC et al (2009) *Agrobacterium tumefaciens* promotes tumor induction by modulating pathogen defense in *Arabidopsis thaliana*. Plant Cell 21:2948–2962

Livaja M, Zeidler D, von Rad U et al (2008) Transcriptional responses of *Arabidopsis thaliana* to the bacteria-derived PAMPs harpin and lipopolysaccharide. Immunobiology 213:161–171

Magori S, Citovsky V (2011) *Agrobacterium* counteracts host-induced degradation of its effector F-box protein. Sci Signal 4:ra69

Morris RO (1986) Genes specifying auxin and cytokinin biosynthesis in phytopathogens. Annu Rev Plant Physiol 37:509–538

Mysore KS, Bassuner B, Deng XB et al (1998) Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. Mol Plant Microbe Interact 11:668–683

Niu X, Zhou M, Henkel CV et al (2015) The *Agrobacterium tumefaciens* virulence protein VirE3 is a transcriptional activator of the F-box gene VBF. Plant J 84:914–924

Pu XA, Goodman RN (1993) Attachment of agrobacteria to grape cells. Appl Environ Microbiol 59:2572–2577

Schrammeijer B, den Dulk-Ras A, Vergunst AC et al (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: Evidence for transport of a novel effector protein VirE3. Nucl Acids Res 31:860–868

Shi Y, Lee LY, Gelvin SB (2014) Is VIP1 important for *Agrobacterium*-mediated transformation? Plant J 79:848–860

Shirasu K, Morel P, Kado CI (1990) Characterization of the *virB* operon of an *Agrobacterium tumefaciens* Ti plasmid: nucleotide sequence and protein analysis. Mol Microbiol 4:1153–1163

Staskawicz BJ, Ausubel FM, Baker BJ et al (1995) Molecular genetics of plant disease resistance. Science 268:661–667

Tao Y, Xie Z, Chen W et al (2003) Quantitative nature of *Arabidopsis* responses during compatible and incompatible interactions with the bacterial pathogen *Pseudomonas syringae*. Plant Cell 15:317–330

Thompson DV, Melchers LS, Idler KB et al (1988) Analysis of the complete nucleotide sequence of the *Agrobacterium tumefaciens virB* operon. Nucleic Acids Res 16:4621–4636

Tie W, Zhou F, Wang L et al (2012) Reasons for lower transformation efficiency in indica rice using *Agrobacterium tumefaciens*-mediated transformation: lessons from transformation assays and genome-wide expression profiling. Plant Mol Biol 78:1–18

Tomlinson AD, Fuqua C (2009) Mechanisms and regulation of polar surface attachment in *Agrobacterium tumefaciens*. Curr Opin Microbiol 12:708–714

Tsuda K, Qi Y, le Nguyen V et al (2012) An efficient *Agrobacterium*-mediated transient transformation of *Arabidopsis*. Plant J 69:713–719

Tzfira T, 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

Tzfira T, Citovsky V (2006) *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. Curr Opin Biotechnol 17:147–154

Tzfira T, Vaidya M, Citovsky V (2004) Involvement of targeted proteolysis in plant genetic transformation by *Agrobacterium*. Nature 431:87–92

Veena JH, Doerge RW, Gelvin SB (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

Vences-Guzman MA, Guan Z, Bermudez-Barrientos JR et al (2013) Agrobacteria lacking ornithine lipids induce more rapid tumour formation. Environ Microbiol 15:895–906

Vergunst AC, Schrammeijer B, den Dulk-Ras A et al (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. Science 290:979–982

Wang Y, Peng W, Zhou X et al (2014) The putative Agrobacterium transcriptional activator-like virulence protein VirD5 may target T-complex to prevent the degradation of coat proteins in the plant cell nucleus. New Phytol 203:1266–1281

Ward DV, Draper O, Zupan JR, Zambryski PC (2002) Peptide linkage mapping of the *Agrobacterium tumefaciens vir*-encoded type IV secretion system reveals protein subassemblies. Proc Natl Acad Sci USA 99:11493–11500

Ward JE, Akiyoshi DE, Regier D et al (1988) Characterization of the *virB* operon from an *Agrobacterium tumefaciens* Ti plasmid. J Biol Chem 263:5804–5814

Watson B, Currier TC, Gordon MP et al (1975) Plasmid required for virulence of *Agrobacterium tumefaciens*. J Bacteriol 123:255–264

Weigel RR, Pfitzner UM, Gatz C (2005) Interaction of NIMIN1 with NPR1 modulates PR gene expression in *Arabidopsis*. Plant Cell 17:1279–1291

Wood DW, Setubal JC, Kaul R et al (2001) The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. Science 294:2317–2323

Zaltsman A, Krichevsky A, Loyter A et al (2010) *Agrobacterium* induces expression of a host F-box protein required for tumorigenicity. Cell Host Microbe 7:197–209

Zhang WJ, Dewey RE, Boss W et al (2013) Enhanced *Agrobacterium*-mediated transformation efficiencies in monocot cells is associated with attenuated defense responses. Plant Mol Biol 81:273–286

Zhang YM, Liu ZH, Yang RJ et al (2016) Improvement of soybean transformation via *Agrobacterium tumefaciens* methods involving α-aminooxyacetic acid and sonication treatments enlightened by gene expression profile analysis. Plant Cell Rep 35:1259–1271

Zhang X, van Heusden GPH, Hooykaas PJ (2017) Virulence protein VirD5 of *Agrobacterium tumefaciens* binds to kinetochores in host cells via an interaction with Spt4. Proc Natl Acad Sci 114:10238–10243

Zhou X, Wang K, Lv D et al (2013) Global analysis of differentially expressed genes and proteins in the wheat callus infected by *Agrobacterium tumefaciens*. PLoS ONE 8:e79390

Ziemienowicz A, Merkle T, Schoumacher F et al (2001) Import of *Agrobacterium* T-DNA into plant nuclei: two distinct functions of VirD2 and VirE2 proteins. Plant Cell 13:369–383

Zipfel C, Robatzek S (2010) Pathogen-associated molecular pattern-triggered immunity: veni, vidi…? Plant Physiol 154:551–554

Zipfel C, Robatzek S, Navarro L et al (2004) Bacterial disease resistance in *Arabidopsis* through flagellin perception. Nature 428:764–767

Zipfel C, Kunze G, Chinchilla D et al (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. Cell 125:749–760