2016

Investigation of the anti-virC2 transcript in Agrobacterium tumefaciens

Abbagail Lauren Johnson
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/etd

Part of the Agricultural Science Commons, Agriculture Commons, Agronomy and Crop Sciences Commons, Biology Commons, and the Botany Commons

Recommended Citation
Johnson, Abbagail Lauren, "Investigation of the anti-virC2 transcript in Agrobacterium tumefaciens" (2016). Graduate Theses and Dissertations. 15943.
http://lib.dr.iastate.edu/etd/15943

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Investigation of the *anti-virC2* transcript in *Agrobacterium tumefaciens*

by

Abbagail Johnson

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Plant Biology

Program of Study Committee:
Kan Wang, Major Professor
Gwyn Beattie
W. Allen Miller

Iowa State University

Ames, Iowa

2016

Copyright © Abbagail Johnson, 2016. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>General Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Importance of <em>Agrobacterium</em></td>
<td>2</td>
</tr>
<tr>
<td>The Process of Plant Transformation</td>
<td>5</td>
</tr>
<tr>
<td>Importance of the <em>virC</em> Operon</td>
<td>9</td>
</tr>
<tr>
<td>Non-coding Regulatory RNA</td>
<td>14</td>
</tr>
<tr>
<td>Research Objectives and Thesis Organization</td>
<td>24</td>
</tr>
<tr>
<td>References</td>
<td>27</td>
</tr>
<tr>
<td>CHAPTER II MATERIALS AND METHODS</td>
<td>45</td>
</tr>
<tr>
<td>Bacterial Strains and Plasmids</td>
<td>45</td>
</tr>
<tr>
<td>Transcript Quantification</td>
<td>51</td>
</tr>
<tr>
<td>Rapid Amplification of cDNA Ends</td>
<td>54</td>
</tr>
<tr>
<td>Western Blot Protocol</td>
<td>55</td>
</tr>
<tr>
<td>AGROBEST Experiments</td>
<td>57</td>
</tr>
<tr>
<td>Tumorigenesis Assays Using <em>Kalanchoe daigremontiana</em></td>
<td>62</td>
</tr>
</tbody>
</table>
# Statistical Analysis

References ................................................................. 68

## CHAPTER III RESULTS

- Selection of a Regulatory RNA in *A. tumefaciens* for Further Investigation ..... 73
- Examination of *anti-vir2* Transcription Start and Stop Sites.......................... 75
- Generation of Deletion Mutants C58 ΔPavc2 and C58 ΔvirC2.......................... 77
- Analysis of Transcript Expression Level Changes in *anti-virC2* Promoter Deletion Mutant C58 ΔPavc2 .......................................................... 78
- Analysis of VirC2 Protein Expression in *anti-virC2* Promoter Deletion Mutant ΔPavc2 .......................................................... 82
- Examination of Virulence in *A. tumefaciens* Mutant Strains C58 ΔPavc2 and C58 ΔvirC2...................................................... 84
- References ............................................................................. 92

## CHAPTER IV DISCUSSION AND CONCLUSIONS

References ............................................................................. 115

## APPENDIX A PRIMERS, FRAGMENTS, PLASMIDS AND STRAINS

## APPENDIX B SUPPLEMENTAL INFORMATION

- Introduction ............................................................................. 127
- Overexpression of Individual *anti-virC2* Regions and Intergenic Region ........ 128
- pAJ025; Constitutive Expression of *virC2* and Inducible Expression of *anti-virC2* .................................................................................................................. 131
- Additional FLAG-tagged Proteins ................................................................ 132
- References ............................................................................. 136
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>The T-DNA transfer process.</td>
<td>43</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Illustration of virulence genes on Ti plasmid in <em>A. tumefaciens</em>.</td>
<td>44</td>
</tr>
<tr>
<td>Figure 3</td>
<td>General strategy for homologous recombination used to alter genomic sequences of <em>A. tumefaciens</em>.</td>
<td>69</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Cloning strategy used to generate pAJ007.</td>
<td>70</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Cloning strategy used to generate pAJ021.</td>
<td>71</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Cloning strategy used to generate pKL1007.</td>
<td>72</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Expression of <em>virC</em> and <em>anti-virC2</em></td>
<td>94</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Illustration of 3’ and 5’ sites for <em>anti-virC2</em> regions 1, 2 and 3.</td>
<td>95</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Illustration of deleted <em>anti-vir2</em> promoter in <em>A. tumefaciens</em> strain C58 ΔPavc2.</td>
<td>96</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Strand specific qRT-PCR shows transcript abundance in C58 and C58 ΔPavc2.</td>
<td>97</td>
</tr>
<tr>
<td>Figure 11</td>
<td>qRT-PCR shows <em>anti-virC2</em> transcript abundance in C58 and C58 ΔPavc2.</td>
<td>98</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Examination of FLAG-tagged VirC2 protein levels.</td>
<td>99</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Scale for qualitative scoring of tumors from <em>Kalanchoe daigremontiana</em> tumorigenesis assays.</td>
<td>100</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Data from all <em>Kalanchoe daigremontiana</em> tumorigenesis assays comparing wild type C58 to C58 ΔvirC2.</td>
<td>101</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Data from all <em>Kalanchoe daigremontiana</em> tumoregensis assays comparing C58 wild type and C58 ΔPavc2.</td>
<td>102</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Levels of GUS are shown in <em>A. thaliana</em> plants infected by various strains of <em>A. tumefaciens</em>.</td>
<td>103</td>
</tr>
</tbody>
</table>
Figure S1 pKL1004 through pKL 1001 were used to overexpress each of the three regions of the anti-virC2 transcript as well as the intergenic region between the virC and virG operons. .............................................. 137

Figure S2 Vir Gene transcript expression is examined in C58 ΔPavc2 strains overexpressing a portion of the anti-virC2 transcript. .............................................. 138

Figure S3 Vector pAJ025......................................................................................................................... 139

Figure S4 Relative abundance of transcript from the virC2/anti-virC2 region is shown for C58(pTF505) (empty vector) and C58(pAJ025).......................................................................................................................... 140

Figure S5 Construction of constructs used to generate A. tumefaciens FLAG-tag mutants. ................................. 141
LIST OF TABLES

Table 1 Summary of 3' and 5' ends for anti-virC2 regions 1, 2 and 3 as observed using RACE.......................................................... 104

Table 2 Strand Specific qRT-PCR was performed to examine transcript levels in promoter deletion mutant C58 ΔPavc2 compared to C58 wild type. .... 105

Table 3 Non-strand specific qRT-PCR was performed on cDNA from wild type c58, c58 ΔPavc2, and c58 ΔvirC2 in order to compare transcript levels .................................................................................................................. 106

Table 4 Data from all Kalanchoe daigremontiana tumorigenicity assays......... 107

Table 5 IntaRNA was used to predict possible mRNA targets of the entire anti-virC2 transcript.................................................................................. 116

Table 6 Primer list........................................................................................ 117

Table 7 Fragments list.................................................................................. 122

Table 8 Plasmid list..................................................................................... 124

Table 9 A. tumefaciens strain list ................................................................. 126

Table S1 Transcript Abundance is shown for overexpression strains C58 ΔPavc2(pKL1001-1004) along with to C58 ΔPavc2 and C58 wild type empty vector for comparison .................................................. 142
ACKNOWLEDGMENTS

First, I would like to gratefully thank my Major Professor and mentor, Dr. Kan Wang, for her encouragement and direction throughout my studies at Iowa State University. Without her support and guidance, completion of this work would not have been possible. I would like to thank my mentor and assistant scientist in the lab Dr. Keunsub Lee. The training I received from Dr. Lee has helped me grow professionally and without his experience and direction this research could not have been performed successfully.

I would also like to thank my POS committee members Dr. Gwyn Beattie and Dr. Allen Miller for their guidance and contributions throughout the course of this research. I deeply appreciate their advice and support.

In addition, I would also like to thank all of my current and former colleagues and friends at the Plant Transformation Facility Lab. It has been a true pleasure working with everyone during my time here. I thank Dr. Jennifer Raji, Ms. Amber Testroet, Mr. Juan Carlos Martinez, Dr. Hartinio Nahampun, Dr. Evangelia Vamvaka, and Dr. Alan Eggenberger for their technical advice as well as their personal support. I also thank the members of the Plant Transformation Facility Ms. Bronwyn Frame, Dr. Marcy Main, Dr. Diane Luth and Mr. Dan Little for their support and encouragement.

I would also like the thank Dr. Yanhai Yin, and Ms. Michelle Guo along with the rest of the Yin lab. Their encouragement helped me begin my journey in graduate school. Their technical support has also been appreciated along the way. I also thank Dr. Bing Yang and all the Yang lab personnel, especially Dr. Dangping Lou, who provided me
with generous technical support. Additional thanks go Ms. Tracey Stewart and Mr. Randall Den Adel at the Imaging Facility for their technical support.

A special heartfelt thanks goes to Ms. Jaci Severson and Ms. Dai Nguyen. Their encouragement and support helped me to complete my degree and they are truly appreciated.

I would also like to thank my family and friends for their love and support throughout my time at Iowa State. I am forever grateful for their emotional support and encouragement.
ABSTRACT

Agrobacterium tumefaciens has long been an important tool for plant genetic transformation. Many important crop plants such as maize, soybean, wheat, and rice, as well as numerous other dicot plant species are able to be transformed using A. tumefaciens. While many potential regulatory RNA have been previously identified in A. tumefaciens, very few have been verified with a functional investigation. With the growing understanding of regulatory RNA, and its importance in gene regulation, it is essential to further classify regulatory RNA in the transformative tool A. tumefaciens. It is possible that regulatory RNA could be involved in virulence and their expression altered to increase transformation efficiency.

In this work, a particular candidate regulatory RNA, anti-virC2, was selected and investigated. While anti-virC2 was identified previous to this research, the function of anti-virC2 has never before been examined. The location of anti-virC2 is on the Ti-plasmid antisense to the virulence gene virC2 and it was initially hypothesized that anti-virC2 acted to regulate the virC2 gene in cis.

Through deletion of the promoter region, expression of the anti-virC2 transcript was reduced by ~39% in mutant strain C58 ΔPavc2. Using this strain, changes in virC2 transcript and protein abundance were able to be observed and compared to wild type. In addition, other possible changes in virulence due to the drop in anti-virC2 expression were examined through Kalanchoe daigremontiana tumorigenesis assays and Arabidopsis thaliana infection experiments.

This work shows that anti-virC2 does not act to regulate the cis encoded virC2 gene but may have some impact on virulence. Additional anti-virC2 targets were also
predicted and may be investigated in future studies. As well as the findings made regarding *anti-virC2*, this study provides an example for the functional study of additional regulatory RNA and may serve as a model for the investigation of other potential regulatory RNAs.
CHAPTER I
GENERAL INTRODUCTION AND LITERATURE REVIEW

1.1 General Introduction

Since its first implementation, the process of plant genetic transformation has become an essential part of agriculture and a necessary component of research. By altering a plant’s genes through plant transformation, plants can gain higher yield, acquire disease resistance, or allow functional analysis of specific genes.

One of the most important tools used for plant transformation is *Agrobacterium tumefaciens*. While a plant pest in nature, the natural DNA transferring process by which *A. tumefaciens* infects plants can be harnessed by researchers to transform a variety of important plants. The process of DNA transfer from bacteria to plants also provides a unique system of interkingdom DNA transfer to study.

A new frontier of research is that of regulatory RNA. Over the past decade the number of identified regulatory RNA identified has been steadily increasing. Regulatory RNAs have also been found to play an important role in the regulation of gene expression.

This research is aimed at studying regulatory RNA within *A. tumefaciens*. Specifically, the role of the transcript opposite the *virC2* gene was examined. By gaining an understanding of this regulatory RNA in *A. tumefaciens* we can come to better understand the general function of regulatory RNA in addition to better understanding *A. tumefaciens*. As understanding of regulatory RNA within *A. tumefaciens* increases it may
be possible to alter expression of these molecules resulting in an improved plant transformation frequency.

1.2 Importance of *Agrobacterium*

The *Agrobacterium* genus is made up of a group of gram-negative, soil-borne, plant pathogenic bacteria (Chilton et al., 1977). In nature, *Agrobacterium* is able to infect plants by inserting a piece of its DNA (T-DNA) into the plant genome where it then incorporates and is expressed in the host (for reviews of this process see Binns and Thomashaw, 1988; Zhu et al., 2000; Gelvin. 2010). Once the bacterial DNA has been incorporated, the encoded opine genes are expressed, leading to the production of short amino acids and sugars the bacteria then utilize (Britton et al., 2008). Formation of a tumor, or gall, is caused by expression of oncogenes which result in increased plant production of the hormones auxin and cytokinin (Morris et al., 1986). The process of plant infection by *Agrobacterium* is an example of interkingdom DNA transfer and provides an interesting phenomenon to study.

*Agrobacterium* is particularly important, not only because of the crown gall disease it causes in plants, but as a tool for genetic manipulation. By exploiting the natural DNA insertion that occurs when *A. tumefaciens* infects plants, it is possible for scientists to insert any DNA sequence into the plant, changing its genetic composition (for reviews see Nester, 2014; Ziemienowicz, 2014).

Within the genus *Agrobacterium*, organisms are divided into species depending on the plants they able to infect and the symptoms produced upon infection. The five main species include *A. rubi* which causes cane gall on raspberries, *A. tumefaciens* which
causes crown gall disease on a wide variety of dicots, *A. vitis* which infects grapes, *A. rhizogenes* which causes a proliferation of roots resulting in hairy root disease, and *A. radiobacter* which is not infectious (Slater et al., 2009). Recently though, it has been proposed that *Agrobacterium* species should be changed to species within the *Rhizobium* genus, though there is still debate (Mousavi and Österman, 2014). While each species of *Agrobacterium* is important for different reasons, this work will be centered on *A. tumefaciens*. *A. tumefaciens* is the species most widely used for plant transformation, and has a very important role in the genetic transformation of several important commercial crops (Gelvin, 2003; Wang, 2006).

### 1.2.1 *Agrobacterium tumefaciens* as a natural plant pest

In nature, *A. tumefaciens* is found either as a plant pathogen or free-living in the rhizosphere (Bouzar and Moore, 1987). Of all known plant pathogens, *A. tumefaciens* has the largest number of host plants, over 600 host plants have been documented (DeCleene and DeLay, 1976). Several of these host plants are of economic importance (Escobar and Dandekar, 2003). Examples of affected plants include grape (Burr and Otten, 1999), stone fruits, nuts (Epstein et al. 2008; Pulawska, 2010), and many dicotyledonous plants (Thomashow, et al. 1980). Crown gall disease can severely damage crops by decreasing yield and growth, however, unless infection occurs when the plant is very young, it is generally not lethal (Escobar and Dandekar, 2003). On average, plants with crown gall tumors are 25% smaller compared to non-infected plants (Tzfira and Citovsky, 2008). The impact of crown gall to agriculture can add up to devastating losses.
Just as the name implies, the main symptom of crown gall disease is the formation of crown gall tumors on the infected area of the plant. The symptoms of crown gall disease were first described in 1853 in a French article (Fabre and Dunal, 1853), but the cause of crown gall disease was unknown. In 1907, it was realized that Agrobacterium* tumefaciens* was the causal agent of crown gall disease in the Paris daisy by isolating the *A. tumefaciens* bacteria from the galls which were formed on daisies (Smith and Townsend, 1907).

The tumors formed can grow up to be 30 cm and contain their own vascularization structures (Tzfira and Citovsky, 2008). The production of crown gall tumors is the result of increased levels of the hormones cytokine and auxin, which are overproduced in cells that have been transformed by *A. tumefaciens* during an infection. However, not all of the cells in a crown gall tumor have been transformed, only 10-26% of cells in a tumor are transformed while the rest of the cells maintain an unaltered genotype (Van Slogteren et al., 1983).

**1.2.2 Agrobacterium tumefaciens as a tool for plant genetic modification**

*Agrobacterium tumefaciens* has been very important to modern biotechnology as far back as the late 1970’s when scientists began to understand the mechanisms of the bacteria’s pathogenesis. As researchers studied *A. tumefaciens*, they learned that during infection, the bacteria transferred a piece of DNA to the plant host. In a groundbreaking study, Chilton et al. (1977) were able to use probes specific to various regions of the *A. tumefaciens* genome to probe crown gall tumors and discover that some piece of the tumor inducing (Ti) plasmid was transferred to the plant during infection.
In later research, scientists defined the exact boundaries of the T-DNA transferred to the plant (Zambryski et al., 1982) as a right border and a left border (Joos et al., 1983; Wang et al., 1984).

Once the concept was established that *A. tumefaciens* is able to transfer a distinct portion of its DNA to be incorporated into the plant genome, scientists began to take advantage of this mechanism to transfer other DNA sequences. In a pioneering study by Zambryski et al. (1983), researchers were able to create a “disarmed” version of the tumor inducing (Ti) plasmid by removing the tumor causing genes and only leaving the left and right border sequences flanking a nopaline synthase gene used to detect the DNA once it was transferred to the plant. This experiment was the first time that plants had been transformed with an altered T-DNA and is one of many studies that have led to the modern day process of plant transformation using *Agrobacterium* to deliver DNA.

Today, using *Agrobacterium* is used to transform many of the world’s most important crops including wheat (Hu, et al. 2003), corn (Frame, et al. 2002), soybean (Trick and Finer, 1998), rice (Hiei, et al. 1997), tomato (Arshad, et al. 2014), as well as many other commercial and ornamental plants (Wang, 2006; Wang, 2015).

1.3 The Process of Plant Transformation by *Agrobacterium tumefaciens*

In order to fully understand the important role of *A. tumefaciens* as a plant pest and genetic engineer, it is vital to understand the biological mechanisms involved with the infection process. While there are many genes involved with this process, the virulence, (*vir*) genes, located on the Ti plasmid, are the most essential.
1.3.1 Genomics of Agrobacterium tumefaciens

The genome of A. tumefaciens is composed of four replicons (Goodner et al., 2001; Wood et al., 2001). First there are the two chromosomes, linear and circular chromosomes are 2,075,560 bp and 2,841,490 bp in size, respectively (Wood et al., 2001). These two chromosomes encode all of the genes necessary for the survival of the organism and are always present in A. tumefaciens. The remaining two replicons are plasmids that may or may not be present, named the At plasmid and the Ti plasmid.

Originally termed the “cryptic plasmid”, the At plasmid has the less important role of the two accessory plasmids. Encoded on this 542,779 bp plasmid are genes important for the metabolism of certain nutrients as well as attachment (Goodner et al., 2001; Baek et al., 2005; Chai et al., 2007). When measuring virulence using qualitative tumorigenesis assays, no effect on virulence was observed upon deletion of the At plasmid (Hooykaas et al., 1977; Rosenberg and Huguet, 1984; Hynes et al., 1985). A more recent study observed a decline in pathogenic virulence upon deletion of the At plasmid by using a qualitative measurement of virulence gene induction (Nair et al., 2003). However, it remained clear that deletion of the At plasmid did not cause total loss of virulence.

The Ti plasmid is the other accessory plasmid that may be present in A. tumefaciens. Primarily responsible for plant infection and tumor induction (Wood et al., 2001), the Ti plasmid is 214,233 bp in size and contains nearly all of the genes important for pathogenic virulence (Wood et al., 2001). The six operons crucial for virulence are \textit{virA}, \textit{virB}, \textit{virC}, \textit{virD}, \textit{virE}, and, \textit{virG} each of which encodes genes with important roles in the plant infection process (Goodner et al., 2001).
1.3.2 Regulation

Regulation of the virulence system is controlled by the VirA/VirG two-component sensor kinase system, which is able to sense phenolic compounds, often produced by wounded plants (Winans et al., 1986, Winans et al., 1994; Lin et al., 2014). Upon sensing these compounds, VirA is able to activate itself through autophosphorylation and subsequently phosphorylate and activate VirG (Jin et al., 1990a; Jin et al., 1990b). VirG then activates expression the of rest of the virulence genes by binding to the vir-box region located in their promoter region (Jin et al., 1990c).

Located on the circular chromosome, the chromosomal virulence (chv) genes, are also involved with regulation of the virulence system (Suzuki et al., 2002). ChvE is able to extend the range of phenolic compounds and sugars detectable by A. tumefaciens. ChvE senses compounds not easily detected by VirA and leads to VirA’s activation (Peng et al., 1998).

In addition to sensing phenolics and sugars, A. tumefaciens is able to sense acidic environmental conditions which are often produced by plants and given off into the soil (Rivoal and Hanson, 1994; Walker et al., 2003; Wang et al., 2006). pH is perceived by the ChvI/ChvG two component system (Mantis and Winans, 1993; Li et al., 2002), which is similar to the VirA/VirG two component system. ChvG detects acidic signals leading to activation of the response-regulator ChvI and downstream activation of VirG and the VirB proteins (Charles and Nester, 1993; Mantis and Winans, 1993; Li et al., 2002).
1.3.3 T-DNA transfer and integration

The activated virulence system then needs to prepare the T-DNA for transfer to the plant. There are several proteins that play a role in this process.

Excising the T-DNA from the Ti plasmid is the first step in preparing the T-DNA to be transferred to the plant host. First, VirD2, along with VirD1, acts to nick the lower strand of the right and left border sequences that directly flank the T-DNA region (Yanofsky et al., 1986; Jayaswal et al., 1987; Porter et al., 1987; Wang et al., 1990), resulting in a single stranded T-DNA molecule (Satchel and Nester, 1986). VirC1 and VirC2 aid in this process by binding to a region near the right border called “overdrive,” though it remains unclear exactly how these proteins function (Lu et al., 2009).

VirD2 binds to the 5’ end of the T-strand in the bacteria and remains attached during transfer to the plant (Herrera-Estrella et al., 1990; Rossi et al., 1993; Herrera-Estrella et al., 1998). VirE2 is also part of this complex, and acts to coat the T-strand in the plant to prevent degradation of the T-strand (Christie et al., 1988; Citovsky et al., 1989; Sen et al., 1989).

A. tumefaciens uses a Type IV secretion system to transfer the T-strand to the plant cell (Christie et al., 1997). The type IV secretion system is composed of the 11 VirB proteins and the VirD4 protein (Ward et al., 2002; Christie et al., 2004), each with a distinctive role. VirD4 directs the T-strand into the secretion channel formed by VirB7, VirB9, and VirB10 (Fronzes et al., 2009), while VirB4 helps to move the complex forward (Berger and Christie, 1993; Kumar and Anath, 2002; Atmakuri et al., 2004). The pilus used to secrete the T-DNA to the plant is composed of VirB2 and VirB5 (Yuan et al., 2005). A simplified graphical representation of this process can be seen in Figure 1.
Once the T-strand complex is inside the plant, the DNA must then integrate with the host genome. Both VirE2 and VirD2 contain nuclear localization signals which guide the T-DNA to the host nucleus (Citovsky et al., 1992; Howard et al., 1992; Rossi et al., 1993; Citovsky et al., 1994). Though integration was originally thought to take place at random locations in the genome, several large studies have found correlations between the site of T-DNA integration and gene locations in the host plant (Brunaud et al., 2002; Alonso et al., 2003). Yet in another study, when T-DNA insertion locations were examined without selection pressure, integration was found to occur in a more random distribution (Kim and Gelvin, 2007).

Once the T-DNA is incorporated in the plant genome it may be expressed stably, transiently, or become silenced depending on the location of insertion and histone or DNA methylation levels. (Matzke et al., 1989; Janssen and Gardner, 1990; Gohlke et al., 2013) The number of insertion events has also been shown to affect gene expression, (Hobbs et al., 1990; Van Der Krol et al., 1990). Genes integrated into the plant’s genome may also be passed down to subsequent generations where the genes can continue to be expressed (Budar et al., 1986), but the pattern of inheritance depends on the gene, plant variety, and the number of insertion events in the plant genome (Yin et al., 2003; Xi et al., 2009).

1.4 Importance of the virC Operon

1.4.1 The host range determinant

The genes contained on the virC were operon first characterized and described as a “host range determinant” during the early 1980’s, but have proven to be an important
part of the process of *Agrobacterium* mediated plant transformation (Klee et al., 1983; Lu et al., 2009).

Like all other *vir* genes of *A. tumefaciens*, the *virC* operon is housed on the Ti plasmid, specifically located between the *virG* and *virD* operons (Figure 2). However, the *virC* operon is different than all other *vir* genes, it is located on the opposing DNA strand (Figure 2). This specific positioning seems to be conserved among *Agrobacterium* strains. When examining nopaline and octopine type Ti plasmids, both have the *virC* operon located on the DNA strand opposite to the other virulence genes (Close et al., 1987; Yanofsky et al., 1985; Komari et al., 1986). Though the alternate strand placement is curious, whether it provides any significant advantage to the organism’s fitness remains to be discovered.

The *virC* operon contains two open reading frames, 696 and 609 nucleotides in length, coding for the VirC1 and VirC2 proteins, respectively (Close et al., 1987). Often referred to as “host range determinants,” the loss of either *virC1* or *virC2* leads to a range of phenotypes from a mild reduction in virulence to full avirulence depending on the host plant infected (Yanofsky et al., 1985; Yanofsky and Nester, 1986).

Specifically, in 1983, it was observed that both *virC1* and *virC2* loss of function mutants produced a severely attenuated phenotype on *Vinca rosea* (periwinkle) and *Kalanchoe daigremontiana*. Yet when these strains were used to infect other plants such as *Datura stramonium* (jimsonweed) and *Nicotiana tabacum* cv. Havana (tobacco), a milder attenuation of virulence was observed (Close et al., 1983).
1.4.2 Function of \textit{virC1} and \textit{virC2}

It is clear that the loss of either \textit{virC1} and or \textit{virC2} causes attenuation to a varying degree, but what are the mechanisms behind this phenomenon? The first clue to the function of these proteins comes from a study by Horsch et al. (1986). For this research, a system was developed to compare rates of T-DNA transfer from a binary vector and from the native Ti plasmid to leaf discs of \textit{Petunia hybrida}. It was possible to determine the source of the DNA transferred to the leaf discs based on specific markers in the DNA.

The system was used to examine rates of T-DNA transfer from several \textit{vir} gene loss of function mutants caused by transposon insertion. For most mutant strains tested, rates of T-DNA transfer uniformly declined from both the Ti plasmid and the binary vector when compared to wild type. However, in \textit{virC1} and \textit{virC2} loss of function mutants, it was observed that while rates of T-DNA transfer from the binary vector and Ti plasmid both declined, T-DNA was transferred much more frequently from the binary vector than the Ti plasmid (Horsch et al., 1986).

The researchers speculated this may occur because the binary vector could act like a natural intermediate formed by the T-DNA during the normal transfer process. The results suggested \textit{virC1} and \textit{virC2} were essential to these early T-DNA processing reactions needed to form an intermediate structure that was not necessary when T-DNA was being transferred from the binary vector (Horsch et al., 1986).

In order to better understand the function of the VirC proteins it is also important to understand the overdrive sequence located on the Ti plasmid, adjacent to the right border (Peralta et al., 1987; van Harren et al., 1987). The overdrive sequence has been shown necessary for full virulence by playing a role in T-DNA transfer efficiency and
possibly in the formation of an intermediate structure (van Harren et al., 1987; Toro et al., 1988). Both VirC1 and VirC2 have been shown to bind the overdrive region through affinity chromatography, supporting the idea that these proteins play a role in early T-DNA processing (Toro et al., 1988).

In 2007 a deeper look at the function of the VirC proteins was taken (Atmakuri et al., 2007). First, researchers examined the effects of the VirC proteins on T-strand production, finding that when VirC1, VirC2, or both were non-functional, the number of T-strand molecules generated in the A. tumefaciens cell was reduced by three- to four-fold. The ability of the VirC1 and VirC2 proteins to bind DNA on the Ti plasmid was demonstrated through co-immunoprecipitation assays. Additionally, VirC1 and VirC2 were reported to work together in early T-DNA processing reactions. Without the presence of VirC2, VirC1 antibodies precipitated fewer copies of the Ti plasmid compared to the wild-type strain. Also, through the use of co-immunoprecipitation assays, researchers were able to show interactions between VirC1 and VirC2. VirC1 was also shown to interact with VirD1, VirD2, and VirD4, other proteins known to be involved in the reactions of early T-DNA formation. Combined, the results suggest VirC1 and VirC2 work together and along with the VirD proteins, stimulate reactions occurring early in the process of T-DNA formation (Atmakuri et al., 2007).

A more recent report from Lu et al. (2009) details the crystal structure and molecular function of the VirC2 protein. A crystal structure was developed depicting the stable C-terminal end of the VirC2 protein from amino acids 82-202. From this crystal structure, researchers were able to detect structural homology between VirC2 and the superfamily of Arc/MetJ repressors (Lu et al., 2009). An important similarity between
VirC2 and this group of repressors is the ribbon-helix-helix (RHH) domain usually used by bacterial proteins to bind DNA (Schreiter and Drennan, 2007). By analyzing VirC2’s crystal structure, researchers were able to find a B-strand-helix-helix motif closely resembling the RHH domain found in Arc/MetJ repressors (Lu et al., 2009).

In the superfamily of Arc/MetJ repressors, two proteins, each containing one RHH domain, work in a dimeric fashion to bind DNA. The VirC2 protein uses two RHH domains within the same protein for DNA binding (Schreiter and Drennan, 2007; Lu et al., 2009).

Lu et al. (2009) examined the ability of VirC2 to bind DNA. Because VirC2 had been shown to bind the overdrive region in the past, researchers used electrophoretic mobility shift assays (EMSA) to determine VirC2’s binding affinity for three different 20 bp regions of double-stranded DNA adjacent to the right border sequence. One of the regions contained the overdrive core sequence. It was observed that full-length VirC2 was able to bind each region of DNA with a higher binding affinity for the DNA segment containing the overdrive sequence. VirC2 was also observed to bind an unrelated 20 bp DNA sequence (used as a control) suggesting a high level of unspecific binding to double stranded DNA. VirC2 was not shown to bind any single-stranded DNA targets (Lu et al., 2009).

To examine the importance of individual amino acids’ role in the DNA-binding process, Lu et al. (2009) generated several mutants containing a missense mutation. These mutants were then assayed for their ability to bind DNA through EMSAs and for virulence by Nicotiana glauca and Kalanchoe daigremontiana tumorigenesis assays. In several of the strains with an altered amino acid located on the surface of the protein
(hypothesized to be involved in DNA binding) a lower binding affinity for DNA and a reduction in virulence was observed, suggesting those particular amino acids were important for DNA binding. By demonstrating the ability of the VirC2 protein to bind to the overdrive region and showing its important for full virulence, Lu et al. (2009) provides further support of hypothesis that VirC2 contributes to early T-DNA formation through its interaction with the overdrive region.

Although research has shown that VirC proteins bind the overdrive region, facilitate early T-DNA processing reactions, and play an important role in virulence, there remain new discoveries to be made. In order to further explore the virC operon, a closer look at the world of regulatory RNA must be taken.

1.5 Non-coding Regulatory RNA

1.5.1 Growing importance of regulatory RNA across life

Regulatory RNAs have been found to be present in every domain of life and often play a role in complex regulatory processes (Gottesman and Storz., 2011; Marchfelder et al., 2012; Barrett et al., 2012). Because A. tumefaciens is such an important organism for biotechnology and genetic engineering, it is important to investigate the possible importance of regulatory RNAs on the biology of A. tumefaciens. In addition, elucidation of novel regulatory RNAs may allow new methods to be developed to modify A. tumefaciens thereby improving the efficiency of plant transformation.

Researchers have long known that RNA has the ability to perform complex functions. Some researchers even hypothesize in the RNA World Theory that at one time,
RNA not only acted as a messenger between DNA and proteins, but also, performed information storage and catalytic functions (Higgs and Lehman, 2015).

The early role of RNA (instead of DNA) was first identified in the 1960’s when researchers realized an all RNA world would be a less complex forerunner to the current model using DNA, RNA, and protein (Crick, 1968; Orgel, 1968). In later years, researchers observed ribonucleotide structures able to act as coenzymes with proteins, further confirming RNA’s ability to perform catalytic functions, and suggesting current coenzyme activity maybe a remnant from the RNA world (White, 1976). However, in the 1960’s and 70’s RNA was still considered to act mainly as a template between DNA and proteins in modern organisms, or in the case of tRNA and rRNA, serve as a structural feature for the construction of proteins (Hoagland et al., 1958; Brenner et al., 1961).

The idea of RNA acting as a regulatory molecule in modern organisms was first proposed by Britten and Davidson in 1969 after taking notice of a discovery previously reported years earlier of a diverse population of heterogeneous nuclear RNA (hnRNA) existing in the nucleus (Warner et al., 1966). The observation of a complex pool of RNA led researchers to propose a model in which RNA acts in a network to regulate processes within the cell (Britton and Davidson, 1969; Britton and Davidson, 1971). However, this hypothesis was short lived as researchers focused on other ideas.

Researchers were eventually led back to the subject in 1993 when a lab studying Caenorhabditis elegan development began taking a closer look at mutants which displayed extra larval molts and exhibited early cell fates at later, inappropriate times in development (Lee, et al. 1993). Researchers already knew these abnormal phenotypes were the cause of a mutation in the lin-4 gene (Chalfie et al., 1981). As research
progressed on the *lin-4* gene, evidence showed *lin-4* coded for an RNA able to regulate target genes by binding their mRNAs at regions of sequence complementarity (Lee et al., 1993). Later, this type of regulatory RNA would become known as microRNA (miRNA) (Lagor-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001).

Using techniques such as size selection and targeted cloning, many additional miRNAs were soon discovered in both plant and animal cells (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001; Reinhart et al., 2002). Another similar class of regulatory RNA named small interfering RNA (siRNA) soon emerged (Timmons and Fire, 1998; Wianny and Zernicka-Gotez, 2000; Xia et al., 2002). SiRNA along with miRNA make up the two main classes of regulatory RNA present in eukaryotic cells (Filipowicz et al., 2005; Carthew and Sontheimer, 2009).

miRNAs transcripts originate from a variety of regions within a eukaryote’s genome and may be found independently, in clusters, or within introns (Bartel, 2004). Once a miRNA is transcribed, excess nucleotides on both the 5’ and 3’ ends are trimmed and then the miRNA is folded into a stem loop structure (Bartel, 2004). This processing step is primarily performed by the Dcl1 protein in plants, while Dicer and Drosha perform these functions in animal cells (Kim, 2005). After processing, one strand of the miRNA duplex structure then associates with the Ago protein along with additional proteins to form the miRNA-induced silencing complex (miRISC) and the other strand is lost (Bartel, 2004; Kim, 2005). The miRISC complex is then able to recognize mRNAs which have complementary to the miRNA and cause translational repression or mRNA cleavage (Wu et al., 2006; Cathew and Sontheimer, 2009).
siRNAs function in a similar manner to miRNA, although there are some notable differences. While miRNAs originate from the organism’s own genome, siRNA usually originate from cleavage of long double stranded RNA (dsRNA) that comes from outside the cell (Mello and Conte, 2004). siRNA can originate in nature when the plant cell detects viral dsRNA and processes the dsRNA into siRNA for defense, but may also be used by researchers to induce silencing of desired genes (Fire et al., 1998; Hannon, 2004; Mlotshwa et al., 2008). When dsRNA enters the cell it is first cleaved by Dicer into 21-24 bp long siRNA (Bernstein et al., 2001; Ketting et al., 2001). Like miRNAs, one strand of the siRNA duplex along with proteins including Ago, form the siRISC complex while the other strand is lost (Liu et al., 2006). The siRISC is then able to cleave RNA strands with sequence homology to the siRNA (Tomari et al., 2007; Wilson and Doudna, 2013).

While siRNA and miRNA are the most well studied small RNA (sRNA) in eukaryotic cells there are several additional classes of RNAs important for gene regulation. Long non-coding RNA (lncRNA) are a class of regulatory RNA that can be easily differentiated by their length. Usually lncRNA are longer than 200nt and do not code for proteins (Morris and Mattick, 2014). For many lncRNA there is no evidence of functionality, yet for other species there are clues suggesting they may serve a purpose (Morris and Mattick, 2014). For example, some lncRNA are differentially expressed depending on tissue and cell type (Mercer et al., 2008; Moràn et al., 2012). Another class of regulatory RNA are piwi-interacting RNAs (piRNA) which are involved with key epigenetic modifications to the chromatin of germ line cells (Bamezai et al., 2012). Circular RNAs (circRNAs) are a recently discovered class of regulatory RNA in animal cells which form a closed loop structure and regulate gene expression in a post-
transcriptional manner (Memczak, et al. 2013). With the research that has already been completed, it is clear that regulatory RNA are an important part of gene regulation. In the coming years, as research in the field progresses, additional classes and functions of regulatory RNA will continue to be discovered.

1.5.2 Non-coding regulatory RNA in bacteria

Non-coding RNA is also important for gene regulation in prokaryotic cells, though the regulatory mechanisms are somewhat different. Specifically, in bacteria there are two main classes of regulatory RNA, riboswitches and sRNA (Waters and Storz. 2009). Riboswitches are portions of mRNA which are able to detect and respond to metabolites or other environmental cues. Riboswitches respond by causing a change in mRNA conformation, decreasing translation and/or stability, resulting in a change in gene expression. However, because riboswitches exist as part of an mRNA, riboswitches are not non-coding regulatory RNA. In this introduction I will focus my attention on sRNA; which unlike riboswitches, bind and alter expression of mRNA and usually have no translational function on their own.

Contrary to their name, sRNA range in size from 50 to 300 bp in length (Storz et al., 2011), notably different from the miRNA and siRNA regulators in eukaryotic cells that average 20-25 bp (Starega-Roslin et al., 2011). sRNA regulators are able to target specific mRNA through base pairing, resulting in modified gene expression (Repiola and Dafreuille, 2009). In some cases sRNA are also able to bind and regulate proteins (Suzuki et al., 2006; Babitzke and Romeo, 2007).
sRNA may be located anywhere in the genome relative to their target mRNAs. When a sRNA is transcribed from the non-coding DNA strand directly opposite its target mRNA it is said to be a *cis*-acting sRNA; if the sRNA is located elsewhere, it is said to be a *trans*-acting sRNA (Waters and Storz, 2009). Both *cis*-acting and *trans*-acting sRNAs are able to form secondary stem loop structures and base pair with target mRNA sequences, similar to the miRNA mechanism of action in eukaryotic cells (Mahmood et al., 2008; Beisel and Storz, 2010). While *cis*-acting sRNA are usually able to directly basepair with target mRNAs, *trans*-acting sRNA sometimes require the RNA chaperone protein Hfq to facilitate binding between the sRNA and target mRNA (Vogel and Luisi, 2011). The nucleotides on the exposed portion of trans-acting RNA are critical for determining if interaction will occur between a sRNA-mRNA pair. This region is named the seed sequence (Brantl, 2007). The correct seed sequence along with facilitation by Hfq allows *trans*-acting sRNA to regulate mRNA that do not have complementary sequences (Kawamoto, et al. 2006). In fact, most *trans*-acting sRNA usually act to regulate more than one target mRNA (Papenfort and Vogel, 2009).

Binding of a target mRNA by a sRNA usually results in a negative regulation of the mRNA, though in some cases the mRNA may be positively regulated (For detailed reviews see Waters and Storz, 2009; Kang et al., 2014). Most commonly, the sRNA will bind to the mRNA and cover the ribosome binding site, leading to a decrease in translation (De Lay et al., 2013). Binding of the sRNA to the mRNA can also lead to degradation of the complex which can occur separately or along with translational repression (Waters and Storz, 2009). In order to positively regulate a target mRNA, sRNAs cause a change in mRNA conformation through binding, making the ribosomal
binding site more accessible, and/or giving rise to a stabilizing effect (Waters and Storz, 2009; Soper et al., 2010).

There are several reasons why using sRNA for regulation is advantageous to bacteria. First, bacteria often need to quickly regulate genes in response to rapidly fluctuating environments. sRNA are able to be both synthesized and degraded more quickly than protein, making regulatory RNA ideal when a quick response is needed (Shimoni, et al. 2007). Secondly, producing a sRNA requires fewer resources when compared to protein production. During protein synthesis there is the additional step of translation, while sRNA synthesis is able to be stopped after transcription. Also, sRNAs can provide an additional level of regulation alongside transcription factors and other proteins, leading to a greater fine tuning of biological responses (Levine et al., 2007; Mehta et al., 2008).

There are many examples in the literature in which bacteria utilize sRNA to regulate various biological processes. For example, in *E. coli*, a recent review estimated between 80 and 100 sRNA are present with roles in many processes including oxidative stress tolerance, iron regulation, sugar metabolism, and cell division (Gottesman, 2004; Gottesman and Storz, 2011)

Other well-known examples of regulatory RNA in bacteria include the sRNA Qrr which is involved in the quorum sensing system of *Vibrio cholerae* (Lenz et al., 2004; Hammer and Bassler 2007). When cell density is low, four Qrr sRNAs are expressed which, along with Hfq, bind the *hapR* mRNA causing destabilization and a low level of HapR protein (Hammer and Bassler, 2007). HapR is a master regulator of quorum sensing and normally acts to repress genes associated with biofilm production and
bacterial colonization, while at the same time activating the gene coding for the hap protease (Jobling and Holmes 1997; Hammer and Bassler, 2003; Zhu and Mekalanos, 2003). When HapR is downregulated, repression targets of HapR are then allowed to be expressed, while genes which are normally activated by HapR are not transcribed (Ng and Bassler 2009). In summary, by acting on HapR the four Qrr sRNAs are able to regulate the quorum sensing system in Vibrio cholerae.

Another example of a well-studied sRNA comes from the bacteria this work is centered on, A. tumefaciens. During the process of infection, plants produce γ-aminobutyric acid (GABA) in wounded tissues as a stress response mechanism (Kinnersley et al., 2000). GABA is then transported into the bacterial cell through an ABC transporter where it disrupts quorum signaling by degradation of the quorum signal, homoserine lactone (Wilms et al., 2011). In a recent study, it was shown that A. tumefaciens possesses a sRNA, AbcR1, which functions by downregulating Atu2422, a periplasmic binding protein essential to the process of GABA uptake, by binding and covering the shine delgarno sequence (Wilms et al., 2011). The group went on to hypothesize that because this sRNA is able to block uptake of GABA, it is acting to preserve bacterial quorum sensing in the presence of plant defense mechanisms (Wilms et al., 2011).

These examples demonstrate a few of the diverse ways bacteria are able to utilize sRNA for gene regulation. As the topic continues to be studied, additional processes involving sRNA will come to light. However, when studying sRNA in bacteria there are many challenges that arise. One of
Using the sequences of mRNA and sRNA matched for complementarity is not enough to predict interaction since base pairing is most often imperfect and could occur at various regions within the transcript. Improving the accuracy of predicting interactions is being tackled most effectively through bioinformatics. There are several programs that are able to predict mRNA targets given bacterial sRNA sequences. These include CopraRNA, IntaRNA, and TargetRNA1 and 2, along with others (Busch et al., 2008; Tjaden, 2008; Wright et al., 2013; Kery et al., 2014). In a recent review paper, researchers detailed different ways these programs were able to predict targets (Pain et al., 2015). Some programs which the group calls “alignment like” search for targets based on sequence complementarity (Pain et al., 2015). Other programs consider the thermodynamics of interactions between specific RNAs and are called “inter-RNA” while another class called “independent fold” uses the way an RNA folds and makes a secondary structure to predict the energy for the two transcripts to bind (Pain et al., 2015). In the same review researchers go on to evaluate various programs using known sRNA/mRNA pairs in *E. coli*. They find that while programs that predict short regions of pairing instead of pairing based on the entire transcript more accurately predict interactions, though there are still some targets that remain difficult to predict regardless of the program. Overall, CopraRNA most accurately identified real targets but was only able to do so for organisms with genomes already in the software. The program also requires the user to input several homologous sRNAs. The researchers found that IntaRNA, RNAplex and RNAup were able to perform well and could be used under circumstances when CopraRNA is unavailable (Pain et al., 2015).
While there remain many challenges in the study of sRNA in bacteria, research is crucial to the full understanding of bacterial gene regulation. As time goes on and research progresses there will undoubtedly be more sRNA uncovered.

1.5.3 Non-coding regulatory RNA in *Agrobacterium tumefaciens*

In *A. tumefaciens*, there is a growing understanding of the function and importance of regulatory RNA. Recently, researchers in three independent studies used RNA-seq in attempts to identify non-coding regulatory RNA in *A. tumefaciens*. Between the three studies, there were a total of 1560 different transcripts identified as potential regulatory RNAs (Wilms et al., 2012; Lee et al., 2013; Dequivre et al., 2015).

There has also been further investigation into the function of several of these regulatory RNAs. As described previously, the sRNA AbcR1 plays a role in regulating GABA uptake and the preservation of quorum sensing by regulating atu2422 (Wilms et al., 2011). Another sRNA that has been characterized in greater detail is repE. Located on the Ti plasmid, repE is involved with the Ti plasmid’s replication (Cho and Winans, 2005). The genes repA, repB, and repC make up the repABC operon and are essential for replication of the Ti plasmid (Tabata et al., 1989). Located inside the intergenic region between repB and repC is the sRNA repE (Cho and Winans, 2005). When expressed, it acts to block replication of the Ti plasmid by preventing the translation of repC mRNA (Cho and Winans, 2005).

Another sRNA, rna1111, was shown to play a role in virulence by targeting the genes, including traA, involved in bacterial virulence (Dequivre et al., 2015).
Though these examples show how sRNA can be characterized, the process often remains a challenge. Currently only a small portion of known sRNAs are fully characterized. This gap in knowledge exists as an area in research where much still needs to be learned.

1.6 Research Objectives and Thesis Organization

1.6.1 Research objectives

This research began with the goal of identifying and characterizing a sRNA in A. tumefaciens. Prior to my joining the lab, experiments to identify potential sRNA had already taken place. A paper previously published (Lee et al., 2013) details the identification of 475 candidate noncoding RNAs from the genome of A. tumefaciens strain C58 after growth under a variety of conditions prior to RNA isolation and sequencing.

Lee et al. (2013) found candidate non-coding RNAs on every replicon in the A. tumefaciens genome. With a large number of non-coding RNA candidates identified, the next step is to validate and characterize these transcripts. Because Agrobacterium plays such an important role in plant transformation the investigation of how sRNAs affect virulence is especially important. It became my research goal to verify and characterize a sRNA involved with pathogenic virulence. Such a sRNA may also play a role in plant transformation.

The logical place to begin the search for a sRNA involved in virulence was on the Ti plasmid. By looking at RNAseq data it became clear that the virC2 antisense strand
may encode a *cis*-acting sRNA; this is evident from the high amount of antisense transcript expression. This sRNA was also previously identified in a similar study characterizing non-coding regulatory RNA in 2012 and named Ti4 by that group (Wilms et al., 2012). There are several reasons to study this candidate sRNA further. First, it is expressed in *cis*, opposite *virC2*. While sRNA expressed in *trans* regulate targets located elsewhere in the genome, which may be difficult to find, sRNA expressed in *cis* regulate targets on the opposite strand. Second, this sRNA has the potential to be involved with bacterial virulence because *virC2* plays an important role in the excision of T-DNA from the Ti plasmid.

Another interesting phenomenon to be noted is the difference in expression between the *virC2* and *virC1* transcripts, which are part of the same operon. Usually genes transcribed from the same operon have a similar level of expression with a slight reduction in expression as the distance from the start site increases and transcription machinery pauses or prematurely stops (Lengeler et al., 1999). Here we are able to see a dramatic difference between the transcript expression of *virC1* and *virC2*. Might the anti-sense RNA opposite *virC2* be somehow influencing the expression of the *virC2* sense transcript?

The goals of this research were to explore the role of the *anti-virC2* transcript as well as that of the *virC2* gene through generation of several mutant strains of *A. tumefaciens*. The mutants were then examined for changes in gene expression, protein expression, and virulence. A variety of experiments were used to accomplish this including qRT-PCR, tumorigenesis assays using *Kalanchea daigromontiana*, and Western blot analysis. *Agrobacterium*-mediated enhanced seedling transformation (AGROBEST)
is a technique previously described in which *A. tumefaciens* is used to transiently infect *Arabidopsis thaliana* seedlings (Wu et al. 2014). AGROBEST was also used in this research to compare the ability of *A. tumefaciens* strains to cause transient GUS expression in *Arabidopsis thaliana* seedlings.

### 1.6.2 Thesis organization

This thesis is organized into four chapters. Chapter I consists of a general introduction, literature review, and an overview of the organization of this thesis. Chapter II presents materials and methods used in this research. Presented in Chapter III are the results while Chapter IV presents general conclusions of this work.

My Major Professor and mentor, Dr. Kan Wang, along with an additional mentor in the lab, Dr. Keunsub Lee, have provided me with advice and guidance along every step of this project. The work for these projects was carried out by myself and Dr. Keunsub Lee. Specifically, Rapid Amplification of cDNA Ends (RACE) experiments and qRT-PCR experiments were done primarily by Dr. Keunsub Lee with assistance from myself. Additional assistance was received from others in the lab to optimize protein extraction and RNA isolation including Dr. Keunsub Lee, Dr. Jen Raji and Dr. Evangelia Vamvaka. Assistance for GUS quantification was provided by Dr. Dangping Luo from Dr. Yang’s lab at Iowa State. My committee, Dr. Allen Miller and Dr. Gwyn Beattie, along with Dr. Kan Wang also provided me with ideas and suggestions to improve this project. Without help from these individuals this work would not have been possible.
1.7 References


Rossi, L., Hohn, B., & Tinland, B. (1993). The VirD2 protein of \textit{Agrobacterlum tumefaciens} carries nuclear localization signals important for transfer of T-DNA to plants. \textit{Molecular and General Genetics}, 239(3), 345-353.


Figure 1. The T-DNA transfer process. (Figure directly from Pitzschke and Hirt, (2010) Figure 1). Illustration of the plant transformation process by Agrobacterium tumefaciens. The T-strand is cleaved from the Ti plasmid and integrated into the plant genome where it may be expressed.
Figure 2. Illustration of virulence genes on Ti plasmid in *A. tumefaciens*. (Figure directly from Cho and Winans. (2005) Figure 1). Graphic shows the arrangement of virulence genes on Ti plasmid from both A) nopaline and B) octopine type Ti plasmids. *VirC* operons are highlighted with the addition of a blue box. Red bars represent protein coding genes, while black bars represent IS elements, and gray bars represent uncharacterized proteins.
CHAPTER II
MATERIALS AND METHODS

2.1 Bacterial Strains and Plasmids

2.1.1 General cloning procedures

2.1.1.1 Extraction of total DNA from *A. tumefaciens*

Bacterial strain C58 from the Wang lab collection was used in all experiments unless otherwise specified. Bacteria were streaked from glycerol stock on fresh YEP plates (5 g/L yeast extract, 10 g/L peptone, 5 g/L NaCl 12 g/L bacto-agar if solid and the pH adjusted to 6.8 with NaOH) (Wang, 2006) and grown for two days at 28 °C. Bacterial colonies were transferred to YEP liquid medium and grown overnight, for approximately 20 hours, shaking at 28 °C at 250 rpm. Bacterial cultures were then centrifuged at 5000 x g for 10 minutes at room temperature (25 °C) and the supernatant was removed. DNA was then extracted following the standard protocol with buffers and solutions from the Qiagen DNeasy Blood and Tissue Kit for gram-negative bacteria (Qiagen Inc., Valencia, CA). Cells were resuspended in 180 μL buffer ATL with the addition of 20 μL proteinase K. Samples were vortexed and incubated at 56 °C until lysed. Then 200 μL of buffer AL was added and samples were vortexed prior to the addition of 20 μL of 100% ethanol. Samples were transferred to the DNeasy mini spin column and centrifuged at 6000 x g for 1 minute. The column was then washed with 500 μL of buffer AW1 followed by 500 μL of buffer AW2. DNA was eluted with 200 μL buffer AE (Qiagen
Inc., Valencia, CA). DNA concentration was measured using a NanoDrop ND1000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA) and confirmed by gel electrophoresis on a 1% agarose gel.

2.1.1.2 Extraction of plasmid DNA from *A. tumefaciens*

Bacteria used for plasmid extraction were grown on fresh YEP plates containing appropriate antibiotics for two days at 28 °C. The antibiotic resistance of each strain can be found in Tables 8 and 9. Kanamycin was used at a concentration of 50 μg/mL, spectinomycin was used at a concentration of 100 μg/mL, and carbenicillin was used at concentration of 100 μg/mL. Bacterial colonies were selected and transferred to YEP liquid and grown overnight, for approximately 20 hours, with appropriate antibiotics, shaking at 28 °C at 250 rpm. Bacterial cultures were then centrifuged at 5000 x g for 10 minutes at room temperature (25 °C) and plasmid DNA was extracted using the QIAprep Spin Miniprep Kit (Qiagen Inc., Valencia, CA) following the manufacturer’s instructions with a few minor alterations to increase yield of plasmid DNA. Twice the recommended amounts of buffers P1, P2 and N3 (Qiagen, Inc., Valencia CA) were used. Briefly, bacteria were resuspended in 500 μL of buffer P1, 500 μL of buffer P2 was then added and tubes were inverted six times. In the next step 600 μL of buffer N3 was added and tubes were mixed six times. Samples were then centrifuged at 13000 x g for 10 minutes at room temperature (25 °C) before transfer to a mini spin column. Because of the additional liquid caused by the increased volumes of buffers used, it was necessary to run the supernatant through the mini spin column in 2-3 aliquots. Spin columns were washed with 500 μL of buffer PB followed by 750 μL of buffer PE. A two-minute centrifuge step
was used to dry the spin column before elution with 50 μL of pure water. DNA concentration was measured using the NanoDrop ND1000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA) and confirmed by gel electrophoresis on a 1% agarose gel.

2.1.1.3 Extraction of plasmid DNA from *E. coli*

Plasmid DNA was extracted from *E. coli* using the QIAprep Spin Miniprep Kit (Qiagen Inc., Valencia, CA). Bacteria used for plasmid extraction were grown on fresh LB plates (25 g/L LB powder, 12g/L bacto agar, pH adjusted to 6.8 with NaOH) containing appropriate antibiotics for one night at 35 °C. Bacterial colonies were selected and transferred to LB liquid and grown with appropriate antibiotics overnight, for approximately 20 hours, shaking at 35 °C. Bacterial cultures centrifuged at 5000 x g for 10 minutes at room temperature (25 °C) and plasmid DNA was extracted using standard protocols for QIAprep Spin Miniprep Kit similar to those described above.

2.1.2 Construction of *A. tumefaciens* mutants using homologous recombination

All primers and plasmids, fragments, and bacterial strains used are listed in Tables 6 through 9 in Appendix A. *A. tumefaciens* strain C58 was used as a wild-type strain to generate all mutant strains. In order to knockout or change desired regions of DNA, the method of homologous recombination was used. This method has been widely used and is a common way used to delete or alter bacterial gene sequences (Reid and Collmer, 1987). A general overview of this process is illustrated in Figure 3.
2.1.2.1 Generation of knockout mutant strains ∆Pavc2, and ∆virC2

Two different *A. tumefaciens* mutants (C58 ∆Pavc2 and C58 ∆virC2) were generated with genetic sequence deletions through homologous recombination. In order to knockout a gene using homologous recombination, up- and down-stream flanking sequences were selected that surround the region to be knocked out. These flanking sequences were then amplified from *A. tumefaciens* total DNA using PCR with primers that have included restriction sites, allowing for the fragment to be easily cloned into the sequencing vector pJET1.2 (ThermoFisher Scientific, USA). An initial sequencing step was used to check for mutations. Flanking sequences were then excised from of pJET1.2 and cloned into the vector pTFsacB. By ligating the flanking sequences into pTFsacB, a sequence was generated that was identical to that in the *A. tumefaciens* genome but excluded the region to be knocked out. Using the natural process of homologous recombination, the chromosomal DNA was allowed to recombine with this sequence present on pTFsacB and the desired knockout mutants were generated. pTFsacB contains the *nptII* gene for kanamycin resistance along with the *sacB* gene, which prevents growth on media containing sucrose. In addition, there pTFsacB lacks an *A. tumefaciens* replication of origin in the plasmid so the plasmid could only persist if it was incorporated into the genome through an initial homologous recombination event, which was screened for using kanamycin resistance. A second recombination was detected by the ability of bacteria to grow on a medium containing sucrose. After two recombination events the native gene sequences were replaced with the gene sequences present on the pTFsacB vectors and knockout mutant will be generated.
The putative promoter region of \textit{anti-virC2} was identified using Softberry bacterial promoter prediction software (Solovyev and Salamov, 2011). A 42 bp region including the identified promoter sequence was targeted for deletion. Primers avcR1-UP-F and avcR1-UP-R were designed to amplify the upstream flanking sequence (Appendix A, Table 6) while primers avcR1-DN-F and avcR1-DN-R were designed to amplify the downstream sequence. Restriction enzyme sites PstI, SalI and XbaI were built into the primers and used for all cloning steps. Total DNA from \textit{A. tumefaciens} was used to amplify up and downstream flanking sequences. Fragments were then cloned into vector pJET1.2 and an initial sequencing step was used to check for possible mutations. A three-piece ligation was then performed to ligate both up- and down-stream flanking sequences into the vector pTFsacB using the PstI/SalI/XbaI restriction sites. The resulting vector was named pAJ007. A detailed diagram of cloning used to generate pAJ007 can be found in Figure 4.

The entire \textit{virC2} gene sequence was also knocked out using homologous recombination. Up-stream and down-stream sequences were identified and amplified using primers VC2KO-UP-F, VC2KO-UP-R, VC2KO-DN-F, and VC2KO-DN-R (Appendix A, Table 6). In addition, primers were designed to contain restriction sites PstI, KpnI and XbaI to be used for cloning steps. After an initial sequencing step to prevent any unwanted mutations using sequencing vector pJET1.2, these fragments could be ligated into the pTFsacB vector. The resulting vector, named pAJ021, was then used to generate the knockout mutant C58 ΔvirC2. A detailed diagram of cloning used to generate pAJ021 can be found in Figure 5.
Each of the vectors used to generate knockout mutants was separately transformed into *A. tumefaciens* using electroporation. Bacteria were then screened for kanamycin resistance, indicating an initial recombination event and incorporation of the knockout vector. A second recombination event was detected by the ability of the bacteria to grow on sucrose. After the second recombination event bacterial colonies were screened for the ability to grow in the presence of sucrose indicating removal of the vector backbone. Presence of the knock-out region was then screened for using PCR primers vc2DR-F, vc2DR-R, VirC2F, and VirC2R (Appendix A, Table 6) and later confirmed by sequencing.

### 2.1.2.2 Construction of FLAG-tagged *A. tumefaciens* mutants

VirC2 protein was FLAG-tagged in several *A. tumefaciens* strains using homologous recombination. Design was carried out by Dr. Keunsub Lee while cloning steps listed in this section were carried out by Dr. Keunsub Lee with additional assistance as needed from myself. A gene fragment named 3xFLAG_virC2 (Appendix A, Table 7) was designed and cloned into vector pTFsacB for use in creating VirC2 FLAG-tagged mutant strains. A 1229 bp DNA fragment was designed to contain the *virC2* coding sequence with an N-terminal 3xFLAG-tag inserted in frame directly after the start codon as well as an additional upstream flanking sequence. HindIII and BamHI restriction sites were also included to allow for cloning into vector pTFsacB. Fragments were ordered from Integrated DNA Technologies (IDT) and following digestion, were directly ligated into pTFsacB. The resulting vector was named pKL1007. A detailed diagram of cloning used to generate pKL1007 can be found in Figure 6.
pKL1007 was transformed into *A. tumefaciens* wild-type strain C58 where LB containing kanamycin was used to screen for an initial recombination event. A second recombination event was identified by screening for the ability of bacteria to grow on LB media containing sucrose. The presence of a FLAG-tag was further confirmed by PCR.

2.2 Transcript Quantification

2.2.1 Bacterial growth conditions

For transcript analysis, bacteria were grown under induction conditions. Bacterial strains were stored in glycerol stock and streaked with a sterile inoculation loop on to fresh YEP plates with appropriate antibiotics and grown for three days at 28 °C. Colonies were selected and transferred to 5 mL YEP medium in a sterile 50 mL falcon tube along with appropriate antibiotics. Bacteria were grown overnight, for approximately 20 hours, in a 28 °C shaker at 250 rpm.

Bacteria were transferred to 5 mL freshly made AB media (described below) from YEP cultures at a dilution of 1:100 and grown overnight, for approximately 20 hours, at 28 °C shaking at 250 rpm. AB medium (one liter) was comprised of 50 mL 20x AB salts solution (20 g/L NH₄Cl, 6 g/L MgSO₄·7H₂O, 3 g/L KCl, 0.2 g/L CaCl₂, 50 mg/L FeSO₄·7H₂O, pH to 7 with KOH), 50 mL 20x AB buffer (60 g/L K₂HPO₄, 20 g/L NaH₂PO₄), and 900 mL sucrose water for a final sucrose concentration of 0.5%. Bacterial cultures were then centrifuged at 4000 x g for 10 minutes at room temperature (25 °C). The supernatant was discarded and bacteria were resuspended in two volumes of induction medium with 0.1mM of acetosyringone (AS) (Thermo Fisher Scientific, USA). Induction media (one liter) was comprised of 50 mL 20x AB salts solution as described
above, 2 mM NaPO₄, 50 mM MES pH 5.6, and 0.5% glucose. Identical cultures were grown without the addition of AS to serve as a non-induced comparison. Bacteria were then grown at 28 °C for 20 hours before isolation of RNA for transcript analysis.

2.2.2 RNA isolation from *A. tumefaciens*

RNeasy Protect Bacteria mini kit (Qiagen Inc., Valencia, CA) was used to isolate total RNA from *A. tumefaciens* cultures. This was carried out according to the manufacturer with minor alterations. All buffers and solutions were from the RNeasy Protect Bacteria mini kit (Qiagen Inc., Valencia CA), unless otherwise specified. In short, bacterial cultures grown with induction by AS were directly mixed with two volumes of RNAprotect bacteria reagent, vortexed, and incubated at room temperature (25 °C) for ten minutes. Cells were then centrifuged at 4000 x g for 10 min at room temperature (25 °C).

The pelleted cells were resuspended in 200 μL TE buffer, containing 1.5 mg lysozyme per mL and incubated at room temperature (25 °C) for 20 minutes with gentle vortexing every two minutes. Then 700 μL of buffer RLT was added and samples were mixed by vortexing. After mixing 500 μL of ethanol was added and samples were further mixed by pipetting. Samples were then centrifuged at 11,000 x g for 30 seconds at room temperature (25 °C) through the RNeasy mini spin column. An on-column DNase digestion was performed using DNase I (Invitrogen, USA) to eliminate DNA contamination followed by two washes with buffer RPE and elution with purified RNase free water.
Concentration of resultant RNA samples was measured using a NanoDrop ND1000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA). Samples were checked for residual DNA contamination and quality by gel electrophoresis. An additional step of DNase I treatment was then performed. The total RNA sample was incubated at 37 °C along with 2 μL of DNaseI (Invitrogen, USA), 1 μL of RNaseOUT RNase inhibitor (Invitrogen, USA), and 10 μL DNase I buffer for a total volume of 100 μL. After 30 minutes, an additional 2 μL of DNase I was added and samples were incubated for an additional 30 minutes. Following DNase treatment, samples were purified by phenol chloroform extraction and were then ready to use in downstream applications.

2.2.3 cDNA preparation

First, RNA was isolated as described above. Next, cDNA was synthesized from purified RNA using SuperScript III First-Strand Synthesis Kit (Invitrogen, USA). In short, primers were used to bind RNA transcripts and synthesize corresponding cDNA using PCR. Random hexamers were used to synthesize cDNA from total RNA transcripts while specific primers were used to synthesize specific cDNA transcripts. Typical reactions consisted of up to 6 μL RNA, 1 μL primer or random hexamer, 1 μL annealing buffer, and RNase free water up to a total volume of 8 μL. Reactions were incubated at 65 °C for five minutes before the addition of 10 μL 2x First-Strand Synthesis Mix and 2 μL SuperScript III Enzyme Mix. Reactions were then transferred to a thermocycler for 10 minutes at 25 °C, followed by 50 minutes at 50 °C, and finally five minutes at 85 °C. cDNA volume was then adjusted to 200 μL and used for qRT-PCR.
2.2.4 Quantification of transcripts using qRT-PCR

qRT-PCR was used to quantify transcript expression levels. SYBR Green PCR Master Mix (applied Biosystems, USA) was used for qRT-PCR reactions. Primers were designed specifically for each amplified transcript and are listed in Appendix A Table 6. Housekeeping gene rpoD was used as an internal standard for comparison. qRT-PCR reactions consisted 5 μL cDNA, 12.5 μL SYBR Green PCR Master Mix, 2.5 μL primer, and 5 μL RNase free water. Reactions were carried out with the following profile: 15 minutes at 95 °C, followed by 40 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C, and 30 seconds at 72 °C. This was followed by a data collection step at the end. Ct values were obtained and used to determine transcript levels relative to those of rpoD.

2.3 Rapid Amplification of cDNA Ends

3’ and 5’ RACE (Rapid amplification of cDNA ends) were both performed on the anti-virC2 transcript in order to determine transcription start and end sites. All RACE experiments were performed primarily by Dr. Keunsub Lee with assistance given by myself as needed.

RNA samples isolated as described previously were used in both 5’ and 3’ RACE. 5’ RACE began with a TAP (tobacco acid pyrophosphate, Epicentre USA) which was needed for ligation. Then, for both 3’ and 5’ RACE a specific adapter sequence was ligated to the end of the RNA sample (specific probes can be found in Appendix A, Table 6) and PCR was used to synthesize first strand cDNA from the ligated RNA. For 5’ RACE, random hexamers were used and for of 3’ RACE, adapter specific primers were used. After first strand cDNA synthesis, transcripts of interest were amplified by PCR.
with gene specific primers. Nested PCR was further used to ensure the correct target was amplified. PCR amplified fragments were then able to be cloned into sequencing vector pJET1.2 and sequenced.

2.4 Western Blot Protocol

2.4.1 Bacterial growth conditions and protein extraction

*A. tumefaciens* strains were streaked from glycerol stocks onto fresh YEP plates and grown for two days at 28 °C before inoculation onto 5 mL of YEP liquid media. Liquid cultures were grown overnight, for approximately 20 hours, at 28 °C shaking at 250 rpm, and pelleted by centrifugation for 10 minutes at 5000 x g at room temperature (25 °C). Pellets were then resuspended in 5 mL induction medium with a final OD$_{600}$ of 0.2 then grown overnight, for approximately 20 hours, at 28 °C shaking at 250 rpm, in the presence of 0.1 mM of AS.

Protein was extracted by use of the freeze and thaw method. Bacterial cells were adjusted to an OD$_{600}$ of 1.0 and equivalent volumes were centrifuged at 5000 x g for 10 minutes at room temperature (25 °C). Pellets were then frozen for three minutes in liquid nitrogen and allowed to thaw on ice for 15 minutes. A phosphate buffered saline solution that was pH 7.2 (PBS) containing the protease inhibitor tablet SIGMAFast (Sigma Aldrich, USA) at pH 7.2 and 2 mg/mL Lysozyme was used to resuspend the bacterial cells. Samples were then allowed to sit on ice for at least 30 minutes. Samples were diluted at a 1:1 ratio with Laemmli sample buffer (BioRad, USA) containing β-mercaptoethanol, boiled for five minutes in H$_2$O and centrifuged at 10,000 x g for 30
minutes at 4 °C. The resulting supernatant contained the protein sample which was used to load into the gel. Total protein concentration was measured using a NanoDrop ND1000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA)

### 2.4.2 Gel electrophoresis and Western blot

Protein samples were run on a 1.5 mm 15% Acrylimide/Bis gel at 200 volts until loading dye from the samples reached the bottom of the gel. Along with samples, Precision Plus Protein Standards (BioRad, USA) were also included. Proteins were then transferred to a 0.2 μm nitrocellulose membrane (BioRad, USA) for further probing. Semi-dry transfer was performed at 25 volts for 30 minutes.

Membrane with transferred protein was blocked overnight, for 16 hours, at 4 °C in 5% non-fat milk in PBS. The membrane was rinsed once with 1x PBS for five minutes followed by incubation with primary antibody for two hours shaking at room temperature (25 °C). The primary antibody was diluted in a solution of 1% non-fat milk in PBS to a concentration of ~0.33 μg/mL. In order to detect the FLAG signal rabbit-antiFLAG antibodies (Sigma Aldrich, USA) were used as the primary antibody.

Following incubation with primary antibody, the membrane was washed three times for five minutes each time with PBS with 0.1% Tween 20 (PBST). The membrane was then incubated with a 1% dry milk in PBS solution containing ~ 0.13ug/mL of secondary antibody shaking at room temperature (25 °C) for one hour. A goat-anti rabbit antibody conjugated with horseradish peroxidase (Thermo Fisher Scientific, USA) was used as the secondary antibody. After this incubation, the membrane was washed three times for five minutes each time in 1x PBST.
SuperSignal West Pico Chemiluminescent Substrate (ThermoFisher Scientific, USA) was then used to detect signal from secondary antibodies. Membranes were incubated with chemiluminescent substrate for one minute then moved into the dark room for development with X-ray film.

2.4.3 Protein quantification with ImageJ software

The free downloadable software ImageJ was used to quantify band intensity in Western Blot experiments (Schneider et al., 2012). X-Ray films were scanned and uploaded before being opened in ImageJ. The software was then used to calculate the area of each band. Using total protein values in each sample, band area over total protein was able to be quantified. Total protein measurements were obtained using the NanoDrop ND1000 Spectrophotometer. These numbers were then used to compare amounts of FLAG-tagged protein between strains of *A. tumefaciens*.

2.5 AGROBEST Experiments

AGROBEST experiments have been previously described by Wu et al. (2014). The group used GUS (β-glucuronidase) gene expression to compare rates of infection by *A. tumefaciens* in *Arabidopsis thaliana*. Graciously, the group assisted by sharing their methods for these experiments. The following methods are heavily credited to Dr. Erh-Min Lai’s Laboratory at the Institute of Plant and Microbial Biology in Taiwan.
2.5.1 Plant material

*Arabidopsis thaliana* seedlings were used in AGROBEST experiments. Wild-type Col-0 seeds were donated from Dr. Yanhai Yin’s lab at Iowa State University. Only Col-0 plants were used for the AGROBEST assays described here.

2.5.2 Plant growth conditions

To grow plants for AGROBEST experiments, seeds were first sterilized for 10 minutes in a 50% bleach, and 0.05% Tween-20 solution by inverting in a 1.5 mL microcentrifuge tube. Seeds were then rinsed five times with sterile water and transferred to 1 mL of ½ MS media (2.17 g/L Murashige and Skoog basal medium pH adjusted to 5.7 with KOH) with 0.5% sucrose in a 6 mm petri dish sealed with parafilm and incubated at 4 °C for two days before they were transferred to the growth chamber. Plants were grown at 24 °C with a 16-hour day length. After four days of growth, seedlings were infected with the desired *A. tumefaciens* strains.

For infection, a sterile inoculation loop was used to transfer seedlings to a new 6 mm petri dish with approximately 20 seedlings per dish. Then 1 mL of bacterial culture with an OD$_{600}$ of 0.02 in ½ induction, ¼ MS media was added. Acetosyringone (AS) was also added to a final concentration of 0.01 mM to induce bacterial virulence. Finally, dishes were sealed with parafilm.

After infection, seedlings continued to grow under the conditions described above for two additional days. On the third day medium was removed from seedling dishes and replaced with 1 mL of new ½ MS media containing 0.5% sucrose and timentin to a final
concentration of 1 mM. Seedlings were then allowed to grow for one additional day before analysis. Methods for plant growth were adapted from Wu et al. (2014).

2.5.3 Bacterial growth conditions

AGROBEST experiments were used to examine virulence of the following bacterial strains C58(pB35S GUS), C58 ΔPavc2R1(pB35S GUS), C58ΔvirC2(pB35S GUS), C58ΔvirG(pB35SGUS) and C58. Each bacterial strain was transformed with the plasmid pB35S-GUS, which expresses the GUS gene driven by the cauliflower mosaic virus 35S (35S) promoter. Because the plasmid also contained right and left border sequences flanking the GUS gene, this fragment of DNA was transferred in a transient fashion to the plant where GUS expression could then be measured.

Bacteria were stored in glycerol stocks at -80 °C and streaked with a sterile inoculation loop on fresh YEP plates containing appropriate antibiotics, which were then sealed with parafilm. Bacteria were grown on plates for three days at 28 °C. Cells from a fresh plate were transferred with a sterile inoculation loop to 5 mL of 523 liquid medium (8 g/L Casein Hydrolysate, 0.0358 g/L MgSO₄·7H₂O, 2 g/L K₂HPO₄, 10 g/L sucrose, and 4 g/L Yeast Extract, pH adjusted to 7.0 with KOH) in a sterile 50 mL falcon tube with appropriate antibiotics. Bacteria were then grown in a 28 °C shaker at 250 rpm overnight, for approximately 20 hours. Next, bacterial cultures were centrifuged at 4000 x g for 10 minutes at room temperature (25 °C). The Supernatant was removed and the bacteria were resuspended in induction medium, and 5 mL of this induction culture was transferred to a new 50 mL falcon tube along with 0.1 mM AS to an OD₆₀₀ of 0.2 This
culture was then grown for 16-20 hours shaking at 28 °C at 250 rpm after which time it could be used for infection.

Directly before the infection process bacteria were again centrifuged as described above. The supernatant was removed and bacteria were then diluted in fresh ½ induction, ¼ MS media to an OD600 of 0.02 at which stage they could be added to Arabidopsis thaliana seedlings. The details of this protocol were adapted from Wu et al. (2014). In this study, AGROBEST was used to examine the expression of GUS, as well as other genes in A. thaliana seedlings.

2.5.4 GUS quantification protocol

The GUS substrate 4-methylumbelliferyl β-D-glucuronide (MUG) was used to quantify the amount of GUS expressed in Arabidopsis thaliana seedlings. When GUS acts as an enzyme on the MUG substrate, MUG is broken into two separate molecules. One of these molecules, 4-methylumbelliferone (4MU) is fluorescent. The level at which fluorescence from 4-MU increases over time corresponds to the level of GUS in the sample.

Ten seedlings were chosen at random from one petri dish to be used as a sample in order to obtain a sample representative of all the plants in the dish. Plants were removed with a sterile inoculation loop and blotted dry on sterile paper towel before being placed in a 1.5 mL microcentrifuge tube and immediately frozen in liquid nitrogen. Once all samples had been collected, frozen plant material was ground into powder with a sterile plastic drill bit followed by the immediate addition of 400 μL extraction buffer (50 mM phosphate buffer pH 7.0, 10 mM EDTA, 0.1% SDS, 0.1% Triton X-100, 10 mM
β-mercaptoethanol, and protease inhibitor tablet SIGMAFast (Sigma Aldrich, USA)). Samples continued to be ground for one to two minutes in extraction buffer and placed on ice till all samples were processed. Samples were then centrifuged at 13000 x rpm for 15 minutes at 4 °C. The resulting supernatant was then transferred to a new 1.5 mL microcentrifuge tube, placed on ice and was analyzed for GUS activity using the MUG substrate.

To analyze for GUS using MUG, 20 μL of sample and 180 μL of 10 mM MUG in extraction buffer were combined in each well of a black 96 well plate with a clear bottom. Three technical replicates were performed per sample. Fluorescent activity was then measured for one hour using a plate reader. A 356 nm excitation and 455 nm emission filter was used with a sensitivity setting of 75. The slope of increasing fluorescence could then be calculated and compared between samples.

In order to standardize experiments, a standard curve was developed using concentrations of 4-MU at 5 mM, 2.5 mM, 1.25 mM, and 0.625 mM of 4-MU. This allowed the GUS activity to be related from one experiment to the next. The slope of each sample could then be reported in terms of pMol 4-MU per minute.

Total protein present in a sample was calculated using Bradford analysis with 50 μL of each sample added to 950 μL of Bradford solution in a 2 mL cuvette. The OD of samples was measured with a spectrophotometer set at 595 nm using plain Bradford solution as a blank. A standard curve was constructed using ten concentrations of bovine serum albumin (BSA) ranging from 0.01 to 1 mg/mL. GUS activity could then be reported as pMol 4-MU/minute/mg protein.
2.5.5 GUS staining protocol

In order to visualize GUS expression within *Arabidopsis thaliana* tissues, 5-bromo-4-chlorohexyl-ammonium salt (X-Gluc) was used as a substrate to stain *Arabidopsis thaliana* seedlings. At three days post infection, seedlings not needed for GUS quantification with MUG were incubated in X-Gluc solution (50 mM NaHPO₄ pH = 7, 5 mM KFerricyanide, 5 mM KFerrocyanide, 0.1% Triton X100, 1% DMSO, 1.5 mM X-Gluc). Seedlings to be analyzed for GUS expression were placed in 1.5 mL microcentrifuge tubes with a sterile inoculation loop and kept on ice until all samples were collected. After collection, samples were taken off ice and a 90% acetone, 10% water solution was added. The tubes were incubated at room temperature (25 °C) for 20 min in order to fix the tissues. Acetone solution was then removed using a pipet and samples were rinsed once with sterile water. Next, X-Gluc solution was added until all tissues were covered (approximately 750 μL per sample). Tubes were then incubated at 37 °C for four hours in the dark.

After the incubation period the X-Gluc solution was removed and replaced with a 70% ethanol, 30% H₂O solution to remove chlorophyll from the tissues. The following day, samples were photographed. Samples were stored for future reference.

2.6 Tumorigenesis Assays using *Kalanchoe daigremontiana*

2.6.1 Plant material

*Kalanchoe daigremontiana* were used for tumor inoculation. Plants were either purchased from an online vendor or were donated from Dr. Stan Gelvin’s group at Purdue University. Plants were grown at room temperature (25 °C) with light for either
16 or 12 hours a day. Plants were watered roughly once a week and were checked frequently for dryness. In order to propagate new plants, small plantlets were taken from older plants that had never been inoculated and were germinated in soil in the same growth chamber.

2.6.2 Bacterial growth conditions

Bacterial strains were stored in glycerol stock and streaked with a sterile inoculation loop on to fresh YEP plates with appropriate antibiotics and grown for two days at 28 °C. In order to grow bacteria in liquid culture, cells were taken from a fresh plate with a sterile inoculation loop and added to 5 mL YEP media in a sterile 50 mL falcon tube along with appropriate antibiotics. Bacteria were then grown overnight, for approximately 20 hours, in a 28 °C shaker at 250 rpm.

For induction, bacteria were transferred to 5 mL fresh AB media from YEP cultures at a dilution of 1:100 and grown overnight, for approximately 20 hours, at 28°C shaking at 250 rpm. Bacterial cultures were then centrifuged at 4000 x g for 10 minutes at room temperature (25 °C). The supernatant was removed and bacteria were resuspended in two volumes of induction media with 0.1mM of AS. Bacteria were then grown at 28 °C for 20 hours before use for plant inoculations.

Regardless of the growth conditions, directly before inoculation, bacteria were centrifuged at 4000 x g for 10 minutes at room temperature (25 °C). The supernatant was removed and the pellet resuspended in 0.9% NaCl solution. Bacteria were centrifuged a second time as described above. The supernatant was discarded and the pellet was
resuspended in 0.9% NaCl solution to an OD$_{600}$ of 1.0. This solution was then used for inoculation of *K. daigremontiana*.

### 2.6.3 *Kalanchoe daigremontiana* inoculation procedures

For inoculation, *Kalanchoe daigremontiana* plants were selected that were approximately six weeks of age. By inoculating plants of various ages, we observed that younger plants consistently had a higher response to inoculation by *A. tumefaciens* while older leaves consistently showed no tumors post inoculation.

At the time of inoculation, a sterile needle was used to create four linear wounds on each side of the leaf’s midrib for a total of eight wounds per leaf. After wounding 5 μL of bacterial solution prepared in 0.9% NaCl solution (as described above) was then applied to each wound site with a sterile pipet tip. In all experiments wild type was inoculated on one side of the midrib while a mutant strain was inoculated on the opposite side in order to generate a direct comparison.

After inoculation, plants were placed back into the growth chamber and continued to grow under normal conditions until tumors were formed (~3 weeks). After the formation of tumors, leaves were photographed and tumors were scored from 0-3 with 0 being no tumor and 3 being a very large tumor.

### 2.7 Statistical Analysis

In order to determine the significance of experimental results, a variety of statistical tests were used. Due to the complicated nature of some sets of data, additional statistical assistance was received from Yinan Fang from the Statistics Department at Iowa State University. Specifically, data from AGROBEST experiments and *K.*
*daigremontiana* tumorigenicity assays were analyzed by Yinan Fang. Data from qRT-PCR experiments and Western blot experiments were analyzed by myself with additional assistance from Dr. Keunsub Lee as needed.

### 2.7.1 Statistical analysis of qRT-PCR data

qRT-PCR data were collected and used to measure transcript levels which were then able to be compared between different strains of *A. tumefaciens*. Values for each transcript were first reported in Ct (cycle threshold) then changed to relative abundance when normalized to housekeeping gene *rpoD*. These values were also able to be normalized to C58 wild-type values. Experiments were performed with biological replicates which were then compared using a two-tailed Student’s paired-sample t-test. In this model the dependent variable (transcript level) is compared between the two bacterial strains. Significance was declared at $P < 0.05$.

### 2.7.2 Statistical analysis of Western blot images

In order to detect changes in amounts of FLAG-tagged VirC2 protein between the two *A. tumefaciens* strains, C58::VirC2FLAG and C58 ΔPavc2::VirC2FLAG, equal volumes of protein samples from each strain were run side-by-side in Western blot. ImageJ (Schneider et al., 2012) was then used to quantify the band area. The amount of total protein in each sample was used to standardize each sample. Final values reported in band size over mg/mL total protein were compared. Three separate paired biological replicates were performed for each strain and compared using a paired-sample two-tailed
Student’s $t$-test. In this model the dependent variable (protein) was compared between the two bacterial strains. Significance was declared at $P < 0.05$.

2.7.3 Statistical analysis of *Kalanchoe daigremontiana* tumorigenesis assays

Statistical analysis was performed on results from *K. daigremontiana* tumorigenesis experiments using a generalized linear mixed model for ordinal variables (GLIMMIX). This model is appropriate when working with a multinominal distribution, in this case, we are working with the tumor score from 0 to 3. Using this model, an F-test was used to test the fixed effects. In this case, fixed effects were the particular bacterial strain used. A significant result meant the bacterial strain used caused a difference in tumor result. A value of less than 0.05 was considered significant. It is important to note that for the data set comparing tumors produced from C58 and C58 ΔvirC2 under nutrient rich conditions, the model used did not fit the data no tumors at all were produced by C58 ΔvirC2 (the data did not converge). However, results were still considered to be statistically significant since the difference could be clearly seen and was greater than other samples found to be statistically significant. In addition, a significant likelihood ratio test was used to compare the distributions with and without factoring in bacterial effect. Because this test was significant ($P < 0.001$) it can be said that the effect of the bacterial strains were indeed significant.

2.7.4 Statistical Analysis of AGROBEST experiments

Results from AGROBEST experiments were analyzed using a linear mixed model. In this case, the response variable was the GUS expression while the specific
bacterial strain used was the fixed effect. The difference of least squares mean was then compared between treatments. This is a common way to examine the difference in means between several groups. A Bonferroni adjustment was also used in order to control for possible type I error (less chance of a false positive) and significance was declared at $P < 0.05$. 
2.8 References


Figure 3. General strategy for homologous recombination used to alter genomic sequences of *A. tumefaciens*. Homologous recombination was used to generate knockout mutant strains C58 ΔPavc2 and C58 ΔvirC2. A homologous recombination strategy was also used to generate FLAG-tagged strains including C58::VirC2-FLAG (based on Reid & Collmer, 1987).
Figure 4. Cloning strategy used to generate pAJ007. Vector pAJ007 was generated for the purpose of deleting the promoter region of anti-virC2. Upstream and downstream sequences flanking the region to be deleted were cloned into vector pTFsacB to generate pAJ007.
Figure 5. Cloning strategy used to generate pAJ021. Vector pAJ021 was generated for the purpose of deleting the entire virC2 gene. Upstream and downstream stream sequences flanking the region to be deleted were cloned into vector pTFsacB to generate pAJ021.
Figure 6. Cloning strategy used to generate pKL1007. 3xFLAG_virC2, a 1229 bp fragment designed to generate A. tumefaciens strains with N-terminally tagged VirC2 protein was ordered from IDT.
3.1 Selection of a Regulatory RNA in *A. tumefaciens* for Further Investigation

### 3.1.1 Many non-coding RNA’s are present in the *A. tumefaciens* genome

In order to identify regulatory RNAs present in *A. tumefaciens*, an assistant research scientist in the lab, Dr. Keunsub Lee, performed large-scale RNA sequencing experiments previously described (Lee et al., 2013). By sequencing RNA samples from *A. tumefaciens* grown under a variety of conditions, Lee et al. (2013) was able identify 385 “novel non-coding RNA’s”. A similar study was performed by Wilms et al. (2012), identifying over 200 potential regulatory RNA transcripts in *A. tumefaciens*. While it is not certain which of these identified transcripts act to regulate expression of other genes, these transcripts form a pool of RNA transcripts worth further investigation. Post-identification, individual non-coding RNAs were then able to be selected and further explored for their specific roles and functionality.

### 3.1.2 Selection of *anti-virC2* transcript for further study

The abundance of candidate regulatory RNAs expressed in *A. tumefaciens* resulted in difficulty when attempting to select a RNA to investigate further. In order to narrow the field, only regulatory RNA’s located in the *vir* gene region of the Ti plasmid were examined and further explored. It was decided to focus on the *vir* region because of the *vir* genes’ essential role in plant transformation. We selected a candidate regulatory
RNA encoded in this region, to increase the probability that the candidate RNA transcript would play a role in pathogenic virulence and plant transformation.

By examining RNA sequencing data previously generated by Lee et al. (2013), a large difference between \textit{virC1} and \textit{virC2} gene expression was observed (Figure 7). As shown by these sequencing data, when bacteria were grown under AS induction conditions, \textit{virC1} had much higher transcript expression levels than \textit{virC2}. This observation was also reported by Wilms et al. (2012), observing \textit{virC1}’s greater increase in upregulation than that of \textit{virC2} under induction conditions.

Because \textit{virC1} and \textit{virC2} are part of the same operon, it makes sense that transcription would produce a polycistronic mRNA with similar transcript abundances for both \textit{virC1} and \textit{virC2}. Why then, is there a large difference in expression between the two genes which make up the \textit{virC} operon? One possible explanation for the lower level of \textit{virC2} transcript compared to those of \textit{virC1} could be the common phenomenon observed through gene strand-specific tiling arrays and transcriptome sequencing in bacterial operons in which each subsequent gene in the operon is transcribed at a slightly lower rate than the preceding gene (Güell et al., 2009). Or perhaps the two transcripts have differing rates of stability with \textit{virC2} quickly degrading. Another possible explanation could involve regulation by regulatory RNA. Perhaps \textit{anti-virC2} is acting to negatively regulate the \textit{virC2} transcript.

Upon examination of previously identified non-coding RNA candidates, an RNA transcript was observed to be expressed from the antisense strand of the \textit{virC2} gene (Lee et al., 2013). Due to this arrangement, the \textit{anti-virC2} transcript would have perfect complementarity with the \textit{virC2} mRNA. Therefore, we hypothesized that the \textit{anti-virC2}
transcript acts in *cis* to downregulate its potential target, *virC2*. If this hypothesis is correct, it would explain the lower *virC2* expression observed when compared to that of *virC1*. With this hypothesis in mind, we decided to further investigate the *anti-virC2* transcript and its potential role in regulation of *virC2* as well as other possible roles in pathogenic virulence.

This candidate regulatory RNA was also identified by Wilms et al. (2012) and named Ti4 though its function was not further explored. For this work this RNA transcript was called *anti-virC2* or *avc2* for short.

### 3.1.3 Expression of the *anti-virC2* transcript

From the RNA sequencing data obtained by Lee et al. (2013), the expression pattern of *avc2* had already been determined (Figure 7). We observed that *avc2* was expressed in *A. tumefaciens* under all conditions, although expression was slightly higher under conditions in which bacteria were induced with AS. An expression pattern consisting of three (possibly four) regions of high expression, separated by valleys of low expression, were also notable and can be observed in Figure 7.

3.2 Examination of *anti-virC2* Transcription Start and Stop Sites

The exact 5’ and 3’ sites of *avc2* were not clear. As seen in the transcript profile, three, possibly four, different, regions of high sequence reads were present on the *anti-virC2* gene (Figure. 7). In order to identify the 5’ and 3’ ends of the *avc2* transcript, as well as to determine whether the transcript was expressed as three separate RNAs or a single RNA, both 3’ and 5’ RACE (Rapid Amplification of cDNA ends) were carried out.
RACE reactions were primarily performed by Dr. Keunsub Lee with additional assistance as needed from myself. Sequencing results from these RACE experiments revealed the 3’ and 5’ sites for each of the three regions of the *avc2* transcript (Figure 8). From examining the resulting sequencing data, it appears that there is a clear 5’ end at nucleotide position 194400 for *anti-virC2* region 1. However multiple 5’ ends were found for regions 2 and 3. Multiple 3’ ends were also observed for each of the three regions (Figure 8).

There are several things that may be indicated by these RACE findings. First, and most importantly, the definite transcription 5’ end of *avc2* region 1 at nucleotide position 194400 indicates the beginning of the *anti-virC2* transcript. Data from these RACE experiments also confirm that *anti-virC2* transcript is indeed expressed.

Second, because 5’ ends for region 2 and 3, and 3’ ends for all three regions were unclear this indicates there may be multiple versions of the *avc2* transcript which are expressed. It is also possible that the *avc2* transcript may be expressed in three separate transcripts or one single transcript depending on environmental conditions. Differential expression of RNA transcripts has been previously observed depending on environmental conditions (Waters and Storz, 2009). Perhaps the different transcript lengths act to regulate biological processes within the cell depending on environmental conditions.

However, one thing is certain, there is gene expression coming from this region opposite *virC2* with one clear start site at nucleotide position 194400 on the Ti plasmid. Because the exact expression of the *avc2* transcript remains unclear, the entire region opposite *virC2* was examined together.
3.3 Generation of Deletion Mutants C58 ΔPavc2 and C58 ΔvirC2

Once it became clear that the avc2 transcript was expressed opposite the virC2 sequence through RACE experiments and sequencing data, the next logical step in elucidating avc2’s function was to alter its expression. Once levels of avc2 were changed, other transcripts could be analyzed for resultant changes in expression.

Because the avc2 transcript is expressed directly opposite the virC2 coding gene, a total knockout mutant was impossible to generate without also removing the virC2 gene. To overcome this obstacle, the promoter region of avc2 was identified and deleted.

3.3.1 Generation of anti-virC2 promoter deletion mutant ΔPavc2

The promoter region of avc2 was identified using Softberry BPROM; Bacterial promoter prediction software (Solovyev and Salamov, 2011). After input of the avc2 region sequence both a -35 box and -10 box region were identified (Figure 9). A 46 base pair region of the A. tumefaciens genome was then selected for deletion. The region included the identified avc2 promoter without including any of the virC2 coding sequence, or any other virulence gene coding sequence (Figure 9).

In order to delete the avc2 promoter region a strategy of homologous recombination was implemented using vector pTFsacB (Figure 3). The resulting avc2 promoter deletion mutant was named C58 ΔPavc2. Additional details concerning the construction of C58 ΔPavc2 can be found in Chapter II Materials and Methods.
3.3.2 Generation of virC2 deletion mutant ΔvirC2

In addition to the avc2 promoter deletion mutant C58 ΔPavc2, a mutant strain of A. tumefaciens was generated in which the entire coding region of the virC2 gene was removed along with the promoter region of anti-virC2. This was accomplished using homologous recombination to delete a 679 bp region of the A. tumefaciens genome. The resulting bacterial strain was names C58 ΔvirC2. Additional details concerning the construction of C58 ΔvirC2 can be found in Chapter II Materials and Methods.

3.4 Analysis of Transcript Expression Level Changes in anti-virC2 Promoter Deletion Mutant C58 ΔPavc2

Upon deletion of the anti-virC2 promoter region, it was necessary to perform qRT-PCR in order to measure changes in the transcript expression levels of anti-virC2 as well as other virulence genes of interest. QRT-PCR experiments described here were performed primarily by Dr. Keunsub Lee with assistance as needed from myself.

3.4.1 Deletion of predicted anti-virC2 promoter leads to reduction of anti-virC2 transcript

Initial qRT-PCR experiments were performed in order to examine transcript expression level in the newly generated mutant C58 ΔPavc2. Both C58 wild type and C58 ΔPavc2 were grown up under induction conditions as described in Materials and Methods Chapter II. RNA was then extracted and cDNA synthesized using strand-specific primers for anti-virC2, virC2 and virC1 transcripts. cDNA for housekeeping
gene rpoD was also synthesized with gene specific primers and used as an internal standard in order to compare levels of transcript between bacterial strains.

Three experiments with a total of eight biological replicates of strand-specific qRT-PCR were performed to measure transcript expression level. Based on the results from these experiments, it was observed that deletion of the avc2 promoter region in mutant C58 ΔPavc2 did indeed cause a reduction in of avc2 transcript expression level (Figure 10, Table 2). When compared to wild type, expression levels of avc2 transcript were reduced by 39% ± 6.5% in C58 ΔPavc2 (P < 0.001).

This reduction of anti-virC2 transcript was also confirmed through regular, non-strand specific qRT-PCR (Figure 11, Table 3). One experiment with three biological replicates was performed to amplify transcripts from virC2/anti-virC2, virG, and rpoD regions in both C58 wild type and C58 ΔPavc2 which had been grown under induction conditions as described in Chapter II Materials and Methods. In contrast to strand-specific qRT-PCR experiments, cDNA for these regular qRT-PCR was synthesized with random hexamers instead of gene specific primers. Because of this, transcripts located opposite each other at the same location in the genome were not able to be differentiated.

Most importantly, expression of anti-virC2 would be impossible to differentiate from that of virC2 since the two genes share the same location.

In order to examine expression levels of only anti-virC2 transcript, bacterial samples which had been grown without induction by AS were examined. As observed in previous sequencing data obtained by Lee et al. (2013), without induction the virC2 mRNA is barely expressed (Figure 7). By examining transcript expression levels from the
*anti-virC2/virC2* region under no induction conditions we are able to assume the vast majority of transcript is from *anti-virC2*.

When comparing relative expression of *anti-virC2* transcript in C58 wild type versus C58 ΔPavc2, an average reduction of 49% ±10.0% was observed (*P* = 0.037) (Figure 11).

Together, both strand-specific and regular qRT-PCR confirm that knockout of the *anti-virC2* promoter region did indeed lead to a decrease in the expression level of *anti-virC2* transcript present in C58 ΔPavc2. Once this reduction had been confirmed, the effect of lower *anti-virC2* abundance was able to be further examined in C58 ΔPavc2 and compared with C58 wild-type.

### 3.4.2 Reduction of *anti-virC2* transcript in C58 ΔPavc2 does not alter *virC2* or *virC1* mRNA transcripts

We had hypothesized that the *anti-virC2* transcript acts to down regulate the mRNA transcript of *virC2*, if true, it would be expected that expression of the *virC2* transcript would be higher in the C58 ΔPavc2 mutant, due to the reduction in *avc2* transcript. From the previously mentioned strand-specific qRT-PCR experiments, expression levels of the *virC2* transcript were examined and compared between the two strains. Instead of observing a reduction in *virC2* transcript expression level, there appeared to be no significant difference between *virC2* transcript in C58 and C58 ΔPavc2 (Figure 10, Table 2).
During strand-specific qRT-PCR experiments, abundance of *virC1* mRNA was also examined. When comparing abundance of *virC1* transcript between C58 and C58 ΔPavc2 there appears to be no significant change between the two strains.

Based on strand-specific qRT-PCR experiments, there does not seem to be any change in the expression level of *virC2* or *virC1* transcript expression level when there is a reduction in *avc2*. Together, these findings suggest the *avc2* transcript does not act in *cis* to regulate the *virC2* transcript encoded on the opposite strand. Nor does *avc2* act to regulate the other gene on the *virC* operon, *virC1*.

### 3.4.3 Additional transcript changes are observed in C58 ΔPavc2 when compared to C58 wild type

While performing regular qRT-PCR to compare *virC2/anti-virC2* transcript expression level in C58 and C58 ΔPavc2, *virG* and *virB1* transcript expression levels were also examined (Table 3). Because there was no alteration in *virG* coding sequence in *A. tumefaciens* mutant C58 ΔPavc2, there was not expected to be any change in *virG* or *virB1* transcript expression level. However, an increase in expression level of both *virG* and *virB1* transcripts were observed in C58 ΔPavc2. Abundance of *virG* transcript nearly doubled (*P* = 0.0386) while *virB1* transcript levels were elevated by ~25% in C58 ΔPavc2 when compared to wild type but not enough to be statistically significant.

There may be several reasons for this higher expression of *virG*. First the deletion of the *avc2* promoter region may have caused a change in the expression of the *virG* transcript. While the *virG* coding sequence is not directly affected, the deletion of a directly downstream region may have caused a change in *virG* transcript expression. It
may also be possible that the avc2 transcript regulates the virG transcript either directly or indirectly. While the reason for the elevated level of virG transcript is unknown, it is important to note, and may be the subject of future investigation.

3.5 Analysis of VirC2 Protein Expression in anti-virC2 Promoter Deletion Mutant

ΔPavc2

Changes in transcript expression level are important. However, it is necessary to check for changes in corresponding protein concentration as well. An observed change in mRNA expression level does not necessarily correspond to a change in the amount of protein. Examining protein concentration in the anti-virC2 promoter deletion mutant was especially important to confirm data from qRT-PCR experiments. Here, expression levels of VirC2 protein are measured and compared between A. tumefaciens strains C58 and C58 ΔPavc2.

3.5.1 Generation of FLAG-tagged VirC2 protein in C58 and C58 ΔPavc2

In order to measure concentrations of the VirC2 protein, an A. tumefaciens strain was generated containing an N-terminal 3x FLAG-tag on the native VirC2 protein. Additional mutant strains were generated which contained an N-terminal 3x FLAG-tag on the native VirC1 and VirG proteins respectively. However, VirC1 and VirG proteins failed to visualize during Western blot after numerous methods were used to extract protein. For the scope of the main thesis, only VirC2 FLAG-tagged protein concentrations will be examined. Additional information regarding A. tumefaciens strains containing VirG and VirC1 FLAG-tagged proteins can be found in Appendix B. Design
of VirC1 and VirC2 FLAG-tag mutants was carried out by Dr. Keunsub Lee, while design of VirG FLAG-tagged mutant was by myself. All cloning steps were performed by Dr. Keunsub Lee or myself.

These bacterial strains containing an N-terminal FLAG tag were named C58 ∆Pavc2::VirC2-FLAG and C58::VirC2-FLAG and were used for comparison of VirC2 protein concentrations. Reassuringly a similar N-terminally FLAG-tagged VirC2 protein was used with success by another group (Atmakuri et al., 2007).

### 3.5.2 VirC2 protein levels are unchanged in C58 ∆Pavc2 when compared to wild type

Western blot was performed to compare concentrations of VirC2 protein. After bands of VirC2-FLAG were developed on film, ImageJ (Rasband, 2012) was used to compare size of the bands. These results could then be standardized to the total amount of protein present in each sample using measurements taken with the NanoDrop ND1000 Spectrophotometer. Taking results from each of three experimental replicates, there does not appear to be a difference in VirC2-FLAG protein expressed in when comparing C58 and C58 ∆Pavc2 (Figure 12C). From these results it can be concluded that deletion of the avc2 promoter region, in C58 ∆Pavc2 does not lead to a decrease in concentration of VirC2 protein when compared to wild type.

This observation confirms the findings of qRT-PCR experiments showing no change in virC2 transcript expression levels between C58 and C58 ΔPavc2. Because amounts of both VirC2 protein and transcript were observed to remain unchanged in C58
ΔPavc2, it is likely that the avc2 transcript plays no role in regulation of the virC2 transcript as previously hypothesized.

3.6 Examination of Virulence in A. tumefaciens Mutant Strains C58 ΔPavc2 and C58 ΔvirC2

Two different methods were used to screen for changes in virulence within the knockout mutant strains C58 ΔPavc2 and C58 ΔvirC2 compared to wild type strain C58. First, the plant species Kalanchoe daigremontiana was used for inoculation experiments in which resultant tumors were then compared. K. daigremontiana inoculations have been commonly used as a way to compare bacterial virulence between strains of A. tumefaciens (Garfinkel and Nester, 1990). Second, AGROBEST experiments were used to quantitatively measure the ability of bacteria to infect and transiently express transferred DNA in Arabidopsis thaliana. AGROBEST experiments are not unique to this thesis and were previously used to transiently express various genes in A. thaliana (Wu et al., 2014).

3.6.1 Measurement of virulence in knockout mutants C58 ΔPavc2 and C58 ΔvirC2 by Kalanchoe daigremontiana tumorigenesis Assays.

As described previously in Chapter I, Introduction and Literature Review, Kalanchoe daigremontiana inoculations are a commonly used method to detect changes in virulence (Garfinkel and Nester, 1990). In fact, inoculation of A. tumefaciens onto K. daigremontiana has specifically been used to measure changes in virulence due to mutations in the virC operon. A. tumefaciens strains containing mutations leading to loss
of function of either VirC1, VirC2, or both proteins showed a severely attenuated phenotype. When compared to wild type, mutant strains exhibited a phenotype consisting of much smaller tumors on *K. daigremontiana* leaves (Yanofsky and Nester, 1986; Klee et al., 1983; Lu et al., 2009).

In this work, *K. daigremontiana* inoculations were used to in order to examine changes in virulence in both C58 ΔPavc2 and C58Δ virC2 mutant strains. First, bacteria to be used for inoculation were grown under either induction or nutrient rich conditions as described in chapter II Materials and Methods. Tumors produced by each strain were scored on a scale of 0 to 3 with a very large tumor scoring a 3 and no tumor scoring a 0 and compared (Figure 13).

Several plant inoculation experiments were performed in order to measure virulence. Eight experiments, one using bacteria grown in induction conditions and seven in nutrient rich conditions, as described in Chapter II Materials and Methods, with a total of 244 inoculations were performed to compare virulence in C58 and C58 ΔPavc2. Three experiments were completed to compare virulence between wild type C58 and C58 ΔvirC2, one experiment from bacteria grown under induction conditions and two from bacteria grown under nutrient rich conditions, as described in Chapter II Materials and Methods, with a total of 64 inoculations. In order to determine whether tumors produced from the inoculation of different bacterial strains were significantly different, a generalized linear mixed model (GLIMMIX) analysis specific for ordinal variables was used as a model for data analysis.
3.6.1.1 *K. daigremontiana* inoculations show *virC2* knockout mutant C58 ΔvirC2 has severely attenuated phenotype

First, let us examine inoculation experiments comparing *virC2* deletion mutant C58 ΔvirC2 with wild-type strain C58. From all experiments performed with the C58 ΔvirC2 strain it was quite clear that virulence in the mutant strain was severely attenuated (Figure 14, Table 4). In fact, when bacterial cultures were grown under nutrient rich conditions, plants inoculated with C58 ΔvirC2 failed to produce any tumors at all. The two strain produced significantly differently sized tumors under both induction and nutrient rich growth conditions (\( P < 0.001 \)).

These results show that removal of the *virC2* coding sequence causes severe attenuation of pathogenic virulence when compared with wild type. These findings also agree with previous studies, all of which observed attenuation upon *virC2* loss of function (Klee et al., 1983; Yanofsky and Nester, 1986; Lu et al., 2009).

3.6.1.2 Changes in virulence observed in C58 ΔPavc2 in *K. daigremontiana* inoculations

*K. daigremontiana* inoculations were also used to compare virulence between wild type C58 and C58 ΔPavc2. Though it was found that deletion of the *anti-virC2* promoter and the resultant reduction in *avc2* transcript expression level did not cause an increase in *virC2* transcript expression or protein concentration, virulence levels in the mutant strain C58 ΔPavc2 were still examined. It may be possible that while the transcript has no effect on levels of VirC2, *anti-virC2* plays some other yet unknown role in *A. tumefaciens* virulence. By looking for changes in virulence when the transcript
expression levels of *anti-virC2* were reduced, we may find evidence of *anti-virC2*’s possible role in pathogenic virulence.

Surprisingly, under both induction and nutrient rich growth conditions, C58 ∆Pavc2 was observed to produce tumors that significantly differed in size from C58 wild type (*P* < 0.001) (Figure 15, Table 4). When bacterial cultures were grown under induction conditions, only 35.9% of C58 wild type inoculation sites did not produce any tumors, while 62.5% of inoculation sites from C58 ∆Pavc2 were observed not to produce any tumor. A higher number of observations with no tumors present were also observed for C58 ∆Pavc2 compared to wild type C58 when bacteria were grown under nutrient rich condition (47.2% for wild type C58 and 58.3 percent for C58 ∆Pavc2). In addition, tumors formed from C58 wild type were more likely to have a higher tumor score than those produced by C58 ∆Pavc2. This result suggests a difference in virulence between the C58 ∆Pavc2 strain compared to C58 wild type.

One hypothesis for this decrease in virulence in the C58 ∆Pavc2 mutant is negative regulation of an alternative virulence gene. Perhaps the *avc2* transcript is acting to stabilize an alternate transcript involved with virulence as previously hypothesized for *virC2*. When expression level of the *avc2* transcript is reduced, abundance of the stabilized virulence gene transcript could also be reduced, leading to an overall reduction in virulence. Alternatively, deletion of the *anti-virC2* promoter region may have disrupted transcription of other genes on the Ti plasmid leading to the observed reduction in virulence.

Another factor to consider when examining results from *K. daigremontiana* inoculations is the nature of the experiment. Because tumors are only measured on a
qualitative scale based on observation by the eye of the researcher, there is room for error. It would be beneficial to conduct additional experiments in order to obtain a quantitative measurement of virulence.

3.6.2 Use of AGROBEST experiments to quantify changes in virulence within A. tumefaciens mutant strains C58 ΔPavc2 and C58 ΔvirC2

While K. daigremontiana tumorigenesis assays have commonly been used to measure pathogenic virulence of A. tumefaciens, this method does not provide quantitative results. In order to obtain a quantitative measurement of virulence, AGROBEST experiments were carried out. Using these experiments as described in a recent 2014 publication, A. tumefaciens virulence can be quantified through transfer, and resultant transient expression of the β-D-glucuronidase (GUS) gene in A. thaliana (Wu et al., 2014). Though in this particular study researchers described AGROBEST experiments as a way to transiently express genes in A. thaliana seedlings, it makes sense that these procedures can also be used to measure the virulence of the infecting bacteria. The only variable changed during these experiments was bacterial strain, so any observed changes in GUS expression in the A. thaliana seedlings can be attributed to the infecting strain of A. tumefaciens.

For AGROBEST experiments, the plasmid vector pB35S GUS was transformed into four strains of A. tumefaciens which could then be used to infect A. thaliana. Because pB35S GUS contains the GUS gene flanked by right and left border sequences, the GUS gene was able to be transferred to the plant during infection where it would then be transiently expressed. Strains of A. tumefaciens transformed with pB35S GUS were
C58, C58 ΔPavc2, C58 ΔvirC2, and C58 ΔvirG (a previously generated virG knockout mutant). These strains were then named C58(pB35S GUS), C58 ΔPavc2(pB35S GUS), C58 ΔvirC2(pB35S GUS), and C58 ΔvirG(pB35S GUS). C58 wild-type was also used as a negative control. These bacterial strains were then used to infect A. thaliana seedlings and GUS expression analyzed by measuring fluorescence in a plate reader using MUG or staining with X-GLUC. C58(pB35S GUS) was used as a standard (Figure 16).

3.6.2.1 virC2 deletion mutant ΔvirC2 shows substantial reduction in virulence compared to wild type using AGROBEST protocol

Virulence among the two negative control strains, C58 and C58 ΔvirG(pB35S GUS), was significantly different than that of wild type C58(pB35S GUS) ($P < 0.05$). These results were to be expected, a wild type strain of A. tumefaciens without a GUS gene should not cause GUS expression in plants. Likewise, a bacterial strain without the essential virulence gene activator virG would not be able to infect plants and therefore no GUS expression would result. Interestingly, it was also observed that C58 ΔvirC2(pB35S GUS) had a significantly lowered virulence when compared to wild type ($P = 0.0072$). This finding is consistent with K. daigremontiana inoculation data where C58 ΔvirC2 was consistently shown to produce absent or very small tumors upon inoculation. It is clear that the virC2 gene is essential for full virulence (Figure 16).
3.6.2.2 *anti-virC2* promoter deletion mutant ΔPavc2 shows no significant change in virulence compared to wild type using AGROBEST protocol

When examining GUS expression in plants infected with strain C58(pB35S GUS) and C58 ΔPavc2(pB35S GUS) no statistically significant difference was observed. While the average expression of GUS in plants infected with C58 ΔPavc2(pB35S GUS) was slightly higher when compared to wild type, this was not the case for every experiment. In fact, large deviations were observed in results from different experiments. When examining data from AGROBEST experiments alone, it is not clear that C58 ΔPavc2 is any more or less virulent than C58 wild type.

Interestingly these results do not reflect the lower amount of virulence observed in C58 ΔPavc2 compared to C58 wild type which was observed from *Kalanchoe daigremontiana* inoculations. There are several hypotheses which may explain these findings. First, AGROBEST may be better at showing large reductions in virulence rather than smaller decreases. The, reduction of virulence in C58 ΔvirC2 was more severe than the reduction of virulence observed in C58 ΔPavc2 as seen in *K. daigremontiana* inoculations. Perhaps AGROBEST was not sensitive enough to catch this difference. In addition, the variability in GUS expression among transformed cells within each individual plantlet may have been greater than the variability due to bacterial strain used for infection when virulence levels were close. Photographs of stained plantlets (Figure 16) showed a large amount of variation in GUS expression within a plant.

It is also important to consider that *A. tumefaciens* to infection of *A. thaliana* plants may have less severe disease levels than on *K. daigremontiana*. *virC1* and *virC2* loss of function mutants had attenuated virulence phenotypes to varying degrees
depending on the host plant which the bacteria was tested (Close et al., 1987). It is possible that C58 ΔPavc2 produces a reduced virulence phenotype in \textit{K. daigremontiana} but not in \textit{A. thaliana}. This difference in tumor phenotype depending on the plant infected may explain the difference in results between AGROBEST experiments and \textit{K. daigremontiana} tumorigenesis assays.

Additionally, it could also be possible that AGROBEST experiments measured a different aspect of virulence as opposed to \textit{K. daigremontiana} inoculation experiments. AGROBEST experiments measured transient expression of transferred DNA rather than stable integration of transferred genes into the genome. It is possible that C58 ΔPavc2 was able to transiently express transferred DNA but unable to stably integrate transferred DNA into the plant genome.
3.7 References


Figure 7. Expression of *virC, anti-virC2*. Depth of coverage on nucleotide positions 194307-195820 on the Ti plasmid. Forward and reverse strands of the *virC* operon region are shown. Regions 1 through 3 of *anti-virC2* are indicated as R1, R2 and R3.
Figure 8. Illustration of 3’ and 5’ sites for *anti-virC2* regions 1, 2 and 3. 3’ and 5’ sites identified by RACE (Rapid Amplification of cDNA Ends) are shown for each of the three regions of *anti-virC2*. 
Figure 9. Illustration of deleted anti-vir2 promoter in A. tumefaciens strain C58 ΔPavc2. The promoter region of anti-virC2 was identified using Softberry Bprom bacterial promoter prediction software. Both a -35 box and a -10 box region were identified. Using homologous recombination, knockout mutant C58 ΔPavc2 was generated with the illustrated 46 bp sequence deleted.
Figure 10. Strand-specific qRT-PCR shows transcript abundance in C58 and C58 ΔPavc2. Strand-specific qRT-PCR was performed to examine transcript abundance in promoter deletion mutant C58 ΔPavc2 compared to C58 wild type. Three experiments were performed with a total of eight separate biological repeats. Significant difference from wild type ($P < 0.01$) is denoted with **. Bars represent standard error.
Figure 11. qRT-PCR shows anti-virC2 transcript abundance in C58 and C58ΔPavc2. Results are shown for qRT-PCR experiments comparing relative abundance of anti-virC2 transcript in A. tumefaciens strain C58ΔPavc2 to that of wild type across three biological replicates. Housekeeping gene rpoD was used as an internal standard. Bars represent Standard Error.
Figure 12. Examination of FLAG-tagged VirC2 protein levels. Western blots comparing protein between C58::VirC2FLAG and C58 ΔPavc2::VirC2FLAG. A. A similar band was observed from FLAG-tagged VirC2 protein in both C58 wild type and C58::VirC2FLAG ΔPavc2::VirC2FLAG. B. Total protein from both bacterial strains was run on a gel for visualization. C. Band intensity was examined using ImageJ software. No statistical difference was observed in protein level based on the three biological replicates.
Figure 13. Scale for qualitative scoring of tumors from *Kalanchoe daigremontiana* tumorigenesis assays. Tumors resulting from *K. daigremontiana* inoculations were scored on a scale from 0 to 3 with 0 being no tumor and 3 being a large tumor.
Figure 14. Data from all Kalanchoe daigremontiana tumorigenesis assays comparing wild type C58 to C58 ΔvirC2. Tumors resulting from K. daigremontiana tumorigenesis assays were scored from 0 to 3 and data was recorded. Tumors produced from C58 wild type were compared to C58 ΔvirC2 when grown under either induction (A) or nutrient rich (B) conditions. A significant difference was detected between tumors generated between C58 wild type and c58 ΔvirC2 when grown under induction conditions. \( P < 0.001 \). Because no tumors were produced from C58 ΔvirC2 under nutrient rich conditions, data did not fit the generalized linear mixed model for ordinal variables to compare bacterial strains. For this data set a significant likelihood ratio test was carried out to compare the data sets with and without the effects of bacterial strain. Using this method, it was determined that bacterial strain was indeed significant \( P < 0.001 \).
Figure 15. Data from all *Kalanchoe daigremontiana* tumorigenesis assays comparing wild type C58 to C58 ΔPavc2. Tumors resulting from *K. daigremontiana* inoculations were scored from 0 to 3 and data was recorded. Tumors produced from C58 wild type were compared to C58 ΔPavc2 when grown under either induction (A) or nutrient rich conditions (B). A significant difference was detected between tumors generated by C58 wild type and C58 ΔPavc2 when grown under nutrient rich of induction conditions. \( P < 0.001 \)
Figure 16. Levels of GUS expression are shown in *A. thaliana* plants infected by various strains of *A. tumefaciens*. AGROBEST experiments were performed to compare virulence between various strains of *A. tumefaciens*. Four-day-old seedlings were infected with bacterial cultures and grown for an additional two days before media was changed and antibiotics added. Plants were harvested the following day and flash frozen before being ground and analyzed for GUS expression. Levels of GUS correspond to the transient expression of the GUS gene transferred by the bacteria. ** denotes the level of GUS is significantly different from that of wild type ($P < 0.05$). Bars represent standard error.
Table 1. Summary of 3’ and 5’ ends for anti-virC2 regions 1, 2 and 3 as observed using RACE.

<table>
<thead>
<tr>
<th>3’ RACE Results</th>
<th>5’ RACE Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcript</td>
<td>Transcript</td>
</tr>
<tr>
<td>194508</td>
<td>193340</td>
</tr>
<tr>
<td>194577</td>
<td>194400</td>
</tr>
<tr>
<td>194660</td>
<td>194400</td>
</tr>
<tr>
<td>194792</td>
<td>194400</td>
</tr>
<tr>
<td>anti-virC2 region 1</td>
<td>anti-virC2 region 1</td>
</tr>
<tr>
<td>194660</td>
<td>193813</td>
</tr>
<tr>
<td>194668</td>
<td>194000</td>
</tr>
<tr>
<td>194739</td>
<td>194402</td>
</tr>
<tr>
<td>194811</td>
<td>194439</td>
</tr>
<tr>
<td>194869</td>
<td>194439</td>
</tr>
<tr>
<td>194885</td>
<td>194455</td>
</tr>
<tr>
<td>194885</td>
<td>194513</td>
</tr>
<tr>
<td>194911</td>
<td>194535</td>
</tr>
<tr>
<td>194912</td>
<td>194542</td>
</tr>
<tr>
<td>194933</td>
<td>194565</td>
</tr>
<tr>
<td>194983</td>
<td>194565</td>
</tr>
<tr>
<td>194985</td>
<td>194588</td>
</tr>
<tr>
<td>195054</td>
<td>194614</td>
</tr>
<tr>
<td>195101</td>
<td>194615</td>
</tr>
<tr>
<td>195106</td>
<td>194615</td>
</tr>
<tr>
<td>anti-virC2 region 2</td>
<td>anti-virC2 region 2</td>
</tr>
<tr>
<td>195323</td>
<td>194692</td>
</tr>
<tr>
<td>195434</td>
<td>194692</td>
</tr>
<tr>
<td>anti-virC2 region 3</td>
<td>anti-virC2 region 3</td>
</tr>
<tr>
<td>194264</td>
<td>194817</td>
</tr>
<tr>
<td>194483</td>
<td>194817</td>
</tr>
<tr>
<td>194614</td>
<td>194817</td>
</tr>
<tr>
<td>194738</td>
<td>194817</td>
</tr>
<tr>
<td>194767</td>
<td>194817</td>
</tr>
<tr>
<td>194845</td>
<td>194817</td>
</tr>
<tr>
<td>194957</td>
<td>194817</td>
</tr>
<tr>
<td>194972</td>
<td>194817</td>
</tr>
</tbody>
</table>
Table 2. Strand Specific qRT-PCR was performed to examine transcript levels in promoter deletion mutant C58 ΔPvc2 compared to C58 wild type. Three experiments were performed with a total of eight separate biological repeats. Levels of anti-virC2 transcript were reduced by 39% ±6.5% in C58 ΔPvc2 ($P < 0.001$).

<table>
<thead>
<tr>
<th></th>
<th>anti-virC2 Relative Abundance</th>
<th>VirC2 Relative Abundance</th>
<th>VirC1 Relative Abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT</td>
<td>avc2</td>
<td>Ratio avc2/WT</td>
</tr>
<tr>
<td>Exp 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.89</td>
<td>1.47</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>2.62</td>
<td>1.58</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>2.39</td>
<td>1.36</td>
<td>0.57</td>
</tr>
<tr>
<td>Exp 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.16</td>
<td>3.16</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>5.39</td>
<td>3.66</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>6.54</td>
<td>4.06</td>
<td>0.62</td>
</tr>
<tr>
<td>Exp 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.45</td>
<td>5.28</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>5.94</td>
<td>2.62</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td></td>
<td>0.61</td>
</tr>
</tbody>
</table>
Table 3. Non-strand specific qRT-PCR was performed on cDNA from wild type C58, C58 ΔPavc2, and C58 ΔvirC2 in order to compare transcript levels. One experiment was performed with three biological replicates for wild type and C58 ΔPavc2 and two biological replicates for C58 ΔvirC2. Housekeeping gene rpoD was used as an internal standard.

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Primer</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rpoD</td>
<td>virG</td>
<td>virB1</td>
<td>Anti-</td>
</tr>
<tr>
<td>+AS</td>
<td></td>
<td></td>
<td></td>
<td>virC2/virC2</td>
</tr>
<tr>
<td>WT</td>
<td>1.00</td>
<td>6.18</td>
<td>36.97</td>
<td>3.19</td>
</tr>
<tr>
<td>ΔPavc2</td>
<td>1.00</td>
<td>12.36</td>
<td>46.42</td>
<td>4.58</td>
</tr>
<tr>
<td>ΔvirC2</td>
<td>1.00</td>
<td>5.43</td>
<td>36.50</td>
<td>No Ct</td>
</tr>
<tr>
<td>-AS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>1.00</td>
<td>0.98</td>
<td>0.00</td>
<td>0.68</td>
</tr>
<tr>
<td>ΔPavc2</td>
<td>1.00</td>
<td>1.47</td>
<td>0.00</td>
<td>0.41</td>
</tr>
<tr>
<td>ΔvirC2</td>
<td>1.00</td>
<td>1.65</td>
<td>0.00</td>
<td>No Ct</td>
</tr>
</tbody>
</table>
Table 4. Data from all *Kalanchoe daigremontiana* tumorigenicity assays. Tumors resulting from *K. daigremontiana* inoculations were scored from 0 to 3 and data was recorded. Tumors produced from C58 wild type were compared to other strains when grown under either nutrient rich or induction conditions. A significant difference was detected between tumors generated between C58 wild type and C58 ΔPavc2 when grown under nutrient rich or induction conditions. (*P* < 0.001). Tumors produced from wild type C58 were also statistically different from those produced from C58 ΔvirC2 (*P* < 0.001). Wild-type versus C58 ΔPavc2 is shown in Figure 15 while C58 ΔvirC2 versus wild-type is shown in figure 14.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Strain</th>
<th>Nutrient Rich</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor Score</td>
<td>Percent</td>
<td>C58 WT n=180</td>
<td>C58 ΔPavc2R1 n=180</td>
<td>C58 WT n=32</td>
<td>C58 ΔvirC2 n=32</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>47.2</td>
<td>17.2</td>
<td>11.7</td>
<td>23.9</td>
<td>58.3</td>
<td>16.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Strain</th>
<th>Induction</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor Score</td>
<td>Percent</td>
<td>C58 WT n=64</td>
<td>C58 ΔPavc2R1 n=64</td>
<td>C58 WT n=32</td>
<td>C58 ΔvirC2 n=32</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>35.9</td>
<td>10.9</td>
<td>23.4</td>
<td>29.7</td>
<td>62.5</td>
<td>20.3</td>
</tr>
</tbody>
</table>
CHAPTER IV
DISCUSSION AND CONCLUSIONS

The scope of this work was to select and further explore the action of a candidate regulatory RNA in *A. tumefaciens*. Because of *A. tumefaciens*’ important role in plant transformation, a potential regulatory RNA on the Ti plasmid that was opposite to a virulence gene was selected. Specifically, the non-coding RNA transcript opposite virulence gene *virC2* was chosen. Initially, we hypothesized that the *avc2* transcript acted to regulate the *virC2* transcript in cis because of their shared location in the *A. tumefaciens* genome. If this hypothesis was true, it could explain the large difference in *virC1* and *virC2* transcript expression level which had been previously observed (Lee et al., 2013). In order to further investigate this hypothesis, mutant strain C58 ΔPavc2 was generated with a deletion in the promoter sequence of *anti-virC2*. As a result of this deletion, expression levels of *avc2* transcript were reduced by approximately 39% ±6.5% in C58 ΔPavc2 as observed by strand-specific qRT-PCR. Mutant strain C58 Pavc2 was then compared with C58 wild type to examine the effects of reduced *avc2* transcript expression level.

The transcript level of several virulence genes including *virC2* were measured for comparison using qRT-PCR in both wild-type C58 and C58 ΔPavc2. In addition, the amount of VirC2 protein present in each strain was examined through Western blotting. In order to detect any additional roles, the *avc2* transcript may play in virulence, both *Kalanchoe daigremontiana* inoculation assays and AGROBEST experiments were carried out. The importance of the virulence gene *virC2* was also confirmed by examining the virulence of *virC2* knockout mutant C58 ΔvirC2.
Initially, before other experiments could be performed, the existence of the *avc2* transcript was confirmed using both 3’ and 5’ RACE. By examining the 3’ and 5’ ends from each of the three regions of *avc2* expression, it was concluded that a clear 5’ start site is present at nucleotide 194400. We also confirmed that the *avc2* transcript is indeed expressed. However, because of the many different 5’ and 3’ sites found for each region of the transcript, we could not determine if the transcript was expressed as one or multiple units. The expression pattern of the *avc2* transcript may even change based on environmental conditions, a phenomenon which has been previously observed in bacteria (Waters and Storz, 2009). In future experiments, the expression pattern of the *avc2* transcript could be further explored using Northern Blots.

After it was confirmed that the *avc2* transcript was expressed from the region opposite *virC2*, *anti-virC2*’s function could be further investigated. Because *anti-virC2* was located directly opposite *virC2*, it was hypothesized that *anti-virC2* may be regulating the *virC2* transcript in *cis*. This would make sense because of the complementarity which the two transcripts shared and may also explain the large difference in *virC1* and *virC2* transcript expression level. If *avc2* acted to bind the *virC2* transcript and target it for degradation, levels of *virC2* transcript would decrease in response to additional *avc2*, and increase as expression level of *avc2* were reduced.

In order to reduce the abundance of *avc2* within the cell and generate a mutant strain of *A. tumefaciens* which could be compared with wild type, the promoter region of *avc2* was deleted using homologous recombination. Through strand-specific qRT-PCR it was found that deletion of the promoter region led to a reduction in anti-virC2 expression level by approximately 39% ($P < 0.001$). This observation was also confirmed using
regular, non-strand specific qRT-PCR. When looking at transcript from the \textit{virC2/anti-virC2} region under conditions of growth without induction by AS, only \textit{avc2} transcript expression level would be observed since \textit{virC2} is not expressed without induction by AS. Regular qRT-PCR performed on samples grown without induction with AS showed a reduction in transcript from the \textit{anti-virC2} region by 49\% (\(P = 0.037\)).

The expression level of \textit{virC2} transcript was then compared using qRT-PCR. In strand specific experiments, it was observed that expression level of the \textit{virC2} transcript remained unchanged in C58 ∆Pavc2. The amount of \textit{virC1} transcript was also measured and compared between C58 wild type and C58 ∆Pavc2. No significant change in \textit{virC1} transcript was observed between the two strains. Ultimately, it was concluded based on qRT-PCR data that a reduction in \textit{anti-virC2} expression does not affect the abundance of either \textit{virC1} or \textit{virC2} transcript.

In addition to \textit{virC2} transcript expression level, VirC2 protein concentration was examined and compared between C58 wild type and C58 ∆Pavc2. Though expression level of the \textit{virC2} transcript showed no change between C58 and C58 ∆Pavc2, transcript expression level may not reflect the actual amount of protein present in each strain. In order to examine protein, three replicates of Western blot experiments were carried out to detect changes in FLAG-tagged VirC2 protein in both C58::C2-FLAG and C58 ∆Pavc2::C2 FLAG. From Western blotting, no change was detected between abundance of VirC2 protein in C58 wild type versus C58 ∆Pavc2 backgrounds. From both Western blot and qRT-PCR experiments it was concluded that a decrease in expression level of the \textit{avc2} transcript did not lead to an alteration in levels of \textit{virC2} transcript expression.
level or protein concentration. Based on this evidence, it is unlikely that avc2 plays a role in regulation of its predicted target, virC2.

Just because the avc2 transcript did not seem to play a role in the regulation of the virC2 gene does not mean it does not play a role within the A. tumefaciens cell. It is possible that the avc2 transcript regulates the transcript of another gene which may be involved with virulence. In order to detect any changes in virulence caused by the reduction of anti-virC2 in C58 ΔPavc2, virulence was compared to C58 wild type using both Kalanchoe daigremontiana inoculations and AGROBEST experiments. A virC2 deletion mutant, C58 ΔvirC2, was also generated and used in virulence assays.

When examining data from K. daigremontiana inoculation experiments a significant difference was observed in tumors produced from inoculations with C58 wild type versus C58 ΔPavc2 ($P < 0.001$). Under both nutrient rich and induction growth conditions, tumors from C58 wild type produced less tumors with a score of 0 and more large tumors with a score of three when compared to C58 ΔPavc2. Based only on K. daigremontiana inoculation results it appears that C58 ΔPavc2 exhibits some attenuated virulence when compared to wild type C58.

This observed decline in virulence is not evident when examining data from AGROBEST experiments. AGROBEST experiments were based on experiments performed by another group in 2014 (Wu et al., 2014). Briefly, strains of A. tumefaciens are transformed with plasmid pB35S GUS after which the bacteria is used to infect A. thaliana plants. The amounts of GUS expression in infected plants corresponds to the virulence of the particular A. tumefaciens strain. GUS production in plants infected by
C58(pB35S GUS) was not significantly different from those infected with C58 ΔPavc2 (pB35S GUS) over the course of four experiments.

There are several possible explanations for the difference in observed virulence. First, AGROBEST experiments and *K. daigremontiana* inoculations are measuring different variables. AGROBEST experiments measure the ability of *A. tumefaciens* to infect and transiently express a gene in *Arabidopsis thaliana* while *K. daigremontiana* inoculations measure the ability of *A. tumefaciens* to infect, stably express genes, and form tumors. It is possible that *avc2* plays a role in stable integration of genes but does not affect transient gene expression. It may also be possible that differences between the two plant species used for these experiments led to different results. It has been previously observed that virulence differs depending on the particular plant species used for inoculation (Close et al., 1987). In another possibility, the AGROBEST experiments may not have been sensitive enough to detect small changes in virulence while *K. daigremontiana* inoculations could detect even subtle differences.

When looking at changes in virulence between C58 ΔvirC2 and C58 wild type there is a more distinct difference. In both AGROBEST experiments and *K. daigremontiana* inoculations, a significant reduction in virulence was observed in C58 ΔvirC2 when compared to wild type when comparing *K. daigremontiana* inoculation results for both induction and non-induction conditions. AGROBEST experiments also showed a highly significant decrease in levels of GUS when comparing plants infected with C58 ΔvirC2(pB35S GUS) to those infected with C58(pB35S GUS) (*P* < .001). This finding demonstrates that both experimental methods are able to show a change in virulence when the difference in virulence between two *A. tumefaciens* strains is large.
The severe decrease in virulence observed in C58 ΔvirC2 also supports previous findings of the importance of virC2 to full virulence.

In future work, additional potential targets of anti-virC2 may be investigated. Using the RNA interaction prediction software IntaRNA several potential targets were identified (Wright et al., 2014). Searches using IntaRNA listed in Table 5 were performed using the entire sequence of the avc2 transcript. Additional searches were performed for each individual region and potential targets overlapping those obtained for the entire region were denoted with highlighted cells in Table 5. Not surprisingly, virC2 is on the top of the list as a potential target, however, based on observations seen in this work, it is not regulated by the avc2 transcript. Future qRT-PCR experiments comparing abundance of these transcripts between C58 and C58 ΔPavc2 may be used to explore the possible regulatory effects of anti-virC2 on these possible target transcripts. In addition, the possibility of a small transcript peptide by anti-virC2 may also be examined. Recently, several groups have found there are several open reading frames along the length of anti-virC2 that could be investigated.

This work showed evidence that avc2 does not influence the expression level of the virC2 transcript or protein concentration as previously hypothesized. However, that is not to say avc2 does not play an alternative role in pathogenic virulence. K. daigremontiana inoculations show a difference in virulence when the abundance of avc2 is reduced in C58 ΔPavc2 when compared to C58 wild type. This work also confirms previous findings of the importance of the virC2 gene in virulence through the methods of K. daigremontiana inoculations and AGROBEST experiments. In addition, work done in this thesis has expanded knowledge of A. tumefaciens and provides evidence for lack
of interaction between a specific non-coding RNA and its hypothesized target. While no specific function was found for anti-virC2, this work provides an example of the study of regulatory RNA and may be useful as a model for other studies.
References


Table 5. IntaRNA was used to predict possible mRNA targets of the entire anti-
virC2 transcript. (Wright et al., 2014) The top 15 targets are listed here. Targets which
were also predicted by searching with anti-virC2 region 2 sequence are highlighted in
blue while those also predicted when searching with anti-virC2 region 3 sequence are
highlighted in purple. One target was also predicted when searching with anti-virC2
regions 1 and 3 and is highlighted in red.

<table>
<thead>
<tr>
<th>p-value</th>
<th>fdr value</th>
<th>Target</th>
<th>Position</th>
<th>Query</th>
<th>Position</th>
<th>Energy</th>
<th>Gene</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>Atu6179</td>
<td>90 -- 148</td>
<td>acv2</td>
<td>583 -- 641</td>
<td>-89.193</td>
<td>virC2</td>
<td>putative crown gall tumor protein VirC2</td>
</tr>
<tr>
<td>0.001</td>
<td>0.963</td>
<td>Atu1319</td>
<td>27 -- 85</td>
<td>acv2</td>
<td>548 -- 604</td>
<td>-22.531</td>
<td>gatA</td>
<td>glutamyl-tRNA amidotransferase subunit A</td>
</tr>
<tr>
<td>0.001</td>
<td>0.963</td>
<td>Atu0050</td>
<td>78 -- 105</td>
<td>acv2</td>
<td>491 -- 517</td>
<td>-21.925</td>
<td>actR</td>
<td>two component response regulator</td>
</tr>
<tr>
<td>0.001</td>
<td>0.963</td>
<td>Atu2266</td>
<td>127 -- 149</td>
<td>acv2</td>
<td>537 -- 565</td>
<td>-21.775</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.001</td>
<td>0.963</td>
<td>Atu2569</td>
<td>102 -- 144</td>
<td>acv2</td>
<td>526 -- 572</td>
<td>-21.084</td>
<td>thiC</td>
<td>thiamine biosynthesis protein</td>
</tr>
<tr>
<td>0.002</td>
<td>0.963</td>
<td>Atu0532</td>
<td>81 -- 145</td>
<td>acv2</td>
<td>543 -- 607</td>
<td>-20.367</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>0.963</td>
<td>Atu4031</td>
<td>101 -- 148</td>
<td>acv2</td>
<td>533 -- 572</td>
<td>-20.101</td>
<td></td>
<td>sugar ABC transporter permease</td>
</tr>
<tr>
<td>0.002</td>
<td>0.963</td>
<td>Atu1979</td>
<td>11 -- 96</td>
<td>acv2</td>
<td>524 -- 611</td>
<td>-20.094</td>
<td></td>
<td>cyclopropane-fatty-acyl-phospholipid synthase</td>
</tr>
<tr>
<td>0.002</td>
<td>0.963</td>
<td>Atu1132</td>
<td>97 -- 144</td>
<td>acv2</td>
<td>521 -- 572</td>
<td>-19.838</td>
<td></td>
<td>glutathione S-transferase related protein</td>
</tr>
<tr>
<td>0.003</td>
<td>0.963</td>
<td>Atu1846</td>
<td>69 -- 149</td>
<td>acv2</td>
<td>532 -- 598</td>
<td>-19.488</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.003</td>
<td>0.963</td>
<td>Atu3684</td>
<td>101 -- 131</td>
<td>acv2</td>
<td>579 -- 607</td>
<td>-19.459</td>
<td></td>
<td>siderophore biosynthesis protein</td>
</tr>
<tr>
<td>0.003</td>
<td>0.963</td>
<td>Atu0545</td>
<td>109 -- 147</td>
<td>acv2</td>
<td>532 -- 565</td>
<td>-19.296</td>
<td>flaA</td>
<td>flagella associated protein</td>
</tr>
<tr>
<td>0.003</td>
<td>0.963</td>
<td>Atu1879</td>
<td>77 -- 120</td>
<td>acv2</td>
<td>534 -- 573</td>
<td>-19.254</td>
<td></td>
<td>ABC transporter substrate binding protein (amino acid)</td>
</tr>
<tr>
<td>0.004</td>
<td>0.963</td>
<td>Atu0837</td>
<td>68 -- 111</td>
<td>acv2</td>
<td>499 -- 536</td>
<td>-18.870</td>
<td></td>
<td>transcrip</td>
</tr>
<tr>
<td>0.004</td>
<td>0.963</td>
<td>Atu1601</td>
<td>129 -- 146</td>
<td>acv2</td>
<td>11 -- 28</td>
<td>-18.680</td>
<td>hemN</td>
<td>oxygen-independent coproporphyrinogen III oxidase</td>
</tr>
</tbody>
</table>
## APPENDIX A

**PRIMERS, FRAGMENTS, PLASMIDS AND STRAINS**

**Table 6. Primer list.**

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Sequence</th>
<th>Restriction Site</th>
<th>Size of amplified region</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>avcR1-UP-F1</td>
<td>CTGCAGGCAATGCTCAGCAATCTTTG</td>
<td>PstI</td>
<td>993bp</td>
<td>Used for PCR amplification of <em>anti-virC2</em> promoter up stream flanking sequence</td>
</tr>
<tr>
<td>avcR1-UP-R1</td>
<td>GTCGACTAGCAATCTGCTACGGTGTTGG</td>
<td>SalI</td>
<td>958bp</td>
<td>Used for PCR amplification of <em>anti-virC2</em> promoter down stream flanking sequence</td>
</tr>
<tr>
<td>avcR1-DN-F1</td>
<td>GTGACAGCACTGCGGAGTGATTGG</td>
<td>SalI</td>
<td>958bp</td>
<td>Used for PCR amplification of <em>anti-virC2</em> promoter down stream flanking sequence</td>
</tr>
<tr>
<td>avcR1-DN-R1</td>
<td>TCTAGACGTAGCAAGCTGAATTTGCAAT</td>
<td>XbaI</td>
<td>958bp</td>
<td>Used for PCR amplification of <em>anti-virC2</em> promoter down stream flanking sequence</td>
</tr>
<tr>
<td>vc2DR-F</td>
<td>AGCTGCCTTTTCTCGGTACT</td>
<td>None</td>
<td>N/A</td>
<td>Used to PCR check deletion of <em>anti-virC2</em> promoter region</td>
</tr>
<tr>
<td>vc2DR-R</td>
<td>GTACCAGGAAAGGCAGCTT</td>
<td>None</td>
<td>N/A</td>
<td>Used to PCR check deletion of <em>anti-virC2</em> promoter region</td>
</tr>
<tr>
<td>VC2-seq-F</td>
<td>TGAGGTGAAACTCAAGCAG</td>
<td>None</td>
<td>495bp</td>
<td>Used for sequencing ∆Pavc2 deletion mutant</td>
</tr>
<tr>
<td>VC2-seq-R</td>
<td>TCGCTAATAGAAGGCTGCTCG</td>
<td>None</td>
<td>495bp</td>
<td>Used for sequencing ∆Pavc2 deletion mutant</td>
</tr>
<tr>
<td>VC2KO-UP-F1</td>
<td>CTGCAGGCAATGCTCAGCAATCTTTG</td>
<td>PstI</td>
<td>993bp</td>
<td>Used for PCR amplification of <em>virC2</em> up stream flanking sequence</td>
</tr>
<tr>
<td>VC2KO-UP-R1</td>
<td>GGTACCTAGCAATCTGCTACGGTGTTGG</td>
<td>KpnI</td>
<td>993bp</td>
<td>Used for PCR amplification of <em>virC2</em> up stream flanking sequence</td>
</tr>
<tr>
<td>VC2KO-DN-F1</td>
<td>GTGACCTTACACCCCTACAAGGCTT</td>
<td>KpnI</td>
<td>1011</td>
<td>Used for PCR amplification of <em>virC2</em> down stream flanking sequence</td>
</tr>
<tr>
<td>VC2KO-DN-R1</td>
<td>TCTAGACGACTGGAGGGCTTACTG</td>
<td>XbaI</td>
<td>1011</td>
<td>Used for PCR amplification of <em>virC2</em> down stream flanking sequence</td>
</tr>
<tr>
<td>VirC2F</td>
<td>TGCAGCTCCGAAGAGTATTAG</td>
<td>None</td>
<td>N/A</td>
<td>Used to PCR check deletion of <em>virC2</em> region</td>
</tr>
<tr>
<td>VirC2R</td>
<td>CGACCGGCTTCCATATTGACGA</td>
<td>None</td>
<td>N/A</td>
<td>Used to PCR check deletion of <em>virC2</em> region</td>
</tr>
<tr>
<td>vc2T-seq-F</td>
<td>GCTGACGTGAGGTGTTTCGTA</td>
<td>None</td>
<td>829bp</td>
<td>Used for sequencing ∆avirC2 deletion mutant</td>
</tr>
<tr>
<td>vc2T-seq-R</td>
<td>CTCCCTGAAAGGAATCTCAG</td>
<td>None</td>
<td>829bp</td>
<td>Used for sequencing ∆avirC2 deletion mutant</td>
</tr>
<tr>
<td>xx-01</td>
<td>GCTGAAGTCCAGCTGCCGAAGA</td>
<td>None</td>
<td>498bp</td>
<td>Used to amplify the BAR gene and screen for the presence of pB35S-GUS</td>
</tr>
<tr>
<td>xx-02</td>
<td>ACCATGAGCCGACAAGCAG</td>
<td>None</td>
<td>498bp</td>
<td>Used to amplify the BAR gene and screen for the presence of pB35S-GUS</td>
</tr>
<tr>
<td>Name</td>
<td>Primer Sequence</td>
<td>Length</td>
<td>Note</td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>--------------------------------------</td>
<td>--------</td>
<td>---------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>virC1-Flag-F1/virC1-Flag-R1</td>
<td>TGGACTACAAGACCCATGACG AGTGGTCGTTTTCATCAG</td>
<td>196bp</td>
<td>Used to screen for N-terminal VirC1 FLAG-tag</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTCCCTGAAACGGAATCTCAG CCCTTGTACCTGACCTT</td>
<td>154bp</td>
<td>Used to screen for N-terminal VirC2 FLAG-tag</td>
<td></td>
</tr>
<tr>
<td>SO501</td>
<td>CGACTCAGCTATAGGGAGAGCGGC</td>
<td>None</td>
<td>Used for sequencing vector pJET1.2 (from ThermoFisher Scientific, USA)</td>
<td></td>
</tr>
<tr>
<td>SO511</td>
<td>AAGAACAGATTTTCCATGGCAG</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5' RACE adapter</td>
<td>ATATTGCAGGTATCTAGGAACGACTAGAAGAAA</td>
<td>None</td>
<td>5' Adapter used during 5' RACE (Vogel et al., 2005)</td>
<td></td>
</tr>
<tr>
<td>B6</td>
<td>GCGCGAATTCTCTTGAAGA</td>
<td>None</td>
<td>5' Adapter Specific Primer used in 3' RACE (Vogel et al., 2005)</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>CAAATCCCTGTAGAACGAACA</td>
<td>None</td>
<td>3' Adapter used during 3' RACE (Vogel et al., 2005)</td>
<td></td>
</tr>
<tr>
<td>3' RACE adapter</td>
<td>TTCCTCGCTTCTTGCAGCCGCCATGCTC</td>
<td>None</td>
<td>3' Adapter Specific Primer used during 3' RACE (Vogel et al., 2005)</td>
<td></td>
</tr>
<tr>
<td>E1</td>
<td>GCTCTTAGCGGCCGACATG</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aveR1-RACE-R1</td>
<td>GGGAGAAGGCAACGACAG</td>
<td>None</td>
<td>Variable Gene Specific Primers for 5' RACE done on anti-virC2 region 1</td>
<td></td>
</tr>
<tr>
<td>aveR1-RACE-R2</td>
<td>CAACGAAACGTCGACTCTCTCA</td>
<td>None</td>
<td>Variable Gene Specific Primers for 3' RACE done on anti-virC2 region 1</td>
<td></td>
</tr>
<tr>
<td>aveR1-RACE-R3</td>
<td>TCAAAAAGATAGGACCCATCCA</td>
<td>None</td>
<td>Variable Gene Specific Primers Used for 5' RACE done on anti-virC2</td>
<td></td>
</tr>
<tr>
<td>aveR1-RACE-R4</td>
<td>AGGACCCATCATCACCCT</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aveR1-RACE-F1</td>
<td>GGAGTGATTGGATGGGTCTCT</td>
<td>None</td>
<td>Variable Gene Specific Primers for 3' RACE done on anti-virC2 region 1</td>
<td></td>
</tr>
<tr>
<td>aveR1-RACE-F2</td>
<td>TGGATGGGGTCATCTTCTTTGA</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aveR1-RACE-F3</td>
<td>CAGCTGTTGTGCTGCTGCCCTCTC</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aveR1-RACE-F4</td>
<td>TCTCCCGAGCAAAGAAACAT</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aveR2-RACE-R1</td>
<td>TATAGGAAGCCGCTCGCAAT</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>aveR2-RACE-R2</td>
<td>TCCCGGTCCTGCTAATAGAA</td>
<td>None</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Primer Sequence</td>
<td>Status</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R11</td>
<td>GTTCAGACCTCCCCGCATGT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R10</td>
<td>CGATCCCTCACACAGCTTTT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R9</td>
<td>TCCGAAGAGTTATCCGATCC</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R8</td>
<td>CGGATGGATCGTTTCGTT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R7</td>
<td>GCGATTTTGAAAAACATGCTG</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R6</td>
<td>CTTTCCAAGTGCTACAAATG</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R5</td>
<td>TGACAACCTGATCTGCAAT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R4</td>
<td>GCGCCTGAAGTGATCGAGAT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R3</td>
<td>CGCCAGAAAGATCCAGGT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R2</td>
<td>GACCGTGATGCTTTGAGTT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-R1</td>
<td>ATTGGGAGCGGGCTTCTATT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F1</td>
<td>TTAGGGAGACCGGGAACAT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F2</td>
<td>GGGAGGTCTGAACGATGATT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F3</td>
<td>GAAAGCTGTGTGAGGGATCG</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F4</td>
<td>GCACGAAAACGATCCATCC</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F5</td>
<td>CCAGCATGTTTTCAAAATCG</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F6</td>
<td>GGCGCAAGATCATTGTTAGC</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F7</td>
<td>GGACTGTATTGCAGATCGAGG</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F8</td>
<td>GGATCAGGTTGTATATCTTCG</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F9</td>
<td>GCTGAAAAGGAAGACCTGGAT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F10</td>
<td>GGATCTTTCTGCGCGCTGT</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F11</td>
<td>CTGTCGTCGAACTCAAAGCA</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
<tr>
<td>avcR2-RACE-F12</td>
<td>ACTCAAGACATCCAGGTC</td>
<td>None</td>
<td>Variable</td>
<td></td>
</tr>
</tbody>
</table>

Gene specific Primers Used for 3’ RACE done on anti-virC2 region 2
<table>
<thead>
<tr>
<th>Primer Set</th>
<th>Primer Sequence</th>
<th>Product Size</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>F13</td>
<td>TGATCAGAGCTGTGACAACG</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>avcR2-RACE-F14</td>
<td>AGGTCGTCGATCTTCTCCTCGT</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>avcR2-RACE-F15</td>
<td>CGAAATCGTCCATCCTTTTT</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>avcR3-RACE-R1</td>
<td>AACAGGCAAAGAAGGATGGA</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>avcR3-RACE-R2</td>
<td>GGACGATTTCCGGTCGAG</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>avcR3-RACE-F1</td>
<td>CTCCCCGACGACAAGCG</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>avcR3-RACE-F1</td>
<td>AGGTCGTCGATCTTCTCCTCGT</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>avcR3-RACE-F2</td>
<td>TGACAAATACCCCGCGAAAAA</td>
<td>None</td>
<td>143bp</td>
</tr>
<tr>
<td>avcR3-RACE-F3</td>
<td>TCCATCTCTCTTGCTGTGTT</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>VirB1 F</td>
<td>TGCAGCTCCGAGAGTTATC</td>
<td>None</td>
<td>102bp</td>
</tr>
<tr>
<td>VirB1 R</td>
<td>CGACGCGCTTCTATAGCGA</td>
<td>None</td>
<td>qRT-PCR Primers for virB1</td>
</tr>
<tr>
<td>VirC2 F</td>
<td>TGCAGCTCCGAGAGTTATC</td>
<td>None</td>
<td>102bp</td>
</tr>
<tr>
<td>VirC2 R</td>
<td>CGACGCGCTTCTATAGCGA</td>
<td>None</td>
<td>qRT-PCR Primers for virC2</td>
</tr>
<tr>
<td>VirD1 F</td>
<td>TCGTCATAAGGATGGAGGCCA</td>
<td>None</td>
<td>133bp</td>
</tr>
<tr>
<td>VirD1 R</td>
<td>ACCGAAAGCGATACGTTCAG</td>
<td>None</td>
<td>qRT-PCR Primers for virD1</td>
</tr>
<tr>
<td>VirG F</td>
<td>TACGCGAGAGGAGGTTATC</td>
<td>None</td>
<td>139bp</td>
</tr>
<tr>
<td>VirG R</td>
<td>TCCACGTCAGCGTCAAAGAA</td>
<td>None</td>
<td>qRT-PCR Primers for virG</td>
</tr>
<tr>
<td>rpoD F</td>
<td>GCCAGGAGCAGCTCGTTACAAAG</td>
<td>None</td>
<td>123bp</td>
</tr>
<tr>
<td>rpoD R</td>
<td>GCCGAGGAGCAGCTCGTTACAAAG</td>
<td>None</td>
<td>qRT-Primers for rpoD</td>
</tr>
<tr>
<td>virC1 F</td>
<td>GAGCGAACTCAAACACAGA</td>
<td>None</td>
<td>250bp</td>
</tr>
<tr>
<td>virC1 R</td>
<td>CTCGTCATCGGAGAAGCTGA</td>
<td>None</td>
<td>qRT-PCR Primers for virC1</td>
</tr>
<tr>
<td>anti-virC2-R1</td>
<td>ATTGCAGACGGCTTCTATTAA</td>
<td>None</td>
<td>231bp</td>
</tr>
<tr>
<td>anti-virC2-R2</td>
<td>GCCGCTGCAACTGAAGATTCAAGAT</td>
<td>None</td>
<td>qRT-PCR Primers for anti-virC2</td>
</tr>
<tr>
<td>avcR1-OE-F1</td>
<td>AAGCTTTAAAAACTCAGCACTGCGAG</td>
<td>HindIII</td>
<td>119bp</td>
</tr>
<tr>
<td>avcR1-OE-R1</td>
<td>CCGCGGAGCTTGGCTACCCGACG</td>
<td>SacII</td>
<td>Primers used to amplify anti-virC2 region 1 for pKL1001</td>
</tr>
<tr>
<td>avcR2-OE-F1</td>
<td>AAGCTTTAACGAGCAGGAGGATCG</td>
<td>HindIII</td>
<td>412bp</td>
</tr>
<tr>
<td>avcR2-OE-R1</td>
<td>CCGCGGCACTTGAACGATAAACAG</td>
<td>SacII</td>
<td>Primers used to amplify anti-virC2 region 2 for pKL1002</td>
</tr>
<tr>
<td>avcR3-OE-F1</td>
<td>AAGCTTAAAAACCGCTTTCTCTGTTC</td>
<td>HindIII</td>
<td>103bp</td>
</tr>
<tr>
<td>avcR3-OE-R1</td>
<td>CCGCGGCGGCGGAGTTGAGTTT</td>
<td>SacII</td>
<td>Primers used to amplify anti-virC2 region 3 for pKL1003</td>
</tr>
<tr>
<td>virGC-IGR-F1</td>
<td>AAGCTTCAAGCGACGAGCGGAGGATCG</td>
<td>HindIII</td>
<td>158bp</td>
</tr>
</tbody>
</table>
|            | CCGCGGAGCTTGGCTACCCGACG | SacII | Primers used to amplify
<table>
<thead>
<tr>
<th><strong>virGC-IGR-R1</strong></th>
<th>CCGGGCCCTGAGGTAGAGGTGCATT</th>
<th>SacII</th>
<th><strong>virC/virG</strong> intergenic region for pKL1004</th>
</tr>
</thead>
<tbody>
<tr>
<td>pTF505F1</td>
<td>GCCTTTATCCATTCTCTATTGAA</td>
<td>None</td>
<td>Variable</td>
</tr>
<tr>
<td>pTF505F2</td>
<td>ACGGCTGCCTGAA</td>
<td>None</td>
<td>Primers for sequencing inserts in pTF505</td>
</tr>
<tr>
<td>Fragment Name</td>
<td>Fragment Sequence</td>
<td>Restriction Site</td>
<td>Size</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>------------------</td>
<td>------</td>
</tr>
<tr>
<td>3xFLAG_vir C1</td>
<td>ATCCaagcttAAAGAAGCCACCAATGCCGCCACCGGCAACCCGATATGGCCCAGAAGCTGCTGATGAGAGAAGT</td>
<td>Hind III/BamH I</td>
<td>1250</td>
</tr>
<tr>
<td></td>
<td>GGGAGATTTGAGTATTGCGTGGCCGATACGCTGATTGCTGAGGAGAAGCCTGCTGAGGAGAAGCCTGCTTTCGAGCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GGCTCCTCCTTCTCGGCCAGTCTACCGGCTGGAACCTGCAACCTGCAACCTGCAACCTGCAACCTGCAACCTGCAACCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GTGCCGACCGCGCGATGTCGGACATGCTCGCAAGCCTTCCAGTTGTACAGTCTCCCATGCACGAGAGAGACGCATTTGCCGCGATGGAAGGAACGTGGCATGTTGCATCTCACATTGCTGAATATGAGAACCGATCCGACAATGCGCCTCCTCGAACGGAATCTCAGAATCGCCATGGAGGAACTCGTCACTATCTCCAAATTGGTTAGCGAAGCCTTGGAGGGGTGAGaatcc</td>
<td>Hind III/BamH I</td>
<td>1229</td>
</tr>
<tr>
<td></td>
<td>CACATCGCAGCGCAGCGATGTCGGCACTACCGGCTGGAACCTGCAACCTGCAACCTGCAACCTGCAACCTGCAACCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TTGGCTACAGCTCTCATTGATGAGTGGGACAGGAGAAGCCTGCTGAGGAGAAGCCTGCTTTCGAGCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fragment used to generate N-terminally FLAG tagged VirG sequence</td>
<td>XbaI/PstI</td>
<td>1250</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>ATCCcttagaAAAAACAGCGCTACCGGGGCGCTAGTTTGAAGATC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmid</td>
<td>Description/Use</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>pTFsacB</td>
<td>Vector for homologous recombination, 5724 bp, Kanamycin resistance</td>
<td>Dr. Kan Wang Lab collections, designed by Dr. Keunsub Lee, (unpublished)</td>
<td></td>
</tr>
<tr>
<td>pJET1.2</td>
<td>Sequencing Vector, 2974bp, Carbenicillin resistance</td>
<td>CloneJET PCR cloning kit (ThermoFisher Scientific, USA)</td>
<td></td>
</tr>
<tr>
<td>pAJ007</td>
<td>Used to generate anti-virC2 promoter knockout mutant ΔPavc2, 7652bp, Kanamycin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>pAJ008</td>
<td>Intermediate Sequencing Vector for pAJ007 Contains Up stream flanking sequence, 3586 bp, Carbenicillin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>pAJ009</td>
<td>Intermediate Sequencing Vector for pAJ007, Contains Downstream Flanking Sequence, 3551 bp, Carbenicillin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>pAJ019</td>
<td>Intermediate Sequencing Vector for pAJ021, Contains up stream Flanking Sequence, 3604 bp, Carbenicillin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>pAJ020</td>
<td>Intermediate Sequencing Vector for pAJ021, Contains downstream Flanking Sequence, 3586 bp, Carbenicillin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>pAJ021</td>
<td>Used to generate virC2 Knockout mutant ΔvirC2, 7705 bp, Kanamycin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>pTF505</td>
<td>Overexpression vector, 6789 bp, Spectinomycin resistance</td>
<td>Dr. Kan Wang Lab collections, Designed by Dr. Keunsub Lee, (Lee et al., 2013)</td>
<td></td>
</tr>
<tr>
<td>pKL1001</td>
<td>Overexpression of anti-virC2 region, 16885 bp, Spectinomycin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>pKL1002</td>
<td>Overexpression of anti-virC2 region 2, 7177 bp, Spectinomycin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>pKL1003</td>
<td>Overexpression of anti-virC2 region 3, 6881 bp, Spectinomycin resistance</td>
<td>This Work</td>
<td></td>
</tr>
<tr>
<td>Vector</td>
<td>Description</td>
<td>Origin</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>pKL1004</td>
<td>Overexpression of <em>virC2/virG</em> intergenic region, 6936 bp, Spectinomycin resistance</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pKL1006</td>
<td>Used to generate VirC2 FLAG-tagged mutant strains, 6905 bp, Kanamycin resistance</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pKL1007</td>
<td>Used to generate VirC1 FLAG-tagged mutant strains, 6926 bp, Kanamycin resistance</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pKL1015</td>
<td>Used to generate VirG FLAG-tagged mutant strains, 6926 bp, Kanamycin resistance</td>
<td>This work</td>
<td></td>
</tr>
<tr>
<td>pB35S-GUS</td>
<td>Used for the expression of GUS during AGROBEST experiments, 11276 bp, Kanamycin resistance</td>
<td>Dr. Kan Wang Lab Collections</td>
<td></td>
</tr>
</tbody>
</table>
## Table 9. *A. tumefaciens* strain list

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>C58</td>
<td>C58 Wild-Type</td>
<td>Dr. Kan Wang Lab Stock</td>
</tr>
<tr>
<td>C58 ∆Pavc2</td>
<td>Deletion mutant of <em>anti-virC2</em> promoter region</td>
<td>This Work</td>
</tr>
<tr>
<td>C58 ∆virC2</td>
<td>Deletion mutant of <em>virC2</em> gene</td>
<td>This Work</td>
</tr>
<tr>
<td>C58 ::VirC2-FLAG</td>
<td>FLAG-tagged VirC2</td>
<td>This Work</td>
</tr>
<tr>
<td>C58 ::VirC1-FLAG</td>
<td>FLAG-tagged VirC1</td>
<td>This Work</td>
</tr>
<tr>
<td>C58::VirG-FLAG</td>
<td>FLAG-tagged VirG</td>
<td>This Work</td>
</tr>
<tr>
<td>C58 ∆Pavc2 ::VirC2-FLAG</td>
<td>FLAG-tagged VirC2</td>
<td>This Work</td>
</tr>
<tr>
<td>C58 ∆Pavc2 ::VirC2 FLAG</td>
<td>FLAG-tagged VirC1</td>
<td>This Work</td>
</tr>
<tr>
<td>C58 ∆virC2 ::VirG-FLAG</td>
<td>FLAG-tagged VirG</td>
<td>This Work</td>
</tr>
<tr>
<td>C58 ∆virC2 ::VirC1-FLAG</td>
<td>FLAG-tagged VirC1</td>
<td>This Work</td>
</tr>
<tr>
<td>C58 ∆virC2::VirG-FLAG</td>
<td>FLAG-tagged VirG</td>
<td>This Work</td>
</tr>
</tbody>
</table>
APPENDIX B
SUPPLEMENTAL INFORMATION

Introduction

While the majority of work performed for this research project is contained in the main body of the thesis, some additional work is described here in Appendix B, Supplemental Information. Projects contained here were either not relevant to the narrative of the *avc2* transcript and its function or did not produce reportable results because of technical errors.

Research is a learning process. By describing all experiments performed whether or not they were successful I am able to better reflect on past mistakes and review what could have been done differently. By including this supplemental work, I hope to demonstrate additional undertakings which have been performed, as well as ways in which I have learned and am able to better plan for the future.

Organization of this supplemental information is divided into sections, each of which describes a different project not included in the main body of the thesis. An overview of each project is included, as well as ways which the project may be improved and used in future research is applicable. Figures and tables included here are referenced as Figure S1, Figure S2, and so on, while figures from the main thesis are simply Figure 1 and Figure 2.
Overexpression of Individual anti-virC2 Regions and Intergenic Region

As previously described in the main body of the thesis, the *avic2* transcript can be divided into 3 (possibly 4) regions based on expression pattern (Figure 7). In order to better understand the function of each individual region, constructs overexpressing each of these regions were generated and transformed into *A. tumefaciens* strains C58 and C58 ΔPavic2. In addition to the regions of *avic2* which were overexpressed, the intergenic region between the *virC* and *virG* operons was overexpressed in a similar way. An illustration depicting overexpressed transcripts and constructs generated can be found in Figure S1.

Construction of overexpression plasmids pKL1001- pKL1004

Four constructs were generated to overexpress each region of the *avic2* transcript as well as the intergenic region between the *virG* and *virC* operons. Regions to be overexpressed were selected based on previous sequencing data generated by Lee et al. (2013) and are illustrated in Figure S1.

First, primer pairs were designed to amplify each of the desired genetic sequences from *A. tumefaciens* total DNA using PCR. Each primer pair was designed to include HindIII/SacII restriction sites for cloning. Once fragments had been amplified they were first cloned into sequencing vector pJET1.2 for sequencing to check for any possible mutations. After the initial sequencing check, fragments were then digested out of pJET1.2 and cloned into vector pTF505. Vector pTF505 is an overexpression vector which contains the constitutive promoter PrrnC from *Sinorhizobium meliloti*, followed by a multiple cloning site, and transcriptional terminator, TpsbANT. A total of four plasmids
were generated and named pKL1001, pKL1002, pKL1003, and pKL1004 to overexpress desired transcript regions as shown in Figure S1.

Each of the four overexpression vectors was transformed into both wild type C58 and C58 ΔPavc2. These bacteria strains were then termed C58(pKL1001-1004) and C58 ΔPavc2(1001-1004). All cloning and design steps were performed by Dr. Keunsub Lee with assistance from myself.

*Vir gene expression in C58 ΔPavc2 overexpression strains*

After overexpression mutants were generated, select transcript levels were then able to be measured using qRT-PCR. Initially, only strains C58 ΔPavc2(pKL1001-1004) were analyzed. This was done to minimize the amount of bacterial cultures handled at one time and avoid human errors which may occur when dealing with an abundance of samples at one time. qRT-PCR was performed using random hexamers to synthesize cDNA as described in Chapter II Materials and Methods. Since only two replicates of qRT-PCR were performed, proper statistical analysis was not carried out. However, observations from the data can still be made. Results for qRT-PCR experiments are shown in Figure S2 as well as Table S1.

When examining transcript levels in overexpression mutants C58 ΔPavc2(pKL1001-1004) there are several observations which can be made. First it can be observed by examining *virC2* levels in C58 ΔPavc2(pKL1002) that the intended region is indeed being overexpressed. Because *virC2* primers used were specific to the region overexpressed by the pKL1002 vector, overexpression of that transcript is able to be observed. This also confirms that the overexpression vector is functioning properly at
least for pKL1002. However, in future studies, primers could be designed specifically to each section in order to confirm the functionality of each individual vector.

Also, it appears that virG transcript levels are elevated in every A. tumefaciens strain which contains a deletion of the avc2 promoter region. This result was previously observed when regular qRT-PCR was performed on C58 and C58 ΔPavc2 as detailed in Chapter III Results. Interestingly, there appears to be elevated expression levels of virG even when the deleted avc2 promoter region or any other region of the avc2 transcript is overexpressed. This demonstrates a failure of any of the pKL1001 through pKL1004 vectors to complement C58 ΔPavc2 and return levels of virG transcript to a similar level as those in wild type C58. This means that the change in nucleotide sequence directly downstream of the virG operon and not lack of avc2 transcript is the cause of this higher observed level of virG. In future studies, homologous recombination could be used to replace the deleted sequence in C58 ΔPavc2. Then, virG expression levels could be examined and compared to wild type to check for complementation.

It is also important to notice the inconsistency of the qRT-PCR data shown here. While transcript levels were compared to that of wild type to generate the graph shown in Figure S2, real numbers are listed in Table S1. By looking at relative transcript abundance as shown in Table S1 it can be seen that the abundance of transcripts between experiments was not at all consistent. Most likely, this inconsistency was caused problems with RNA extraction and handling when I was the one performing the extraction. Many repeated attempts were made which resulted in completely degraded RNA samples which could not be used for cDNA synthesis and when RNA was extracted and cDNA synthesized, results were not consistent. In future studies, an alternate
researcher may be able to perform additional qRT-PCR in order to more fully examine the effect of overexpression of various regions of the *avc2* transcript.

**pAJ025; Constitutive Expression of *virC2* and Inducible Expression of *anti-virC2***

As stated earlier, it was initially hypothesized that the *avc2* transcript somehow acted to regulate its potential target and *cis*-encoded gene *virC2*. While we can now see that there is a lack of evidence to link the two, at the beginning of this research a connection was hypothesized.

A vector is designed to test regulation of *virC2* by *anti-virC2*

In order to test for the possible regulation of *avc2* on *virC2*, a special vector named pAJ025 was designed (Figure S3). On this vector, which was synthesized at Genscript using pTF505 as starting material, *virC2* is expressed using the constitutive promoter Prrnc while *avc2* is inducibly expressed by the *virB* promoter. The process of pAJ025 design and ordering was assisted by Dr. Keunsub Lee and Dr. Jennifer Raji. Once received from Genscript, the vector was first transformed into *A. tumefaciens* C58 wild type to check for proper expression.

pAJ025 does not properly express *virC2* and *anti-virC2* transcripts

Regular qRT-PCR was performed to examine expression levels in C58(pAJ025). Four biological replicates of qRT-PCR were performed in order to examine levels of transcript from the *virC2/anti-virC2* region (Figure S4). It is important to remember that because these two genetic sequences are complementary, they will both be detected as
one when performing non-strand specific qRT-PCR. Because the pAJ025 vector is constructed to constitutively overexpress *virC2* and inducibly express *avc2*, transcript levels were expected to be greatly elevated in qRT-PCR experiments. Unfortunately, this was not the case. Based on the four replicates of qRT-PCR, no statistical difference was detected between C58 transformed with pAJ025 compared to C58 transformed with the empty backbone vector pTF505. When bacteria were grown under induction conditions, levels of *virC2/avc2* transcript were slightly elevated in C58(pAJ025) when compared to C58(pTF505) but not enough to be considered statistically significant (Figure S4). Under non-induction conditions there was barely any difference in transcript expression level between C58(pAJ025) and C58(pTF505). Based on these results, it appears there is something wrong with the pAJ025 vector.

One possible explanation for the lack of *virC2* and *avc2* expression observed in C58(pAJ025) comes from the design of the vector itself. Because both *avc2* and *virC2* sequences are encoded in pAJ025, the vector has two large regions which are complementary to one another. It could be possible that the vector is base-pairing with itself while in the bacterial cell, which is leading to its lack of function. Because of the evidence already produced showing a lack of interaction between the *avc2* and *virC2* transcript, this project was put aside. This may be one part of the project which is not worth picking up again in the future.

**Additional FLAG-Tagged Proteins**

As described in the main body of this thesis, *A. tumefaciens* strains were generated in which an N-terminal FLAG tag was inserted into the genomic sequence of
the virC2 gene. In addition, A. tumefaciens strains were generated in a similar manner which contained an N-terminal FLAG tag in the coding sequence of either the virC1 or virG gene. Because VirG-FLAG and VirC1-FLAG proteins were unable to be visualized during Western blot, their construction and attempted visualization with Western blot will be described here in the supplemental chapters.

Construction of A. tumefaciens strains containing VirC1-FLAG and VirG-FLAG tagged proteins

The method of homologous recombination (Ried and Collmer. 1987) was used to generate mutant A. tumefaciens strains which contained an N-terminal FLAG-tag in the coding sequence of either the virC1 or virG genes. First, oligo fragments were ordered from IDT which contained a sequence including upstream and downstream regions surrounding the FLAG-tag to be inserted. Restriction sites were also added for cloning. Fragments 3xFLAG-virG and 3xFLAG-virC1 were digested with appropriate restriction enzymes and then cloned into vector pTFsacB to generate pKL1015 and pKL1006 respectively. Construction details for all vectors used to generate FLAG-tagged proteins are shown in Figure S5.

After construction of vectors pKL1006 and pKL1015 were complete, vectors were transformed into A. tumefaciens strains C58 wild type and C58 ΔPavc2 to generate strains C58::virC1-FLAG, C58::virG-FLAG, C58 ΔPavc2::virC1-FLAG, and C58 ΔPavc2::virG-FLAG. Selection for homologous recombination was then performed as described in Chapter 2 Materials and Methods and illustrated in Figure 3. Briefly, LB containing kanamycin was used to screen for an initial recombination event. A second
recombination event was found for by screening for the ability of bacteria to grow on LB media containing sucrose. The presence of a FLAG-tag was further confirmed by PCR.

**VirC1-FLAG and VirG-FLAG proteins fail to visualize during Western blot**

Western blot was then performed on *A. tumefaciens* strains to compare VirC1-FLAG and VirG-FLAG tagged proteins between C58 wild type and C58 ΔPavc2 as described in Chapter 2 Materials and Methods.

Bacterial protein extracts containing VirC1-tagged or VirG-tagged proteins failed to visualize on film. Samples were run along with VirC2-FLAG tagged protein which was able to be used as a positive control. Additionally, the design of sequence of fragments ordered from IDT were checked for accuracy and no error was found.

There are several possible reasons why these proteins were unable to be visualized. First, it is possible that VirG-FLAG and VirC1-FLAG proteins were more fragile than VirC2-FLAG and were degraded during the extraction process. However, unless the FLAG-tag epitope was completely degraded a band would still be expected to be seen at an somewhere on the blot, or perhaps as a smear (Mahmood and Yang, 2012). It is also possible that the amount of FLAG-tagged protein present in the sample was not present in a concentration high enough to be detected (Mahmood and Yang, 2012). The problem may also have to do with the 3xFLAG tag interfering with protein expression.

In order to overcome this problem and visualize VirG and or VirC1 FLAG-tagged proteins using Western blot in the future, alternative *A. tumefaciens* mutants could be generated which contain C-terminally tagged proteins instead of the current N-terminal tag. This could easily be done by ordering additional fragments through IDT which could
then be cloned into vector pTFsacB followed by homologous recombination to generate a C-terminal FLAG tag.
References


Figure S1 pKL1004 through pKL 1001 were used to overexpress each of the three regions of the anti-\textit{virC2} transcript as well as the intergenic region between the \textit{virC} and \textit{virG} operons. Expression of each region expressed is shown in under induction and non-induction conditions. Specific regions expressed in each construct are denoted in green arrows and are aligned to their specific place along the \textit{A. tumefaciens} genome sequence.
Figure S2. *Vir* Gene transcript expression is examined in C58 ∆Pavc2 strains overexpressing a portion of the *anti-virC2 transcript*. C58 wild type and C58 ∆Pavc2 both containing empty vectors are used for comparison. Induction conditions were used for bacterial growth conditions.
**Figure S3. Vector pAJ025.** This vector was designed with the goal of constitutively overexpressing the *virC2* transcript while inducibly expressing the *avc2* transcript. The vector backbone comes from the pTF505 which was sent to Genscript where pAJ025 was constructed and sent back for transformation into *A. tumefaciens*. 
Figure S4. Relative abundance of transcript from the virC2/anti-virC2 region is shown for C58(pTF505) (empty vector) and C58(pAJ025). There is no statistical difference between transcript levels when comparing C58(pTF505) and C58(pAJ025) under either induction or non-induction conditions ($P = 0.1333$ and $0.9928$ respectively by paired sample students t-test).
Mutant strains were generated which contain N-termianally tagged VirC2, VirC1 and VirG proteins. While VirC2-FLAG proteins were able to be detected and visualized using Western blot, VirC1-FLAG, and VirG-FLAG proteins failed to do so.
Table S1. Transcript Abundance is shown for overexpression strains C58 ΔPavc2(pKL1001-1004) along with C58 ΔPavc2 and C58 wild type empty vector for comparison.

<table>
<thead>
<tr>
<th>eDNA</th>
<th>rpoD Rep 1</th>
<th>rpoD Rep 2</th>
<th>virG Rep 1</th>
<th>virG Rep 2</th>
<th>virC2 Rep 1</th>
<th>virC2 Rep 2</th>
<th>virC1 Rep 1</th>
<th>virC1 Rep 2</th>
<th>virB1 Rep 1</th>
<th>virB1 Rep 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>C58 (pTF505)</td>
<td>1</td>
<td>1</td>
<td>121.10</td>
<td>15.89</td>
<td>60.27</td>
<td>2.21</td>
<td>11.66</td>
<td>16.68</td>
<td>70.36</td>
<td>86.62</td>
</tr>
<tr>
<td>C58 ΔPavc2R1 (pTF505)</td>
<td>1</td>
<td>1</td>
<td>220.30</td>
<td>19.16</td>
<td>58.08</td>
<td>0.91</td>
<td>11.11</td>
<td>22.37</td>
<td>92.20</td>
<td>112.73</td>
</tr>
<tr>
<td>C58 ΔPavc2R1 (pKL1001)</td>
<td>1</td>
<td>1</td>
<td>194.01</td>
<td>22.06</td>
<td>73.86</td>
<td>1.14</td>
<td>15.38</td>
<td>17.55</td>
<td>75.76</td>
<td>119.43</td>
</tr>
<tr>
<td>C58 ΔPavc2R1 (pKL1002)</td>
<td>1</td>
<td>1</td>
<td>166.96</td>
<td>29.18</td>
<td>79.52</td>
<td>5.86</td>
<td>11.99</td>
<td>12.04</td>
<td>60.27</td>
<td>134.67</td>
</tr>
<tr>
<td>C58 ΔPavc2R1 (pKL1003)</td>
<td>1</td>
<td>1</td>
<td>144.34</td>
<td>38.14</td>
<td>70.69</td>
<td>1.88</td>
<td>10.31</td>
<td>12.52</td>
<td>75.58</td>
<td>90.09</td>
</tr>
<tr>
<td>C58 ΔPavc2R1 (pKL1004)</td>
<td>1</td>
<td>1</td>
<td>238.86</td>
<td>20.97</td>
<td>78.61</td>
<td>0.58</td>
<td>11.96</td>
<td>8.61</td>
<td>79.89</td>
<td>84.64</td>
</tr>
</tbody>
</table>