Spring June 10th, 2017

Epigenetics of Host-Pathogen Interactions: The effect of acetosyringone on Ti Plasmid methylation patterns in Agrobacterium tumefaciens C58

Genevieve Roberts
Alumni

Follow this and additional works at: http://digitalcommons.spu.edu/honorsprojects

Part of the Molecular Genetics Commons

Recommended Citation
http://digitalcommons.spu.edu/honorsprojects/88

This Honors Project is brought to you for free and open access by the University Scholars at Digital Commons @ SPU. It has been accepted for inclusion in Honors Projects by an authorized administrator of Digital Commons @ SPU.
Epigenetics of Host-Pathogen Interactions:

The effect of acetosyringone on Ti Plasmid methylation patterns in *Agrobacterium tumefaciens* C58

By

Genevieve Roberts

Faculty Advisor: Dr. Derek Wood

Second Reader: Dr. Tim Nelson

A project submitted in partial fulfillment

of the requirements of the University Scholars Program

Seattle Pacific University

2017

Approved_______________________

Date___________________________
Table of Contents

Title Page.................................................................1
Abstract .................................................................3
Table of Contents .......................................................2
Introduction .............................................................4-11
  Figure 1 4
  Figure 2 9
  Figure 3 9
  Figure 4 11
Methods ...............................................................12
Results .................................................................13-18
  Table 1 13-15
  Figure 5 16
  Figure 6 17
  Figure 7 18
Discussion .............................................................18-24
References ............................................................25-37
Abstract

The plant pathogen *Agrobacterium tumefaciens* C58 can transfer a portion of its tumor-inducing (Ti) plasmid to plant hosts in response to the plant wound signals, including Acetosyringone. The portion transferred is aptly titled Transfer DNA (T-DNA) which encode genes involved in tumor production and biosynthesis of a unique bacterial food source called opines that provide an advantage to the inducing *Agrobacterium*. The Ti plasmid contains a number of genes, including the virulence region that enables T-DNA transfer. Epigenetics investigates how chemical modifications to DNA that don’t alter sequence are used to control gene expression (for example, genes involved in pathogen virulence). Epigenetic modifications can be detected by a next generation sequencing technology called Single Molecule Real Time (SMRT) sequencing. SMRT detects these modifications by tracking kinetic shifts during DNA synthesis. Given the unique inter-kingdom DNA transfer, and the importance of epigenetic regulation in other bacteria and plant species, a comparative exploration of methylation patterns of the Ti plasmid under conditions that induce virulence was undertaken. Over a dozen genes with a variety of purposes (virulence regulation, ion transport, DNA replication, etc.) lost methylation following exposure to the virulence inducing molecule acetosyringone, suggesting increased transcription. Three genes, *VirD5, TraM* and a phosphate/sodium symporter gained methylation throughout the gene, suggesting down regulation. The patterns discovered, while intriguing, are limited by possible methodological flaws in SMRT sequencing due to incongruities between reported findings and those described in the literature. A further examination of the expression profiles of these genes is warranted given these findings.
Introduction

Epigenetics

Epigenetics, as defined by a 2008 convention at Cold Springs Harbor, encompasses “stably heritable phenotypes resulting from changes in a chromosome without alteration to the DNA sequence” (Berger et al., 2009). This definition is used to describe a variety of long and short term regulatory concepts in genetics, including transient epigenetic modification that occurs during mammalian development, that describe altered gene expression via chemical modifications of DNA while maintaining the same gene sequence. Epigenetic modification can often be reversed by enzymes that are sensitive to environmental cues or by factors present during distinct phases of growth. In Eukaryotes, these modifications include histone modification, brought about by acetylation or methylation. These modifications can lead to either increased or decreased gene expression, depending upon the nature of the modification (Histone acetylation can often act to upregulate while histone methylation can down regulate) (Cohen et al., 2011). Additional epigenetic modifications include small RNA molecules that regulate gene expression and DNA methylation. The modifications can be inherited, common in plants but rare in other Eukaryotes, or arise anew during early development, common for animals (Quadrana & Colot, 2016).

In prokaryotes, epigenetic modification occurs through DNA methylation (Figure 1). The enzymes responsible for these processes include DNA Adenine Methylase (Dam) or DNA Methyl Transferase (DNMT). Dam works by rotating an adenine base 180° and transferring a methyl group to the N6 Nitrogen via S-adenosylmethionine (Pogolotti et al., 1988). DNA chemical modification can prohibit or diminish protein binding by physically blocking key residues from binding their targets. Conversely, hyper-methylation can attract proteins to bind. Small RNAs can target and bind mRNA transcripts for degradation (Willbanks et al., 2016).

Figure 1: N6-Methyladenine
Epigenetic processes may be influenced by external environmental cues. While many of the following examples are documented in eukaryotic species, they serve to illustrate mechanisms that may help us better understand prokaryotic epigenetics. Lead exposure has been shown to cause DNA methylation, histone modification and changes in microRNA (miRNA, short pieces of regulatory RNA produced transiently in the cell). Carcinogen exposure-- from cigarette smoke for example-- can lead to DNA damage, some of which requires DNA methylation to repair; methyltransferin is used to identify the proper strand to repair (Russo et al., 2016). Benzene exposure can lead to both hypo- and hypermethylation of the p15 tumor suppressor in some cancers. DNA methylation and other types of modifications are known or strongly considered to play a role in early mammalian development, aging, diabetes, cardiovascular health, Alzheimer’s, multiple sclerosis, rheumatoid arthritis and viral infections (Villeneuve et al., 2011; Sanchez-Mut & Gräff 2015; Paschos & Allday 2010). In prokaryotes, the amount of DNA methylation near the origin of replication alters SeqA binding which in turn controls replication rates by physically blocking protein-binding sites. In prokaryotes, methylation the parental DNA strand, guiding repair enzymes to correct mismatches or damage on a newly formed daughter strand. Further, by quickly turning gene expression on or off and allowing for acclimatization to varying climate conditions. It also aids in offspring adaptation by enabling expanded use of the genome. (Willbanks et al., 2016)

**Bacterial epigenetics**

Methylation systems can target two of the four nucleotides in the genome, Cytosine and Adenine. Adenine can be methylated by DNA adenine methylase (Dam) which targets 5’-GATC-3’ rotating the adenine base out 180° and adding a methyl group via the methyl donor S-adenosyl-L-methionine (Malone et al., 1995; Horton et al., 2005; Liebert et al., 2004; Urig et al., 2002). Dam is thought to play a role in chromosome replication in bacteria species with two or more chromosomes (*Agrobacterium tumefaciens* C58 has two chromosomes) (Egan et al., 2006). Dam does this by both
methylating the origin of replication to ensure replication occurs only once per cell cycle and recruiting other methylation proteins (Demarre & Chattoraj, 2010). Adenine can also be targeted via a cell cycle regulated DNA methyltransferase (CcrM) that targets 5’-GANTC-3’ sites. CcrM strongly prefers hemi-methylated DNA, part of a process that can signal to a cell which DNA strand is parental or daughter. CcrM also plays role in maintaining methylation and in other parts of DNA regulation since it interacts with other gene regulatory elements (Robertson et al., 2000; Casadesús & Low, 2006; Gonzalez et al., 2014).

Methylation systems originally evolved as a part of restriction enzyme modification systems. They act by methylating native DNA and targeting and cleaving foreign, invading DNA. Many of the known prokaryotic methylation systems, however, have no known associated role in restriction enzymes, and are dubbed “orphan methyltransferases”. These orphan methyltransferases are more active than typically thought, accounting for nearly half of the methylations across a diverse sample of prokaryotic species (Blow et al., 2016). A large array of unique methyltransferases exist within a multitude of phyla. These methyltransferases target adenine and cytosine residues at a variety of motifs suggesting that there are a variety of mechanisms, each of which may be controlled by different environmental cues, that can drive methylation.

The hemi-methylated state of newly replicated DNA is short-lived (0.5 to 4 minutes) since restriction enzymes present in the cell target and degrade unmethylated DNA (Campbell and Kleckner, 1988). Hemi-methylated DNA has a variety of functions including marking the correct template strand (Marinus 1996; Casadesús & Low, 2006), controlling rates of transposon insertion by regulating the binding of transposase (Roberts et al., 1985) and regulating transcription rates by blocking protein binding, particularly for genes with a GATC site in their promoter (Plumbridge and Soll, 1987; Sternberg et al., 1986). The regulation of methylation, particularly for gene promoters, occurs in part through protein binding at specific sites that block the activity of methylating enzymes (Braaten et al.,
These sites may be modified in response to growth phase and food source (Calvo & Matthews, 1994; D’Ari et al., 1993; Newman et al., 1993). DNA Adenine Methylase targets typically remain methylated throughout the lifetime of the organism and any loss of methylation can usually be traced to specific environmental or life cycle cues. Other sites remain almost entirely nonmethylated throughout the life of the organism and therefore play a near constant role in physiological processes (Blomfield 2001; Casadesús & Low, 2006).

Variances in methylation control virulence in some organisms (Casadesús & Low 2006). These include Salmonella typhimurium where the expression of pili on the cells surface can switch on and off rapidly within a population in a process known as phase variation. These pili are critical for cell adhesion and are key epitopes recognized by the immune system. This variation is controlled by methylation of two GATC sites upstream of the gene. When the ‘on’ site is methylated, pili are expressed. When the sites switch, the pili are not expressed (Casadesús & Low, 2006). The sites are methylated via Dam and methylation of the ‘off’ site blocks the binding of leucine regulatory proteins (Lrp) responsible for activating the expression of the pili genes. Reversal of this state is accomplished by Lrp binding at adjacent non-methylated sites producing what is referred to as “mutual exclusion” where binding at the second site (the ‘off’ site) decreases binding at another by more than 10-fold due to an unknown mechanism. Lrp expression is controlled via cyclic AMP and catabolite activator/ repressor protein (CAP/CRP), which inhibit Lrp binding (Forsman et al., 1989) (Casadesús & Low, 2006). To switch phases, DNA replication must first generate a hemi-methylated state with non-methylated sites. Dam then methylated one of the two GATC sites (on/off) which inhibits the methylation of the other site. Methylation at regulatory motifs flanking the ‘on/off’ GATC sites is controlled in two ways. First by reducing the processivity of Dam to allow other DNA binding proteins to bind and block methylation, And second by natural resistance to methylation. Resistance occurs at sites near palindromic repeat sequences or when DNA is in the H form, a naturally-occurring triplex form of DNA that plays a
regulatory role in many diseases (Allers & Leach, 1995; Peterson & Reich, 2006; Wang & Vasquez, 2004).

Virulence is controlled by both hypo and hypermethylation (Marinus & Casedesús, 2009). Additional prokaryotic systems where methylation is a virulence determinant include: inhibited motility via altered expression of flagellar genes in *Salmonella enterica* (Balbontin et al., 2006), envelope instability leading to protein leakage (Pucciarelli et al., 2002), overexpression of fimbriae leading to virulence attenuation by interfering with signal exchange (Jakomin et al., 2008) and hypersensitivity to bile salts (Prieto et al., 2004). Pathogenesis is tightly linked to environmental phenomena and occurs primarily at a population level: death befalls a limited subset of individuals which die during the production of virulence factors. These factors enable survival for the rest of the population.

*Agrobacterium*

*Agrobacterium tumefaciens* C58 are Gram negative soil-dwelling Alphaproteobacteria that infect dicotyledonous plants in one of the few instances of trans-kingdom DNA transfer in nature (there are a few reported cases under laboratory conditions) (Lacroix & Citovsky, 2016). Upon plant wounding, the plant wound signal acetosyringone activates the virulence (*vir*) regulon of *Agrobacterium* (Figure 3). The *vir* regulon controls the cleavage, transfer and integration of the Transfer-DNA (T-DNA). C58 will then migrate toward the plant and attach itself with the help of vir genes along with chromosomal genes. Successful insertion of the T-DNA into the plant hosts chromosomal DNA occurs up to 90% of the time, depending on the location of insertion (Wang, 2008). Integrated T-DNA produces plant hormones resulting in tumor formation and generates unique nitrogenous compounds called (Tzifra & Citovsky, 2003). Opines, such as Octopine and nopaline, are produced by differing sub-types of Ti plasmids and are amino acid derivatives of N-carboxyalkyl amino acids. These are synthesized from arginine, pyruvate for octopine and 2-ketoglutarate for nopaline (Zanker et al., 1994). Plant hormones
produced by The T-DNA, termed oncogenes include auxin and cytokinin, that cause formation of tumors, a hallmark of C58 infection (Valentine, 2003; Gohlke & Deeken, 2014) (Figure 2).

**Figure 2:** Shows transfer of T-DNA from pTi into plant host upon Acetosyringone detection.

T-DNA transfer is mediated by a suite of proteins termed *vir* (Virulence) proteins that comprise a regulon under the control of the VirA/G two component system. The *vir* regulon is stimulated by a combination of plant-derived monosaccharides with conserved hydroxyl groups, including acetosyringone (Palmer & Shaw, 1992). It’s thought that upon cell damage, and particularly cell wall degradation by enzymes, the wounded plant produces phytoalexins which stimulate the production of phenols, including Acetosyringone (Ankenbauer & Nester, 1990.) The phenols are detected by a protein termed chvE, a virulence protein involved in sugar binding and transport and is thought to promote chemotaxis of the bacterium to the wound site (Kemner et al., 1997). C58 also produces auxin which promotes cellular degradation of the plant to produce Acetosyringone. In addition to these phenolic signals, low pH is essential for *vir* induction (Ankenbauer
& Nester, 1990). Phenols, like acetosyringone, interact with the transmembrane portion of VirA and promote its phosphorylation. VirA then transfers this phosphate to VirG enhancing its ability to bind to target promoters of the *vir* regulon and induce target gene expression. Without the ability to phosphorylated, VirA and VirG are unable to promote chemotaxis or T-DNA transfer (Palmer & Shaw, 1992).

The first *vir* proteins involved in T-DNA transfer is VirD2, a restriction endonuclease that nicks the Ti plasmid at 25-bp border repeat sequences that bracket the T-DNA (Herrera-Estrella, et al., 1988). VirD2 also serves as a guide for the T-DNA leaving the cell, trafficking it into the plant nucleus. VirD2 also has a nuclear localization signal which is thought to aid T-DNA integration in the plant nucleus (Mysore et al., 1998). Nuclear targeting is assisted by importin α and β proteins which move from the cytoplasm to the nucleoplasm (Terry et al., 2007; Gelvin 2012). VirD2 is a phosphoprotein whose expression is regulated by phosphorylation/de-phosphorylation by cyclin dependent kinase activating kinases (Bakó et al., 2003; Tao et al., 2004).

VirE2 binds the single stranded T-DNA to prevent degradation by endonucleases in the cell upon transport. Without VirE2, either fewer strands enter the plant cell or more are degraded, leading to a substantial attenuation in virulence (Yusibov et al., 1994). Recent works shows that VirE2 facilitates the exit of T-DNA by associating with VirE1, forming a hole in the plasma membrane of the plant host. This complex interacts with plant host chaperone proteins to enable T-DNA to enter the nucleus and facilitates integration by interacting with plant proteins and chromatin (Ward & Zambryski 2001; Zupan et al., 1996; Abu-Arish et al., 2004; Gelvin, 2012).
Other members of the *vir* regulon assemble and control the type IV secretion apparatus responsible for transporting T-DNA to the plant host. VirB4 and VirD4 are ATPases that aid the secretory motor protein VirB11. VirB11 is thought to play a role in both pili biogenesis and DNA transport (Ripoll-Rizada et al., 2013). VirB3 and VirB5 play a role in pilus formation and require VirB6 to stabilize their dimerization (Hapfelheimer et al., 2000); VirB2 and VirB6 are also channel proteins (Al-Khedery et al., 2012). VirB1, VirB2 and VirB5 play a role in cellular attachment (Figure 4) (Christie & Vogel, 2000) while VirG is a transcriptional activator (Pazour & Das, 1990).

**Figure 4**: Shows the role of several Vir proteins in T-DNA processing and transfer, adapted from Lai and Kado, 1990.
Methods

*Agrobacterium tumefaciens* C58 were grown in Mg/L overnight while vigorously shaking at 28°C. The culture was transferred to induction media (AB salts, NaPO₄, MES (pH 5.6), 0.5% glucose, 100 µM acetosyringone (Gelvin)) which turns on the *vir* regulon. Cells were then divided into six aliquots, three of which were to be exposed to Acetosyringone (50 µM in dimethylsulfoxide, pH 5.5) and three were exposed to Dimethylformamide (DMF), as a control (Manfroi et al., 2015). Cultures were grown for an hour at 30°C. Genomic DNA was isolated using the QIAGEN DNeasy Blood & Tissue Kit (Germantown, MD) and sent to the Joint Genome Institute (Walnut Creek, CA) for sequencing using Pacific Biosciences (Menlo Park, CA) Single Molecule Real Time (SMRT) sequencing. SMRT sequencing is able to detect over 20 nucleotide chemical modifications such as N6- or N4-methyladenine, 5-methylcytosine and 5-hydroxymethylcytosine by measuring reproducible changes in polymerase kinetics after incorporation of a fluorophore-bound nucleotide onto the parent strand (Blow et al., 2016). Notably, SMRT is only moderately sensitive to m5c methylation (Blow et al., 2016).

Shifts in polymerase kinetics are measured by the interpulse duration (IPD)-or the time between nucleotide addition and fluorophore cleavage by DNA Polymerase (Pirone-Davies et al., 2012). Typically a methylated base takes longer than the average 1-3 bases per second. The IPD of the strands are compared to the IPD of the control template by running a two-tailed t-test. The p-value was used to generate a Quality Value score (QV=-10log (p-value)). If the QV score ≥ 30, then there is a 99.9% confidence that the base is stably methylated (Flusberg et al., 2010).

These SMRT sequence data files (.csv) were opened in R Studio (version 1.0.136, Boston, MA), sorted by IPD ratio and bases with an IPD>30 were identified. The findings reported here were restricted to modifications found on the Ti plasmid. Nucleotide location reference was then compared to the Ti plasmid map (IMG Genome Number: 639279302) found on IMG (Integrated Microbial Genome, Joint Genome Institute, Department of Energy, Walnut Creek, CA, USA). The location of
each modified base was examined to determine if the nucleotide was located in a gene, in a promoter region or an intergenic region. Operons were examined to determine if the methylation markers had the potential to modify their expression.

Results

Methylation patterns varied between the bacteria exposed to acetosyringone and controls. Only genes that showed statistically significant results, as determined by a paired samples t-test were considered. The location of the genes that were methylated across all three trials are shown below (Table 1.) Between trials, the data varied widely: some genes were not methylated in one trial and then methylated over a dozen times in the next.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Name</th>
<th>Location of Methylation sites in genes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AS                              noAS</td>
</tr>
<tr>
<td>Atu6093</td>
<td>DnaE</td>
<td>T  111650                        A  111494</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A  111333                        G  112344</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A  111601                        G  112344</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G  112344                        A  111601</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T  112401                        C  111781</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T  112413                        A  112051</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C  112608                        A  112565</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T  112905                        A  113089</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A  113186                        T  113434</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T  113625                        T  113659</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A  113659                        C  114022</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A  114187                        A  114239</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A  114348                        C  120994</td>
</tr>
<tr>
<td>Atu6101</td>
<td>helicase, SNF2 fam</td>
<td>T  121089                        C  120994</td>
</tr>
<tr>
<td></td>
<td>Atu6127</td>
<td>Atu6136</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>traA</td>
<td>phosphate Na+ symporter</td>
</tr>
<tr>
<td>T</td>
<td>121298</td>
<td>A 156015</td>
</tr>
<tr>
<td>A</td>
<td>123703</td>
<td>A 156095</td>
</tr>
<tr>
<td>A</td>
<td>124832</td>
<td>A 156331</td>
</tr>
<tr>
<td>C</td>
<td>122592</td>
<td>T 156376</td>
</tr>
<tr>
<td>A</td>
<td>125209</td>
<td>A 159148</td>
</tr>
<tr>
<td>T</td>
<td>125227</td>
<td>A 160667</td>
</tr>
<tr>
<td>A</td>
<td>125958</td>
<td>C 155818</td>
</tr>
<tr>
<td></td>
<td>T 124547</td>
<td>A 156015</td>
</tr>
<tr>
<td>A</td>
<td>144227</td>
<td>A 156095</td>
</tr>
<tr>
<td>A</td>
<td>144658</td>
<td>A 156331</td>
</tr>
<tr>
<td>T</td>
<td>145390</td>
<td>T 156376</td>
</tr>
<tr>
<td>A</td>
<td>145762</td>
<td>A 159148</td>
</tr>
<tr>
<td>G</td>
<td>145849</td>
<td>A 160667</td>
</tr>
<tr>
<td>A</td>
<td>146013</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>147210</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144124</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144334</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144386</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145477</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144227</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144658</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>145390</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145762</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>145849</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>146013</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>147210</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144124</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144334</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144386</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145477</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144227</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144658</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>145390</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145762</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>145849</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>146013</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>147210</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144124</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144334</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144386</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145477</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144227</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144658</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>145390</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145762</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>145849</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>146013</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>147210</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144124</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144334</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144386</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145477</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144227</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144658</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>145390</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145762</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>145849</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>146013</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>147210</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>144124</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144334</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>144386</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>145477</td>
<td></td>
</tr>
</tbody>
</table>
Table 1: Locations of each methylated gene (QV > 30) for all six trials for each gene that is greater than 2 SDs. All modifications here are within a gene.
* > 2 SD; ** > 3 SD

Paired Samples t-test: p = 0.021

**Figure 5:** Shows the methylation changes of genes with the most significant changes between exposures. Protein products here are only for modifications within a gene.
Figure 6: Shows methylation changes for genes with methylated promoters that lost or gained more than one base methylation, on average, between exposures.

Of the 163 genes that have been identified on the Ti plasmid, on average only eight genes showed statistically significant methylation changes across the whole gene upon Acetosyringone exposure (Figure 5 and Table 1). Since promoter methylation was noted as being strongly linked to gene expression regulation, the genes that showed the loss or gain of more than one base methylation in their promoters are noted in Figure 6 above. Only nine genes are affected in this manner. Further, the only gene in common between the two groups is virA. The remaining genes of the Ti plasmid are methylated, though not in a statistically significant manner.
Discussion

A variety of genes found on the Ti plasmid are methylated in the promoter region (Figure 6) or across the whole gene (Figure 5 and Table 1) in response to acetosyringone treatment. The following sections describe the known functions of these genes and provide insight into possible implications of this DNA methylation.

Figure 7: The locations of the genes from Figure 5 are noted here along with the location of T-DNA. The orange lines indicate T-DNA borders.
**dnaE**

DnaE is a core catalytic subunit of the DNA Polymerase enzyme that has a role in the trans-splicing of foreign protein content *in vivo* as well as peptide-bond cleavage in laboratory experiments (Evans et al., 2000; Scott et al., 1999). Further, DnaE has been shown to be essential for chromosomal replication in *Escherichia coli* and *Staphylococcus aureus* (Inoue et al., 2001). *A. tumefaciens* C58 has an extra copy of dnaE on its Ti plasmid, a duplication not seen in other organisms. Its role in pathogenesis of C58 has not been studied, however it’s possible that it may work to maintain high copy number of the plasmid or to quickly amplify T-DNA for enhanced transfer success. DnaE was also shown to be able to perform protein cyclization by ligating peptides to proteins. (Evans et al., 2001; Scott et al., 1999). Its role in protein cyclization may ensure that DnaE has a role in the processing of *vir* proteins during pathogenesis.

**helicase; SNF2 family**

Members of the SNF2 family play roles in transcriptional activation and contain a helicase-like function, although it has also been described as a DNA-activated ATPase. Members are involved in DNA replication, recombination and repair activities, in part by facilitating and stabilizing binding of proteins and DNA (Laurent et al., 1992). The family is highly conserved across many Eukaryotes, from yeast to fruit flies (Laurent et al., 1992; Laurent et al., 1993; Pazin and Kadonaga, 1997). It is not clear why this gene is a target of epigenetic modification.

**traA**

In *A. tumefaciens* C58 TraA, stimulated by the regulatory protein TraR, is a large multi-faceted protein. Studies have shown it to behave as an endonuclease, relaxase and helicase. Further, TraA from the Ti plasmid, when acting with a suite of other Tra proteins, autorepresses two different promoters known to be involved in the transfer of the Ti plasmid. Without TraA, conjugation of the Ti plasmid
diminishes. Although these seem contradictory, TraA was shown to repress itself by binding to its promoter, acting in a negative feedback loop (Cho & Winans, 2007). TraA regulates conjugation by binding not only to its own promoter by binding to part of the origin of transfer and by blocking parts of the type-IV secretion system (Kurenbach et al., 2006). The methylation patterns observed suggest that it may become suppressed after acetosyringone exposure in order for C58 to maximize conjugation.

\[ \text{PO}_4^{3-}/\text{Na}^+ \text{ symporter} \]

Little is known about the role of these proteins, but one study in Staphylococcus bovis suggested that the control of sodium/phosphate gradients regulate the uptake of some amino acids like alanine and serine (Russell et al., 1988). Enhanced uptake of these amino acids might provide a competitive advantage in quickly changing environments (Wood, personal communication). De-methylation and the potential increase in expression of this gene may work to enhance C58 survival during its pathogenesis.

\textit{accC}

Acetyl-CoA is a critical part of a cell’s metabolism and \textit{AccC} is one of two subunits for Acetyl-CoA carboxylase. \textit{AccC} has also been shown to play a role in down-regulating biotin biosynthesis in \textit{E. coli} enabling pathogen survival (Abdel-Hamid & Cronan, 2006; Cheng et al., 2009). It is possible that methylation and subsequent repression of \textit{AccC} would promote more biotin binding to proteins; this binding may enhance or enable protein function. However, it is notable that the production of biotin would diminish since \textit{AccC} plays a direct role in this process (Marini et al., 1995). It therefore may work to limit the supply of total biotin but ensure that it is directed at essential proteins. Both contradictory possibilities may be possible and more work is needed to make clear what methylation might do.
*virA & virG*

It comes as little surprise that *virA* and *virG* had both significant promoter methylation and significant changes in methylation in response to methylation. These findings make sense in light of the following facts: the *virA* promoter is inducible by acetyosyringone and the *virG* promoter is inducible by acetyosyringone as well as starvation conditions (Winans et al., 1988). *VirA* responds to acetyosyringone by phosphorylating *VirG* which in turn binds to promoters of the *vir* regulon to activate gene expression (Melchers et al., 1989; Jin et al., 1990a; Jin et al., 1990b). Both have been shown to be necessary for tumorigenesis and the pathogenesis of *A. tumefaciens* C58 (Chang & Winans, 1992).

*virD5*

*VirD5* stabilizes *VirF*. The latter is an F-box protein that works by targeting other *vir* proteins for degradation, and is itself the target of rapid proteolysis. *VirD5* maximizes *VirF*’s lifespan and function, in turn promoting effective host-pathogen interaction (Magori & Citovsky, 2011). *VirD5* is exported into the host cell along with four other *vir* proteins. It appears to have activate transcription of T-DNA in coordination with *VirE2*. This supresses host gene expression and prevents degradation of T-DNA by blocking host ubiquitin degradation mechanisms (Wang et al., 2014). It’s possible that *VirD5* methylation would regulate to prevent protein degradation in C58 and bolster C58’s defenses against host susceptibility to invading pathogens.

*D-nopaline dehydrogenase*

Nopaline is an amino acid produced by C58-infected host tumors that is used predominantly by the infecting *A. tumefaciens* C58. The nopaline synthesis pathway includes D- nopaline dehydrogenase, an enzyme that helps produce the final product from intermediates. Production of nopaline is energetically costly for the plant (Montoya et al., 1977); the bacterium would keep this gene active throughout the virulence stage to ensure successful survival and sequestration of resources.
**phzF**

Phenazines are cyclic nitrogen-containing compounds with anti-microbial behavior. The protein product of *phzF* plays an essential role in the last steps of phenazine synthesis (Blankenfeldt et al., 2004). In other bacteria such as *Pseudomonas fluorescens*, phenazines keep other microbial species at bay, thus providing a competitive advantage. Phenazines might also aid in virulence because *phzF* often closely resembles amino acid biosynthesis genes found in plant species, possibly allowing it to augment substrate flow for nopaline (Pierson et al., 1995). Since methylation decreased in response to acetosyringone, it’s possible that this gene is upregulated to make plant amino acid synthesis more potent. Phenazine synthesis doesn’t occur in C58 and it isn’t exported to the plant so it’s unclear how this would directly aid in C58 survival.

**ssuD**

This gene allows for organisms to utilize a variety of sulfur-containing substrates as food source by first converting them into simple aldehydes when typical sulfur sources are less abundant (Eichhorn et al., 2002; Ellis 2011). This gene’s methylation decreased upon acetosyringone exposure, suggesting that energy for growth was not a priority or that cells were not sulfur-starved during virulence.

**LysR**

The LysR family encompasses many genes that work to positively regulate gene expression in prokaryotes. For *A. tumefaciens* C58 some members of the LysR family are involved in octopine biosynthesis (Schell 1993) while in other organisms these genes are involved in regulation of the entire Type-IV secretion systems. The LysR systems are controlled by quorum sensing in *E. coli* (Sperandio et al., 2002). It is unclear what effect methylation would have on this system. Since LysR is one of the largest regulatory systems in Prokaryotes, the regulation of LysR might be related to energy consumption and regulation to promote sustained virulence.
**ImuB**

In *Mycobacterium tuberculosis*, ImuB plays a role in inducing mutagenesis and damage tolerance and simultaneously acts in increasing DNA polymerase activity, and increasing Ti plasmid copy number (Warner et al., 2010). Under acetosyringone exposure, ImuB lost methylation by more than two-fold, suggesting that during virulence mutation-accrual is favored, although ImuB is regulated by a highly conserved regulatory family LexA (Erill et al., 2006). Work has not been done to uncover these how ImuB and LexA are involved (if at all) in virulence and so the effect of methylation on this system remains elusive.

**ardC**

*ardC* is an anti-restriction protein found on an array of plasmids that enable plasmid evolution and survival (Fernandez-Lopez et al., 2006). It becomes more methylated in response to acetosyringone, suggesting that more DNA restriction systems are activated. *ardC* might also play a role in protein export in *A. tumefaciens* C58’s secretion system (Cabezón et al., 2014) although more studies are need to reach substantial conclusions on this topic.

**traM**

*traM* works to regulate the Tra region-activating protein *TraR* by decreasing rates of conjugation (Fuqua et al., 1995), although *traM* remains essential for successful conjugation (Penfold et al., 1996). It promotes conjugation by binding multiple origin of transfer on DNA to allow binding of the F factor (Disqué-Kochem & Dreiseikelmann, 1997). *TraM* becomes more methylated upon acetosyringone exposure suggesting it is repressed to promote successful and controlled conjugation.

These results suggest that methylation of genes and their promoters may play a role in the pathogenesis in *A. tumefaciens* C58. A paired samples t-test was run on all genes to determine that global methylation decreased significantly after acetosyringone exposure. Only genes that lost or gained methylation more than two standard deviations with a 99.9% confidence after exposure were
considered. All results must be considered critically, however, given the divergent methylation patterns seen between treatments in this study (Table 1). The findings might tell a different story: that methylation, while regulatory, is in fact one of many ways to promote and establish initial virulence but that other mechanisms are more important for the sustenance of C58’s pathogenesis.

Surprisingly, thymine appeared to be methylated in this study (Table 1) which has never been seen previously in the literature. Other reports using SMRT to detect methylation patterns in Prokaryotes have reported novel methylation motifs. This makes the above findings, while unprecedented, possible (Pirone-Davies et al., 2015). Additionally, the location often cited for regulatory methylation in many other Prokaryotic species (GATC sites) are not the sites methylated in my findings. Instead, many new motifs have emerged as either possible candidates for regulation or simply flukes of SMRT detection mechanisms. Therefore, while these findings are interesting, they must be better understood in a larger biological context including that would include validation studies to define the methylation status of these genes in vivo gene regulation studies such as Q-RT-PCR to determine if these alterations impact gene expression. Although C58 divides very rapidly, it is worth noting that these cells were not synced. It is therefore possible that some results are maintenance methylation which is not involved in pathogenesis; although this is unlikely since there was little consistency between trials, a hallmark of regulatory methylation. The coverage scores (each trial had each base read, on average, more than 30 times) would mitigate some of these concerns although it is impossible to be thoroughly confident in the findings.

*Agrobacterium tumefaciens* C58 remains a well-studied organism for its remarkable ability to transfer and integrate part of its DNA into a plant host. The findings above suggest that methylation might be one of the ways the organism quickly responds to environmental stimuli during pathogenesis.
References


Casadesus, J., and D. Low. "Epigenetic Gene Regulation in the Bacterial World." Microbiology and


Disque-Kochem, C., and B. Dreiseikelmann. "The Cytoplasmic DNA-binding Protein TraM Binds to


Forsman, K., M. Goransson, and B. E. Uhlin. "Autoregulation and multiple DNA interactions by a


Marinus, Martin G., and Josep Casadesus. "Roles of DNA Adenine Methylation in Host-pathogen Interactions: Mismatch Repair, Transcriptional Regulation, and More." *FEBS Microbiology*


Urig, Sabine, Humaira Gowher, Andrea Hermann, Carsten Beck, Mehrnaz Fatemi, Andreas Humeny, and Albert Jeltsch. "The Escherichia coli Dam DNA Methyltransferase Modifies DNA in a


Willbanks, Amber, Molly Greenshields, Camila Tyminski, Sarah Heerboth, Karolina Lapinska,

