USING A MAMMALIAN VIRUS TO CREATE PLANTS FOR SITE-SPECIFIC TRANSGENE INSERTION

by

William J. Zabaronick

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

Biology

APPROVED:

_____

Muriel Lederman, Chair

_____ _____

Brenda S.J. Winkel **Elizabeth A. Grabau**

April 23, 2001
Blacksburg, Virginia

Keywords: Plant Transformation, *Agrobacterium tumefaciens*, Vacuum Infiltration, Adeno-associated Virus

USING A MAMMALIAN VIRUS TO CREATE PLANTS FOR SITE-SPECIFIC TRANSGENE INSERTION

by

William John Zabaronick

Committee chairwoman: Muriel Lederman, PhD.

Department of Biology

(ABSTRACT)

A novel strategy for site-specific DNA transformation of plants has been proposed and the first component of the system developed. The proposed method overcomes the limitations of current techniques by providing a specific integration site for the insertion of transgenes using features of the adeno-associated virus (AAV) life cycle. In the absence of helper virus, AAV integrates into a specific location on human chromosome 19, the AAVS1 locus. The sequence for AAV integration was introduced into the model plant *Arabidopsis thaliana* using *Agrobacterium tumefaciens*-mediated transformation. A portion of the human AAVS1 sequence, including the Rep binding site (RBS) and terminal resolution site (TRS), was cloned between T-DNA borders of the *Agrobacterium* Ti plasmid. The reporter gene, β-glucuronidase (GUS) was inserted proximal to AAVS1 in the plasmid for use in screening for the presence of T-DNA. In addition, it will serve as an indicator of the expression level expected for transgene inserted into AAVS1 by recombinant AAV. PCR amplification, dideoxy sequencing, GUS expression assays and genomic Southern blots were performed to examine putative transgenic plants for the presence of the AAVS1 sequence.

ACKNOWLEDGMENTS

It would have been impossible to produce this thesis without the help and support of the many people that have affected my life. Primarily, I would like to thank Dr. Muriel Lederman for not only providing me with an environment to work in, and a project to do, but also many of the character traits that I now possess. I would also like to thank Dr. Brenda Winkel and Dr. Beth Grabau for all their help, suggestions and input. I would like to thank former committee member Dr. Bruce Turner for all his time and effort.

I would like to also thank Dr. Mike Fisher and Dan Sisco for all their help during various aspects of this process. I would like to thank Dr. Muffazer Cicek, for listening to my off the wall ideas about how to modify procedures/experiments and Cathie Aime for reading and rereading my thesis and being there as a good friend.

I would like to thank my parents, John and Kathy Zabaronick for being who they are and for never once trying to get me to change my mind and stop pursuing this degree. Finally, I would like to say thanks to my best friends whom would be embarrassed if I listed them in my acknowledgments. Thank you for being who you are.

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

Introduction

This thesis documents the successful completion of the first part of a strategy to develop plants capable of site-specific integration and high-level expression of transgenes. This strategy exploits the natural biological process by which the mammalian parvovirus AAV-2 integrates into a specific site on human chromosome 19. I have transferred a 409 base pair portion of human chromosome 19 that is critical for AAV integration into the genome of *Arabidopsis thaliana*.

Biotechnology and Agriculture.

The goal of contemporary plant biotechnology is to introduce genes for desirable qualities into plants. The use of molecular genetics and molecular biology has opened a new realm of possibilities by allowing such manipulations. This technology has been used to engineer plants for such wide-ranging purposes such as to produce polyhydroxybutyrate, a biodegradable thermoplastic (Poirier *et al*., 1992), to make a plant sweeter by expressing the protein monellin (Peñarrubia *et al*., 1992), to produce monoclonal antibodies (Khoudi *et al*., 1999), to confer pest resistance (*Bacillus thuringiensis cry* genes, Perlak *et al*., 1990), or to confer herbicide resistance (Qing *et al*., 2000). All of these transgenic plant lines were developed using classic transformation methodology (See: **Current Transformation Methods**). This project outlines the first part of a strategy which is designed to overcome specific limitations of the classic methodologies. This novel strategy would allow for the site-specific integration of transgenes into plants by inserting the human sequence for integration of the mammalian

virus adeno-associated virus (AAV). Additionally, expression of the AAV rep gene and the recombinant AAV will be required as a vector for gene delivery.

Current Transformation Methods

There are a several methods of generating transgenic plants. The success of each of these methods depends upon the survival of the cell or tissue after the procedure and the ability to regenerate transformed plants. Biolistics (Seki *et al.*, 1999), direct injection (Holm *et al.*, 2000), use of the Ti-plasmid of *Agrobacterium tumefaciens* (Matzke and Chilton, 1981) or derivatives thereof, and electroporation (Arencibia *et al.*, 1998) are the main types of transformation techniques. Biolistics, the acceleration of microspheres of gold or tungsten coated with DNA by gas pressure or by gunpowder, can deliver DNA into nearly any tissue type. Direct injection of DNA involves piercing the nucleus of a cell with a microscopic needle that contains a solution of the transgene(s). Ti-plasmid transformation is the most 'natural' of any of the transformation methods because it is based upon the normal infection of a plant by a bacterium. Unfortunately, *A. tumefaciens* only infects dicotyledonous plants which limits the range of agronomically important plants that can be transformed by this method. Electroporation uses an electrical discharge or pulse to open transient holes in the cellular membrane that allow DNA to enter the cell.

Current *Agrobacterium tumefaciens* Transformation Technology.

Because of its simplicity and consistency, the tumor-inducing plasmid of the bacterium *A. tumefaciens*, the Ti-plasmid, has been used as the standard to transfer DNA

into plants (Willmitzer, et al., 1983). The *A. tumefaciens* Ti-plasmid (shown in Figure 1) found in nature is the causative agent of crown gall tumors in dicotyledonous plants (Zupan and Zambryski, 1995). A native Ti-plasmid contains two border sequences, an origin of replication, and a variety of genes involved in the infection process. The coding regions are the *vir* loci as well as genes responsible for auxin, cytokinin, and opine production and catabolism. The *vir* region contains genes that are responsible for the transfer of the T-DNA (Zambryski, 1992).

Five virulence operons (*virA, virB, virD, virE, and virG)* located in a 30-kilobase pair region outside the T-DNA comprise the *vir* region (Zambryski, 1992). VirA is a constitutively expressed chemoreceptor that detects the small phenolic compounds associated with plant wounding. *VirB*-encoded proteins are responsible for creating transmembrane pores between the bacterial cell and the plant cell. *VirD1* encodes a single strand-specific endonuclease that binds the border sequences and introduces site-specific nicks in one strand of the double-stranded DNA (dsDNA) Ti-plasmid; the virD2 protein binds to the 5' end of the newly formed single-stranded DNA (ssDNA) molecule termed the T-strand. *VirE* encodes a single stranded DNA binding protein that encases the T-strand. *VirG* induces the expression of the *vir* genes upon activation by *virA* by phosphorylation. The T-DNA region contains left and right border sequences that are recognized by products of genes in the *vir* locus. There has been no positional specificity observed with regard to the insertion of the T-DNA fragment (Ten Hoopen, *et al*., 1999). Auxin and cytokinin are plant growth regulators; tumor growth results from the synthesis of excess auxin and cytokinin in the infected plant tissue from genes in the T-DNA region that has been integrated into the host plant genome. The natural T-DNA also

contains genes responsible for the synthesis of a class of conjugated amino acids called opines (Van Montagu *et al*., 1980). These opines are metabolized by *A. tumefaciens* via opine catabolism genes contained on the Ti plasmid (Van Montagu *et al*., 1980). Essentially, the bacterium transforms the plant tissue into an undifferentiated cell mass that produces opines that the bacterium can then use as a food source.

Researchers have used the fact that the T-DNA is transferred from the Ti-plasmid to the plant genome to introduce novel genes into plants (Willmitzer *et al*., 1983). Currently there are two major plant transformation vector systems that utilize a modified Ti-plasmid: the binary vector system and the cointegrate vector system. In the binary vector system, the genes of interest, along with selectable marker(s) and reporter gene(s) are placed within the T-DNA borders on one plasmid, which is missing the *vir* locus. The second plasmid, harbored within *A. tumefaciens* contains the entire *vir* locus, but not the T-DNA itself. The binary vector system will be used in the work reported here.

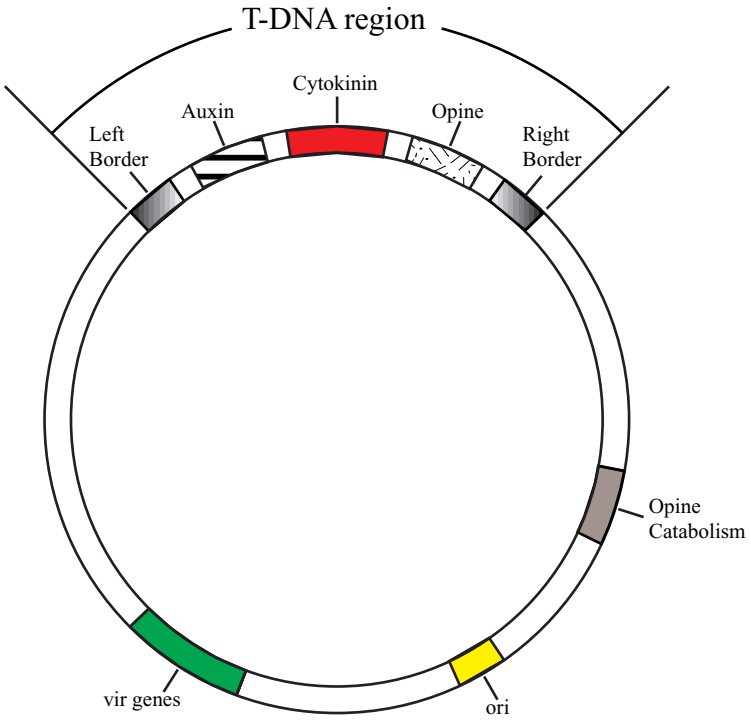

Figure 1. The Ti-Plasmid of *Agrobacterium tumefaciens*, adapted from Glick and Pasternak, 1994.

Problems with Current Technologies/Methods

Despite the widespread use of plant transformation technology, there are important limitations to the current methods. Insertional mutagenesis and position effects are the two primary concerns (Gelvin, 1998). Insertional mutagenesis occurs when the T-DNA disrupts the normal expression of a host gene. This may occur by disrupting the coding region or promoter of the gene or altering the gene in such a way as to disrupt its normal regulation. Position effects are a class of events that affect the transgene and are based on its position in the host genome. These events are characterized by the lack of a protein product from a functional gene and are due to chromatin condensation at the insertion site or through hypermethylation of the gene, promoter or terminator (Meyer, 1995; Stam *et al.*, 1998).

Another concern is the lengthy screening process itself. Researchers must screen large numbers of putative transgenic plants to obtain plants that not only are transgenic, but also express the transgene at an appropriate level (Gelvin, 1998). While this has been a means to generate a number of transgenic plants, it is far from efficient (for example, 1% to 9% of screened seeds (after selection) from biolistic transformation are transgenic) (Finer *et al*., 1999). Hansen and Wright (1999) outlined several criteria to be met by an efficient, effective plant transformation vector. They state that the vector needs to be able to target tissues that are competent for propagation or regeneration after the transformation procedure; that the vector needs to deliver DNA efficiently; it has to include agents that enable selection for transgenic tissues; that the vector has to be part of a technique that allows the recovery of transgenic plants at a reasonable frequency and that the transformation needs to be a simple, efficient, reproducible, genotype-

independent and cost-effective process, and that the process needs to have a tight timeframe in culture to avoid somaclonal variation and possible sterility.

While these are good standards, they omit one of the most critical parts of generating a transgenic organism, the specificity by which the transgene is integrated into the host genome. Problems with transgene expression begin with the lack of control over the region where the transgene is to be inserted. If such control were possible, it would provide a better route for ensuring higher and more consistent levels of expression.

Targeted Integration Methods

The transfer of the T-DNA from the *A. tumefaciens* Ti-plasmid to the plant genome is an elegant and natural process. There are problems, specifically the random nature of the integration and the possibility of multiple insertions. There have been few solutions found for increasing the site-specificity of the plant transformation process. Two methods of note are the Cre/*lox* system (Albert *et al*., 1995) from bacteria and a chloroplast transformation method (Daniell *et al*., 1998). The Cre/*lox* system uses a dual plasmid method of integration. One plasmid that carries a *lox* (locus of crossing over) sequence is randomly integrated into the genome. The other plasmid contains a cassette of the Cre recombinase, the transgene and another *lox* site. The Cre recombinase, a 38kDa protein, directs integration into *lox* sites randomly integrated into the genome (Sauer and Henderson, 1990). Often there are multiple insertion sites and when recombination occurs, the dynamic may favor excision versus integration (Odell and Russell, 1994).

Chloroplasts are maternally inherited prokaryote-like organelles that contain their own chromosome and replicate autonomously in the cell. Daniell *et al*. (1998) used the universal chloroplast-transforming vector, pSBL-RD-EPSPS, to introduce glyphosphate resistance into tobacco plants via homologous recombination within the chloroplast genome. The vector contains highly conserved sequences (*trnA* and *trnI* genes) as flanking sequences for homologous recombination. Using this method to transfer a plasmid containing markers for spectinomycin and glyphosphate resistance, results from Southern analysis indicated that the chloroplast genome recombined with the vector in a site-specific manner via targeted disruption of the *trnA* and *trnI* genes. All seeds collected were 100% resistant to spectinomycin and all mature plants were resistant to 5 mM glyphosphate, while 0.5 mM glyphosphate was lethal to untransformed control plants (Daniell *et al*., 1998).

Another strategy for directed integration is the use of homologous recombination. Two instances of Ti-plasmid assisted homologous recombination in *Arabidopsis* have been reported. The neomycin phosphotransferase (NPTII) gene was used to disrupt the TGA3 locus in the first instance (Maio and Lam, 1995) and the AGL5 gene in the second (Kempin *et*. *al*, 2000). Both methods used the same plasmid and methodology, which was to disrupt the coding regions of these genes (TGA3 and AGL5) with NPTII, resulting in NPTII flanked by wild-type genomic DNA. This cassette, residing between the T-DNA borders inside pBI101, was introduced via *Agrobacterium* mediated transformation. The TGA3 locus disruption produced 2 transgenic plants out of 2580 transgenic plants total (less than 0.1%), and the AGL5 disruption yielded 1 homologous recombination event out of 750 transgenic plants generated (0.1%).

Parvovirus and Site-Specific Integration

Currently, certain viruses are being exploited as gene therapy vectors because these viruses integrate into the host genome. Most integrating viruses do so in a random manner. There are several reasons why this can be detrimental to the survival of the host cell. One consequence of random insertion is the possibility of disrupting the expression of a gene essential for the viability of the cell. A randomly integrated virus may be inserted into any portion of the gene (the promoter, the coding region, terminator or enhancers) in such a way as to disrupt normal expression. This may increase, decrease or totally eliminate the gene product. Another consequence of random integration is that the integration event may place an essential host gene under the control of a viral promoter. As an example, the long terminal repeat (LTR) of a retrovirus contains an extremely strong promoter and enhancer. These can activate or increase the expression of any gene that lies downstream or upstream from its integration site. This is called promoter or enhancer insertion (Lodish *et al*., 1999). Using viruses as transformation agents poses the same problems as any of the standard techniques described (See: **Current Transformation Methods**). However, there is a virus that integrates site-specifically into the genome of its host, the dependent parvovirus adeno-associated virus-type 2 (AAV2).

Parvoviruses, with an average capsid diameter of 22 nm, are some of the smallest animal DNA viruses known. There are two types of parvoviruses, autonomous and dependent. Autonomous parvoviruses require passage through the S phase of the cell

cycle for a productive infection. The dependent parvoviruses (dependoviruses) need functions provided by a helper virus to complete the cycle of infection.

A number of diseases have been associated with the autonomous parvoviruses. The diseases caused by bovine, porcine, and rodent parvoviruses usually result in fetal abortion and teratogenesis. Canine parvovirus and feline panleukopenia virus cause death by myocarditis and enteritis, respectively, in young animals. Human parvovirus, B19 causes erythema infectiosum, arthralgia, and arthritis. In immunocompromised patients, B19 can result in chronic infections. In individuals with sickle cell disease, B19 may cause aplastic crisis. Infection during pregnancy can cause fetal hydrops and it has been implicated in fetal death (The Wadsworth Center, 2000). On the other hand, dependoviruses have not been shown to cause any type of disease due to their life cycle (for a review see Berns and Linden, 1995).

Genome Organization of AAV-2

Members of the family Parvoviradae have a genome that is composed of a single DNA strand between 4.5 and 6 kilobase-pairs in length. Parvovirus genomes contain two large open reading frames (ORFs). The left ORF encodes the nonstructural or replication proteins and the right ORF encodes the structural or capsid proteins. The viral proteins are expressed from this small genome through alternative splicing of the transcripts and post-translational modifications of the proteins. The ends, or termini, of most parvoviruses are self-complementary and form regions of highly base-paired secondary structure. Adeno-associated virus-type 2 (AAV-2) is the best studied of the dependent parvoviruses. It received its name when it was found as a virus-like particle in an

adenovirus preparation. Srivastava *et al*. (1983) determined the sequence of the entire 4680 bp AAV-2 genome. The termini of AAV-2, like all parvoviruses, have a high degree of secondary structure. As shown in Figure 2, the terminal 145 nucleotides at both ends of the AAV-2 genome are paired into a T-shaped inverted terminal repeat (ITR) (Berns, 1996).

Proteins Encoded by AAV-2

The AAV-2 genome encodes four nonstructural proteins, Rep78, Rep68, Rep52 and Rep40 (Srivastava *et al*., 1983; Trempe *et al*., 1987), as illustrated in Figure 3. The Rep78 and Rep68 proteins are transcribed from a promoter located at map unit 5 (p5) and are required for the replication of the viral genome. The Rep52 and Rep40 proteins are transcribed from a second promoter located at map unit 19 (p19), and are necessary for the production of infectious viral progeny (Trempe *et al*., 1987). An mRNA produced by splicing between nt 1907 and nt 2227 is the smaller of each set of transcripts whereas the larger mRNA transcript is unspliced. All the AAV-2 transcripts are co-terminal at a polyadenylation signal near the right end of the genome beginning at nucleotide 4322 (Srivastava *et al*., 1983).

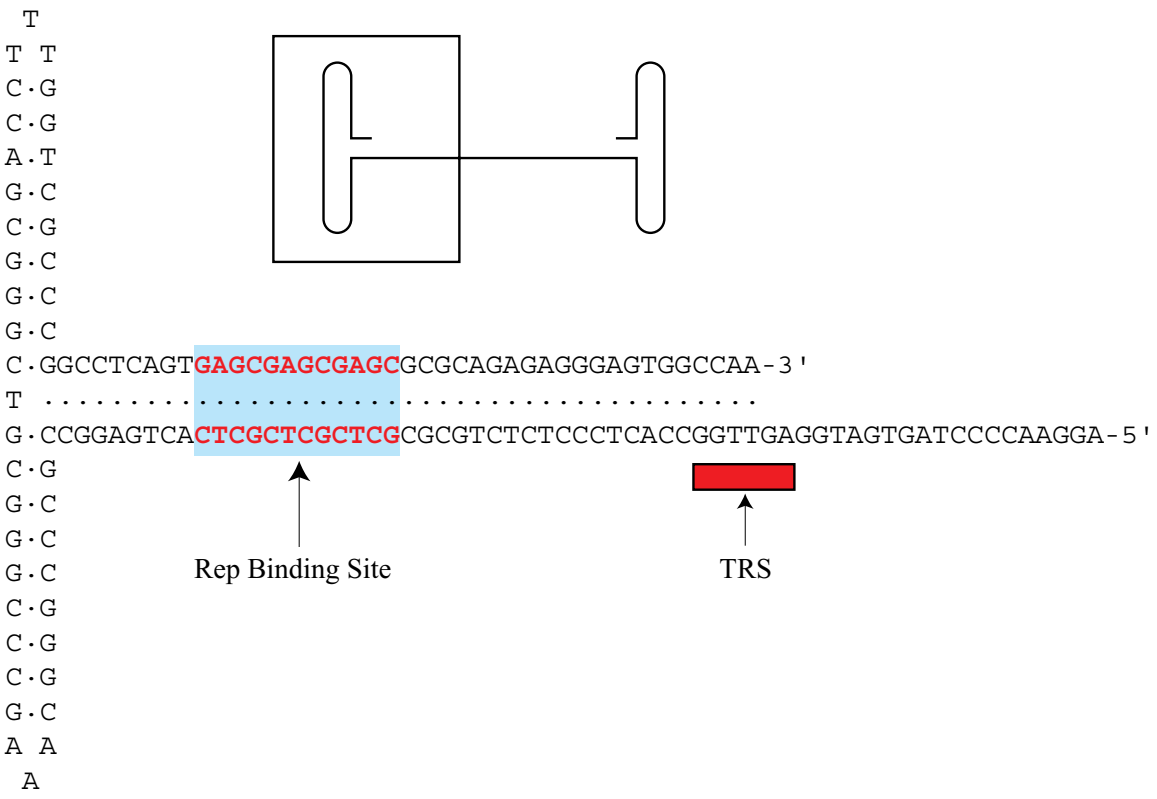

Figure 2. An inverted terminal repeat (ITR) of AAV.
The ITR contains the Rep binding site, the region where Rep78 binds the DNA, and the region where Rep78 nicks the DNA prior to the integration, the terminal resolution site (TRS). Region in box enlarged to show sequence of 3' (left end) ITR of AAV.

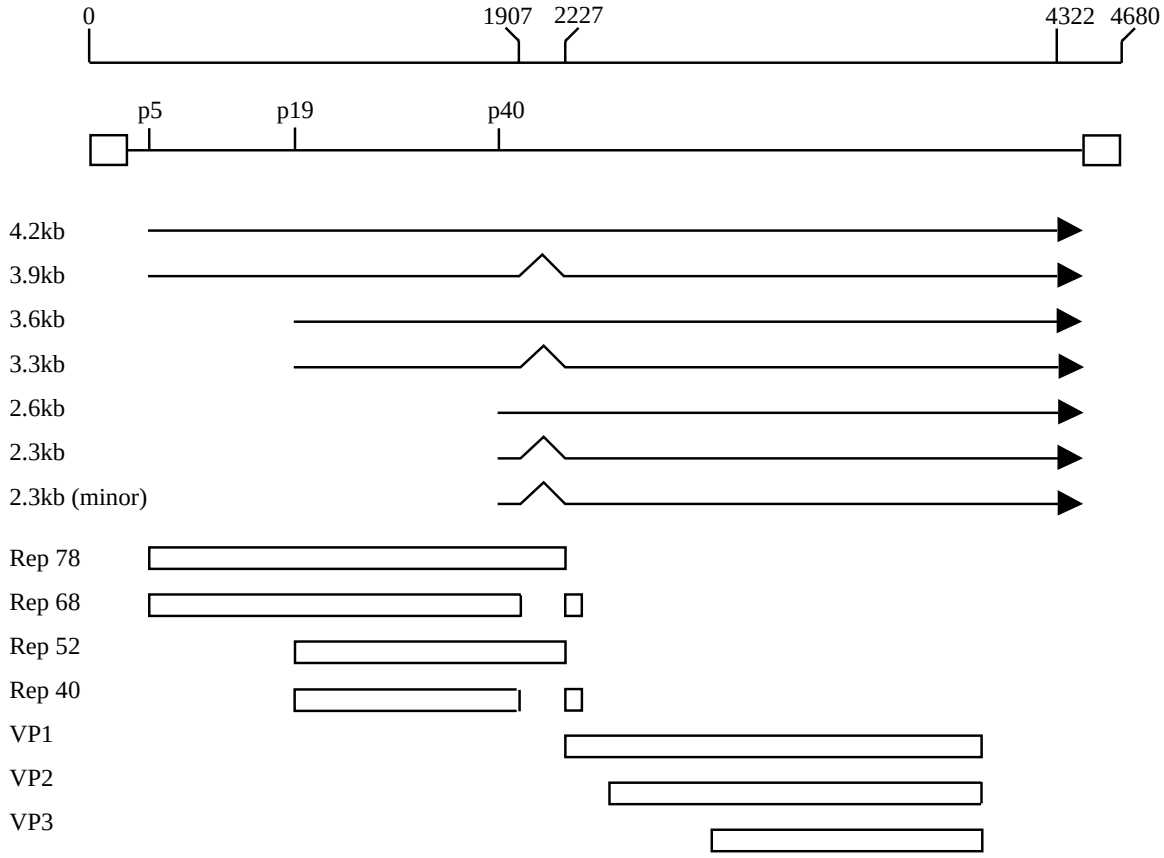

Figure 3. AAV-2 transcription map (modified from Redmann *et al*., 1989). Viral mRNA species are shown as arrows below the genome. The caret indicates the intron and the arrowheads indicate the common poly-adenylation signal. The size of each mRNA transcript is shown on the left in kilobases. The intron of the minor 2.3kb mRNA molecule is 27 nucleotides shorter than the major 2.3 kb intron. Viral proteins are diagrammed below the mRNA molecules as open boxes, with the name of each protein indicated on the left. The uppermost line identifies the nucleotide locations of the splicing event (nt. 1097 to nt. 2227) and the polyadenylation signal (nt. 4322).

The Rep78/68 proteins contain a site-specific, strand-specific DNase activity that nicks the genome at nucleotide 124. These proteins also have an ATP-dependent helicase activity that enables unwinding of the terminal hairpins. The Rep52/40 proteins do not contain any DNA binding domains or properties (Im and Muzyczka, 1990), and therefore may not play a role in AAV-2 DNA replication. Studies show that Rep52 or Rep40 is required for viral ssDNA progeny accumulation (Chejanovsky and Carter, unpublished data cited in Redmann *et al*., 1989). Genetic analyses have shown that when capsid protein synthesis is inhibited, single stranded viral progeny DNA do not accumulate (Labow *et al*., 1986; Tratschin *et al*., 1984). These data suggest that Rep52/40 may be involved in shuttling the ssDNA genomes to an assembling capsid. In addition to performing the replication functions, the Rep proteins positively and negatively regulate gene expression (Labow *et al*., 1986).

The capsid is assembled in the nucleus. All the capsid proteins of AAV-2 [VP1 (90 kDa), VP2 (72 kDa), and VP3 (60 kDa)] are transcribed from the p40 promoter (p40). There are three mRNA species that are generated from p40 but only 2 are responsible for the production of the capsid proteins. All mRNA species for the capsid genes (VP1, VP2 and VP3) have a 5' end that begins at nt 1852 and a 3' end at nt 4322. Protein VP1 (90kDa) is translated from a 2.3 kb mRNA that has a splice site from nt 1907 to nt 2200. The initiation codon for VP1 is an ATG that resides at nt 2203. Proteins VP2 (72kDa) and VP3 (60kDa) are translated from a different 2.3 kb mRNA that has a splice site from nt 1907 to nt 2227. The initiation codon for VP2 is a unique initiation codon ACG at nt 2615. The initiation codon for VP3 is an ATG that is located at nt 2810 (Becerra *et al*., 1988). The ratio of VP1:VP2:VP3 has been found to be 1:1:20; each virion is composed

of 3-4 molecules of VP1 and VP2, and 60 molecules of VP3 (Rose *et al*., 1971). The viral capsid is self-assembling, with protein VP2 specifically required for self-assembly, and VP1 and VP3 required for infectivity (Ruffing *et al*., 1992).

DNA strands are designated either plus or minus with respect to mRNA. The nucleic acid strand that has the same sequence as the mRNA is designated plus, and the complementary strand is designated minus. An AAV-2 infection generates virions that each contain a single strand of one type of DNA (plus or minus) packaged into a viral capsid (Berns and Adler, 1972).

The Cycle of Infection of AAV

Actively dividing cells are required for autonomous parvovirus replication. During a productive infection, the virus is adsorbed onto the cellular membrane of a host cell and then transported to the nucleus. During the S-phase of the cell cycle, the cellular machinery responsible for DNA replication and protein synthesis is sequestered for the production of viral progeny. The cell will continue to produce virus particles until it eventually lyses and dies, releasing the progeny.

The dependoviruses have a different method of viral reproduction. In the presence of a helper virus, a dependovirus infection will proceed along a pathway similar to the autonomous parvoviral relication pathway. In the absence of a helper virus, the dependovirus infection will proceed along another pathway that leads to the integration of the viral genome into the host genome.

The role of the helper virus for AAV-2 replication is becoming more defined. Ward *et al*. (1998) showed that Adenovirus (Ad) encodes a DNA-binding protein (Ad-

DBP) that positively affects the processivity of replicating AAV DNA, presumably by reducing template strand switching, leading to increased full length AAV genome synthesis. With Herpes Simplex Virus (HSV) as the helper, AAV-2 seems to use the HSV enzymes as opposed to using the cellular equivalents (Mishra and Rose, 1990; Weindler and Heilbronn, 1991). In the absence of the helper virus, the dependovirus will integrate into the genome of the host cell's DNA.

There is no clear-cut distinction between the autonomous state and the dependent state of dependant parvoviruses. It is known that genotoxic agents (chemical carcinogens, UV irradiation, heat shock and metabolic inhibitors) will allow for AAV replication in the absence of helper virus (Yakobson *et al*., 1987, Yakinoglu *et al*., 1988). AAV-2 replication is also highly dependent upon events that occur during the S phase of the cell cycle; host cells must pass through S-phase for AAV-2 to complete its infectious cycle. After hydroxyurea treatment, reversal of polyamine depletion or physical mitotic detachment, a small percentage of cells will be permissive for AAV-2 replication. Each of these treatments arrests the cell at the S-phase checkpoint. When removed, the infected cell then proceeds through S-phase, replicating its own DNA, as well as allowing AAV-2 to replicate in the absence of a helper virus.

Figure 4. Alignment of the Rep binding site (in blue) and the terminal resolution site (in red) of the inverted terminal repeat (ITR) of the AAV genome and the integration site in human chromosome 19 (AAVS1).

AAV Integration

The site-specific insertion site of AAV-2 is found on the q arm of human chromosome 19, at 19q13.3-qter (Kotin *et al*., 1990) and has been termed AAVS1 (Giraud *et al*., 1994). The most critical portion of this integration sequence is the 510 base pairs at the 5' end as determined by deletional analysis (Giruad *et al*., 1995). Within this region there are signal sequences, the dodecamer (GCTC)3, which has been identified as a binding site for Rep78/68 (Weitzman et al., 1994), and the terminal resolution site, the site that is nicked by the Rep 78/68 proteins (Berns and Linden,1995). These sequences (shown in Figure 4) are arranged in the same orientation in the integration sequence and in the viral genome (Berns and Linden, 1995). The viral genomes are inserted into the AAVS1 region in multiple copies, in a head to tail configuration.

Applications of AAV

Recent studies have shown the efficacy of using AAV-2 as a viral vector for gene therapy. Preliminary research has shown that recombinant AAV-2 can be used to transfer DNA into mammalian cell culture (including primary skin fibroblasts and hematopoietic cells) using reporter genes under the control of the p40 promoter in place of the capsid genes (Hermonat and Muzyczka, 1984; Tratschin *et al*., 1985; and Lebkowski *et al*., 1988; Flotte *et al*., 1992). Researchers have successfully transferred DNA into human cells in culture, restoring their aberrant phenotypes to normal by introducing functional genes into the cells (NAPH oxidase in Thrasher *et al*., 1995; IL-2 in Vieweg *et al*., 1995; CFTR in Zhang *et al*., 1998). AAV has also been reengineered to reverse or treat human

diseases in tissue culture to set the stage for applications in human patients (human apolipoprotein AI and lecithin-cholectrol acteyltransferase in Fan *et al*., 1998; NADPH oxidase in Li and Dinauer, 1998).

Research Justification

The development of transgenic plants using current technology is time and labor intensive. Researchers must screen hundreds of plants to determine if the transformation was successful. Representative transformation frequencies for plants that make it through screening are Arabidopsis: 0.48% (Clough and Bent, 2000), Tobacco: 90%-100% (personal communication; Fabricio Medina-Bolivar), Wheat: 10-20% (Wheat Transformation Program, MSU, 2000). However, transgenic plants containing the transgene often exhibit variable levels of expression. The data suggest that the problem has been due to either the transgenes causing gene silencing (Assaad *et al*., 1993), gene insertion into a region inaccessible to the transcriptional machinery, or hypermethylation (Meyer, 1995). In each of these cases the result is lower than desired expression. It is also important to have a predictable or stable level of expression for some experiments, such as comparing expression levels from different constructs as in the case of a promoter deletion series. Our system has been designed to overcome these problems.

This proposed system is tripartite, with the following requirements: 1) the AAVS1 integration site to coordinate the integration event from the genome side, 2) the transgene contained within the ITRs of AAV to coordinate the integration from the vector side and 3) Rep78 to catalyze the integration event using the AAVS1 sequence and the ITRs. Our approach was to utilize the AAV integration site, AAVS1, to provide a locus for site-

specific integration. The reporter gene β-Glucuronidase (GUS) was included in the AAVS1 vector. By randomly placing the GUS/AAVS1 cassette into the plant genome, I introduced both the site specificity of the AAV integration event and a reporter gene. GUS expression will give us an indication of how accessible the region is to the transcription machinery. This should give an indication of the relative expression level that might be expected for a transgene subsequently inserted into the AAVS1 region by AAV or an AAV based vector.

Materials and Methods

Reagents

Molecular biology reagents and biochemicals were purchased from Sigma (St. Louis, MO), Fisher (Pittsburgh, PA), Roche Molecular Biochemicals (Indianapolis, IN), and Promega (Madison, WI), except where noted. Primers 3Eco and 4Eco were obtained from Integrated DNA Technologies (Coralville, IA). Primer3 and Primer4 were obtained from Ransom Hill Biosciences (Ramona, CA). Reagents for bacterial media were purchased from Fisher or Sigma, except where noted. Reagents for seed germination media were purchased from Fisher, Sigma and Agri-Bio (North Miami, FL).

Definition of the AAVS1 CORE Sequence.

A region of 394 base pairs (bp) containing the terminal resolution site (TRS) and the Rep-binding site (RBS) of the AAVS1 sequence was chosen as the CORE sequence. Plasmid pRVK (a gift of Dr. R. M. Linden) contains 3500 bases at the 5' end of the AAVS1 sequence, including the 394 bp CORE sequence. pRVK was derived from an 8.2 kilobase (kb) λ EMBL human library clone and is contained within pBluescript(KS+) (Kotin *et al*., 1992).

A. Construction of Plasmids Containing the CORE Sequence.

A1. Production of Plasmids.

pBI121 (Clontech, Palo Alto, CA) is a plasmid containing the left and right border sequences of the *Agrobacterium tumefaciens* Ti-plasmid in the binary vector pBIN19 (Bevan, 1984). Contained within the left and right border sequences are the neomycin phosphotransferase II (NPT II) gene and the β-glucuronidase (GUS) gene under the control of the Cauliflower Mosaic Virus (CaMV) 35S promoter.

To introduce the CORE sequence into the EcoRI site of pBI121 to produce the plasmid pCORE, it was necessary to amplify the CORE sequence to produce a 409 bp PCR product that contains terminal restriction sites for EcoRI (shown in bold). This was done using primers 3Eco (5'-AC**GAATTC**GTACTTGCTAGTATGCCGTGGG-3') and 4Eco (5'-AC**GAATTC**GTCATCCTCTCCGGACATCG-3') in a 50 µl reaction volume using MasterAmp Mix G (Epicentre, Madison, WI). The cycling conditions are outlined in Table 1.

Final primer and template concentrations were 0.6 µM and 2 ng/µl, respectively. The 409 bp PCR product was purified using the QIAQuick purification kit (Qiagen, Santa Clarita, CA) and then digested with EcoRI (Promega). The resulting fragment was ligated into pBI121 that had been digested with EcoRI and dephosphorylated with calf intestinal alkaline phosphatase (Roche Biochemicals). A molar ratio of 13:1 (vector:insert) was used in a 20 µl reaction containing 1 unit of T4 DNA ligase (Gibco-BRL) for 12-14 hours (overnight) at 16°C.

Temp	Time	Num. Cycles
94°C	45 sec	30
60°C	45 sec	
72°C	90 sec	
72°C	10 min	1

Table 1. PCR cycling conditions for amplifying the CORE sequence from pRVK.

A2. Preparation of Competent cells.

E. coli DH10B cells (Gibco-BRL) were prepared for electroporation by the following procedure (adapted from Dower, 1988). A single, fresh colony of cells was harvested from a plate of LB agar (Luria-Bertani (LB) Agar: 10 g/l Bacto-Tryptone, 5 g/l Bacto-Yeast extract, 10 g/l NaCl, 15 g/l Bacto-Agar, pH 7.0), and grown in 5 ml LB broth overnight at 37°C. This culture was then used to inoculate 500 ml of LB broth and allowed to grow to an OD_{600} of 0.6. The cells were collected and centrifuged in two 250 ml centrifuge bottles at 5000 × *g* for 15 min at 4°C in a JA-14 rotor (Beckman Coulter, Fullerton, CA) and washed twice in 1 volume (250 ml) ice-cold 10% glycerol. After the second centrifugation, the cells were resuspended with as little agitation as possible in a total volume of 250 ml ice-cold 10% glycerol, combined into one bottle and centrifuged again. The supernatant was decanted and the cells were resuspended with as little agitation as possible in ice-cold 10% glycerol at 1% of the original culture volume (5 ml). This suspension was divided into 200 µl aliquots in screw top tubes prechilled in a –70°C, 70% ethanol solution and stored at –70°C until needed.

A. tumefaciens was prepared for heat-shock transformation by culturing a single colony of bacteria in 5 ml of 2xYT (16 g/l Bacto-Tryptone, 10 g/l Bacto-Yeast Extract, 5

g/l NaCl, pH 7.0) supplemented with 15 µg/ml Gentamycin (Sigma) for 2 days at 28°C. 1 ml of the 5 ml bacterial culture was combined with 49 ml of 2xYT + 15 µg/ml Gentamycin in a 500 ml flask and incubated at 28°C with rotation at 300 rpm until the OD_{600} reached 0.5 (about 4 hours). After the reaching the appropriate OD_{600}, the flask was incubated on ice for 10 min. The 50 ml culture was then centrifuged at 1000 × g at 4°C for 10 min. The supernatant was removed and the cells gently resuspended in 50 ml ice cold 10 mM $CaCl_2$ (Sigma), and incubated on ice for 10 min. Following this incubation, the cells were centrifuged as above and resuspended in 2 ml 10 mM $CaCl_2$ + 15 % Glycerol (Sigma). The cells were then aliquoted in 200 µl fractions, frozen in liquid nitrogen and stored at –70°C.

A3. Transformation of *E. coli* and *A. tumefaciens*.

Following the procedure of Dower (1988), electroporation of *E. coli* cells was performed in an ice-cold 0.2 cm cuvette (Bio-Rad, Hercules, CA) at 2.5 kV, 25 µF, 200 ohms. 40 µl of competent cells were added to 1-5 µl of ligation reaction. Typical pulse times were 4.5-4.7 msec. The cells were then incubated in one ml SOC broth (20 g/l Bacto-Tryptone, 5 g/l Bacto-Yeast Extract, 0.5 g/l NaCl, 2.5 mM KCl, and 20 mM glucose) at 37°C for 45 min to 1 h. 100 µl of the electroporation mixture was spread on plates of 2xYT medium (2xYT broth, 15 g/l Bacto-Agar, pH 7.0, 30 µg/ml kanamycin).

A. tumefaciens was transformed by incubating the cells at 37°C for 5 minutes with 1 µl of ligation reaction. 1 ml of 2xYT media was added to the transformation mixture after the 5 min incubation, the minute was incubated at 28°C for 2 h with moderate shaking (150 rpm). The mixture was centrifuged to collect the cells and the pellet was

resuspended in 100 µl 2xYT media and plated on a 2xYT agar plate supplemented with 30 µg/ml kanamycin and incubated at 28°C overnight.

A4. Growth and Extraction of Plasmid DNA.

Cells containing pCORE plasmid were grown on 2xYT agar medium and in 2xYT liquid cultures at 37°C for *E. coli* and 28°C for *A. tumefaciens*, both supplemented with kanamycin at 30 µg/ml. The procedure of Birnboim and Doly (1979) was used for all plasmid extractions from both *A. tumefaciens* and *E. coli*. The only modification was that the cellular debris was pelleted at 50,000 × *g* in a TLA 100.3 rotor (Beckman Coulter, Palo Alto, CA) at 4°C, rather than at 16,000 × *g* in a microcentrifuge at room temperature.

A5. Analysis of Potential Bacterial Clones.

Recombinant clones were identified by EcoRI digestion in a 20 µl reaction volume using 5 units of EcoRI and 0.5 µg plasmid DNA. The reactions were incubated at 37°C overnight and the DNA fractionated on a 0.8% agarose (Gibco-BRL) TAE (0.04 M Tris-acetate, 0.001 M EDTA) gel containing ethidium bromide at a concentration of 0.3 µg/ml with a voltage of 5-10 V/cm. The stain was not able to detect the excised CORE fragment without overloading the gel, so diagnostic PCR was performed. The reaction was performed, as described in section 1 above, using Primer3 (5'-GTACTTGCTAGTATGCCGTGGG-3') and Primer4 (5'-GTCATCCTCTCCGGACATCG-3'). The product was fractionated in a 1% TAE agarose gel as above with an expected product size of 394 bp. This product is fifteen bp

less than the product using primers 3Eco and 4Eco because primer 3 and 4 do not contain the EcoRI restriction sites. The locations of the products on the AAVS1 sequence are shown in Figure 5. After the PCR identified positive clones, a Southern blot (as described in Materials and Methods, A7) in EcoRI digested pCORE was performed to visualize the fragment excised by the digestion.

A6. Generation of Digoxigenin labeled probes.

The 394 bp probe for the CORE sequence was generated by PCR using primer3 and primer4 as described in section A1, with a DIG-11-dUTP (Roche Biochemicals) to dTTP (Roche Biochemicals) ratio of 1:1 in the 50 µl PCR reaction. Alternatively, the 33 bp oligonucleotide, AAVS1 core sequence (ACS, shown in Figure 6) (5'-GGCGGTTGGGGCTCGGCGCTCGCTCGCTCGCTG-3', Integrated DNA Technologies, Coralville, IA) was labeled at the 3' end using the Genius System (Roche Biochemicals). A 20 µl reaction containing 100 pmoles of the oligonucleotide, 5 mM $CoCl_2$, 0.05 mM DIG-dUTP, 0.5 mM dATP and 2.5 units of terminal transferase in 1× reaction buffer was incubated at 37ºC for 15 min. After the incubation, the reaction was placed on ice and terminated with 1 µl of 200 mM EDTA, pH 8.0.

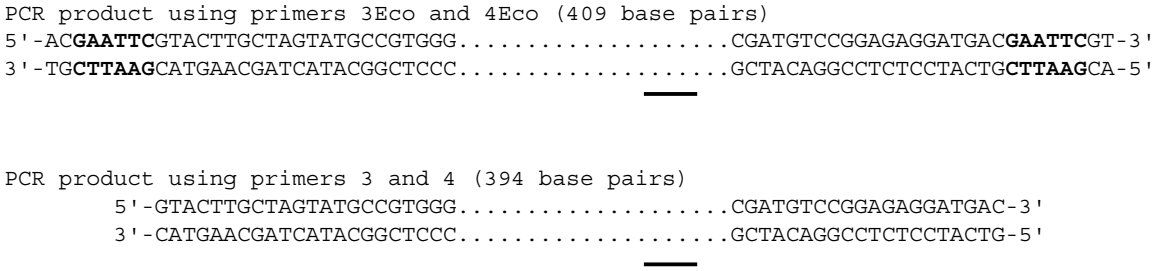

Figure 5. Diagram of primer locations within the pCORE plasmid. The Rep binding site and the terminal resolution site are contained within both PCR products, towards the right hand end of each product as indicated by the horizontal line. EcoRI restriction sites are shown in bold.

```
AAVS1    - 5'-TGGCGGCGGTTGGGGCTCGGCGCTCGCTCGCTCGCTGGGCGGGCGGGCGG-3'
ACS oligo- 5'-    GGCGGTTGGGGCTCGGCGCTCGCTCGCTCGCTG-3'
```

Figure 6. Alignment of an oligonucleotide probe with the AAVS1 sequence.

A7. Southern Analysis

The detection of the CORE sequence in both pCORE and *Arabidopsis thaliana* was performed using a Southern blotting procedure adapted from Church and Gilbert (1984). DNA samples were fractionated on TAE agarose gels (1-2% for PCR reactions, 0.75% for genomic DNA). Following electrophoresis, the gels were placed in a glass tray containing depurination solution (0.25 M HCl) on an orbital shaker at 75 rpm for 10 min followed by two rinses with deionized water. The gel was placed in denaturation solution (0.5 M NaOH, 1.5 M NaCl), shaken slowly for 20 min, and then rinsed twice with deionized water. The gel was then equilibrated twice for 20 min in neutralization solution (3 M NaCl, 0.5 M Tris, pH 7) on the shaker and rinsed with deionized water.

The DNA was transferred from the agarose to an uncharged nylon membrane (0.2 micron, Biotrans, ICN, Irvine, CA) in 25 mM $NaHPO_4$ buffer, pH 6.5 (prepared fresh from a stock solution of 1 M $NaHPO_4$) via capillary action. The DNA was transferred overnight (12-14 h). The DNA was covalently attached to the membrane in an UV-crosslinker (Fisher Scientific, FB-UVXL-1000).

The membrane was prehybridized in Church buffer (1% BSA fraction V (Sigma), 1 mM EDTA pH 8.0, 0.5 M $NaHPO_4$ pH 7.2, 7% SDS) at a ratio of 20 ml solution per 100 cm^2 membrane for l h at 60°C. The DIG-labeled probe was denatured by microwaving in a microcentrifuge tube for 1.5 min and quickly cooling on ice. The membrane was hybridized with 5 ng probe/ml (using the same volume to surface ratio as above) at 60°C for 12-14 h for the 394 bp probe, 4-6 h for the oligonucleotide probe.

Following hybridization, the membrane was washed twice with 200 ml wash/SDS buffer (1 mM EDTA pH 8, 40 mM $NaHPO_4$, 1% SDS, preheated to 60°C) at room

temperature for 15 min. After the second wash, the membrane was rinsed in wash buffer (1 mM EDTA pH 8, 40 mM $NaHPO_4$) and equilibrated in Na-malate buffer (100 mM maleic acid, 150 mM NaCl, pH 7.5) for 1 min.

The membrane was incubated in blocking buffer [1% w/v blocking reagent (Roche Biochemicals) in Na-malate buffer] for 30 min then with anti-DIG antibody Fab fragments (Roche Biochemicals) at 1:10000 in blocking buffer for 30 minutes. The membrane was then washed twice with Na-malate wash buffer (Na-malate buffer containing 3% v/v Tween 20) for 15 min at 75 rpm, each wash. The membrane was incubated in 20 ml detection buffer (100 mM Tris pH 9.5, 100 mM NaCl) for 5 min, followed by a 5 min incubation in 10 ml detection buffer containing 100 µl CDP-Star (Roche Biochemicals). The membrane was placed between two sheets of clear acetate and used to expose X-Ray film (Kodak Biomax MR) for 30-60 min. The film was processed in the Medical Film Processor (Model QX-70, Konica Corporation, Tokyo, Japan).

B. Transformation of *Arabidopsis thaliana* with pCORE

B1. Growth of Plants.

Arabidopsis thaliana plants, ecotype Columbia (Col), were grown in soil (commercial potting soil, Metro Mix 200, Wetsel, Harrisonburg, VA) with a 12 h light/dark cycle exposed to light of 120 microeinsteins m^{-1} s^{-1} at 22°C in three inch pots. Plants were watered weekly by soaking the pots in tap water for 15 min and then allowing the excess water to drain from the pots for 15 min.

B2. Transformation of *Arabidopsis thaliana by* Vacuum Infiltration with pCORE.

Arabidopsis thaliana was vacuum infiltrated via the *Agrobacterium tumefaciens*-mediated vacuum infiltration method of Clough and Bent (1998). At week 3, plants were trimmed to remove any growing flowers, buds or siliques. The plants were infiltrated 4-8 days after clipping. The day before infiltration, the plants were soaked in water overnight to allow for the stomata to remain open during infiltration. The next day, the plants were allowed to drain for 15-20 min before infiltration. *A. tumefaciens* carrying the pCORE construct were grown in 2xYT, 30 μg/ml kanamycin, to an OD_{600} of 0.8. After centrifugation at 5000 × g, the bacteria were resuspended in 1 volume (400 ml) of infiltration medium (0.5× Murashige and Skoog Salts, 1 × B5 Vitamins (10 μg/ml thiamine HCl, 1 μg/ml nicotinic acid, 1 μg/ml pyridoxine HCl, 100 μg/ml myoinositol, 5% sucrose, 0.044 μM benzylamino purine).

The plants in pots (with soil) were then inverted in the bacterial suspension and exposed to 25 mm Hg vacuum for 15 min in a vacuum oven (Precision, Chicago, IL). The vacuum was released rapidly to force the bacteria into the submerged tissues. The plants were then returned to the growth chamber and grown as described above (**Materials and Methods, B1**) until T1 seeds were harvested 4-5 weeks later.

B3. Selection of Putative Transformed Seeds.

T1 seeds were arranged in a grid pattern on plates of selective medium (0.8% agar, 1× Murashige and Skoog Salts, 1× B5 vitamins, 1% sucrose, with 75 μg/ml kanamycin, 100 μg/ml timentin (SmithKline Beecham, Philadelphia, PA). The plates were incubated at 4°C in the dark for 3-4 days, then placed in a growth chamber at 22°C

under constant light (40 microeinsteins m^{-1} s^{-1}). After 7-10 days, seedlings that remained green (see **Results** pg. 6) were transferred to a fresh selection plate. After another 7-10 day period on selective medium, the remaining green seedlings were transferred to soil and placed in the growth chamber under the conditions described in **Materials and Methods, B1**.

B4. Histochemical Staining of Transgenic Seedlings or Leaves.

Both one-week old transgenic seedlings and mature leaves were stained to assay for the expression of the GUS gene as monitored by its enzymatic activity. The biochemical reaction catalyzed by GUS using 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-Gluc, Clontech) as substrate is a hydrolytic cleavage (Martin, *et al*., 1992). The X group, (5-bromo-4-chloro-3-indolyl) forms a precipitate as a positive indicator for the presence and activity of GUS.

Staining was carried out by placing the tissue in a microcentrifuge tube with 500 µl of staining solution [1 mM X-Gluc; 50 mM Tris, pH 7.81; 0.1% Triton X-100] and subjected to a vacuum in a warm (45°C) Speed-Vac Concentrator (Savant) 3 times for 2 min each. This vacuum infiltration was followed by an overnight incubation at 37°C (Stomp, 1992). Following this incubation, the stained tissue was soaked in 95% ethanol to remove the chlorophyll in order to improve visualization of the blue precipitate indicative of GUS expression.

B5. Extraction of Genomic Plant DNA.

Genomic DNA was extracted from 3 week old plants using the DNeasy™ Plant Mini Kit (Qiagen, Valencia, CA). 80-100 mg of plant leaf tissue was frozen and ground to a fine powder under liquid nitrogen. 400 µl of kit buffer AP1 and 4 µl of kit RNase A solution was added to the ground plant tissue and vortexed vigorously to remove any clumps. The mixture was then incubated at 65°C for 10 min. After the incubation, 130 µl of kit buffer AP2 was added to the mixture and this was placed on ice for 5 min. This was then applied to the Qiashredder spin column and centrifuged for 5 minutes at 16000 × g in a microcentrifuge. Half a volume of kit buffer AP3 and one volume of 100% ethanol was added to the flow through fraction in a clean microcentrifuge tube. This mixture was applied to the DNeasy spin column and centrifuged for 60 seconds at 6000 × g discarding the flow through fraction. 500 µl of kit buffer AW was added to the DNeasy spin column and centrifuged for one min as above. Another 500 µl of kit buffer AW was added to the column and centrifuged for two min as above to dry the membrane. After transferring the spin column to a clean microcentrifuge tube, 50 µl of preheated (65°C) kit buffer AE was added to the spin column and centrifuged at 6000 × g for one minute. This entire step was repeated with an additional 50 µl of preheated (65°C) kit buffer AE.

B6. Screening of Transgenic Plants by PCR & Southern Blot Analysis.

100 ng of genomic DNA in 50 µl of MasterAmp Mix G (Epicentre, Madison, WI) with 1% DMSO (Sigma) was used to amplify the core sequence from transgenic plants using primers 3 and 4. Final primer and template concentrations were 0.6 µM and 2

ng/µl, respectively. The expected fragment size was 394 bp. The cycling conditions are outlined in Table 2.

Temp	Time	Num. Cycles
94°C	45 sec	35
58°C	45 sec	
68°C	60 sec	
68°C	10 min	1

Table 2. PCR cycling conditions for the amplification of the CORE sequence from transgenic plants.

Additionally, genomic DNA was digested with either EcoRI or XbaI and analyzed with the Southern blotting procedure in **Materials and Methods, A7**, using DIG labeled 394 bp PCR product as probe, **Materials and Methods, A6**.

B7. Sequence of Transgenic *Arabidopsis thaliana* PCR products.

The sequence of the AAVS1 CORE sequence in putative transgenic plants was determined using automated cycle sequencing with 1.6 pmol primer3 or primer4 and 100 ng of template in 10 µl Big dye terminator mix (PE Applied Biosystems, Foster City, CA). The reactions were performed using an MJR-100 thermalcycler (MJ Research Inc., Waltham, MA) following cycling conditions shown in Table 3, and run on an ABI Prism 310 sequencer (Perken Elmer, Foster City, CA) in the Biology Department at Virginia Tech. The results were interpreted using DNA-Star programs MapDraw, Seqman and EditSeq (DNASTAR, Madison, WI).

Temp	Time	Num. Cycles
95°C	30 sec	25
50°C	15 sec	
60°C	4 min	

Table 3. Cycling conditions for cycle sequencing of the CORE sequence from pCORE, pRVK and transgenic plants.

Results and Discussion

The overall purpose of this project was to create a collection of transgenic plants each containing the site-specific integration sequence (AAVS1), derived from the human genome, for the mammalian parvovirus AAV. These plants will be the first part of a system for delivering transgene into plant nuclear genomes in a site-specific manner to address the problems associated with random integration events in transgenic plants. The other parts of the system are developing a means to express the AAV-coded Rep proteins within plants containing AAVS1. Rep proteins are required for the integration of recombinant AAV, the third part of the system.

A. Production and Analysis of pCORE

To create plants capable of expressing transgenes carried in recombinant AAV, it was necessary to insert the specific site for this viral integration into the model plant, *A. thaliana.* Linden *et al*. (1996) described the 33 bp minimum sequence of AAVS1 necessary for the integration of AAV into human chromosome 19. Figure 7 shows the Rep binding site (RBS) (blue) and the terminal resolution site (red) within the minimal sequence necessary for integration of the AAV-2 genome (yellow), contained within flanking human DNA sequences. A 394 bp fragment of AAVS1 harboring the minimal integration sequence and flanking sequences was termed the CORE sequence. This region was selected on the basis of size, to balance ease of amplification by PCR with the ability to insert the amplified fragment into appropriate plasmids.

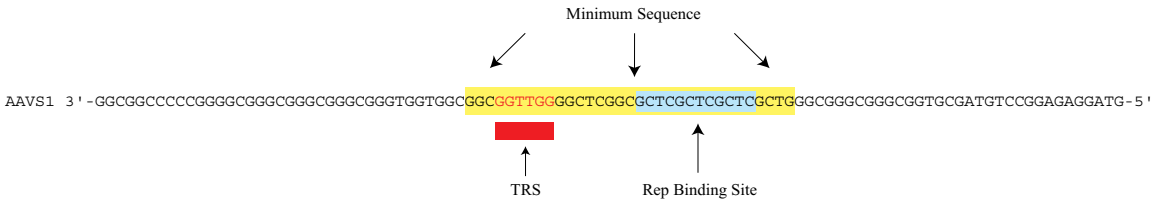

Figure 7. The AAVS1 sequence from human chromosome 19. The minimum sequence sufficient for targeted integration was determined by Linden *et al*. (1996). This minimum sequence encompasses the Rep binding site (RBS) and the terminal resolution site (TRS). The RBS is the region where the Rep proteins bind and the TRS is the sequence where the Rep proteins nick the DNA during the integration event. The minimum sequence (33 nucleotides) was used as a DIG-labeled probe in Southern blots to confirm the presence of the sequence in putative clones (see **Materials and Methods, A6**).

The plasmid pCORE was constructed to move the CORE sequence into *A. thaliana*. CORE sequences were inserted into pBI121, a plant transformation vector containing the borders of the T-DNA derived from a Ti-plasmid, a neomycin phosphotransferase II gene (which confers kanamycin resistance) and the β-glucuronidase (GUS) gene as a reporter for effective expression of proteins proximal to the AAVS1 CORE sequence. This vector is half of a binary system (see **Introduction**) and is complemented by another plasmid, carried in *A. tumefaciens*, that contains the *vir* genes necessary for chromosomal insertion of sequences between the T-DNA borders.

The plasmid pBI121 (Figure 8A) contains a unique EcoRI site adjacent to the GUS reporter construct, which was used for insertion of the CORE sequence. Primers were designed to include terminal EcoRI recognition sequences. Primers 3Eco (5'-AC**GAATTC**GTACTTGCTAGTATGCCGTGGG-3') and 4Eco (5'-AC**GAATTC**GTCATCCTCTCCGGACATCG-3') with these sites (shown in bold) were used to amplify a 409 bp fragment containing the 394 bp CORE sequence via PCR. The template for this amplification was pRVK (a gift from R.M. Linden). pRVK contains the 5' most 3500 bases of human DNA including the AAVS1 sequence as an EcoRI-KpnI fragment in pBluescript(KS+). The 409 bp product was ligated into the unique EcoRI site of pBI121 immediately downstream of the β-GUS expression cassette to create pCORE, shown in Figure 8B. GUS expression in plants was used as an indicator of potential expression level of integrated transgenes. In addition, detectable GUS expression indicates that the region of the plant genome into which AAVS1 is integrated is accessible to the transcriptional machinery. If the GUS gene were successfully expressed, it is likely that a transgene integrated into AAVS1 (directly adjacent to GUS

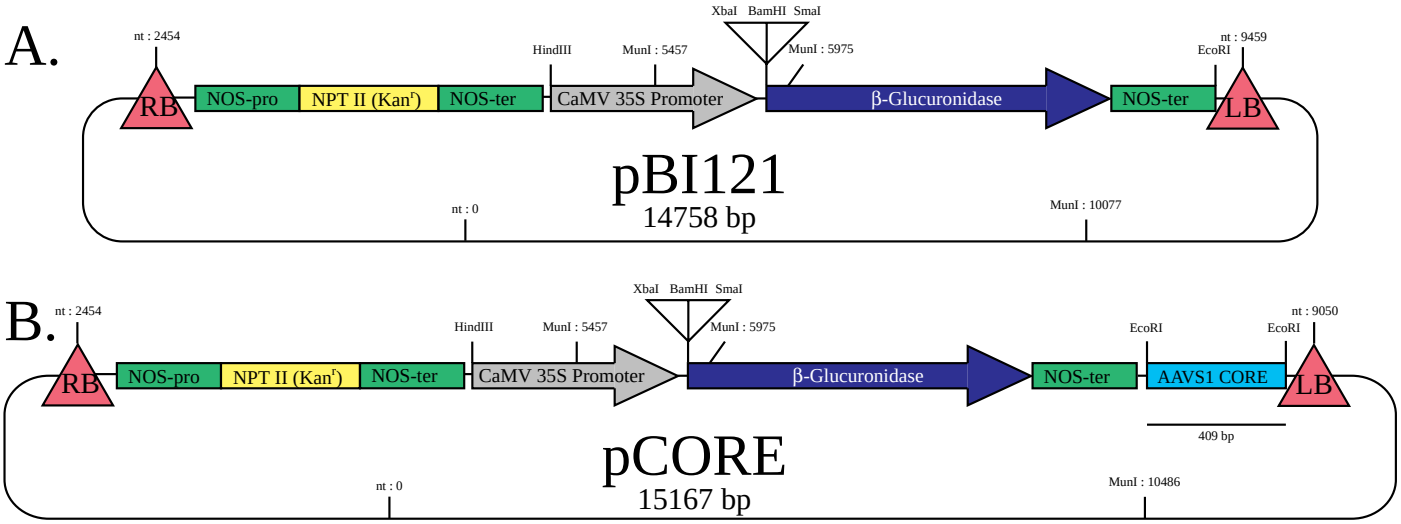

Figure 8. Maps of pBI121 and pCORE. A. Schematic diagram of pBI121, the plasmid used as the basis for the construction of pCORE. The plasmid contains the neomycin phosphotransferase II (NPT II) gene driven by the nopaline synthase promoter (NOS-pro), and the β-GUS gene under the control of the cauliflower mosaic virus promoter (CaMV 35S). The nopaline synthase terminator is present at the end of each gene, providing a poly A signal in the plant. B. Schematic diagram of pCORE, the plasmid used to transform *A. tumefaciens* and *A. thaliana*. For the construction of pCORE, the 409 bp AAVS1 CORE sequence (described in the text) was inserted into a unique EcoRI site at the terminus of the NOS-ter flanking the β-GUS gene.

in pCORE) would also be expressed. In the absence of GUS expression in transformants, we can safely assume that the region of insertion is inaccessible to the transcriptional machinery; plants lacking the ability to express the reporter gene would be eliminated from further characterization.

A1. Screening of Bacterial Transformants

Bacterial transformants were screened to identify those harboring pCORE. Bacterial colonies were grown on selective media and their plasmid DNA harvested. Routinely, plasmids are screened for the presence of an insert by restriction enzyme digestion and visualization of the desired fragment after electrophoresis and ethidium bromide staining of the gel. If pCORE were digested with EcoRI, the expected sizes for the plasmid backbone is about 13 kb and the size of the insert is 409 bp. The mass ratio of these fragments precludes visualizing CORE when electrophoresing reasonable amounts of DNA (Fig 10, lane A1).

Instead, the presence of the CORE sequence in pCORE was confirmed by PCR. PCR results using various plasmids as templates with primers 3Eco and 4Eco are shown in Figure 9. pRVK (AAVS1 in pBluescript, lane 2) and pCORE (AAVS1 in pBI121, lane 4) produce fragments of the predicted size while pBI121 did not produce any products (lane 3). Lane 1 is the negative control that does not contain any DNA template. The PCR results show that pBI121 does not contain any sequence similarity with the CORE sequence, and that pCORE contains a sequence that produces a fragment of appropriate size upon amplification using primers 3Eco and 4Eco.

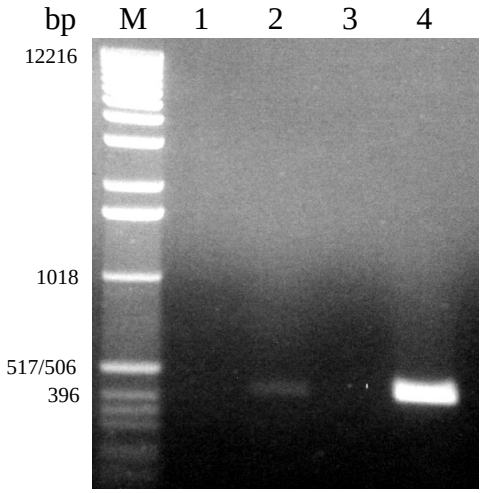

Figure 9. PCR of AAVS1 from various templates; Lane1 : Negative control (no DNA), Lane 2 : pRVK, Lane 3: pBI121, Lane 4: pCORE (pBI121 w/AAVS1). The size standards are the 1 kb DNA ladder from Gibco-BRL, indicated at the left.

A2. Characterization of pCORE via Southern Blotting

The presence of the CORE sequence in pCORE was additionally confirmed by Southern blotting. The experiment shown in Figure 10 was performed using the putative pCORE construct probed with a DIG-labeled 33 nt oligonucleotide corresponding to the minimal integration sequence (See **Materials and Methods, A6, and** Figure 7). Samples were digested with EcoRI or XmaI, separated by gel electrophoresis and stained with ethidium bromide (A) or probed (B). Lane 1 contains pCORE digested with EcoRI, which liberates the 409 bp CORE fragment from the 13 kb vector, lane 2 contains pCORE digested with XmaI to linearize the plasmid. Figure 10 shows that the 33 nt probe hybridizes to the 409 bp CORE sequence whether excised from (lower band in lane 1) or present within pCORE (only band in lane 2), but not to the pBI121 backbone of pCORE (upper band in lane 1), indicating that the construct contains the AAVS1 region of the human genome.

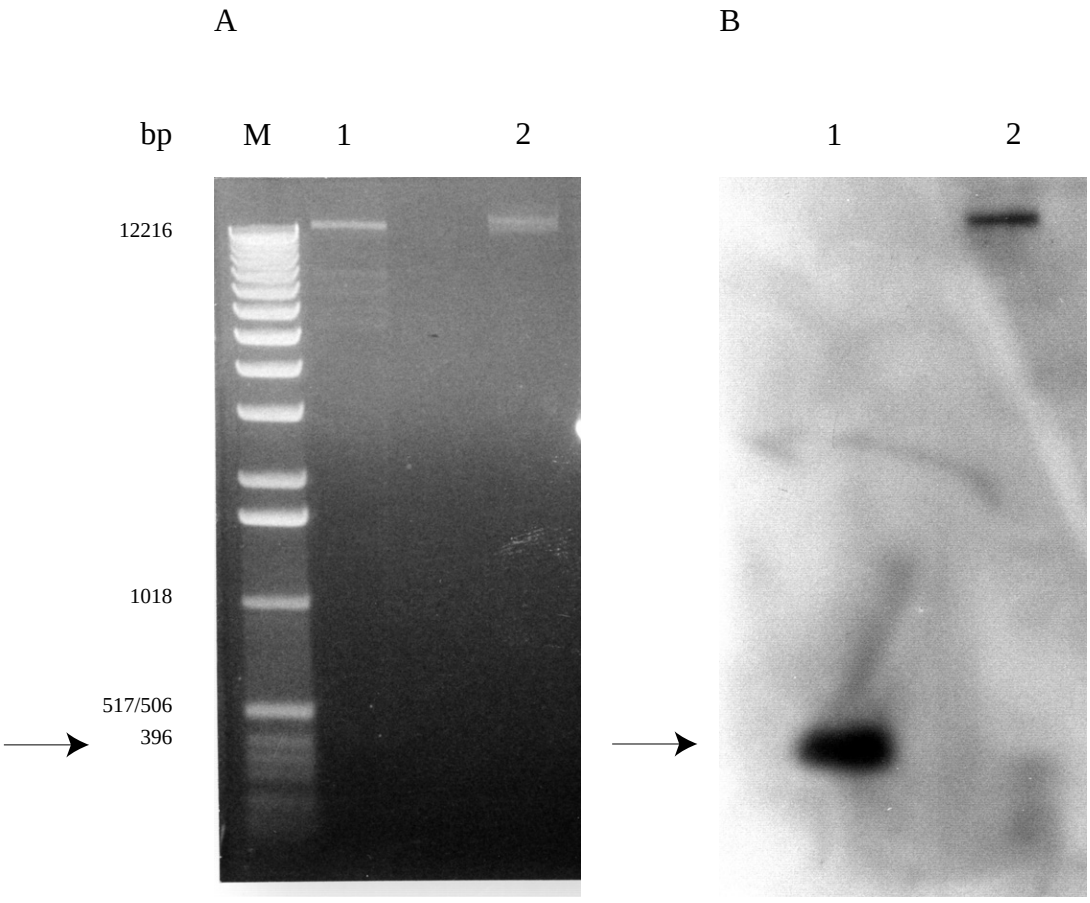

Figure 10. pCORE probed with DIG-labeled oligonucleotide. Gel (A) and Southern Blot (B) of digested pCORE. Lane 1: EcoRI digest; Lane 2 : XmaI digest. The arrow indicates where the CORE fragment appears in both the gel and corresponding southern blot. Probe is DIG-labeled 33 nt oligonucleotide corresponding to the minimal integration sequence (See **Materials and Methods, A6, and Figure 7)**.

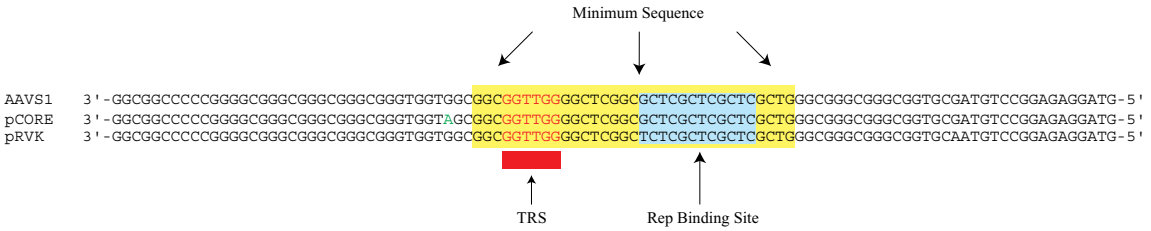

Figure 11. Alignment of sequences from AAVS1, pCORE and, pRVK.

A3. Characterization of pCORE via Sequencing

In order to confirm that the sequence present in pCORE is the same as that present in pRVK, dideoxy-sequencing was performed on the PCR product generated with primers 3 and 4 (See **Materials and Methods, A7**) in both the 5' and 3' directions. Shown below in Figure 11 is the sequence of the AAVS1 CORE fragment from human chromosome 19 (GENBANK Accession Number S51329) and the sequence from pCORE. There is a single G to A base pair substitution (shown in green) when comparing pCORE sequence with GENBANK sequence. pRVK has been re-sequenced in our laboratory and this alteration is not present. However, as the change is not contained within the minimum sequence necessary for integration (shown in yellow), it is unlikely to be problematic in future applications of the construct.

B. Production and Characterization of Transgenic Plants

Transgenic plants were generated using vacuum infiltration (Bechtold *et al*., 1993). This involved subjecting maturing plants (ecotype Columbia) to a vacuum while submerged in a solution of *Agrobacterium tumefaciens* containing the pCORE construct. Releasing the vacuum forced the bacteria into the plants and introduced pCORE into growing cells. Existing flowers and emerging siliques (seed pods) were removed prior to the infiltration process to ensure that there were no pre-existing seeds. Arising from the apical meristem, the newly growing germ tissue (flowers and siliques) should contain the integrated T-DNA; seeds from these plants should harbor our sequences of interest flanked by T-DNA.

Seeds (300 per selection plate) harvested from these plants were placed on agar plates containing kanamycin and allowed to grow for 2-4 weeks. Survivors (1%) were transplanted into soil and allowed to grow and produce seed. Leaves were harvested from transplanted seedlings for genomic DNA isolation and GUS staining when the plantlets were large enough to survive the loss of 5 or 6 leaves.

B1. Characterization by Gene Expression

The GUS reporter gene was used as a rapid visual screen for insertion of the intact T-DNA cassette into a region of the genome capable of transcription. As well, the level of GUS expression in plants with the T-DNA region contained in pCORE is an indicator of the level of expression expected for a transgene inserted into the AAVS1 portion of the T-DNA using AAV as a vector. Plants were stained by vacuum infiltration of the colorless substrate of the enzyme; a blue color after incubation indicates GUS activity. This color is shown to advantage after removing the chlorophyll by soaking in 95% ethanol. Figure 12 shows that three clonal progeny (the F_4 generation) from the F_3 generation of transgenic plant 37 used in Figures 13 and 14, transformed with pCORE, are stained to approximately the same level as a plant transformed with pBI121. No color is seen with a nontransformed Columbia wild-type plant, the ecotype used for all transformations with pCORE. These results show that the T-DNA insertion allows transcription of the GUS transgene. We would expect that insertion into the adjacent AAVS1 sequence would also allow transcription.

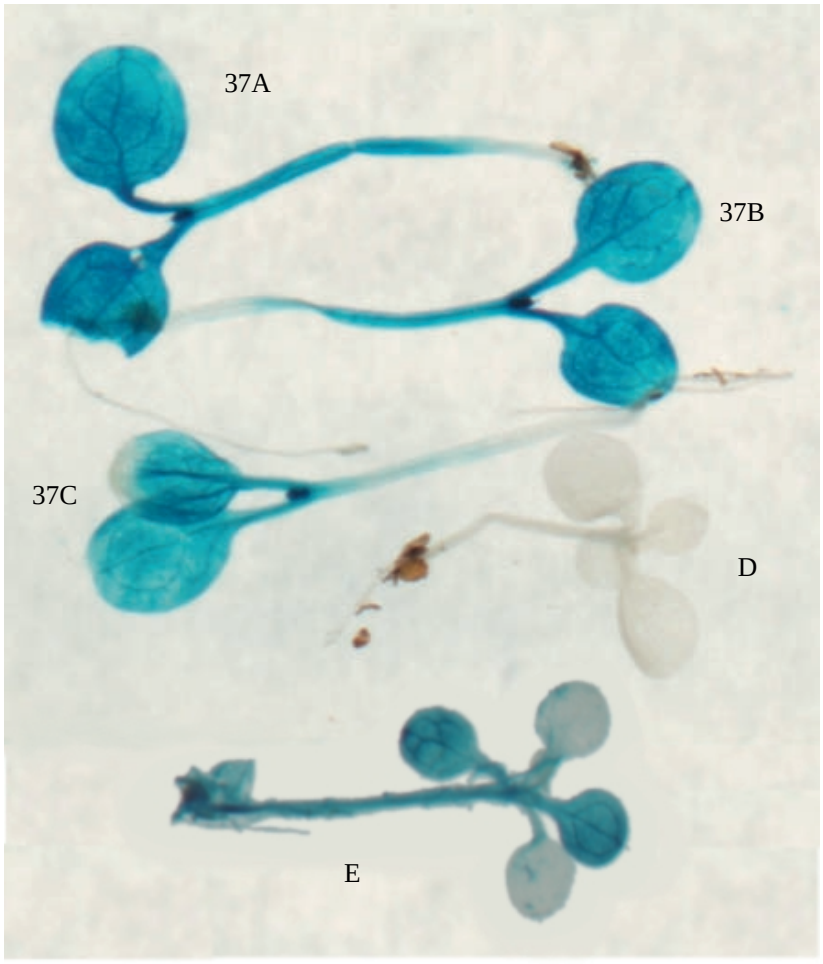

Figure 12. GUS expression in *A. thaliana*. All seedlings have been stained with GUS stain and then destained with 95% Ethanol (See **Materials and Methods, B4**). Plants labeled 37A-C are transgenic progeny (F_4) from plant 37 (F_3 generation) that exhibit GUS activity. Seedling D is non-transformed Col, seedling E is Col transformed with pBI121.

B2. Characterization of Transgenic Plants by the Polymerase Chain Reaction

In order to determine if the F_3 generation of transgenic plants contained the intact CORE sequence, 100 ng of genomic DNA was harvested from several lines of F_3 plants all of which were derived from the same primary transformant. These DNAs were used as template for PCR reactions using primers 3 and 4 (see **Materials and Methods, A5**). The primers lack terminal EcoRI restriction sites, and were used to increase specificity of binding for the reactions. The PCR product obtained should be 394 bp if the sequence were inserted into the genome or present in some other state. The products of these reactions are shown in Figure 13. Each reaction produced a product of the expected size of 394 bp. Lanes 1-7 contain products from amplification of genomic DNA by PCR from the F_3 generation of plants; plants 36, 37, 38, 42, 44, 46 and 47. Lane 8 contains PCR product from pCORE, lane 9 contains PCR products using water as the no-template control in the PCR reaction. Lane 10 contains PCR products using wild-type *A. thaliana* ecotype Columbia (Col) as template, used as the recipient ecotype for the vacuum infiltration to construct the transgenic plants. (see **Material and Methods, B1 and B2**). The results show that primers designed for the amplification of the CORE fragment generate an appropriately sized fragment when the DNA recovered from putatively transgenic plants is used as template in a PCR reaction. No such product was seen when no template is present or Col DNA is used as template. Therefore, the CORE sequence is present in these plants.

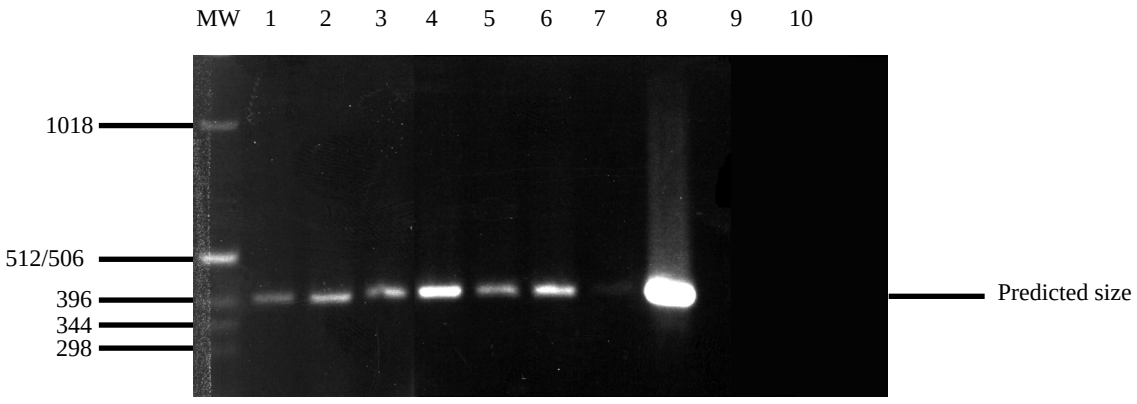

Figure 13. Agarose gel of PCR products using genomic DNA from the F_3 generation of transgenic plants. Each reaction used primers 3 and 4. Lanes 1-7 contained template from plants 36, 37, 42, 44, 46, 47, respectively and lane 10 contained DNA from a nontransformed Columbia control plant. The positive control (lane 8) is the AAVS1 sequence amplified from the original clone pCORE. The negative control (lane 9) lacks DNA in the reaction. Columbia wild type was the original plant ecotype that was transformed with pCORE.

B3. Characterization of Transgenic Plants by Sequencing

In order to confirm that the CORE sequence in the plants was the same as that in the pCORE plasmid used to produce the plants, the PCR products were sequenced. Sequence data from the PCR products are shown in Figure 14. Each of the clones contained the TRS, RBS and the single G to A base pair substitution found in pCORE when its sequence is compared to the reported sequence and to pRVK. The sequence data confirmed that the plants contain the intact CORE sequence without any errors in the minimum sequence required for AAV integration.

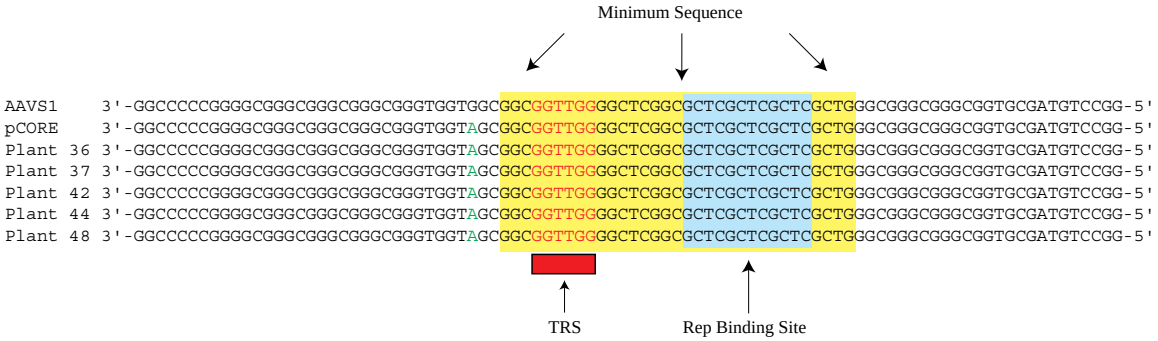

Figure 14. Alignment of CORE sequences from transgenic plants. Each plant exhibits the Rep binding site, the terminal resolution site and a single G to A base pair substitution.

B4. Characterization of Transgenic Plants by Genomic Southern Blotting

Genomic Southern blotting was used to demonstrate that the CORE sequence is contained within the genomes of the plants, rather than harbored extrachromosomally within the cells. Digestion of genomic DNAs with a restriction enzyme that has a site within the T-DNA upstream of the CORE sequence and a site downstream of the CORE sequence outside the T-DNA will generate hybridizing fragments with sizes greater than 409 bp. The hybridizing bands should be the same for clonal lines of plants and different for each independent transformant. MunI digestion of the vector pBI121 yields fragments of 518 bp, 4102 bp, and 10,138 bp while MunI digestion of pCORE generates fragments of 518 bp, 4511 bp (4102 bp + 409 bp), and 10,138 bp since the CORE fragment has been introduced within the 4102 bp fragment (Figures 8 and 15). Figure 16 shows a Southern hybridization of genomic DNAs from several F_4 transgenic plants from a single transformation event (lanes 37, 38, 43, 44, 36 and 47) that were digested with MunI and probed with the DIG-labeled PCR product of 394 bp that encompasses the TRS and RBS. The banding pattern for the genomic DNAs is different than the banding pattern of pCORE digested with the same enzyme. MunI digestion of the plant genomes displayed two hybridizing bands at approximately 8.7 and 4.5 kb. No bands are seen in an untransformed plant (Col), in a plant transformed with pBI121 (35S) and in one transformed plant (probably due to insufficient recovery of DNA). The presence of two hybridizing bands could result from two single copy inserts or, from a tandem insertion in a head to head or head to tail orientation (see Figure 17).

Figure 17 illustrates the predicted sizes for tandem T-DNA inserts. Integration in a head to tail configuration would yield two hybridizing fragments, one of 6 kb and one

greater than 3 kb. Integration in a head to head configuration would yield two hybridizing fragments greater than 3 kb in size. In the event of a tail to tail configuration one should expect one hybridizing band of 6.1 kb (see Figure 17).

If the observed 4.5 kb hybridizing band is derived from non-integrated plasmid, it appears that there is a single insertion of the T-DNA in the transformed plant, resulting in a hybridizing fragment of 8.7 kb. If the 4.5 kb band represents an authentic insertion, it would mean that there is a second insertion of the T-DNA within the plant genome. An 8.7 kb fragment cannot be generated via MunI digestion of the plasmid pCORE even if the plasmid has been maintained episomally across generations. The 4.5 kb restriction fragment could be generated from a MunI restriction site within the T-DNA and a second MunI site in the *A.thaliana* genome. The most probable explanation is that this fragment is derived from a plasmid harbored extrachromosomally within the plant tissue. Infrequently, flanking vector sequences are transferred with the T-DNA (Kononov, *et al*., 1997) and the possibility exists that the 4.5 kb fragment could represent an abberant insertion of the CORE sequence.

In summary, since the 8.7 kb fragment was observed in transgenic plants, there is at least one copy of the CORE sequence integrated into the transgenic plant genome. If the 4.5 kb fragment is on a episome, there is only one CORE sequence in the genome. If the 4.5 kb is not episomal, the genome of each transgenic plant contains two copies of the CORE sequence. Since the probe binds to the inserted CORE sequence at high stringency, this indicates that the transgenic plants contain the intact CORE sequence.

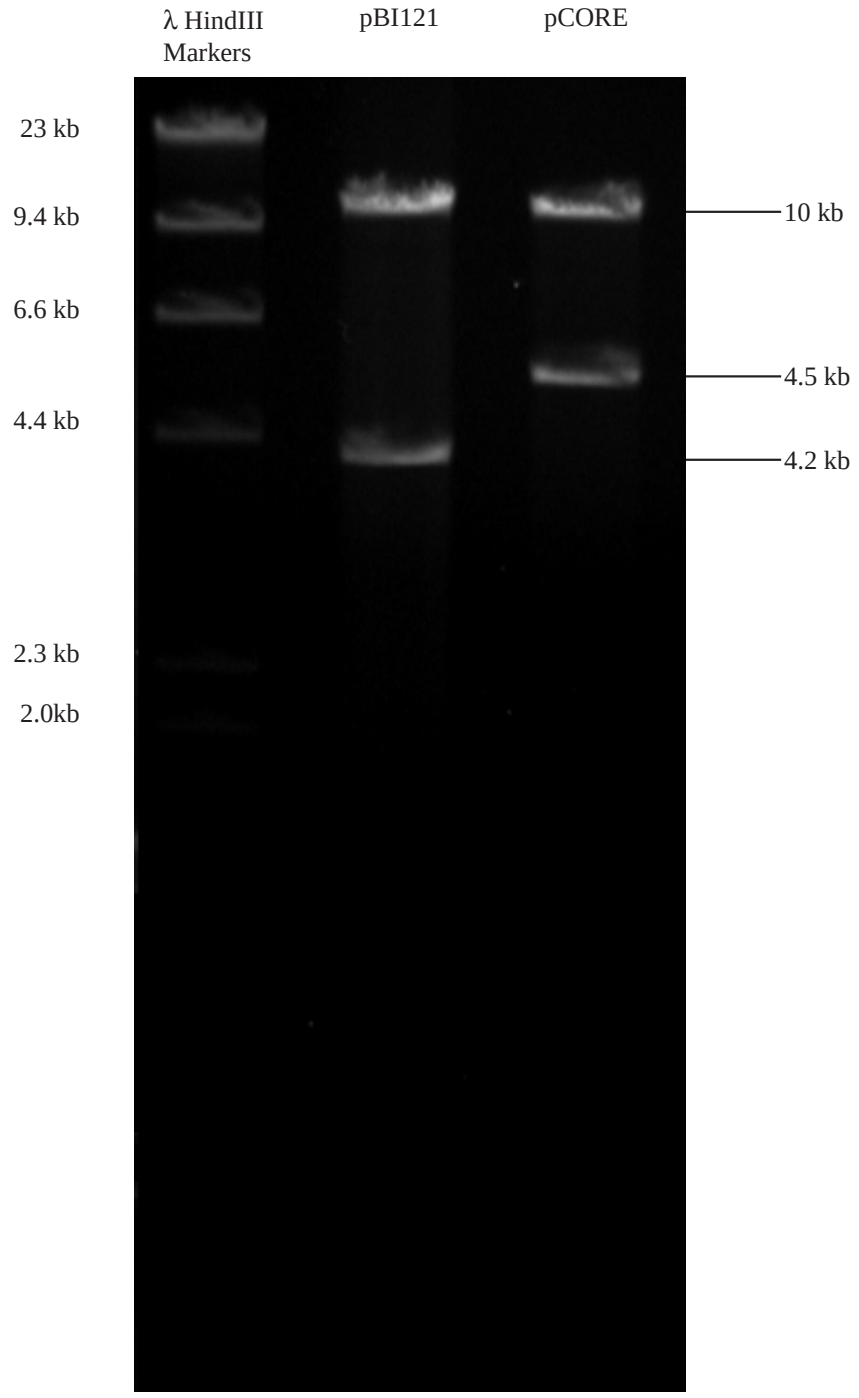

Figure 15. Digestion of pBI121 and pCORE with MunI. MunI digestion of pBI121 yields fragments of 518 bp, 4102 bp, and 10138 bp while MunI digestion of the pCORE plasmid generates fragments of 518 bp, 4511 bp (4102 bp + 409 bp), and 10138 bp.

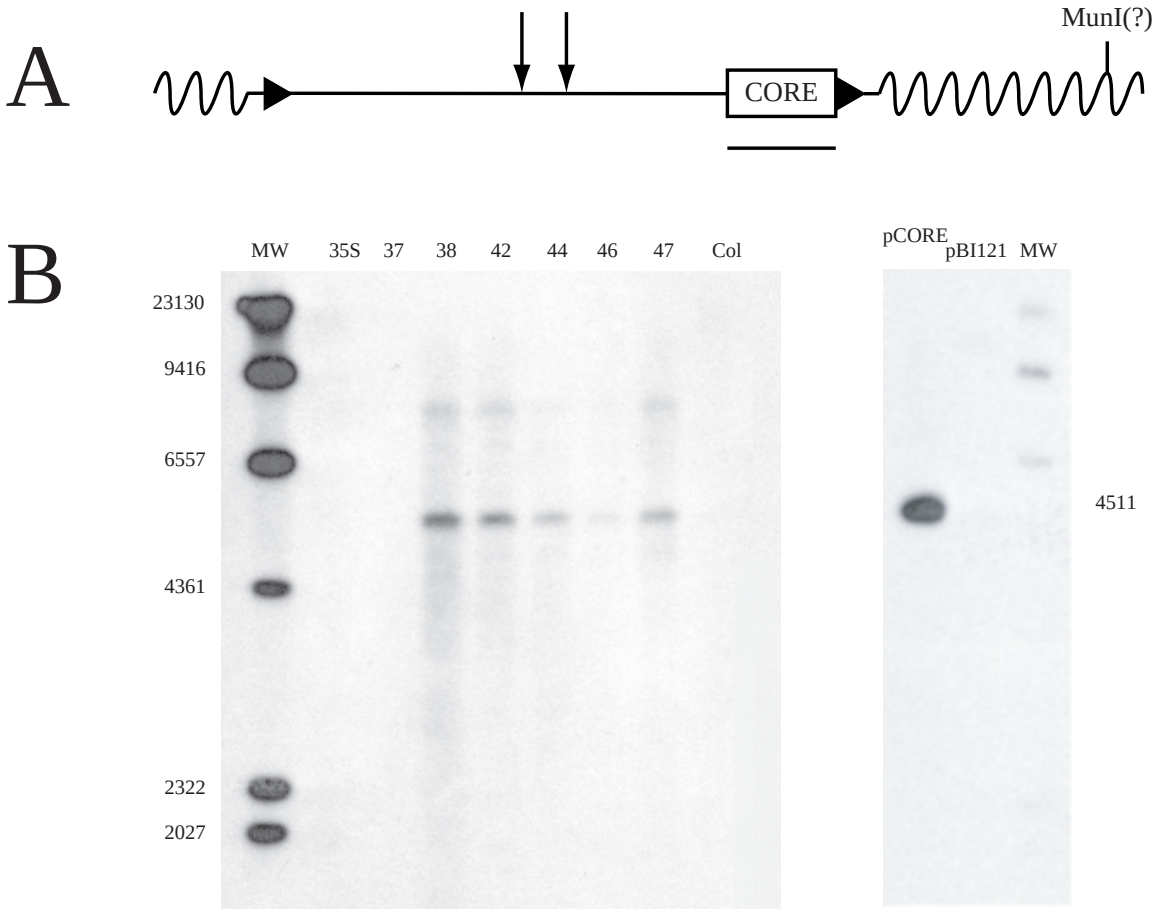

Figure 16. Southern blot analysis to detect insertion of CORE sequence. A: Predicted integration pattern for a single CORE sequence T-DNA insert into the plant genome. Black triangles represent the left and right border sequences; probe binding is indicated by the black line; genomic sequences are indicated with wavy lines. B: Southern blot analysis of clonal transgenic plants line carrying the pCORE construct. One microgram of pCORE plasmid and two micrograms of genomic DNA digested with MunI were analyzed with the Southern technique. Numbers at the top of the lanes represent the plant line from which DNA was harvested. Col is the untransformed Columbia wild type plant genome, the 35S lane is DNA from Col transformed with pBI121. Vertical arrows indicate known MunI sites. pCORE and pBI121 lanes are loaded with their respective plasmid, digested with MunI. The size standards are the DIG labeled Lambda/HindIII molecular weight marker (Roche Biochemicals).

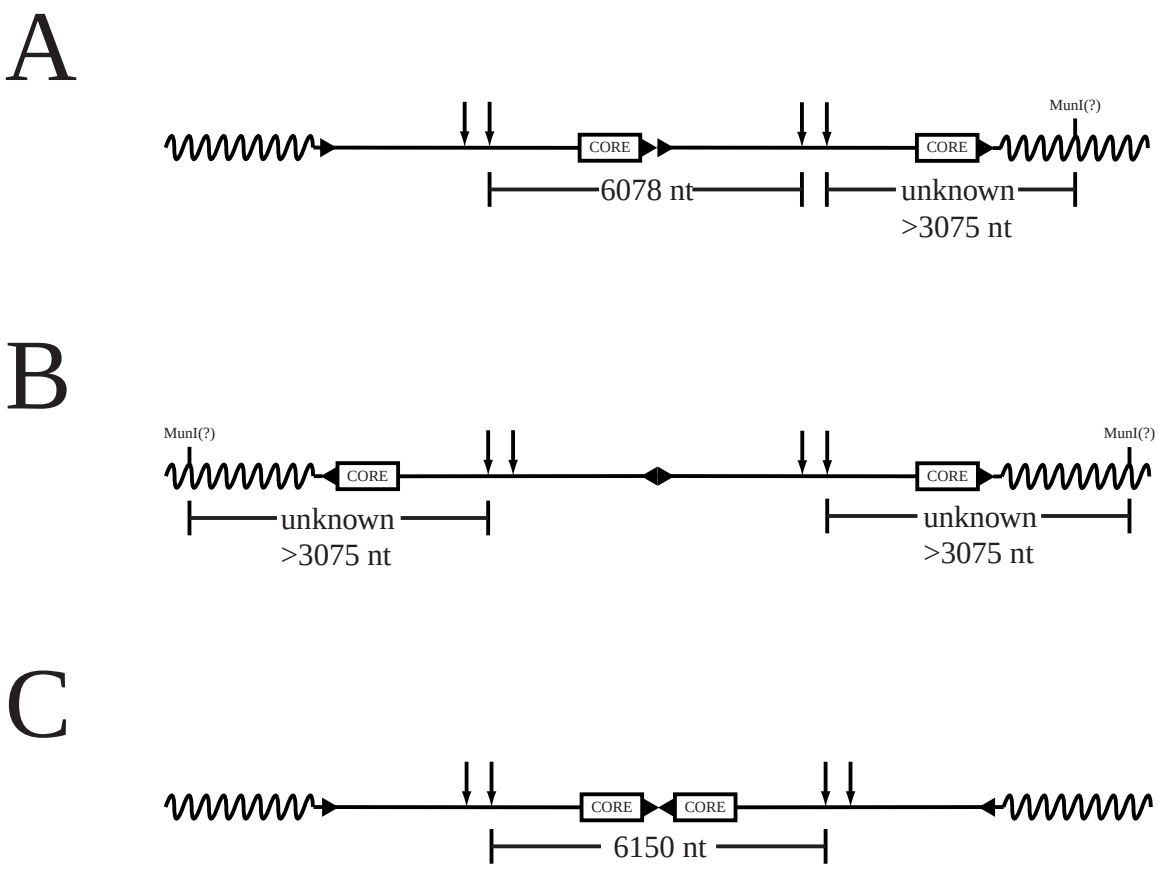

Figure 17. Possible orientations of a tandem insertion. A: Head to tail orientation that would generate a two hybridizing bands of 6 kb and greater than 3kb by Southern blot analysis. B: Tail to tail orientation would generate to two hybridizing bands greater than 3 kb upon Southern blot analysis. C: Head to head orientation that would generate a single hybridizing band of 6 kb upon Southern blot analysis. Arrows indicate known MunI sites, question marks indicate the presence of a MunI site at an undefined position in the genome, triangles represent the left and right borders, and wavy lines symbolize genomic DNA strands. The size of the digested fragment that will appear on Southern analysis is shown below each tandem insertion.

Conclusions

Linden *et al*. (1996) described the minimum sequence required for the integration of AAV into the human chromosome 19, AAVS1. The most critical portion of this integration sequence is the 5'-most 510 base pairs of the originally discovered λ EMBL clone as determined by deletion analysis (Giruad *et al*., 1995). The Rep binding site (Weitzman *et al*., 1994), and the terminal resolution site, the site that is nicked by the Rep78 protein (Berns and Linden, 1995), are contained within this region. During a normal infection process in the absence of a helper virus, these sequences direct the integration of adeno-associated virus. Within the context of this project, these sequences are being harnessed to design a new method of generating transgenic plants. This method is being developed to provide a site-specific alternative to current plant transformation methods as well as provide a linked GUS gene as an indicator of transgene expression potential.

The critical portion of the AAVS1 region, termed CORE, has been introduced into the plasmid pBI121, thus generating pCORE (Figure 8). Further, it has been demonstrated that the GUS gene, adjacent to the CORE sequence, is expressed and has enzymatic activity within the plant (Figure 12). PCR evidence suggests that the CORE sequence is within the plant tissue (Figure 13), although this result remains to be confirmed by Southern blot analysis of multiple transgenic plants resulting from unique transformation events; an effort to generate these plants is underway. Sequencing data revealed that the sequence is consistent from plant to plant (Figure 14). A procedure has been outlined detailing the construction of transgenic plants to contain the CORE sequence. Once plants are identified that contain an intact AAVS1 sequence and exhibit

high levels of reporter gene expression, it is likely that any transgene integrated into the AAVS1 region within these plants will be expressed in a consistent manner in the presence of AAV-coded Rep protein. This assumption will be tested in future experiments to determine the suitability of using this system to generate plants containing transgenes.

Future Directions

Future developments of this technology will require the construction of a recombinant AAV vector able to deliver a transgene into plants. In addition, *rep78*, necessary for site-specific integration, can be delivered along with the transgene (bordered by the ITRs of AAV). Alternatively, *rep78* can be introduced into the plant containing the AAVS1 sequence. Disadvantages of this strategy are that breeding will be required to produce plants homozygous for both genes and the diminished choices of selective markers. Current systems for moving transgenes into mammalian cells using recombinant AAV do not result in site-specific integration since *rep78* is not included in the construct or in the target cell line (Tattersall, 2000). A hybrid adenovirus/adeno-associated virus vector that achieves site-specific integration using the Rep68/78 + ITR formula in mammalian cell culture has been recently reported (Recchia *et al*., 2000). The ability to engineer plants to express Rep is an advantage of our system.

Our goal is to construct the AAV-based plant transformation vectors from the infectious genomic clone of AAV, psub201(+) (Samulski, *et al.*, 1987) carried in pBI121. For transformation of plants expressing Rep, the transgene will be flanked by the ITRs. For plants containing only AAVS1, the vector will contain Rep as well as the transgene,

flanked by the ITRs. In both cases, Rep will be expressed under control of the CaMV35S promoter and will utilize the NOS terminator. Depending upon the nature of the transgene and the particular application, we will select an appropriate promoter that could be tissue specific, constitutive or inducible.

Introduction of the vector could be as simple as taking a leaf disk from CORE plants and submerging the disk in a solution of recombinant vector. The leaf disks would then be incubated on agar plates with appropriate selection and hormones to generate plants containing the transgene (Horsch, *et al.*, 1985). One could also construct a protoplast suspension of tobacco harboring the CORE sequence and electroporate the recombinant vector into the suspension and grow the protoplasts into mature plants to screen them for the presence of the transgene.

Potential uses of this technology include the production of proteins that are critical to manage or cure genetic diseases, to create plants that produce protein based pharmaceuticals, antibiotics, antibodies or to change the flavor or nutritional content of agronomically important crops. This method is faster than current technologies being used to achieve these ends since only one line of plants, already known to express transgenes at a desired level, need be utilized.

References

Albert, H; Dale, EC; Lee, E; Ow, DW. 1995. Site-specific integration of DNA into wild-type and mutant *lox* sites placed in the plant genome. The Plant Journal 7:649-659.

Arencibia, A; Gentinetta, E; Cuzzoni, E; Castiglione, S; Kohli, A; Vain, P; Leech, M; Christou, P; Sala, F. 1998.Molecular analysis of the genome of transgenic rice (Oryza sativa L.) plants produced via particle bombardment or intact cell electroporation. Molecular Breeding, 4:99-109.

Assaad, FF; Tucker, KL; Signer, ER. 1993. Epigenetic repeat-induced gene silencing (RIGS) in Arabidopsis. Plant Mol Biol 22:1067-1085.

Becerra, SP; Koczot, F; Fabisch, P; Rose, JA. 1988. Synthesis of adeno-associated virus structural proteins requires both alternative mRNA splicing and alternative initiations from a single transcript. J. Virol. 62:2745-2754.

Bechtold, N; Ellis, J; Pelletier, G. 1993. *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. C R Acad Sci Paris Life Sci 316:1194-1199.

Berns, KI; Adler, S. 1972. Separation of two types of adeno-associated virus particles containing complementary polynucleotide chains. J. Virol. 9:394-396.

Berns, KI; Linden, RM. 1995. The cryptic life style of adeno- associated virus. BioEssays 17:237-245.

Berns, KI. 1996. *Parvoviridae* : The Viruses and Their Replication, *In* Fields, B.N., D. M. Knipe, P.M. Howley *et al*.(ed.), Fundamental Virology, 3rd ed. Lippincott-Raven Publishers, Philadelphia, PA. pp. 1026-1027.

Bevan, M. 1984. Binary Agrobacterium vectors for plant transformation. Nucleic Acids Research 12:8711-8721

Birnboim, HC; Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Research 7: 1513-1523.

Church, GM; Gilbert, W. 1995. Genomic Sequencing. Proc. Natl. Acad. Sci. USA, 81:1991-1995.

Clough, SJ; Bent, AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J., 16:735-743.

Daniell, H; Datta, R; Varma, S; Gray, S; Lee SB. 1998. Containment of herbicide resistance through genetic engineering of the cholorplast genome. Nature Biotechnol 16:345-348.

Dower, WJ. 1988. Transformation of *E.coli* to extremely high efficiency by electroporation. Mol. Biol. Reps. 6:3-4.

Fan, L; Dunckleyt, MG; Owen, JS; Dickson, G. 1998. Efficient coexpression and secretion of anti-atherogenic human apolipoprotein AI and lecithin-cholesterol acyltransferase by cultured muscle cells using adeno-associated virus plasmid vectors. Gene Ther 10:1434-1440.

Flotte, TR; Solow, R; Owens, RA; Afione S; Zeitlin, PL; Carter BJ. 1992. Gene expression from adeno-associated virus vectors in airway epithelial cells. Am J Respir Cell Mol Biol 7:349-356.

Finer, JJ; Finer, KR; Ponappa, T. 1999. Particle bombardment-mediated transformation, in *Current Topics in Microbiology and Immunology (Vol. 240), Plant Biotechnology: New Products and Applications* (Hammond, J., McGarvey, P.B. and Yusibov, V., eds), Springer-Verlag.

Gelvin, SB. (1998). The introduction and expression of transgenes in plants. Curr. Opin. Plant Biol. 1:142-148.

Giraud, C; Winocour, E; Berns, KI. 1994. Site-specific integration by adeno-associated virus is directed by a cellular DNA sequence. Proc. Natl. Acad. Sci. USA 91:10039-10043.

Giraud, C; Winocour, E; Berns, KI. 1995. Recombinant junctions formed by site-specific integration of adeno-associated virus into an episome. J. Virol 69:6917-6924.

Hansen, G; Wright, MS. 1999. Recent advances in the transformation of plants. Trends In Plant Sci 4:226-321.

Hermonat, PL; Muzyczka, N. 1984. Use of adeno-associated virus as a mammalian DNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells. Proc. Natl. Acad. Sci. U S A 1984 81:6466-6470

Holm, PB; Olsen, O; Schnorf, M; Brinch-Pedersen, H; Knudsen, S. 2000. Transformation of barley by microinjection into isolated zygote protoplasts. Transgenic Research, 9:21-32

Horsch, RB; Fry, JE; Hoffman, NL; Eichholtz, D; Rogers, SG; Fraley, RT. 1985. A simple and general method for transferring genes into plants. Science 227:1229-1231.

Im, DS; Muzyczka, N. 1990. Partial purification of adeno-associated virus Rep78, Rep68, Rep52, and Rep40 and their biochemical characterization. J. Virol. 66:1119-1128.

Kempin, SA; Liljegren, SJ; Block, LM; Rounsley, SD, Yanofsky, Mf and Lam, E. 2000. Targeted disruption in *Arabidopsis*. Nature. 389:802-803.

Kononov, ME; Bassuner, B; Gelvin, SB. 1997. Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. Plant J 11:945-957

Khoudi H; Laberge, S; Ferullo, JM; Bazin, R; Darveau, A; Castonguay, Y; Allard, G; Lemieux, R; Vzina, LP. 1999. Production of a diagnostic monoclonal antibody in perennial alfalfa plants. Biotechnol. Bioeng. 64:135-143.

Kotin, RM; Siniscalco, M; Samulski, RJ; Zhu, XD; Hunter, L; Laughlin, CA; McLaughlin, S;Muzyczka, N; Rocchi, M; Berns, KI. 1990. Site-specific integration by adeno-associated virus. Proc. Natl. Acad. Sci. USA 87:2211-2215.

Kotin, RM; Linden, RM; Berns, KI. 1992. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. EMBO J, 11:5071-5078.

Labow, MA; Hermonat, PL; Berns, KI. 1986. Positive and negative autoregulation of the adeno-associated virus type 2 genome. J. Virol. 60:515-524.

Lebkowski JS; McNally MM; Okarma TB; Lerch LB. 1988. Adeno-associated virus: a vector system for efficient introduction and integration of DNA into a variety of mammalian cell types. Mol Cell Biol 8:3988-3996

Li, LL; Dinauer, MC. 1998. Reconstitution of NADPH oxidase activity in human X-linked chronic granulomatous disease myeloid cells after stable gene transfer using a recombinant adeno-associated virus 2 vector. Blood Cells Mol Dis 4:522-538.

Linden, RM; Ward, PJ; Giraud, C; Winocour, E; Berns, KI. 1996. Site-specific integration by adeno-associated virus. Proc. Natl. Acad. Sci. USA, 93:11288-11294.

Linden, RM; Winocour E; Berns, KI. 1996. The recombination signals for adeno-associated virus site-specific integration. Proc. Natl. Acad. Sci. USA, 93:7966-7972.

Linden, RM; Berns, KI. 1997. Site-specific integration by Adeno-associated virus: a basis for a potential gene therapy vector (Editorial). Gene Therapy, 4:4-5.

Lodish, H; Berk, A; Zipursky, SL; Matsudaira, P; Baltimore, D; Darnell, JE. 1999. Cancer. from Molecular Cell Biology 4th Ed. W.H. Freeman and Company, NY, New York, pp 1270-1280.

Martin, T; Wohner, R-V; Hummel S; Willmitzer L; Frommer G. 1992. Using GUS gene as a reporter of gene expression. *In* GUS protocols: using the GUS gene as a reporter of gene expression. (Gallagher, SR; ed.) Academic Press, San Diego, CA, pp. 23-29.

Matzke, AJ; Chilton, MD. 1981. Site-specific insertion of genes into T-DNA of the Agrobacterium tumor-inducing plasmid: an approach to genetic engineering of higher plant cells. J. Mol. Genet 1:39-49.

Meyer, P. 1995. Understanding and controlling transgene expression. Trends in Biotechnology, 13:33-37.

Maio, Z; Lam, E. 1995. Targeted disruption of the TGA3 locus in *Arabidopsis thaliana.* The Plant J 7(2), 359-365.

Mishra, L; Rose, JA. 1990. Adeno-associated virus DNA synthesis. Virology 179:632.

Odell, JT; Russell, SH; in J. Paszkowski, ed. 1994. *Homologous Recombination and Gene Silencing in Plants.* Kluwer Academic Publishers, Dordrecht.

Peñarrubia, L; Kim, R; Giovannoni, J; Kim, SH; Fischer, RL. 1992. Production of the sweet protein monellin in transgenic plants. Bio/Technology 10:561-564.

Perlak FJ; Deaton RW; Armstrong TA; Fuchs RL; Sims SR; Greenplate JT; Fischhoff DA. 1990. Insect resistant cotton plants. Biotechnology 8:939-943.

Poirier, Y; Dennnis, DE; Klomparens, K; Somerville, C. 1992. Polyhydroxybutyrate, a biodegradable thermoplastic, produced in transgenic plants. Science 256:520-523.

Qing, CM; Fan, L; Lei, Y; Bouchez, D; Tourneur, C; Yan, L; Robaglia, C. 2000. Transformation of Pakchoi (*Brassica rapa L. ssp. chinensis*) by Agrobacterium infiltration. Mol. Breed. 6:67-72.

Recchia, A; Parks, RJ; Lamartina, S; Toniatti, C; Pieroni, L; Palombo, F; Ciliberto, G; Graham, F; Cortese, R; La Monica, N; Colloca, S. 2000. Site-specific integration mediated by a hybrid adenovirus/adeno-associated virus vector. Proc. Natl. Acad. Sci. USA 96:2615-2620.

Redmann, BE; Mendelson, E; Carter, BJ. 1989. Adeno-associated virus Rep protein synthesis during productive infection. J. Virol. 63:873-882.

Ruffing, M; Zentgraf, H; Kleinschmidt, JA. 1992. Assembly of virus-like particles by recombinant structural proteins of adeno-associated virus type 2 in insect cells. J. Virol. 66:6922-6930.

Sauer, B; Henderson, N. 1990. Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. New Biol. 2:441-449.

Seki M; Iida A; Morikawa H. 1999. Transient expression of the beta-glucuronidase gene intissues of Arabidopsis thaliana by bombardment-mediated transformation. Mol Biotechnol. 11:251-255.

Stam, M; Viterbo, A; Mol, JNM; Kooter, JM. 1998. Position-dependent methylation and transcriptional silencing of transgenes in inverted T-DNA repeats: Implications for posttranscriptional silencing of homologous host genes in plants. Molecular and Cellular Biology, 18:6165-6177

Stomp, AM. 1992. Histochemical localization of B-Glucronidase. *In* GUS protocols: using the GUS gene as a reporter of gene expression. (Gallagher, S.R., ed.) Academic Press, San Diego, CA, pp. 103-112.

Srivastava, A; Lusby, EW; Berns, KI. 1983. Nucleotide sequence and organization of the Adeno-associated virus 2 genome. J. Virol. 45:555-564.

Tattersall, P. 2000. The persistence of alien genomes. Proc. Natl. Acad. Sci. USA 97:6239-6241.

Ten Hoopen, R; Montijn, BM; Veuskens, JTM; Oud, OJL; Nanninga, N. 1999. The spatial localization of T-DNA insertions in petunia interphase nuclei: Consequences for chromosome organization and transgene insertion sites. Chromosome Research, 7:611-623.

Thrasher, AJ; de Alwis, M; Casimir, CM; Kinnon, C; Page, K; Lebkowski, J; Segal, AW; Levinsky, RJ. 1995. Functional reconstitution of the NADPH-oxidase by adeno-associated virus gene transfer. Blood 86:761-765.

Tratschin, JD; Miller, IL; Carter, BJ. 1984. Genetic analysis of adeno-associated virus: property of deletion mutants constructed in vitro and evidence for an adeno-associated virus replication function. J. Virol. 51:611-619.

Tratschin JD; Miller IL; Smith MG; Carter BJ. 1985. Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. Mol Cell Biol 5:3251-60.

Trempe, JP; Mendelson, E; Carter, BJ. 1987. Characterization of adeno-associated virus *rep* proteins in human cells by antibodies against rep expressed in *Escherichia coli*. Virology 161:18-28.

Van Montagu, M; Holsters, M; Zambryski, P; Hernalsteens, JP; Depicker, A; De Beuckeleer, M; Engler, G; Lemmers, M; Willmitzer, L; Schell, J. 1980. The interaction of Agrobacterium Ti-plasmid DNA and plant cells. Proc. R Soc. Lond. 210:351-365.

Vieweg, J; Boczkowski, D; Roberson, KM; Edwards, DW; Philip, M; Philip, R; Rudoll, T; Smith, C; Robertson, C Gilboa, E. 1995. Efficient gene transfer with adeno-associated virus-based plasmids complexed to cationic liposomes for gene therapy of human prostate cancer. Cancer Res. 55:2366-2372.

Ward, P; Dean, FB; O'donnell, ME; Berns, KI. 1998. Role of Adenovirus DNA-Binding Protein in In Vitro Adeno-Associated Virus DNA Replication. J. Virol. 72:420-427.

Weindler, FW; Heilbronn, R. 1991. A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication. J. Virol. 65:2476.

Weitzman MD; Kyostio SR; Kotin RM; Owens RA. 1994 .Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. Proc Natl Acad Sci U S A. 91:5808-5812.

Willmitzer, L; Depicker, A; Dhaese, P; De Greve, H; Hernalsteens, JP; Holsters, M; Leemans, J; Otten, L; Schroder, J; Schroder, G; Zambryski, P; van Montagu, M; Schell, J. The use of Ti-plasmids as plant-directed gene vectors. 1983. Folia Biol (Praha). 29:106-114.

Yakinoglu, AO; Heilbronn, R; Burkle, A; Schlehofer, JR; zur Hausen, H. 1988. DNA amplification of adeno-assicoated virus as a respoinse to cellular genotoxic stress. Cancer Res 48:3123-3129.

Yakobson, B; Koch, T; Winocour, E. 1987. Replication of adeno-associated virus in synchronized cells without the addition of a helper virus. J. Virol. 61:972-981.

Zambryski, P. 1992. Chronicles from the *Agrobacterium*-Plant cell DNA transfer story. Annu. Rev. Plant Physiol. Plant Mol. Biol. 43:465-490

Zhang, L; Wang, D; Fischer, H;. Fan, PD; Widdicombe, JH; Kan, YW; Dong, JY. 1998. Efficient expression of CFRT function with adeno-associated virus vectors that carry shortened CFTR genes. PNAS 95:10158-10163.

Zupan, JR; Zambryski, P. 1995. Transfer of T-DNA from Agrobacterium to the plant cell. Plant Physiol. 107:1041-1047.

Internet Sites:

The Wadsworth Center - http://www.wadsworth.org/databank/b19virus.htm

Curriculum Vitae

William J. Zabaronick
Department of Biology, Virginia Tech, Blacksburg, VA 24061-0406
117 Hearthstone Dr., Apt E., Blacksburg, VA 24060 (540) 951-5111, zab@vt.edu

Education

Master of Science; Biology Department, VPI&SU, May 2001
M. Lederman, Advisor
Title: Using A Mammalian Virus To Create Plants For Site-Specific Transgene Insertion

Bachelor of Science; Biochemistry, June 1993
Minor: Chemistry
VPI&SU, Blacksburg, Virginia

Experience

Graduate Research Assistant / Server Administrator
VPI&SU, Blacksburg, VA
Fall 1997 – Present
Server Administrator for General Biology Online Manual; duties include writing HTML, CGI, managing user accounts, maintaining server, overseeing two undergraduate work-study students

Graduate Teaching Assistant
VPI&SU, Blacksburg, VA
Freshman Biology Labs, Fall 96-Spring 97
Senior Level Virology, Fall 1998

Analytical Chemist, Quality Control
Wyeth-Lederle Vaccines and Pediatrics, Marietta, PA
February 1995 to June 1995
Analyzed various pharmaceuticals and in-process samples using HPLC,GLC, AA, IR and UV/VIS spectroscopy

Research Technician with Diego Restrepo, Ph.D.
Monell Chemical Senses Center, Philadelphia, PA
November 1993 to January 1995.
Expression profile of mRNA indicative of olfactory neurons using Southern blotting
Cycle sequencing of novel genes using ^{32}P and ^{33}P-ATP
Prepared media and other solutions, ordered supplies, maintained lab equipment, attended group meetings, taught coworkers new techniques

Funding

Sigma Xi Foundation Grant Proposal, Funded, July 1997

Awards

GTA Excellence Award, Spring 2000