

Results and Discussion

rep78 plant vector construction

Initial strategies to express *rep78* in *A. thaliana* involved cloning *rep78* into a plant expression cassette in plasmid pRTL (Topfer et al., 1987). Under the regulation of the CaMV 35S promoter and the Nos terminator, the *rep78* cassette was inserted into the Ti plasmid pBIB-HYG (Becker et al., 1992). This plant vector does not contain a reporter gene, and plants are selected by hygromycin resistance. However, no transformed plants were obtained, and an alternative expression strategy was employed.

The pCAMBIA series of plant vectors is designed for *Agrobacterium*-mediated transformation (Hajdukiewicz et al., 1994). pCAMBIA-3302 contains *GFP* with 5 restriction sites, and a GFP fusion protein may be produced via in-frame cloning into one of these upstream sites. This allows for quick characterization of gene expression and/or protein location. Regulatory sequences within the plasmid include the cauliflower mosaic virus 35S promoter and the plant Nos terminator for constitutive expression of *rep78-GFP*. As disarmed vectors, chromosomally-located *Agrobacterium* genes orchestrate insertion of the plasmid T-DNA into the plant genome. pCAMBIA-3302 was chosen to express AAV-2 *rep78* in *A. thaliana* (Fig. 4). This vector allows plant transformants to be selected by resistance to phosphinothricin -based herbicides (see Materials and Methods, *rep78* Plant Vector Construction), while bacterial transformants are selected by kanamycin resistance.

Samulski et al. (1987) created a recombinant plasmid carrying the infectious AAV-2 genome termed pSUB-201. No convenient restriction sites exist in pSUB-201 for *rep* excision, although a *Bgl*III site is present in the cloning region of pCAMBIA-3302.

Using *rep*-specific primers sRep3 and sRep4 that incorporate terminal *Bgl*III sites (Fig. 4), pSUB-201 was used as a template for *rep78* amplification. Figure 6 shows the expected 1866 bp *rep78* PCR product (Fig. 6, lane 2) with a 1 kb DNA standard (Fig. 6, lane 1). This amplification also incorporates three 5' nucleotides so as to provide in-frame cloning into the plant vector pCAMBIA-3302. Cloning of this *rep78* product into *Bgl*III-digested pCAMBIA-3302 yields pSdDan (Fig. 4). This strategy places *rep78* in frame with *GFP*, ensuring expression of a functional Rep78-GFP fusion protein. A glycine residue is introduced between Rep78 and GFP in the process.

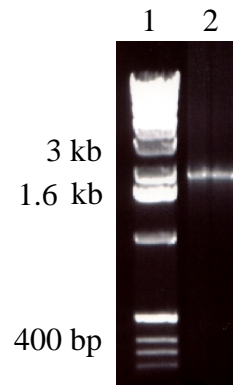


Figure 6. Amplification of *rep78* from pSUB-201. Lane 1: 1 kb DNA standards. Lane 2: *rep78*.

Internal *rep* primers Rep5 and Rep6 were used to confirm insertion of *rep78* into pCAMBIA-3302. Figure 7 shows the expected 330 bp *rep78* PCR product with pSdDan as template (Fig. 7, lane 4). The figure also shows the results of the same PCR using empty plasmid (Fig. 7, lane 2) and pSUB-201 (Fig. 7, lane 3) as template. With positive and negative controls in place, this amplification demonstrates that *rep78* has indeed been cloned into the vector. Lane 1 contains a 1 kb DNA standard (Fig. 7, lane 1).

GFP and *rep* primers GFPL and Rep5 were used to determine the orientation of *rep* in pCAMBIA-3302. Figure 7 shows PCR products with empty plasmid (Fig. 7, lane 5) and pSdDan (Fig. 7, lane 6) as template. The 1900 bp product in lane 6 consists of a 1500 bp *rep78* fragment and 400 bp *GFP* fragment amplified from the *rep78-GFP* fusion cassette. Were *rep78* inserted in the wrong orientation (3' -5') no product would result, as with the negative control (Fig. 7, lane 5). This amplification using a *rep78* primer and a *GFP* primer demonstrates that *rep78* has been inserted in the correct orientation upstream of *GFP* in the plant vector.

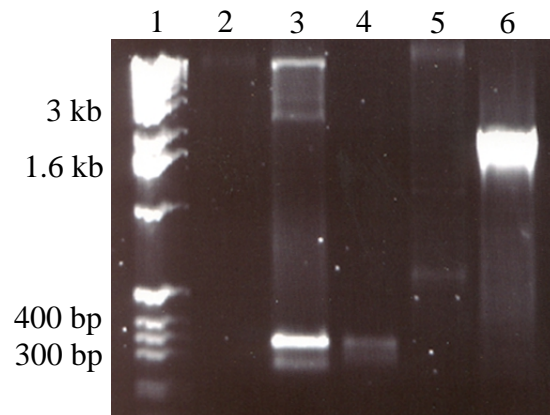


Figure 7. Amplification of *rep78* (lanes 2-4) and *rep78-GFP* (lanes 5-6) fragments. Lane 1: DNA standards. Lane 2: negative control; pCAMBIA-3302. Lane 3 positive control; pSUB-201. Lane 4: pSdDan. Lane 5: negative control; pCAMBIA-3302. Lane 6: pSdDan.

rep78* in *A. thaliana

Wild type and AAVS1-engineered *Arabidopsis* plants were transformed with *A. tumefaciens* GV3101 carrying pSdDan. Approximately 0.2-1% of the seedlings produced from these transformed plants survived herbicide selection. This phenotype suggests genomic integration of pSdDan T-DNA carrying the herbicide-resistance gene.

Plants that survived selection were tested for the presence of *rep78* in chromosomal DNA by PCR. Genomic DNA from this first generation of plants (T1) was extracted and used as template with *rep*-specific primers (see Materials and Methods, PCR Analysis of Transformed Plants). Two out of 14 WT plants (#2 and #11) tested positive for *rep* amplification. The data presented here is from plant #2, though the results apply to both plants. No AAVS1-*rep* plants were obtained. The full 1866 bp *rep* gene was amplified as five ~400 bp fragments. Figure 8 shows these PCR products with wild type (Fig. 8, lanes 2, 4, 6, 8, and 10) and plant #2 (Fig. 8, lanes 3, 5, 7, 9, and 11) genomic DNA as template. All amplified fragments were the expected size based on the location of the primers within *rep78*. There were no products formed in the wild type reactions (Fig. 8, lanes 2, 4, 6, and 8), demonstrating that the products formed in the reactions using plant #2 DNA as template (Fig. 8, lanes 3, 5, 7, and 9) are in fact *rep78* fragments and not *Arabidopsis* sequence. This PCR data suggests that this plant has indeed been stably transformed with *rep78*.

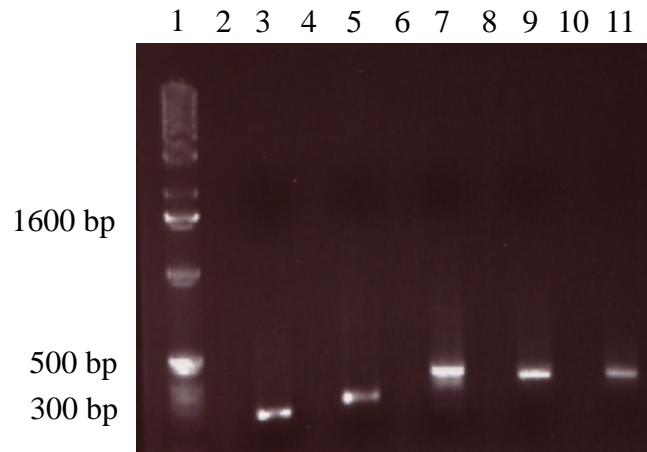


Figure 8. Amplification of *rep78* fragments from transgenic plant #2. Lane 1: DNA standards. Lanes 2, 4, 6, 8, and 10: wild type genomic DNA with primers Rep3 and Rep7, Rep5 and Rep6, Rep8 and Rep9, Rep10 and Rep11, and Rep12 and Rep4, respectively. Lanes 3, 5, 7, 9, and 11: plant #2 genomic DNA with the same primer sets.

Both transformed plants were also tested for *GFP* by PCR. Using *GFP* primers GFPU and GFPL, plants testing positive for *rep* also tested positive for a ~400 bp *GFP* fragment. Figure 9 shows PCR products with wild type (Fig. 9, lane 2) and plant #2 (Fig. 9, lane 3) genomic DNA as template. The ~400 bp product in lane 3 is of the correct size based on the location of the *GFP* primers. No product was formed in the amplification reaction using untransformed plant DNA as template (Fig 9, lane 2), demonstrating that the product in lane 3 is indeed *GFP*, and not *Arabidopsis*, sequence. The presence of *GFP* in the plant genome further confirms successful transformation with pSdDan.

It should be noted that the older leaves of plant #2 T1 generation exhibited curious mottling (Fig. 10). The T2 generation did not exhibit mottling. Normally associated with nutritional deficiency or infection, this lesion-like chlorosis was not observed with any other plants during the study. The possible pleiotropic effects of *rep78* expression in *Arabidopsis* are not known.

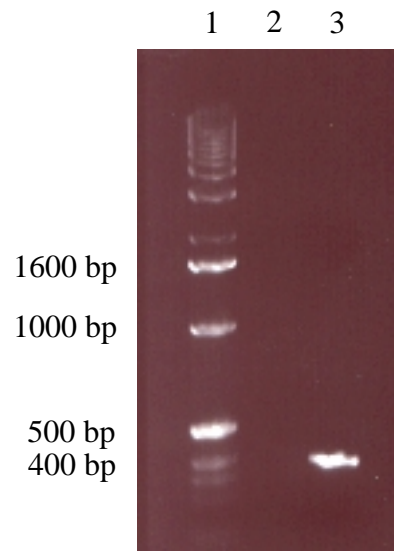


Figure 9. Amplification of *GFP* fragment from transformed plant #2. Lane 1: DNA standards. Lane 2: wild type genomic DNA with primers GFPL and GFPU. Lane 3: plant #2 genomic DNA with the same primers.



Figure 10. Chlorotic lesions as observed in the older leaves of the T1 generation of transformed plant #2.

rep78* sequence in *A. thaliana

The *rep78* fragments amplified from plant #2 genomic DNA (Fig. 8) were purified and used as templates for nucleotide sequencing reactions. 1856 out of the 1866 amplified bp (~99%) were sequenced and shown to have 100% identity to AAV-2 *rep78* (Fig. 11). This information confirms the presence of the complete *rep78* gene in the plant genome.

AAV-2 <i>rep78</i>	5 -ACGATCCCCAATATGCGGCTTCCGTCTTTCTGGGATGGGCCACG-3
Plant 2 sequence 1	5 -ACGATCCCCAATATGCGGCTTCCGTCTTTCTGGGATGGGCCACG-3
Plant 2 sequence 2	3 -ACGATCCCCAATATGCGGCTTCCGTCTTTCTGGGATGGGCCACG-5

Figure 11. Alignment of *rep78* (930 to 973 bp shown) and plant #2 sequence data.

Rep78 expression in *A. thaliana*

rep transcription

Our transformation strategy relies on a functional Rep enzyme in the plant system. Rep expression was therefore characterized in Rep-transformed plants. Qualitative mRNA assays such as RT-PCR indicate transcription of specific genes. RNA was isolated from plant #11 (T2 generation) and used as template with *rep*-specific primers. If *rep* is being transcribed in the plant, the message should reverse transcribe and amplify via RT-PCR; wild type and non-*rep* transcribing plants should not yield RT-PCR product.

Figure 12 shows these RT-PCR products with wild type (Fig. 12, lanes 2, 4, and 6) and plant #11 (Fig. 12, lanes 3, 5, and 7) RNA as template. The sizes of the products in lanes 5 and 7, and the lowest molecular weight product in lane 3, correspond to those

of the expected amplified *rep* fragments (~400 bp, see Fig. 8) based on the location of the primers within *rep78*. There were no products formed in the wild type reactions (Fig. 12, lanes 2, 4, and 6), demonstrating that the products formed in the reactions using plant #11 RNA as template (Fig. 12, lanes 3, 5, 7) are in fact based on *rep* transcript and not *Arabidopsis* messages. This RT-PCR data suggests that the *rep-GFP* fusion is being transcribed in plant #11.

This hypothesis is further confirmed as plant #11 also tested positive for *GFP* transcription by RT-PCR. Figure 13 shows RT-PCR products with *GFP*-expressing plant (Haseloff et al., 1997; a gift of Brenda Winkel; Fig. 13, lane 2) and transformed plant #11 (Fig. 13, lane 3) RNA as template. The size of both products corresponds to the expected size of the amplified *GFP* fragment (400 bp, see Fig. 9). The product of the reaction with *GFP*-expressing plant RNA (Fig. 13, lane 2) is of the same size as the transformant RNA product (Fig. 13, lane 3), demonstrating that the product in lane 3 is in fact based on *GFP* transcript and not *Arabidopsis* transcript. This RT-PCR data further supports *rep-GFP* transcription in transformed plant #11.

RT-PCR is sensitive to genomic contamination. In order to show that the products were amplified from reverse transcribed RNA, and not DNA, standard PCR reactions with Taq polymerase were also performed with the isolated RNA. Without reverse transcriptase, no products should form if the RNA sample is devoid of genomic DNA. No PCR products formed in reactions containing plant #11 RNA template and *Rep* or *GFP* primer pairs (Fig. 14 A, lanes 3 and 5; B, lane 3).

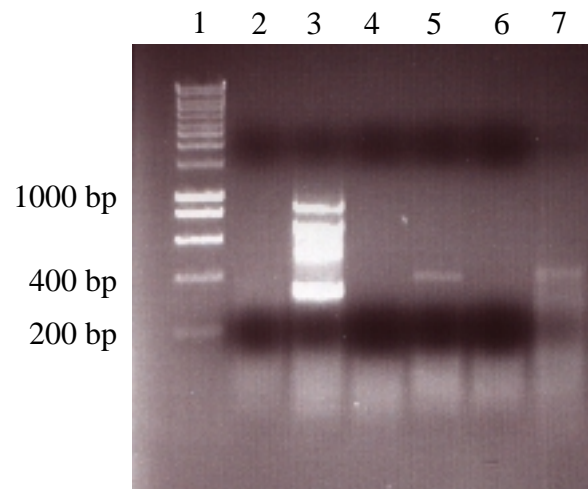


Figure 12. *rep* RT-PCR from transgenic plant #11. Lane 1: DNA standards. Lanes 2, 4, and 6: wild type RNA with primers Rep5 and Rep6, Rep10 and Rep11, and Rep12 and Rep4, respectively. Lanes 3, 5, and 7: plant #11 RNA with the same respective primer sets.

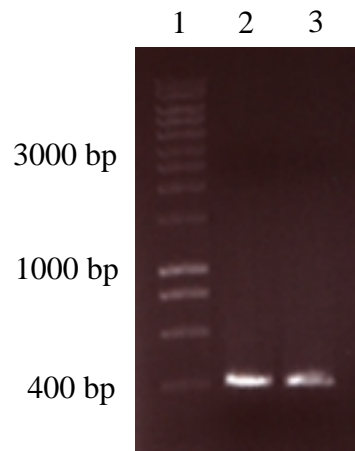


Figure 13. *GFP* RT-PCR from transgenic plant #11. Lane 1: DNA standards. Lane 2: *GFP*-expressing plant RNA with primers GFPU and GFPL. Lane 3: plant #11 RNA with the same primers.

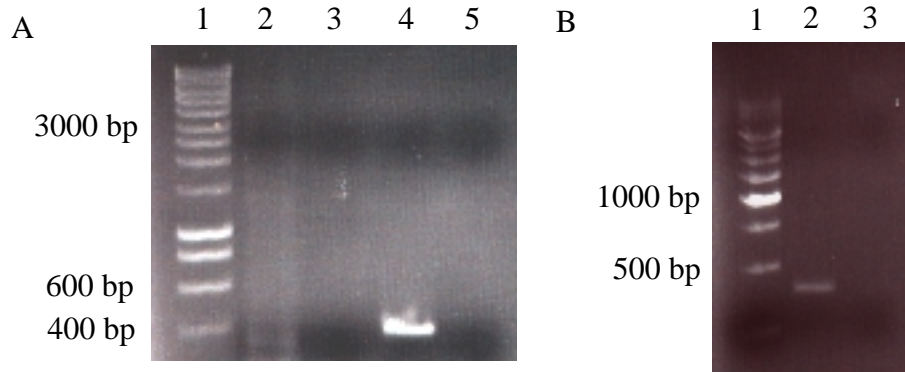


Figure 14. *rep* and *GFP* PCR from transgenic plant RNA and positive templates. A. Lane 1: DNA standards. Lane 2: *E. coli* containing pSdDan with primers Rep5 and Rep6. Lane 3: transgenic plant #11 RNA with the same primers. Lane 4: *E. coli* containing pSdDan with primers Rep10 and Rep11. Lane 5: transgenic plant #11 RNA with the same primers. B. Lane 1: DNA standards. Lane 2: *E. coli* containing pSdDan with primers GFPU and GFPL. Lane 3: transgenic plant #11 RNA with the same primers.

Rep translation

Western blots using Rep and GFP-specific antibodies were performed in an effort to detect the Rep-GFP fusion protein in the T2 generation of transformed plants. The encouraging RT-PCR data suggest *rep* is being transcribed, therefore Western blots might reveal detectable levels of Rep protein being translated. However, neither anti-Rep nor anti-GFP antibody could detect the fusion protein in transformed plant #2 or #11.

Figure 15 shows the GFP-specific antibody reaction with purified GFP (Fig. 15, lane 2), wild type *Arabidopsis* lysate (Fig. 15, lane 3), transformed plant #2 lysate (Fig. 15, lane 4), and transformed plant #11 lysate (Fig. 15, lane 5). Shown also is the Rep78/68-specific antibody reaction with AAV-infected HeLa cell lysate (Fig. 15, lane 6), wild type *Arabidopsis* lysate (Fig. 15, lane 7), transformed plant #2 lysate (Fig. 15, lane 8) and transformed plant #11 lysate (Fig. 15, lane 9). The appropriately sized GFP and Rep bands (Fig. 15, lanes 2 and 6, respectively) indicate that the antibody system was able to detect target protein (the Rep protein is deceptively large, as the lower numbered lanes electrophoresed ahead of the higher lanes). However, no protein bands were evident in the transformed plant lysate lanes (Fig. 15, lane 4, 5, 8, and 9). GFP fluorescence was not observed by fluorescent microscopy in the transformed plants through the T2 generation (data not shown). Consistent with the Western blot data, this suggests that the Rep-GFP fusion protein is not produced at detectable levels. This suggests that, while the *rep-GFP* fusion gene may be transcribed, it is not being translated to a detectable level in the transformed plants (see Conclusions for discussion).

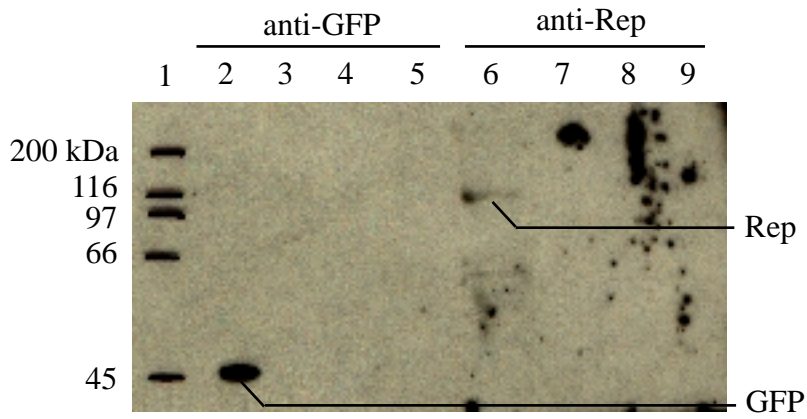


Figure 15. Western blot of *rep/GFP* transformed plants. Lanes 2-5 use primary antibody against GFP. Lanes 6-9 use primary antibody against Rep78/68. Lane 1: protein standards. Lane 2: purified GFP. Lane 3: wild type *Arabidopsis* lysate. Lane 4: transformed plant #2. Lane 5: transformed plant #11. Lane 6: AAV-2+Ad-5 infected HeLa cell lysate. Lane 7: wild type plant lysate. Lane 8: plant #2. Lane 9: plant #11.

Rep78 expression *in vitro*

Without access to purified Rep protein, a positive control was needed for use with transformed plant Western blot data (Fig. 15, lane 6). Therefore, AAV-infected HeLa cells were used to ensure the efficacy of Rep antibodies. A productive AAV-2 HeLa cell infection can be achieved by co-infection with adenovirus type 5. Doubly infected cells were harvested at 8 hours (Fig. 16, lane 3), 12 hours (Fig. 16, lane 4), 24 hours (Fig. 16, lane 5), and 36 hours (Fig. 16, lane 6) post infection. Cells infected with Ad-5 only were also harvested at 36 hours and were used as a negative control for AAV-2 Rep78 synthesis (Fig. 16, lane 2). Cells infected with Ad-5 alone show diffuse reaction to the Rep78/68-raised monoclonal antibody and one band at ~130 kDa. Most importantly, the 78 and 68 kDa Rep bands were not visible in this Ad-5 only sample, while they were clearly visible in the Ad-5 and AAV-2 infected sample (Fig. 16, lane 6). While nonstructural Rep proteins have been detected as early as 4 hours post infection (Redemann et al., 1989), figure 16 shows the detection of Rep78/68 by monoclonal antibody 36 hours post infection (Fig. 16, lane 6). The antibody's ability to detect Rep78 is demonstrated here, and it will be used to characterize Rep78 expression in transformed plants.

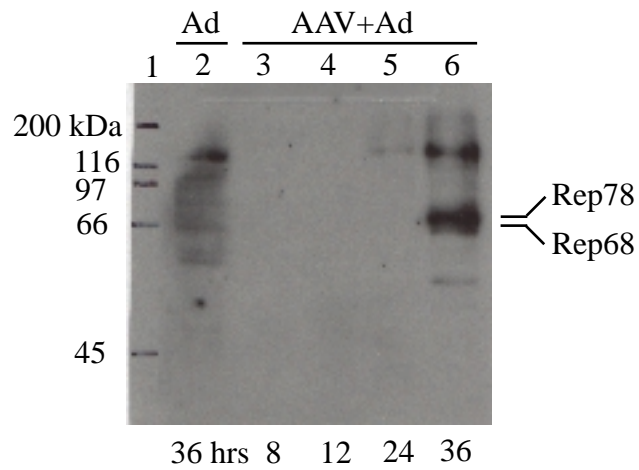


Figure 16. Western blot of AAV-2 and Ad-5 infected HeLa cell lysate. Lane 1: protein standards. Lane 2: Ad-5 alone, 36 hours post infection. Lane 3: AAV-2+Ad-5, 8 hours post-infection. Lane 4: AAV-2+Ad-5, 12 hours post infection. Lane 5: AAV-2+Ad-5, 24 hours post infection. Lane 6: AAV-2+Ad-5, 36 hours post infection.

Conclusions

In the absence of helper virus and during subsequent low levels of replication, adeno-associated virus type 2 selectively integrates into a region termed AAVS1 on human chromosome 19 (Berns and Linden, 1995). Rep78/68 is a viral enzyme that catalyzes this reaction (Linden et al., 1996). Since AAVS1 shares homology with AAV-2 termini, Rep binds and nicks each strand (Im and Muzyczka, 1990). Integration is believed to occur as the cellular replication machinery switches templates between the proximal AAV-2 and AAVS1 (Linden et al., 1996). Consistent with this function are Rep's biochemical activities, including site-specific endonuclease activity, DNA binding, and helicase activity (Berns and Linden, 1995). This mechanism has been employed for non-homologous, targeted integration in mammalian cell culture and for gene therapy studies (Ponnazhagan et al., 2001; Larson et al., 2001). We hope to develop a novel plant transformation strategy using this recombination technology. To this end, the goal of this project is to express AAV-2 Rep78/68 in plants.

rep78 was introduced into plant vector pCAMBIA-3302, producing a *rep78-GFP* fusion cassette (pSdDan, Fig. 4). *A. tumefaciens* harboring pSdDan was used to transform wild type *A. thaliana*. The presence of *rep78* in the plant genome of two plant lines was demonstrated by PCR (Fig. 8). Sequence analysis further confirmed the introduction of the complete *rep78* gene into *Arabidopsis* (Fig. 11). RT-PCR demonstrated *rep* transcription in one transformed plant line (Fig. 12), but Western analysis failed to detect Rep-GFP protein in the T2 generation of either transformed plant line.

There are a number of potential explanations for the lack of detectable protein in the plants. The *rep-GFP* gene may simply be expressed at levels too low to detect. This may be due to the location of the gene(s) within the plant genome, yielding the sequence susceptible to position effects and inefficient expression (Jin et al., 2002). The location(s), orientation, and copy number of *rep* within the plant genome have not been investigated.

The codon bias of *A. thaliana*, may be a factor in the low *rep* expression levels. Transgenes are often reconstructed with preferred codons to accommodate the host's translational machinery (Pan et al., 1994). CUG (Leu) and CCC (Pro) are used four times less frequently in *Arabidopsis* than in humans, where AAV-2 Rep is normally synthesized. GCC (Ala) is used nearly three times less frequently. All other codons are used at comparable frequencies. Each of these three codons comprises about 6% total of those in Rep78. Therefore, it seems unlikely that a lack of available tRNAs in the plant would account for undetectable levels of Rep, though each individual codon bias could theoretically compound and make for inefficient translation.

Post-transcriptional gene silencing would also explain the lack of Rep protein in the transformed plants. Believed to be a defense mechanism against pathogens such as RNA viruses, post-transcriptional gene silencing is often a problem with transgene expression in plants (Matzke et al., 2002). Small plant RNAs mediate homology-dependent dsRNA formation and degradation, and methylation of the foreign DNA. This commonly observed but poorly understood silencing mechanism offers an attractive explanation (Chicas et al., 2001), as transcript was present but no Rep protein was detected.

This problem of Rep expression in *Arabidopsis* should be approached several ways. Firstly, the two existing transformed plant lines should be maintained in order to develop plants homozygous for the *trans* sequence. The plants are currently second generation, and re-activation of silenced transgenes has been observed upon passage through multiple generations (Reddy et al., 2000), or under non-selective conditions (Dominguez et al., 2002). Secondly, continued screening for optimal transformants may yield a competent Rep-expressing plant. If the complication is not *rep* sequence dependent, a plant lacking aberrant position effects may have yet to be identified.

It is assumed that the *rep-GFP* fusion would yield a biochemically functional Rep enzyme, as many GFP fusion proteins have been successfully expressed in plants (Goodin et al., 2002). However, it is entirely possible that the Rep-GFP protein did not fold correctly. It is possible, too, that the fusion protein is simply not stable in the plant system. The fact that only non-expressing *rep* plant transformants were obtained suggests that Rep78 may be cytotoxic. Rep78/68 expression has been shown to disrupt the cell cycle and induce apoptosis in animal cell culture (Schmidt et al., 2000; Zhou et al., 1999). How Rep induces apoptosis is currently under investigation. Though there are distinct differences in terms of how plants and animals regulate programmed cell death, there are also homologous key players within each system (Lam et al., 2001), and it is worth noting the possibility of a conserved cytotoxic response to Rep. This issue may be addressed by employing a transient or an inducible expression assay in which Rep's effect on plant cells may be observed directly. If Rep toxicity is indeed established, these expression systems may also be used to generate non-toxic Rep protein levels sufficient for integration.

Wild type and AAVS1 *Arabidopsis* were transformed with *rep* alone, but no transformants were obtained (data not shown). Rep is not believed to have caused chromosomal rearrangement at AAVS1 loci and plant death, though continued transformation and screening of wild type and AAVS1 plants seems sensible.

Future directions

The work presented here is but a piece of a much larger, exciting project. Currently, no targeted transformation methods exist to circumvent variable transgene expression levels in plants, and we aim to devise a recombinant AAV-2 based system that will dissolve such problems. Three simple components are required for site-specific integration: 1) AAV termini (ITR) flanked transgene, 2) Rep78 or 68, and 3) AAVS1 engineered into the plant genome.

We have previously engineered plants containing AAVS1 (Zabaronick, personal communication) and an ITR-flanked *GFP* transgene (Fisher, personal communication). Crossing the AAVS1 plants with the newly engineered Rep78-expressing plants would yield an *Arabidopsis* suitable for experimental transformation with the ITR-flanked transgene. Multiple AAVS1 and Rep plant self-crosses will yield homozygous lines. These two lines may then be crossed and selected through several generations to generate a line homozygous for each trait. Rep should insert the transgene into AAVS1, and experiments would be designed to confirm this. Initially, the ITR-transgene cassette may be introduced any number of ways, perhaps by standard *Agrobacterium* infiltration or, in order to avoid complications caused by random integration, by a viral plant vector transient expression system (Shiboleth et al., 2001). Demonstration of targeted insertion would be the primary objective at this developmental stage.

One foreseeable problem of our strategy involves the excision, or rescue, of the integrated transgene. Normal AAV-2 infection sees Rep78/68 down-regulating AAV promoters and producing a latent infection in the absence of helper virus, with replication-dependent rescue of AAV possible with future co-infection with helper virus. It is theoretically possible that an unregulated Rep may nick the ITR regions in the plant system, resulting in chromosomal rearrangement or, in a very unlikely event, excision and replication of the AAV and transgene sequence. In this case, Rep expression may be controlled via an inducible promoter or a transient expression system.

We are optimistic as to the efficacy and reliability of our proposed plant transformation strategy. The Rep enzyme should be properly processed and functional in plants. We foresee the target site AAVS1 being engineered into optimal expression regions of plant genomes, with simple, site-specific insertion of any gene of interest. Plant research experimental design and agricultural biotechnology will hopefully benefit from this novel approach to site-specific integration.