THE CONSTRUCTION OF AN EXPRESSION VECTOR FOR THE TRANSFORMATION OF THE GRAPE CHLOROPLAST GENOME

Julia Robson

Thesis presented in partial fulfillment of the requirements for the degree of Master of Science at the Department of Genetics, University of Stellenbosch.

December 2003

Supervisor: Prof JT Burger

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

J Robson

ABSTRACT

The genetic information of plants is found in the nucleus, the mitochondria, and the plastids. The DNA of plastids is comprised of multiple copies of a double-stranded, circular, prokaryotically–derived genome of ~150 kb. The genome equivalents of plastid organelles in higher plant cells are an attractive target for genetic engineering as high protein expression levels are readily obtained due to the high genome copy number per organelle. The resultant proteins are contained within the plastid organelle and the corresponding transgenes are inherited, in most crop plants, uniparentally, preventing pollen transmission of DNA.

Plastid transformation involves the uniform modification of all the plastid genome copies, a process facilitated by homologous recombination and the non-Mendelian segregation of plastids upon cell division. The plastid genomes are in a continuous state of inter- and intra-molecular exchange due to their common genetic complement. This enables the site-specific integration of any piece of DNA flanked by plastid targeting sequences, via homologous recombination. The attainment of homoplasmy, where all genomes are transformed, requires the inclusion of a plastid-specific selectable marker. Selective pressure favouring the propagation of the transformed genome copies, as well as the random segregation of plastids upon cell division, make it feasible to acquire uniformity and hence genetic stability. From this, a complete transplastomic line is obtained where all plastid genome copies present are transgenic, having eliminated all wild-type genome copies.

The prokaryotic nature of the chloroplast genetic system enables expression of multiple proteins from polycistronic mRNAs, allowing the introduction of entire operons in a single transformation. Expression cassettes in vectors thus include single regulatory elements of plastid origin, and harbour genes encoding selectable and screenable markers, as well as one or more genes of interest. Each coding region is preceded by an appropriate translation control region to ensure efficient translation from the polycistronic mRNA.

The function of a plastid transformation vector is to enable transfer and stable integration of foreign genes into the chloroplast genomes of higher plants. The expression vector constructed in this research is specific for the transformation of the grape chloroplast genome. *Vitis vinifera* L., from the family, Vitaceae, is the choice species for the production of wine and therefore our target for plastid transformation. All chloroplast derived regulatory elements and sequences included in the vector thus originated from this species.

OPSOMMING

Die genetiese inligting van plante word gevind in die kern, die mitochondria, en die plastiede. Die DNA van plastiede bestaan uit veelvuldige kopieë van 'n ~150 kb dubbelstring, sirkulêre genoom van prokariotiese oorsprong. Die genoomekwivalente van plastiede in hoër plante is 'n aantreklike teiken vir genetiese manipulering, aangesien die hoë genoom kopiegetal per organel dit moontlik maak om gereeld hoë vlakke van proteïenuitdrukking te verkry. Hierdie proteïene word tot die plastied beperk, en die ooreenstemmende transgene word in die meeste plante sitoplasmies oorgeërf, sonder die oordrag van DNA deur die stuifmeel.

Plastied transformasie behels die uniforme modifikasie van al die plastied genoomkopieë, 'n proses wat deur homoloë rekombinasie en die nie-Mendeliese segregasie van plastiede tydens seldeling gefasiliteer word. As gevolg van die gemeenskaplike genetiese komplement, vind aanhoudende inter- en intra-molekulêre uitruiling van plastiedgenome plaas. Dit maak die setel-spesifieke integrasie, via homoloë rekombinasie, van enige stuk DNA wat deur plastied teikenvolgordes begrens word, moontlik. Vir die verkrying van homoplasmie, waar alle genome getransformeer is, word die insluiting van 'n plastiedspesifieke selekteerbare merker benodig. Seleksiedruk wat die vermeerdering van die getransformeerde genoomkopieë bevoordeel, en die lukrake segregasie van plastiede tydens seldeling, maak dit moontlik om genetiese stabiliteit en uniformiteit van die genoom te verkry. Dit kan op sy beurt tot die verkryging van 'n volledige transplastomiese lyn lei, waar alle aanwesige plastiedgenome transgenies is, en wilde tipe genoomkopieë geëlimineer is.

Die prokariotiese aard van die chloroplas genetiese sisteem maak die uitdrukking van veelvuldige proteïene vanaf polisistroniese mRNAs moontlik, wat die toevoeging van volledige operons in 'n enkele transformasie toelaat. Uitdrukkingskassette in vektore bevat dus enkel regulatoriese elemente van plastied oorsprong, gene wat kodeer vir selekteerbare en sifbare merkers, asook een of meer gene van belang (teikengene). Voor elke koderingsstreek, is daar ook 'n toepaslike translasie beheerstreek om doeltreffende translasie vanaf die polisistroniese mRNA te verseker.

Die funksie van 'n plastied transformasie vektor is om die oordrag en stabiele integrasie van transgene in chloroplasgenome van hoër plante moontlik te maak. Die uitdrukkingsvektor wat in hierdie studie gekonstrueer is, is spesifiek vir die transformasie van die druif chloroplasgenoom. *Vitis vinifera* L., van die familie Vitaceae, is die voorkeur species vir die produksie van wyn, en daarom die teiken vir plastied transformasie. Alle chloroplast-afgeleide regulatoriese elemente en volgordes wat in hierdie vektor ingesluit is, het hul oorsprong vanaf *Vitis vinifera* L.

ABBREVIATIONS

A	adenine
aadA	spectinomycin resistance gene
badh	betaine aldehyde dehydrogenase gene
bp	base pairs
C	cytosine
cDNA	complementary DNA
ctDNA	chloroplast deoxyribonucleic acid
dH_20	distilled water
DNA	deoxyribonucleic acid
G	guanine
g	gram
gfp	green fluorescent protein gene
LB	Luria Bertani
LF	grape left targeting sequence
M	molar
MCS	multiple cloning site
ml	millilitre
mM	millimolar
mRNA	messenger ribonucleic acid
PCR	polymerase chain reaction
P*rrn*	grape 16S ribosomal RNA operon promoter
ptDNA	plastid deoxyribonucleic acid
rbcL 5'TCR	rubisco large subunit 5'translation control region
rbs	ribosome binding site
RE	restriction endonuclease
RF	grape right targeting sequence
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal ribonucleic acid
T	thymine
T_A	annealing temperature
T*psb*A	grape photosystem II terminator
TSP	total soluble protein
U	uracil
VCT	*Vitis* chloroplast transformation

ACKNOWLEDGEMENTS

I would like to thank all those who assisted me in this research, particularly my supervisor, Prof JT Burger for his constant support and encouragement.

Thank you to the members of the *Vitis* lab, MN Gardner and M-J Freeborough for technical assistance.

Thanks to the Harry Crossley Foundation, Winetech and THRIP for financial support. Thanks also to the Rockefeller Foundation for awarding me an international travel grant to attend the Plant and Animal Genome X conference in San Diego, California, in January last year. It was there that I met Prof P Maliga and I have since relied on his invaluable guidance throughout my project.

CONTENTS

1. LITERATURE REVIEW

1.1 INTRODUCTION

The genetic information of plants is distributed among three cellular compartments: the nucleus, the mitochondria, and the plastids. The latter two are believed to have originated from formerly free-living eubacteria: mitochondria from α-proteobacteria and plastids from cyanobacteria. The evolutionary conversion of prokaryotes into contemporary cellular organelles can be explained by the endosymbiosis theory: a bacterium was engulfed by a pre-eukaryotic host cell and, instead of being digested, became integrated into the host cell's metabolism. The gradual integration of the acquired endosymbionts was accompanied by the loss of dispensable or redundant genetic information and the translocation of genetic information, particularly from the endosymbiont to the host genome. Mitochondrial and plastid genomes are therefore greatly reduced and contain only a small proportion of the genes that their free-living ancestors had possessed (Hager & Bock, 2000; Bock, 2001; Heifetz & Tuttle, 2001).

The plastid genome is a circular, double-stranded chromosome of 120-160 kb in size and contains approximately 130 genes (Sugiura, 1992; Bock, 2001; Ruf *et al.*, 2001; Maliga, 2002). The genome contains an inverted repeat, the segments of which are separated by one large and one small single-copy region. All plastid differentiation types distributed ubiquitously throughout the differentiated plant cells contain identical copies of this genome. This is due to the organelles' shared biogenesis from undifferentiated proplastids present predominantly in the meristematic tissue. Leaves and green tissues contain photosynthetic chloroplasts; mature fruits and flowers contain pigmented chromoplasts; tubers and other storage organs contain amyloplasts or elaioplasts; and other non-green tissues including roots, contain leucoplasts (Heifetz, 2000; Bock, 2001; Heifetz & Tuttle, 2001). An important feature of plastid organelles is the extraordinarily high copy number of genomes per cell. A single leaf cell may contain dozens or even hundreds of chloroplasts, each harbouring several copies of the genome, which can amount to a ploidy level of up to 10 000 genome equivalents per cell (Bock & Hagemann, 2000; Bock, 2001). Chloroplasts evolved to become the energy transduction and metabolic centres of the plant cell, and of all plastid types, possess the highest rates of transcription and translation (Heifetz & Tuttle, 2001). This feature makes leaf chloroplasts suitable for the production of large amounts of recombinant proteins.

The development of transgenic technologies enabling the genetic transformation of the plastid genome has made feasible the introduction and expression of novel coding regions from engineered plastid genomes (Daniell & Dhingra, 2002; Daniell *et al.*, 2002). The characteristics of plastid transformation

reflect both the prokaryotic origin of plastid organelles and offer several advantages over the conventional method of nuclear transformation (Bock & Hagemann, 2000; Hager & Bock, 2000; Heifetz, 2000; Bock, 2001; Heifetz & Tuttle, 2001; Daniell & Dhingra, 2002; Daniell *et al.*, 2002; Maliga, 2002; Maliga, 2003).

1.2 PLASTID TRANSFORMATION IN HIGHER PLANTS

Plastid transformation involves the targeting of foreign genes to the plastid's double-stranded, circular DNA genome instead of the plant nuclear chromosomal DNA. Genetic engineering of the plastid genome is a particularly challenging task because of the enormously high copy number of genomes per organelle, and the absolute requirement to uniformly transform all genome copies to attain homoplasmy, and therefore genetic stability (Bock & Hagemann, 2000; Hager & Bock, 2000; Heifetz, 2000; Bock, 2001; Heifetz & Tuttle, 2001; Daniell & Dhingra, 2002; Daniell *et al.*, 2002; Maliga, 2002; Maliga, 2003).

The successful transformation of the plastid genomes of higher plants is dependent on a number of factors. Firstly, a method is required to enable delivery of foreign DNA through the cell wall, the plasma membrane, and the double membrane of the plastid. Following delivery into the plastid organelle, a mechanism is required to enable integration of that foreign DNA into the plastid genome. The attainment of homoplasmy and the ability to propagate homoplasmic cells requires a plastid-specific selectable marker gene and a highly regenerable tissue culture system, respectively.

1.2.1 Delivery of foreign DNA into plastid genomes

The development of a ballistic device termed the particle gun enabled the first successful chloroplast transformation of *Chlamydomonas reinhardii*, a unicellular green alga that possesses a single, large chloroplast (Boynton *et al.*, 1988). This system involves the coating of inert microscopic heavy metal particles (gold or tungsten) with nucleic acids, and the subsequent helium-driven acceleration of these particles into recipient target cells. Boynton *et al.* (1998) used photosynthetically incompetent *C. reinhardii* mutants carrying defective alleles of the chloroplast *atpB* gene in order to complement this allele with the cognate wild-type *atpB* allele, thus restoring photoautotrophic growth. Stable chloroplast transformants were obtained in which the mutant *atpB* allele had been replaced by the wild-type gene, as present in the transformation vector, via homologous recombination. Homologous recombination has since been the principle mode of insertion of foreign sequences into the plastid genome of higher plants. Therefore, any piece of DNA to be inserted into the plastid genome is flanked by sequences derived from the target region ptDNA.

The same particle gun-mediated transformation commonly referred to as the biolistic (biological + ballistic) technique was later used for the transformation of plastids in the higher plant, *Nicotiana tabacum* (Svab *et al.*, 1990; Hager & Bock, 2000). Following biolisitic bombardment, sterile leaf discs are placed on selective plant regeneration medium enabling exclusive propagation of those cells harbouring transformed plastid genome copies. After 4-6 weeks of incubation, formation of small green calli can be observed from which resistant shoots will form. Initially these resistant shoots (primary chloroplast transformants) are heteroplasmic, and are therefore subjected to additional rounds of selection and regeneration, to eliminate all wild-type genome copies. Homoplasmic transplastomic plants are then rooted on phytohormone-free medium, and subsequently transferred to the soil and maintained under greenhouse conditions.

The numerous particles involved in biolistic bombardment make this method relatively efficient. The ability to use different explant material in biolistic bombardment, and the relatively rapid regeneration of transformed tissues are advantages that make this the method of choice, despite the high costs involved. An alternative method involving the chemical treatment of protoplasts with polyethylene glycol, in the presence of transforming DNA, has also proved successful in generating stable chloroplast transformants in tobacco (O'Neill *et al.*, 1993). This technique, however, appears less efficient than particle bombardment, despite the small chance of a particle hitting a chloroplast in a nondestructive manner (Daniell, 1999b; Heifetz & Tuttle, 2001; Van Bel *et al.*, 2001).

A novel approach in the delivery of foreign DNA to the plastid genome involves direct microinjection into plastid organelles using the galinstan expansion femtosyringe (Daniell, 1999b; Knoblauch *et al.*, 1999; Van Bel *et al.*, 2001). This technique relies on the heat-induced expansion of galinstan, which results in the controlled expulsion of DNA into the plastid through a capillary tip with a diameter of about 0.1 μm. This tip diameter reduces damage to the cell in that cellular contents are not forced back into the capillary upon delivery, a process that results in the instantaneous release of turgor pressure. Individual chloroplasts in tobacco mesophyll cells have successfully been injected using plasmid DNA carrying the GFP gene under control of a chloroplast ribosomal RNA promoter (Knoblauch *et al.*, 1999).

1.2.2 Integration of foreign sequences into the plastid DNA

Plastid transformation is based on the integration of transgenes via homologous recombination between the vector and endogenous plastid nucleotide sequences (Bock & Hagemann, 2000; Hager & Bock, 2000; Bock, 2001; Heifetz & Tuttle, 2001). Homologous recombination in plastids is a highly efficient process that facilitates the targeted integration of transgenes flanked by native plastid target sequences, and the subsequent exchange to other wild-type genome copies. Initially, only one (or at

most a few) out of several thousand copies of the plastid genome in a single leaf cell is successfully transformed. The primary plastid transformant therefore contains a mixture of both transformed and wild-type genome copies, a situation termed heteroplasmy. The ability to attain complete homoplasmy, a situation characterized by the uniform transformation of all genome copies, is made feasible due to the non-Mendelian inheritance of plastid organelles. Plastid DNA is organized into nucleoids each harbouring 5-10 genome copies, contributing to the total ~100 copies per organelle. Segregation of plastids upon cell division and the partitioning of the enclosed plastid nucleoids (the segregating units upon organelle division), are both random processes. Heteroplasmy is therefore genetically unstable, and in the absence of selective pressure, typically results in rapid and dramatic changes in the relative ratios of the two genome types. In order to obtain a stable transplastomic plant, all residual wild-type genome copies are eliminated through repeated cycles of plant regeneration on synthetic media containing high concentrations of the selecting antibiotic. In this way, a strong selective pressure is maintained which favours high expression levels of the transformed plastid DNA molecules, thereby gradually driving genome sorting to the accumulation of the transplastomes.

1.2.3 Plastid-specific selectable marker genes

Plastid transformation is dependent on the use of plastid-specific selectable marker genes that facilitate complete segregation of transformed genome copies at the expense of non-transformed ones (Bock & Hagemann, 2000; Hager & Bock, 2000; Bock, 2001). The first chloroplast experiments in tobacco used a mutant allele of the plastid 16S ribosomal RNA gene as a recessive selectable marker (Svab *et al.*, 1990). This gene was engineered by introducing two point mutations that conferred resistance to the aminoglycoside antibiotics, spectinomycin and streptomycin; both of which act as prokaryotic translational inhibitors and are thus plastid-specific. However, only those plastid genome copies that incorporated this mutant allele, in place of the wild-type copy, transcribed ribosomes resistant to the respective antibiotics. Consequently, the transformation frequencies were very low.

Dominant selectable marker genes encoding antibiotic-inactivating enzymes have led to higher chloroplast transformation frequencies (Maliga, 2003). In theory, a single transformed genome copy is sufficient to render the entire organelle resistant to that antibiotic. The most extensively used selectable marker in plastid transformation is the bacterial spectinomycin resistance gene, *aadA*, which confers resistance to spectinomycin and streptomycin in plants (Svab & Maliga, 1993; Khan & Maliga, 1999; Bock & Hagemann, 2000; Hager & Bock, 2000; Bock, 2001). The encoded enzyme, aminoglycoside 3'-adenylyltransferase, catalyzes the covalent transfer of an AMP residue from ATP to aminoglycoside-type antibiotics, thereby inactivating them. In order to convert the *aadA* gene from *Escherichia coli* into a plastid-specific selectable marker, the coding region is fused to plastid expression signals. These include a 5'-DNA segment providing promoter, 5'-untranslated region

(UTR), and Shine-Delgarno sequence; and a 3'-plastid DNA segment providing a 3'-UTR which is required to confer transcript stability *in planta*. Engineering the plastid genome with a gene of interest imparting a desired trait is facilitated by co-introduction of that gene with this chimeric *aadA* gene. The spectinomycin resistance gene is a nonlethal marker allowing differentiation by colour (Svab *et al.*, 1990; Khan & Maliga, 1999). Resistant clones are green on selective media and sensitive clones are white due to the inhibition of chlorophyll accumulation. Nonlethal selection in higher plants is critical in obtaining transplastomic lines, as it allows sufficient time for the resistant plastid genome copies to increase in numbers to allow phenotypic expression. Spectinomycin selection in tobacco, on average, yields one transplastomic line in a bombarded leaf sample (Svab & Maliga, 1993).

A chimeric *nptII* gene, encoding neomycin phosphotransferase that confers resistance to kanamycin, has also been used as a selectable marker in tobacco plastid transformation (Carrer *et al.*, 1993). However, recovery of kanamycin resistance plastid transformants was 20- to 30-fold less efficient than for selection with spectinomycin (Carrer *et al.*, 1993). Since kanamycin inhibits cell division while spectinomycin does not, it was speculated that lack of phenotypic resistance in early kanamycin plastid transformants prevented the potential accumulation of transformed genome copies. It was thus predicted that the introduction of kanamycin after cultivation in the absence of selection, might allow sufficient time for increased transformed plastid genome replication, thereby improving efficiency.

An alternative selection process involves the use of a plant-derived selectable marker, the *badh* gene, encoding the enzyme betaine aldehyde dehydrogenase (BADH) (Daniell *et al.*, 2001a). The selection process involves the conversion of toxic betaine aldehyde to the beneficial osmoprotectant glycine betaine. The *badh* gene is present in a few plant species adapted to dry and saline environments and is naturally expressed in the nucleus in response to osmotic stress. The encoded enzyme is then targeted to the chloroplast where the conversion takes place. This selectable marker gene has been used to transform tobacco plastids and is reportedly more efficient than the *aadA* gene, obtaining rapid regeneration and increased transformation frequencies (Daniell *et al.*, 2001a). Success with this selectable marker gene has shown that nonlethal selection is not necessarily a requirement for chloroplast transformation, since betaine aldehyde is toxic to plant cells and inhibits growth. The limitation to using this marker gene is the inability to distinguish between control and transgenic seeds during germination, probably due to an active endogenous BADH or similar enzymatic activity in non-green plastids during this stage (Daniell *et al.*, 2001a). The selection process is thus specific to chloroplasts. Genetic modification of the plastid genome using a plant-derived selectable marker will ease concern regarding the use of antibiotic resistant genes in genetically modified crops.

A novel marker gene construct incorporating both a selectable and screenable marker has been developed, enabling direct visual identification of transplastomic sectors present in chimeric tissues

(Khan & Maliga, 1999). Wild-type sectors present in chimeric tissue exhibit a resistant phenotype due to phenotypic masking by the transgenic cells. The ability to eliminate these sectors will improve the efficiency of obtaining homoplastomic clones. The construct, termed FLARE-S, was obtained by translationally fusing the *aadA* gene to the *Aequorea victoria* green fluorescent protein gene. The expressed bifunctional protein prevents separation of the two genes, thereby simplifying engineering. Bombardment of tobacco leaves (chloroplasts) and rice embryogenic cells (non-green plastids) with the FLARE-S vector enabled visualization of individual transformed and wild-type plastids under selection for spectinomycin and streptomycin, respectively.

Selectable marker genes are essential in the attainment of homoplasmy, however, following successful transformation, are no longer necessary and may permit the horizontal transfer of antibiotic resistance genes to microbes (Corneille *et al.*, 2001). High-level expression of these genes may also be an unnecessary metabolic burden to the plant. Corneille *et al.* (2001) used the P1 bacteriophage site-specific recombination system for the elimination of a negative selectable marker gene encoding cytosine deaminase. This gene was introduced into the tobacco plastid genome flanked by two directly orientated *lox* sites and excised by the subsequent introduction of a nuclear-encoded, plastid-targted CRE recombinase. This nuclear gene was later removed by segregation in the seed progeny. The CRE-*lox* system tested in this study will lead to the generation of transplastomic plants free of selectable marker genes.

1.2.4 Regenerable tissue culture systems

Successful recovery of transplastomic cells is dependent on efficient selection and plant regeneration protocols that require highly optimized cell and tissue culture systems. At present, tobacco is the only higher plant for which efficient plastid transformation protocols are available (Bock, 2001). However, successful plastid transformation has recently been reported in tomato and a wild oilseed species, and although currently far less efficient than tobacco, the results indicate good progress in new plant species (Maliga, 2001; Ruf *et al.*, 2001; Skarjinskaia *et al.*, 2003). Limitations in the currently available tissue culture systems are considered to be the main obstacle in the extension of the transplastomic technology to other species, and most importantly, to major crops.

1.3 PLASTID VERSUS NUCLEAR TRANSFORMATION

Transformation of plants via the plastid genome offers several advantages over the conventional method of nuclear transformation.

1.3.1 High levels of transgene expression

The polyploid nature of plastids provides exceptionally high transgene expression levels due to the high transgene copy number present in a given transplastomic cell; levels that far exceed reported nucleocytoplasmic transgenic expression levels (Khan & Maliga, 1999; Bock & Hagemann, 2000; Hager & Bock, 2000; Bock, 2001; De Cosa *et* al., 2001; Kuroda & Maliga, 2001a). Nuclear transformation experiments generally result in a few copies of the transgene per cell. High protein accumulation in plastids can also be attributed to the lower degradation capacity for foreign proteins in the organelle.

1.3.2 Absence of epigenetic effects

The presence of gene-silencing in nuclear transformation experiments has detrimental effects on potential expression levels owing to the inactivation or increased decay of mRNAs. Gene-silencing phenomena have not been observed in plastids, making the outcome of transformation experiments in plastids more predictable and reliable than in the nucleus (Bock & Hagemann, 2000; Hager & Bock, 2000).

1.3.3 Absence of position effects

Nuclear transformation in higher plants is a random process where transgenes are integrated unpredictably into genomic locations via non-homologous recombination. This generates transgenic lines with different copy numbers and integration sites of the respective transgene leading to variable gene expression, largely dependent on the genomic context and eukaryotic-like chromatin structure at the site of integration. Consequently, large numbers of nuclear transformants are required for screening and identification of a line displaying reasonably high expression levels. Transgene integration into the plastid genome, however, occurs via homologous recombination facilitating the targeting of foreign genes to a precise, predetermined location in the plastid DNA. Expression levels are thus exclusively determined by the signals fused to the coding region driving transgene transcription, mRNA stability, and translation. Resultant transgenic lines are therefore uniform with respect to transgene expression levels compared to the variable expression levels observed in nuclear transgenic plants (Bock & Hagemann, 2000; Hager & Bock, 2000; Bock, 2001).

1.3.4 "Gene-stacking": expression of operons as polycistronic mRNAs

The prokaryotic nature of gene expression in plastids allows efficient translation of proteins from polycistronic mRNAs (Daniell & Dhingra, 2002). This enables the concerted expression of several ribosome binding site (rbs)-coding region segments, driven by a single promoter and stabilized by a single 3'-UTR. The ability to express multiple proteins in a single transformation raises the possibility of expressing entire bacterial operons encoding novel biosynthetic pathways in the plastid compartment (Bock & Hagemann, 2000; Hager & Bock, 2000; Bock, 2001; De Cosa *et* al., 2001).

1.3.5 Ecological safety: containment

In most crop species, plastid DNA is inherited maternally, preventing pollen transmission of DNA (Daniell, 1999a). Cytoplasmic inheritance in higher plants is determined by the presence or absence of organellar DNA in the generative or sperm cells of mature pollen grains (Nagata *et al.*, 1999). The presence of organellar DNA corresponds to a biparental/paternal inheritance, whereas the absence of organellar DNA corresponds to a maternal inheritance. During pollen development, mitochondrial and plastid DNA is either synthesized or degraded, leading to a biparental/paternal or maternal inheritance, respectively.

In tobacco, maternal inheritance of plastids was illustrated using a line that was homoplastomic for the spectinomycin resistance gene (Svab & Maliga, 1993). The resultant progeny following pollination of the transplastomic flowers with wild-type pollen were uniformly resistant to spectinomycin, whereas the reciprocal cross generated progeny sensitive to that antibiotic. In grape, restriction fragment length polymorphisms of chloroplast DNA between the parents were used to trace the origin of plastids in the progeny (Strefeler *et al.*, 1992). Inheritance was maternal for both the crosses examined.

1.4 SUCCESSFUL PLASTID TRANSFORMATION IN HIGHER PLANTS

1.4.1 *Nicotiana tabacum*

In 1990, the first stable plastid transformation in a higher plant was achieved in *N. tabacum*, using the pZS148 plastid transformation vector (Svab *et al.*, 1990). This vector contained the 16S rRNA gene from the *N. tabacum* SPC2 line. SPC2 cells are resistant to both spectinomycin and streptomycin due to point mutations 278 bp apart in the 16S rRNA-encoding DNA. A *Pst* I site was incorporated downstream of the spectinomycin gene in this construct. The spectinomycin gene was thus flanked by the streptomycin gene and *Pst* I site, linked markers that enabled the distinction of spontaneous mutants from the transgenic lines. Leaf tissue was bombarded with tungsten particles (1 μm) coated

with the pZS148 plasmid DNA. Following spectinomycin selection, resistant clones from the calli were regenerated and selected for the *Pst* I linked marker by Southern blot analysis. The transgenic lines obtained were all homoplasmic for both the selected markers, demonstrating replacement by homologous recombination of the wild-type 16S rRNA gene with the mutated version from the pZS148 plasmid. Since the streptomycin marker was unselected, spectinomycin resistant lines were heteroplasmic for this trait, a result of random sorting of plastids in the absence of selection. Therefore, resultant seed progeny, after selfing, were uniformly resistant to spectinomycin, but heteroplasmic for streptomycin. This trait was inherited, as expected, for a plastome-encoded, heteroplasmic trait. These results highlighted the necessity to uniformly transform all plastid genome copies, prior to release of selective pressure, to ensure homoplasmy and hence genetic stability.

1.4.2 *Arabidopsis thaliana*

In 1998 the plastid transformation technology was extended to the higher plant, *A. thaliana*, a model species for plant research (Sikdar *et al.*, 1998). The plastid transformation vector used, pGS31A, included the spectinomycin resistance gene (*aadA*) flanked by *A. thaliana* plastid sequences of 1 kb in length to enable targeting between the *trnV* and *rps12/7* genes. The spectinomycin gene was controlled by regulatory elements of tobacco origin that included the rRNA operon promoter fused to a synthetic ribosome binding site and the plastid *psbA* 3'-untranslated region, for transcription and transcript stability, respectively. The pGS31A vector DNA was introduced into leaf chloroplasts on the surface of microscopic tungsten particles (1 μm) using a helium-driven biolistic gun. The plants regenerated from spectinomycin resistant cell lines were screened by Southern blot analysis to distinguish resistance due to plastid transformation from resistance due to spontaneous mutation. All the transplastomic plants flowered, but were sterile. This was thought to be due to extensive polyploidy of the leaf tissue, a consequence of plant regeneration from tissue culture. The efficiency of plastid transformation in *A. thaliana* was significantly lower than in tobacco. This may have been due to an inefficient homologous recombination system in *A. thaliana* chloroplasts or the relatively short plastid targeting sequences flanking the spectinomycin gene.

1.4.3 *Solanum tuberosum*

In 1999, stable chloroplast transformation was reported in potato, *S. tuberosum* (Sidorov *et al.*, 1999). Two tobacco-specific plastid vectors were used, pZS197 and pMON30125. These were designed to integrate into the large single copy and invert repeat regions of the plastid genome, respectively. The pZS197 expression cassette was analogous to that used in the pGS31A vector used to transform *A. thaliana* in 1998. The pMON30125 construct included the spectinomycin resistance gene under the control of the *psbA* expression signals, and the green fluorescent protein gene (*gfp*) under the control

of the rRNA operon promoter and *rps16* terminator region. Vector DNA was introduced into leaf explants of the potato line FL1607 by bombardment using gold particles (0.6 μm). Since potato and tobacco are both sensitive to the antibiotic streptomycin, selected spectinomycin resistant lines were subjected to streptomycin selection to eliminate those resistant to spectinomycin as a result of spontaneous mutation of the plastid *rrn16* gene. The putative potato transformants from the pZS197 vector were confirmed to be homoplasmic for the transformed plastid genome by Southern blot analysis. The pMON30125 transformants were screened visually for GFP fluorescence during all selected events. Three of the fourteen spectinomycin-selected lines from the pMON30125 transformants were positive for GFP fluorescence. These three lines were also resistant to streptomycin, a result that illustrates the advantage of using a screenable marker in the selection of plastid transformants. Since GFP fluorescence could be detected during callus formation, this enabled the elimination of spontaneous mutants early in the selection process without further selection for streptomycin resistance.

1.4.4 *Oryza sativa*

Later in 1999, plastid transformation was demonstrated in rice, *O. sativa*, in the first experiment of this kind in a cereal crop (Khan & Maliga, 1999). The pMSK49 plastid transformation vector was rice-specific, targeting integration between the *trnV* and *rps12/7* genes and encoding the bifunctional protein, FLARE-S (fluorescent antibiotic resistance enzyme, spectinomycin and streptomycin). The promoter consisted of the plastid rRNA operon promoter and the leader sequence of the T7 phage gene 10, a combination known to mediate high levels of protein accumulation in plastids. The *aadA* and *gfp* genes were translationally fused with an 11-mer peptide. In rice transformation, plants are regenerated from non-green embryogenic cells. Earlier studies had shown expression of a similar FLARE-S construct in non-green tobacco plastids found in the petal and root cells (Khan & Maliga, 1999). These results prompted the transformation of embryogenic rice tissue culture cells. pMSK49 plasmid DNA was bombarded into these cells on the surface of gold particles, targeting the proplastids. Since rice and most cereals are naturally resistant to spectinomycin, selection of transplastomic lines was carried out using streptomycin, which inhibits growth of embryogenic rice cells. Resistant, regenerated green shoots were rooted and grown into plants, where integration of the FLARE-S cassette was confirmed using PCR. GFP fluorescence was visualized by confocal laser scanning microscopy and enabled the distinction of transformed and wild-type plastids. Only a small fraction of the chloroplasts expressed FLARE-S, thus creating heteroplasmic rice plants. These plants would still have to undergo further streptomycin selection prior to the attainment of complete homoplasmy. However, plant regeneration from cultured rice cells by standard protocols occurs too fast, thus not allowing sufficient time for the attainment of genetically stable transplastomic plants

(Maliga, 2001). The ability to attain homoplastomic rice plants will rely on the modification of existing plant regeneration protocols.

1.4.5 *Glycine max* (L.) Merr.

In 2001, plastid transformation was attempted using cultured soybean embryogenic tissues as target for biolistic bombardment (Zhang *et al.*, 2001). The aim of this work was to replace the plastid-encoded large subunit of Rubisco, *rbcL*, with that from *C. reinhardii*, and to include the nuclear-encoded algal small subunit (*rbcS*) of this enzyme into the soybean chloroplast. Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase) is the key enzyme of carbon dioxide (CO_2) fixation during photosynthesis. Since *C. reinhardii* Rubisco has been shown to have a higher net photosynthetic rate under high CO_2 conditions, transplastomic soybean plants were expected to show improved photosynthetic ability under those conditions. The soybean plastid transformation vector, pZVII, thus included the spectinomycin resistance gene driven by the 16S rRNA operon promoter from tobacco, flanked by the *C. reinhardii rbcL* and *rbcS* genes, both driven by the soybean *rbcL* promoter region. This expression cassette was flanked by soybean target sequences directing integration into the large single copy region between the *atpB* and the *rbcL* 3'-UTR, encouraging recombination of the *rbcL* genes. pZVII-coated gold particles (0.6 μm) were used for bombardment of about 100 000 embryogenic clumps. Only three spectinomycin-resistant clumps were recovered and only one proliferated sufficiently for further PCR analyses. PCR screening confirmed the integration of the cassette present in the pZVII construct, however, not all plastids were transformed so the tissue was heteroplastomic. Unfortunately, none of the resistant embryos survived subsequent plant regeneration, thus preventing the attainment of homoplastomic plants. It was thought that factors such as the low plastid copy number per cell, and the low rate of embryo proliferation and plant regeneration may have contributed to the low plastid transformation efficiency in soybean.

1.4.6 *Lycopersicon esculentum*

Later in 2001, stable plastid transformation was achieved in tomato, *L. esculentum*, with notable expression levels from the marker gene in the fruit chromoplasts (Maliga, 2001; Ruf *et al.*, 2001). Young tomato leaves were bombarded with gold particles (0.6 μm) coated with the pRB70 plastid transformation vector. This vector included the spectinomycin resistance gene flanked by the rRNA operon promoter and a synthetic *rbcL*-derived Shine-Delgarno sequence, and the plastid *psbA* 3'-UTR regulatory elements. The expression cassette was targeted to the large single copy region between the *trnfM* and *trnG* genes. The selection and plant regeneration process was greatly optimized in the generation of transplastomic tomato plants. Primary spectinomycin-resistant calli were selected after

three to four months' incubation under low-light conditions, more than three times the selection phase for tobacco. These calli were screened by PCR to confirm successful transgene integration, and to eliminate spontaneous mutants. After one to two additional rounds of callus propagation on selective media, chloroplast transformation and homoplasmy was confirmed by restriction fragment length polymorphism (RFLP) analyses. The selected calli were further tested for resistance to streptomycin, and following shoot and root induction respectively, were grown to maturity under greenhouse conditions. Maternal inheritance of the transgene further confirmed stable plastid transformation and homoplasmy. The significance of this work was the relatively high expression levels obtained in the ripe tomato fruit, approximately half that expressed in green leaves and green fruit. Most genes present in the plastid genome are involved in photosynthesis and are drastically down-regulated in non-photosynthetic tissues. In transplastomic potato plants, transgene expression levels were 100-fold lower in microtubers than in green leaves (Sidorov *et al.*, 1999). It was speculated that the relatively high foreign protein accumulation in ripe tomato fruits might be due to the actively photosynthetic tissues present in the green fruit prior to fruit ripening. It was then assumed that active plastid gene expression is maintained during the conversion from chloroplasts to chromoplasts. This work will initiate the possibility of producing edible vaccines and improved nutritional quality of a consumable fruit.

1.4.7 *Brassica napus*

In 2003, plastid transformation was reported in the oilseed rape, *B. napus* (Hou *et al.*, 2003). The vector used in this work, pNRAB, included two expression cassettes for spectinomycin and the insect resistance gene, *cry1Aa10*, respectively. Both cassettes were flanked by the 16S rRNA operon promoter and *psbA* 3'-UTR, all of tobacco origin except for the rice *psbA* 3'-UTR element in the *cry1Aa10* cassette. Cotyledon petioles were used as explants for bombardment, and following spectinomycin selection, 36 putative transformants were obtained from ~1 000 petioles. Spontaneous mutants were eliminated following PCR screening, from which only 4 were identified as heteroplastomic. Southern blot analysis confirmed site-specific integration of both cassettes into the *B. napus* plastid genome. Although homoplasmy was not accomplished, leaf bioassays were conducted against *Plutella xylostera*, a major pest of oilseed rape. Leaf material from all four transplastomic lines was toxic to the test insects, causing 33-47% mortality. These results illustrate the potential for generating stably transplastomic, *P. xylostera*-resistant oilseed rape.

1.4.8 *Lesquerella fendleri*

Also in 2003, successful plastid transformation was reported in *L. fendleri*, a wild oilseed species (Skarjinskaia *et al.*, 2003). Since *Lesquerella* and *Arabidopsis* both belong to the Brassicaceae family,

the pZS391B vector used in this work was derived from the *Arabidopsis*-specific pGS31A vector (Sikdar *et al.*, 1998). The left targeting region was extended using the 16S rRNA gene from tobacco, and the FLARE-S marker gene from the pMSK57 construct was included (Khan & Maliga, 1999). The *aadA* and *gfp* genes in the pMSK57 construct were translationally fused with a 16-mer peptide and driven by the 16S rRNA operon promoter. The *rbcL* gene leader sequence and the coding region N terminus, including the first 14 amino acids fused with FLARE-S, were included to enhance expression (Kuroda & Maliga, 2001b). Following biolistic bombardment using pZS391B-coated tungsten particles, the spectinomycin-resistant green shoots were further selected using streptomycin, eliminating spontaneous mutants. Only 2 of the 110 clones tested were also resistant to streptomycin. Southern blot analysis was used to confirm transgene integration and homoplasmy prior to plant regeneration, where only one of the two plants was morphologically normal. The seed progeny from a cross using the transplastomic line as female parent, segregated for streptomycin resistance indicating heteroplasmy. The resistant seed progeny, however, were shown to be homoplastomic during subsequent crosses that confirmed maternal transmission of the transgenes.

To date, stable transformation of only *N. tabacum, L. esculentum* and *L. fendleri* has been achieved via plastid engineering; efforts involving the other plants mentioned above have only been partly successful. The difficulties encountered include the inability to attain homoplasmy (*O. sativa, G. max* and *B. napus*), and the regeneration of fertile transplastomic plants (*A. thaliana* and *S. tuberosum*). Since *S. tuberosum* is largely propagated vegetatively by tuber, plant regeneration may be less important in this species.

Plastid transformation has also been achieved in the water plants, *Lemna* and *Spirodela*, however, due to support from industry, this work has not yet been published (pers. comm., A Perl, Agricultural Institute, Bet-Dagan, Israel). The vector used for *Lemna* transformation was specific to that water plant so as not to risk using heterologous sequences that might reduce the transformability of the construct. Selection using *badh* proved successful in this work.

1.5 TRANSGENIC PLASTIDS IN PLANT BIOTECHNOLOGY

1.5.1 Biopharmaceuticals

Chloroplasts have great potential for the high level expression of proteins of biopharmaceutical importance such as recombinant therapeutic proteins, as well as antibodies and vaccines for oral immunization (Daniell *et al.*, 2001b). Transplastomic plants will enable large biomass scale-up with relatively low maintenance and cost requirements. Specific advantages will be reflected in the examples that follow.

Staub *et al.* (2000) successfully expressed the human secretory protein, somatotropin (hST), in a soluble, biologically active, disulphide-bonded form in tobacco chloroplasts. This protein is used primarily in the treatment of hypopituitary dwarfism in children. Naturally, hST is produced in the pituitary gland and targeted to the secretory system where removal of the signal peptide takes place. This modification leaves phenylalanine as the N-terminal amino acid rather than methionine. In order to mimic this process, yeast ubiquitin was fused to hST such that cleavage with ubiquitin protease would generate an N-terminal phenylalanine on hST. In one cassette, this gene was driven by the *psbA* promoter, and in another, the 16S rRNA operon promoter including the T7 phage gene 10 leader sequence, yielding ~1% and ~7% of TSP, respectively. An expression cassette was also tested using the *psbA* promoter and the full-length hST cDNA encoding methionine as the first amino acid, yielding only ~0.2% of TSP. All three chimeric genes contained the 3'-UTR from the ribosome small subunit 16 gene, *rps16*. Interestingly, the hST mRNA levels in the three transplastomic lines were similar, whereas expression levels were quite different. This illustrates the importance of translation control signals such as the T7 phage gene 10 leader that led to hST accumulation of greater than 300-fold higher than reported hST expression following nuclear transformation (Staub *et al.*, 2000). The transplastomic line expressing hST to ~7% of TSP, generated predominantly hST with an N-terminal proline residue following ubiquitin processing. The majority of hST expressed from the full-length cDNA, had N-terminal alanine residues suggesting a methionine aminopeptidase activity in chloroplasts. Both the unexpected hST species were, however, shown to be biologically active.

In a recent study, Tregoning *et al.* (2003) reported expression of the non-toxic Fragment C domain of the tetanus toxin (TetC) in tobacco chloroplasts for the development of a plant-based, mucosal tetanus vaccine. Due to the AT-rich nature of the *Clostridium tetani* TetC coding sequence, expression in *E. coli* and yeast both required codon optimization for increased expression levels. Expression in these organisms had further concerns. In *E. coli*, expressed proteins may contain toxic cell wall pyrogens, whereas TetC expressed in yeast was inactive as an immunogen due to protein glycosylation. Both these difficulties could be circumvented in chloroplasts. In this study, the TetC polypeptide was expressed from three different genes, all flanked by the 16S rRNA operon promoter and *rbcL* 3'-UTR. The native bacterial AT-rich (72.3% AT) gene included the T7 phage gene 10 leader sequence in one cassette, and the *atpB* 5'-UTR in the other, with TetC accumulation of ~25% and ~10% of TSP, respectively (Kuroda & Maliga, 2001a; Kuroda & Maliga, 2001b). The synthetic higher-GC (52.5% AT) gene included the T7 phage gene 10 leader sequence and resulted in ~10% of TSP. The increased translation from the high-AT TetC gene compared to the higher-GC gene in the same cassette reflects the plastid genome's AT-rich codon bias, however, the differences in the TSP levels were relatively insignificant. In higher plant plastids, protein accumulation is primarily controlled post-transcriptionally; therefore, the translation control signals are the main focus in optimization of expression levels. Unfortunately, the plants that expressed TetC to ~25% of TSP grew unusually

slowly and had a chlorotic phenotype characterized by chlorophyll deficiency. The reason for this seemed to be linked specifically to the high-level TetC accumulation. However, TetC expression in chloroplasts proved effective for the mucosal immunization of mice against the tetanus toxin.

1.5.2 Resistance management

The maternal inheritance of plastid DNA in most crops is particularly favourable when the transgenes involved encode herbicide or disease resistance proteins since pollen-mediated escape of transgenes to related weeds or neighbouring crops is prevented (Daniell, 1999a). Also, when transplastomic plants have been engineered to encode insecticidal proteins, transgenic pollen no longer poses a threat to non-target insects (Daniell, 1999a).

Plastid transformation offers the ability to express multiple genes in a single transformation step. De Cosa *et al.* (2001) achieved this by expressing the *Bacillus thuringiensis* (*Bt*) *cry2Aa2* operon in tobacco chloroplasts, which led to the highest level of foreign gene expression reported in transgenic plants to date. The *cry2Aa2* gene is the distal gene in a three-gene operon. The open reading frame directly upstream, encodes a putative chaperonin that facilitates folding of the Cry2Aa2 protoxin to form proteolytically stable cuboidal crystals. The *cry2Aa2* operon and *aadA* were controlled by the 16S rRNA operon promoter and *psbA* 3'-UTR, and targeted to the plastid inverted repeat regions in the *trnI* and *trnA* intergenic region. This enables read-through transcription from the upstream, endogenous 16S rRNA operon promoter, facilitating expression of the heterologous proteins (Maliga, 2002). Following the confirmation of homoplasmy in transgenic plants, Cry2Aa2 protein accumulation in young, mature, and old leaves was quantified as 34.9%, 45.3%, and 46.1% of TSP, respectively. Insect bioassays were successful using the tobacco budworm, the cotton bollworm and the beet armyworm. Consumption of transgenic leaves led to 100% mortality amongst all tested insects, and due to the absence of insecticidal protein in transgenic pollen, was not toxic to monarch butterflies. The exceptionally high *Bt* toxin accumulation in tobacco chloroplasts will reduce the probability of resistance development in pests, a possible risk in commercial nuclear transgenic crops, due to sub-optimal toxin production.

The transplastomic technology has also been successful in generating tobacco plants resistant to the herbicide, glyphosate (Kota *et* al., 1999; Ye *et al.*, 2001). Glyphosate is a broad-spectrum herbicide that inhibits the production of aromatic amino acids through competitive inhibition of the enzyme, 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS). EPSPS is naturally a nuclear-encoded chloroplast-localized enzyme. The most successful of the chimeric genes tested in the research done by Ye *et al.* (2001), included the EPSPS gene from the *Agrobacterium* strain CP4, the most highly glyphosate-resistant EPSPS identified to date. This gene was controlled by the strong, constitutive

plastid *rrn* phosphoenolpyruvate (PEP) promoter and *rps16* 3'-UTR. To enhance expression levels, the bacteriophage gene 10 leader sequence was included (Staub *et al.*, 2000), and the 14 N-terminal amino acids of the GFP gene were N-terminally fused to the EPSPS gene, leading to >10% of TSP in leaves. Glyphosate resistance was however shown to be consistently lower in the reproductive tissues compared to the vegetative tissues. The lower plastid copy number or lower transcriptional or translational activity in reproductive cells could explain this. Since glyphosate generally targets meristematic tissue, these cells may be particularly susceptible due to lower plastid numbers, despite high expression levels found in leaves. However, plastid transformation has prevented the possibility of pollen transmission of herbicide resistance transgenes to related weeds or neighbouring non-transgenic crops.

De Gray *et al.* (2001) have extended the plastid transformation technology to the development of transplastomic tobacco plants conferring resistance against phytopathogenic bacteria and fungi, through expression of the antimicrobial peptide (AMP), MSI-99. AMP's target prokaryotic membranes by aggregating to the outer membrane surface and forming pores that result in the loss of essential metabolites and disruption of vital cellular functions. The high concentration of AMP's required for this process is released from the chloroplast during infection upon organelle lysis, preventing further spread. The gene cassette used in this work included the 16S rRNA promoter, the MSI-99 gene with the consensus chloroplast ribosome binding site, GGAGG, the *aadA* selectable marker gene, and the *psbA* terminator. The flanking sequences were of petunia chloroplast origin and targeted the expression cassette to the inverted repeat regions between the *trnV* and *rps12* genes. Transformation frequencies in this work were unusually low, a result presumed to be linked to the less than 100% homology between the petunia flanking sequences and the tobacco chloroplast genome. Cell-free extracts of the transgenic plants were shown to strongly inhibit the growth of the bacteria, *Pseudomonas syringae* pv. *tabaci*, and leaf extracts inhibited the growth of pregerminated spores of three fungal species, *Aspergillus flavus*, *Fusarium moniliforme*, and *Verticillium dahliae*. In an *in planta* assay, mature leaves of the transgenic plants were inoculated with *P syringae* pv. *tabaci*, resulting in no necrotic symptoms, whereas the wild-type control plants were affected. This demonstrates the viability of expressing MSI-99 in tobacco chloroplasts, as the peptide was biologically active and not harmful to the chloroplast organelle.

1.6 AIM OF THIS STUDY

The South African wine industry is of huge importance to our economy. The area under vines in South Africa in 2002, according to SAWIS (Wait, 2003), was 120 384 hectares. The SA wine industry employs an estimated 348 500 people with a total producers' income of R 2 088.484 million. In 2002, 1 079 875 tons of grapes were harvested, yielding 834.156 million litres of wine, including

non-alcoholic products. The SA wine industry is the world's 8th largest in terms of wine production, producing 2.5% of the total world production in 2000. The genetic modification of grapevines has great potential in the development of disease resistance and improved berry quality, thereby ensuring maximum production and top quality of an economically important crop.

Currently, grapevine transformation protocols have only targeted the nucleus (Iocco *et al.*, 2001). The aim of this study was therefore to construct an expression vector for the transformation of the grape chloroplast genome to enable high-level expression of proteins in the leaf chloroplasts or berry chromoplasts for resistance management or improved nutritional quality, respectively. In order to achieve this, regulatory elements of grape chloroplast origin to control expression of a selectable and a screenable marker gene, and ultimately the gene(s) of interest, would be required. A selectable marker would enable the selective propagation of transformed genome copies in the attainment of homoplasmy. A screenable marker would assist direct visualization of transformed tissue sectors for possible improvement in plastid transformation efficiency. Each coding region in the expression cassette would require appropriate translation control signals to ensure efficient translation from the polycistronic mRNA. Finally, the expression cassette would have to be flanked by grape targeting regions to ensure targeted integration into the grape plastid genome. The successful construction of such a vector, employing either spectinomycin (pVCT09a) or betaine aldehyde dehydrogenase (pVCT09b) as selectable markers, is described in the chapters that follow.

2. MATERIALS AND METHODS

2.1 Chloroplast DNA isolation

Grapevine leaf material was collected from the ARC-Infruitec-Nietvoorbij experimental farm. Approximately 90 g of young leaves were washed in cold dH_20 and chopped into 1 cm^2 squares. The leaf material was then homogenized in a chilled blender in 200 ml cold isolation buffer (0.5 M sucrose, 50 mM Tris-HCl pH 8.0) for 60 secs at 20 000 rpm. The chloroplast mixture was strained through 2 layers of cheesecloth, followed by 2 layers of Miracloth (Calbiochem®) to remove cellular debris, and the filtrate transferred to six 40 ml centrifuge tubes. The filtrate was centrifuged for 5 mins at 4 000 rpm, and the supernatant discarded. The chloroplast-containing pellets in each tube were gently resuspended in 2 ml resuspension buffer (0.2 M sucrose, 5 mM Tris-HCl pH 8.0, 2 mM $MgCl_2$, and 0.5 mg/ml bovine serum albumin [BSA]). The intact grape chloroplasts were then lysed by addition of 4 volumes extraction buffer (1.4 M NaCl, 0.5 M Tris-HCl pH 8.0, 2% N-cetyl-N,N,N-trimethyl ammonium bromide [CTAB], 20 mM ethylene diamine tetra-acetic acid di-sodium salt [EDTA], 0.5% β-mercaptoethanol), and incubated at 60°C for 60 mins. An equal volume of chloroform:isoamyl alcohol (24:1) was added and mixed well prior to centrifugation at 10 000 rpm for 10 mins. The ctDNA in the supernatant was precipitated by addition of 0.1 volumes 3 M NaOAc pH 5.2, followed by 0.8 volumes isopropanol and incubated overnight at –20°C. Centrifugation at 10 000 rpm for 10 mins pelleted the ctDNA, which was subsequently washed with 70% ethanol, and resuspended in dH_20.

2.2 Primer design

Degenerate oligonucleotide primers were designed using chloroplast sequence data from related plants, as the grape sequence is not yet available. This data was obtained from the National Centre for Biotechnology Information (NCBI), Bethesda, USA. Sequences from *Spinacia* (NC_002202), *Nicotiana* (NC_001879), *Oryza* (NC_001320), *Zea* (NC_001666), *Pinus* (NC_001631), *Lotus* (NC_002694), and *Arabidopsis* (NC_000932) species were initially considered for primer design, but due to the highly conserved nature of chloroplast genomes, only *Spinacia* and *Nicotiana* were necessary for further design. Primers were designed with the aid of Primer Designer – Version 1.01, and synthesized by Inqaba Biotech, Pretoria, SA. The melting temperatures (T_m) for all primers were calculated on the PE Applied Biosystems GeneAmp® PCR System 9700.

2.3 Polymerase Chain Reaction (PCR)

The Expand High Fidelity (EHF) PCR System (Roche Biochemicals) was used when amplifying genetic elements to be included in the final vector, where high fidelity was a requirement. The final reagent concentrations included: 0.2 mM dNTPs, 0.3 µM forward primer, 0.3 µM reverse primer, 0.1-

0.75 µg template DNA, 1x EHF buffer with 1.5 mM $MgCl_2$, and 1.3 Weiss units (U) EHF enzyme mix/50 µl reaction. The dNTPs, primers, and template DNA were made up to 25 µl with dH_20, and in a separate mix, the buffer, $MgCl_2$, and enzyme also made up to 25 µl with dH_20, and kept on ice. The 2 mixes were combined for the PCR. The conditions included 1 cycle for the denaturation of the template DNA at 94°C for 2 mins; 10 cycles of denaturation at 94°C for 15 secs, annealing at 45-65°C for 30 secs, elongation at 72 or 68°C for 45 secs-8 mins; 15-20 cycles as the previous 10 cycles, however, including an additional cycle elongation of 5 secs for each cycle during the elongation step; and a final prolonged elongation at 72°C for 7 mins. The annealing temperatures (T_A) varied according to the calculated T_m for the specific primers used. The elongation temperature was 72°C when amplification products were less than 3 kb, and 68°C for products greater than 3 kb. The elongation time was dependent on the PCR fragment length, with 45 secs for up to 0.75 kb, 1 min for 1.5 kb, 2 mins for 3 kb, 4 mins for 6 kb, and 8 mins for 10 kb.

The Bioline BIOTAQ™ DNA Polymerase was used for all colony PCR-screening. The final reagent concentrations included: 1x NH_4 Buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.5 µM forward primer, 0.5 µM reverse primer, and 0.1 U enzyme/10 µl reaction. The PCR conditions included 1 cycle for denaturation at 94°C for 5 mins; 20 cycles including denaturation at 94°C for 30 secs, annealing at 55°C for 30 secs, elongation at 72°C for 60 secs; and 1 elongation cycle at 72°C for 7 mins.

The *PfuTurbo*™ DNA polymerase (Stratagene) was used in the PCR-based site-directed mutagenesis and restriction site-free cloning protocols (see below).

2.4 Gel electrophoresis

DNA fragments were separated on 0.8% agarose D1 LE (Hispanagar) gels in TAE (Tris-acetate/EDTA) electrophoresis buffer (40 mM Tris-acetate pH 8.0, 1 mM EDTA) at 120 V for 60 mins. The 1 Kb Plus DNA Ladder (GibcoBRL, Life Technologies) was used as a molecular size marker, and Lambda (λ) DNA standards (Promega) were used for DNA quantification.

2.5 Gel purification

The Qiagen QIAquick® Gel Extraction Kit (Southern Cross Biotechnology) was used to extract and purify DNA from agarose gels. The DNA fragments were excised from the gel, solubilized in the supplied Buffer QG, applied to a QIAquick column, and centrifuged. The bound DNA was washed with Buffer PE, and eluted with Buffer EB.

2.6 Ligation systems

The Ligafast™ Rapid DNA Ligation System (Promega) was used when available, and included in a 10 µl reaction: 5 µl 2x Rapid Ligation Buffer (60 mM Tris-HCl pH 7.8, 20 mM $MgCl_2$, 20 mM dithiothreitol [DTT], 2 mM adenosine triphosphate [ATP], 10% polyethylene glycol [PEG]), 3 U T4

DNA ligase, vector and insert DNA. A 1:10 molar ratio of vector:insert DNA was used, using the following equation to convert molar ratios to mass ratios: ng of vector x kb size of insert / kb size of vector x molar ratio of insert:vector = ng of insert. The reactions were incubated at room temperature for 1 hr.

The NEB T4 DNA Ligase ligation reaction of 10 µl included: 1x T4 DNA Ligase Reaction Buffer (50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 µg/ml BSA), 80 U ligase, vector and insert DNA (quantities as described above), and was incubated at 16°C for 8 hrs.

The pGEM®-T Easy Vector System (Promega) ligation reaction of 10 µl included: 5 µl 2x Rapid Ligation Buffer, 3 U T4 DNA ligase, 12.5 ng pGEM®-T Easy Vector, and PCR product insert (quantity as described above). The reactions were incubated at room temperature for 1 hr, or if the maximum number of transformants was required, at 4°C for 16 hrs. All ligation products were transformed into ultra-competent *Escherichia coli* cells (see below). Successful cloning of an insert into the pGEM®-T Easy Vector interrupts the coding sequence of β-galactosidase; recombinant clones can therefore be identified by colour screening when colonies are grown on plates containing X-Gal (Promega). Clones that contain PCR products, in most cases, produce white colonies, and those that do not, produce blue colonies.

2.7 Escherichia coli transformation

Ultra-competent DH5α cells (UCC) were prepared according to the protocol of Inoue *et al.*, 1990. One millilitre aliquots were stored at –80°C and thawed on ice when required. Five microlitres of a ligation reaction was added to 100 µl of UCC in 1.5 ml microfuge tubes, and gently flicked to mix. The cells were incubated on ice for 20 mins, followed by a 42°C heat-shock for 50 secs, and immediately returned to ice for 2 mins. Nine hundred microlitres of LB (Merck) broth was added to each tube and incubated at 37°C, 150 rpm. One hundred microlitres of these cells were plated out onto LB plates (agar bacteriological, Biolab) including ampicillin (Roche Biochemicals, 120 µg/ml unless otherwise specified) as the selective antibiotic. When low transformation efficiencies were predicted, the 900 µl transformation mix was centrifuged and the cell pellet resuspended in 100 µl LB broth, for all potential transformants to be plated out. The plates were incubated at 37°C for 16-24 hrs.

2.8 Plasmid DNA purification

The Wizard® *Plus* SV Minipreps DNA Purification System (Promega) was used for all plasmid purifications. A single colony was inoculated in 5 ml LB medium (containing antibiotic) and grown for 12-16 hrs at 37°C, 225 rpm. For high-copy-number plasmids, 5 ml of culture was processed, and for low-copy-number plasmids, 10 ml was processed. The bacterial cells were pelleted, resuspended in Cell Resuspension Solution (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 100 µg/ml RNase A), and lysed in Cell Lysis Solution (0.2 M NaOH, 1% SDS). Alkaline Protease Solution was added

following lysis to inactivate endonucleases and other proteins released during the lysis that could adversely affect the quality of the isolated DNA. Following the addition of the Neutralization Solution (4.09 M guanidine hydrochloride, 0.759 M potassium acetate, 2.12 M glacial acetic acid pH 4.2), the bacterial lysate was centrifuged to precipitate proteins, cell debris and chromosomal DNA. The cleared lysate was then transferred to a Spin Column and washed with Column Wash Solution (162.8 mM potassium acetate, 22.6 mM Tris-HCl pH 7.5, 0.109 mM EDTA). The bound plasmid DNA was eluted with 50 µl Nuclease-Free Water.

2.9 Restriction endonuclease (RE) digestions

A typical RE digest of 20 µl included: 2 µl of the provided SuRE/Cut buffer (Roche Biochemicals), up to 1 µg DNA, and 5 U of RE, and was incubated at 37°C for 90 mins. All RE's used were from Roche Biochemicals, except for *Xcm* I and *Sma* I (25°C) that were from New England Biolab (NEB). Restriction endonuclease digests were electrophoresed and the relevant fragments gel-purified for ligation.

2.10 Dephosphorylation of DNA 5'-termini

Shrimp Alkaline Phosphatase (SAP, Promega) treatment prevents the recircularization and religation of linearized cloning vector DNA by removing phosphate groups from both 5'-termini. The reaction included: SAP (3 U/µg DNA), the restriction-digested vector (restriction enzyme and buffer removed, DNA in water), and 1x SAP reaction buffer (50 mM Tris-HCl pH 9.0, 10 mM $MgCl_2$) to a final volume of 30-50 µl. This was incubated at 37°C for 15 mins. The SAP was heat-inactivated at 65°C for 15 mins before 2-4 µl was used for the ligation reaction.

2.11 Klenow blunt-ending

The Klenow enzyme (Roche Biochemicals) was used to blunt cohesive termini to facilitate cloning. The 30 µl reaction included: 0.25 mM dNTP's, 1x A buffer (Roche Biochemicals), 4 U Klenow enzyme and ± 500 ng DNA. This was incubated at 25°C for 30 mins. The Klenow enzyme was heat-inactivated at 75°C for 10 mins before 5-10 µl was used for ligation.

2.12 PCR-based site-directed mutagenesis

PCR-based site-directed mutagenesis was achieved according to the protocol of Fisher & Pei, 1997. Adjacent, non-overlapping primers were designed on opposite DNA strands with the mutation encoded at the 5'-end of one primer. Before PCR, 50 pmol/µl of each primer was phosphorylated using 1 U/µl polynucleotide kinase (usb™) in NEB ligation buffer, and incubated for 30 mins at 37°C. Long PCR was carried out in a 50 µl reaction using 1.25 U *PfuTurbo*™ DNA polymerase (Stratagene), 10-50 ng template plasmid DNA, 50 pmol of each phosphorylated primer, and 0.2 mM

dNTPs in supplied *Pfu* reaction buffer. The PCR conditions included 1 cycle for denaturation at 94°C for 4 mins; 16 cycles including denaturation at 94°C for 1 min, the annealing temperature for 1 min, and elongation at 72°C for 2.5 mins/kb plus a cycle elongation of 5 secs for each cycle. After PCR, 5 µl of the reaction was electrophoresed on a 0.8% agarose gel to confirm amplified product formation. The remaining PCR reaction was purified using the Qiagen QIAquick® Gel Extraction Kit (Southern Cross Biotechnology), and ligated using the NEB T4 DNA ligase. Following ligation, the ligase was heat-inactivated at 65°C for 10 mins. This reaction was digested with 60 U *Dpn* I for 2 hrs. One microlitre of the digest was transformed into ultra-competent DH5α cells as described above.

2.13 Plant RNA isolation

Total RNA was isolated from 3 *Amaranthus* species for the amplification of the betaine aldehyde dehydrogenase gene (*badh*). *Amaranthus hypochondriacus*, *A. hybridus*, and *A. tricolor* seeds (obtained from M Spreeth, ARC-VOPI, Pretoria) were germinated and grown under greenhouse conditions. Following ± 4 weeks of growth from the seedling stage, leaves from all 3 species were detached using a scalpel blade, cutting obliquely at the base of the petioles, under water. BADH expression was induced by submerging the leaves in a 0.5 M NaCl solution for 2-3 hours (Legaria *et al.*, 1998). The leaves were then rinsed in dH_2O for use in a total RNA extraction protocol using the Qiagen RNeasy® Plant Mini Kit. Approximately 100 mg of plant material was ground in liquid nitrogen and then lysed with the supplied Buffer RLT. Insoluble material was removed by centrifugation through a QIAshredder homogenizer. Ethanol was added to the cleared lysate to promote selective binding of RNA to the RNeasy silica-gel membrane. The sample was then applied to the RNeasy mini column and washed with Buffers RW1 and RPE. The bound RNA was eluted with RNase-free water.

2.14 Reverse transcription (RT) nested PCR

The Improm-II™ Reverse Transcriptase (Promega) was used for reverse transcription of *A. tricolor* RNA for the amplification of the *badh* gene. Approximately 1 µg of RNA and 20 pmol of reverse primer were combined and made up to 5 µl. This mix was heat-denatured at 70°C for 5 mins, and then chilled immediately on ice for 5 mins for the primer to anneal. A 15 µl mix that included: 1x Improm-II™ Reaction Buffer, 3 mM $MgCl_2$, 0.5 mM dNTPs, 1 U/µl rRNasin® Ribonuclease Inhibitor, and 1 µl Improm-II™ Reverse Transcriptase, was added to the RNA/primer mix. The 20 µl RT reaction was first incubated at 25°C for 5 mins, followed by first-strand cDNA synthesis at 42°C for 1 hr. The Improm-II™ Reverse Transcriptase was heat-inactivated at 70°C for 15 mins before proceeding with the nested PCR. The nested PCR involved the use of an external and a gene-specific primer pair, for the first and second PCR, respectively. (The reverse external primer was used in the RT reaction.)

2.15 Restriction site-free cloning

Restriction site-free cloning involved the targeted integration of a PCR product into a vector without the use of restriction endonucleases or DNA ligase (Geiser *et al.*, 2001; Chen *et al.*, 2000; Tseng *et al.*, 1999; pers. comm., M Appel, Shimoda Biotech, SA). The primers designed for this technique had 2 portions. The 5'-portions of the forward and reverse primers corresponded to the DNA sequence directly upstream and downstream from the point of desired insertion, respectively. The 5'-portion of the reverse primer was derived from the complementary DNA strand. The 3'-portions isolated the inserts by PCR. The EHF PCR system was used to amplify the fragment to be cloned. Long PCR was carried out in a 50 µl reaction using 1.25 U *PfuTurbo*™ DNA polymerase (Stratagene), 20-40 ng recipient plasmid, ~250 ng purified PCR product, and 0.2 mM dNTPs in supplied *Pfu* reaction buffer. The PCR conditions included 1 cycle for denaturation at 95°C for 2 mins; 18 cycles including denaturation at 95°C for 30 secs, the annealing temperature for 30 secs, and elongation at 68°C for 2.5 mins/kb plus a cycle elongation of 5 secs for each cycle. Twenty microlitres of the PCR reaction was digested with 25 U *Dpn* I and incubated for 3 hrs. One microlitre of the digest was transformed into ultra-competent DH5α cells as described above.

2.16 Plasmids

The following plasmids were used during vector construction:

pT-NOT (obtained from N Dear, Max Planck Institute for Immunology, Freiburg, Germany),

pGEM.Ubi1-s*gfp*S65T (obtained from H Nel, Institute for Plant Biotechnology, Stellenbosch University),

pMECA (Thomson & Parrot, 1998),

pDG1726 (obtained from M Sathik, Dept. of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai, India),

pBluescript®SK(+/-) (Stratagene),

pGEM®-T Easy (Promega),

pJB004 (obtained from E Bremer, Dept. of Microbiology, Philipps-University of Marburg, Germany), and

pGL10 (A Toukdarian, University of California, San Diego, USA; obtained from MN Gardner, Dept. of Microbiology, Stellenbosch University).

2.17 Protein expression in Escherichia coli

Spectinomycin adenyltransferase: Bacterial cells harbouring the pMECA/(+) and pMECA/(–) constructs were grown in 1 ml cultures on variable concentrations of the antibiotic spectinomycin, and a constant concentration of the antibiotic ampicillin (Roche Biochemicals, 100 µg/ml). Four cultures of each were grown on the following concentrations of spectinomycin dihydrochloride (Sigma-

Aldrich): 0 μg/ml, 50 μg/ml, 100 μg/ml, and 200 μg/ml, for 16 hours at 37°C, 200 rpm. Optical densities were measured at 600 nm (OD_{600}) for all 8 cultures using the Pharmacia Ultrospec III® spectrophotometer and recorded graphically.

Green fluorescent protein: Bacterial cells harbouring the pVCT09a and pVCT09b constructs were visualized for GFP fluorescence following growth in darkness, overnight at 37°C. LB-agar plates were first stored at 4°C for 48 hours. *Escherichia coli* colonies were then illuminated using the UVP (Upland, CA 91786) Black-Ray Long Wave UV Lamp Model B100 AP (pers. comm., P Maliga, Waksman Institute, Rutgers, USA).

Betaine aldehyde dehydrogenase: MC4100 [Δ(*betTIBA*)] cells were prepared for transformation following the $CaCl_2$ method as described in Sambrook & Russel (2001), making 1 ml of competent cells. The MC4100/pJB004/pGL10_9bEC and MC4100/pJB004 bacteria were selected on ampicillin (Roche Biochemicals) and streptomycin sulfate (ICN Biomedicals Inc.) at 120 μg/ml each. MC4100/pJB004/pGL10_9bEC bacteria were also selected on 50 μg/ml kanamycin sulfate (Roche Biochemicals). These strains were grown in minimal media (pers. comm., E Bremer, Dept. of Microbiology, Philipps-University of Marburg, Germany). This included 60.3 mM K_2HPO_4, 33.0 mM KH_2PO_4, 7.6 mM $(NH_4)_2SO_4$ and 1.7 mM $Na_3C_6H_5O_7.2H_2O$ and was first autoclaved, then cooled prior to addition of sterile 1.0 mM $MgSO_4$, 0.2% glucose, 0.05% vitamin B_1 hydrochloride (Fluka BioChemika), and 40μg/ml Casamino acids (all from Saarchem, unless otherwise specified). Solid media was made up using agarose D1 LE (Hispanagar). MC4100/pJB004/pGL10_9bEC and MC4100/pJB004 bacteria were grown in minimal media with 1 mM betaine aldehyde chloride (Sigma-Aldrich) and 200 mM NaCl at 37°C, 225 rpm. Optical densities were measured daily at 578 nanometres (OD_{578}) using the Pharmacia Ultrospec III® spectrophotometer and recorded graphically.

2.18 Vector diagrammes

All vector diagrammes were constructed using Vector NTI 5.0 Deluxe. The RE sites and vector sizes depicted in the results and discussion section are approximations; however, the cloned elements were all drawn to scale.

2.19 Sequencing

CJ van Heerden at the Central Analytical Facility, Stellenbosch University, performed all sequencing reactions. The sequence results were analyzed for restriction endonuclease sites using Webcutter (Heiman, 1997), and homology comparisons using other published sequences on NCBI, using the BLAST algorithm (Altschul *et al.*, 1990) to aid primer design.

3. RESULTS AND DISCUSSION

3.1 GRAPE CHLOROPLAST TRANSFORMATION VECTOR CONSTRUCTION

A stepwise approach was followed during the construction of a grapevine-specific chloroplast transformation expression vector pVCT (*Vitis* chloroplast transformation) in this study. A series of intermediate plasmids, each containing an additional genetic element was constructed in this process. These plasmids are the following:

pVCT01: Grape 16S ribosomal RNA operon promoter (~P*rrn*)
pVCT02: ~P*rrn* + Green fluorescent protein gene_nopaline synthase terminator (*gfp*_nos)
pVCT03: Shortened ~P*rrn* (P*rrn*) + Multiple cloning site (MCS) + *gfp*_nos
pVCT04: P*rrn* + MCS + Spectinomycin resistance gene (*aadA*) + *gfp*
pVCT05: P*rrn* + MCS + *aadA* + *gfp* + Grape photosystem II terminator (T*psb*A)
pVCT06a/b: P*rrn* + MCS + Ribosome binding site (rbs) + *aadA/badh* + *gfp* + T*psb*A
pVCT07a/b: Grape left targeting sequence (LF) + P*rrn* + MCS + rbs + *aadA/badh* + *gfp* + T*psb*A
pVCT08a/b: LF + P*rrn* + MCS + rbs + *aadA/badh* + *gfp* + T*psb*A + Grape right targeting sequence (RF)
pVCT09a/b: LF + P*rrn* + MCS + rbs + *aadA/badh* + Rubisco large subunit 5'-translation control region (*rbcL* 5'-TCR) + *gfp* + T*psb*A + RF

The betaine aldehyde dehydrogenase gene (*badh*) replaced the *aadA* gene in constructs pVCT06a – 9a, creating the constructs pCVT06b – 9b.

3.1.1 Grape chloroplast DNA isolation

Grape ctDNA was isolated from intact grape chloroplasts to use as template for the amplification of all plastid derived regulatory elements and sequence included in the vector (Fig. 3.1.1, Appendix B). These included the 16S ribosomal RNA operon promoter (P*rrn*), the photosystem II terminator (T*psb*A), regulatory regions required for the transcription and stability of messenger RNA's, respectively; the left and right targeting sequences to enable precise integration into the plastid genome via homologous recombination; and finally, the Rubisco large subunit 5'-translation control region (*rbcL* 5'-TCR), included to enable efficient translation of the green fluorescent protein gene, *gfp*.

3.1.2 Cloning of the 16S ribosomal RNA operon promoter

The plastid ribosomal operon promoter was chosen for driving transcription of transgenes included in the expression cassette of the vector. This is one of the strongest and therefore most widely used promoters in plastid transformation vectors (Maliga, 2002). Transcription from P*rrn* has been shown to be active, not only in green chloroplasts, but also in non-green plastids such as rice embryogenic cells and tobacco roots, trichomes, petal chromoplasts (Khan & Maliga, 1999), and tomato fruit chromoplasts (Ruf *et al.*, 2001).

Grape chloroplast (ct) DNA was used as template for the amplification of the 1 166 bp region spanning the 16S ribosomal RNA operon promoter (~P*rrn*), using primers, trnV_left_for and 16SrRNA_left_rev (Fig. 3.1.2a). Primer sequences are given in Appendix A. The resultant PCR products were first sequenced and then ligated into the *Xcm* I-digested pT-NOT vector (Fig. 3.1.2b). *Xcm* I digestion of pT-NOT releases the stuffer fragment of ± 500 bp and provides T-overhangs for the facilitated cloning of PCR products generated by certain thermostable polymerases. The Expand High Fidelity PCR system generates PCR products that resemble a mixture of blunt-ended and 3'-single A overhang products analogous to Taq DNA polymerase, and are therefore suitable for cloning into this vector. The sequence data were used for comparison with related sequences on NCBI to verify the area amplified, and analyzed for restriction endonuclease sites. Approximately 300 bp covered the promoter of interest, however, the additional sequence in this area was required for the targeting regions ultimately flanking the expression cassette.

Following transformation, colonies were PCR-screened using the vector-specific primers, T7 and T3, to determine the presence of the insert, ~P*rrn* (results not shown). Plasmid DNA from 4 positive colonies was isolated for digestion with *Ksp* I to determine the orientation of the insert. The desired T7→T3 orientation resulted in the release of a ± 700 bp fragment, representing the 5'-end of the promoter, from the remaining ± 3 465 bp vector and 3'-end ~P*rrn* (Fig. 3.1.2c). The selected clone was named pVCT01 (Fig. 3.1.2d).

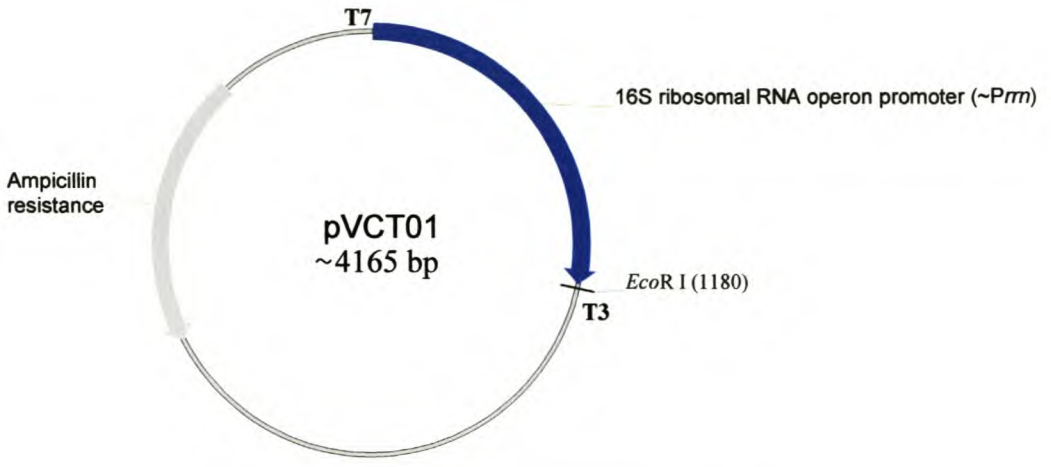


Figure 3.1.2d pVCT01: Grape 16S ribosomal RNA operon promoter (~P*rrn*) in pT-NOT

3.1.3 Cloning of the green fluorescent protein gene into pVCT01

Due to the high plastid genome copy number per cell, transformation initially yields chimeric tissues, which requires the identification of transplastomic cells for the regeneration of plants. The inclusion of the green fluorescent protein screenable marker gene will enable direct visual identification of transplastomic sectors present in chimeric tissues (Khan & Maliga, 1999). Wild-type sectors present in chimeric tissue exhibit a resistant phenotype due to phenotypic masking by the transgenic cells, conferred by the included selectable marker gene. The ability to eliminate these sectors will improve the efficiency of obtaining homoplastomic clones.

The green fluorescent protein gene, *gfp*, and nopaline synthase (nos) terminator were released together as a 1 010 bp fragment (*gfp*_nos) from the pGEM.Ubi1-s*gfp*S65T construct, by digestion with *Eco* RI (Fig. 3.1.3a). The *Eco* RI sites on pGEM.Ubi1-s*gfp*S65T flanked *gfp*, but also included the nos terminator. This was removed at a later stage during vector construction, but was temporarily transferred at this step into pVCT01. The *gfp*_nos fragment was ligated to the *Eco* RI-linearized pVCT01. Prior to ligation, the phosphate groups from both 5'-termini were removed from pVCT01 using shrimp alkaline phosphatase to prevent religation of the linearized vector. Three colonies were PCR-screened for the presence of both the ~P*rrn* and *gfp*_nos fragments using the vector–specific primers, T7 and T3 (Fig. 3.1.3b). The resultant ± 2.5 kb PCR products depicted in lanes 2 and 3 confirm *gfp*_nos incorporation, whereas the ± 1.5 kb fragment in lane 4 indicates religation of pVCT01, amplifying only ~P*rrn*. Plasmid DNA from both positive colonies was digested with *Sal* I to determine the orientation of the *gfp*_nos fragment. One *Sal* I site is present 3'-of *gfp*_nos, and another, 3'-to the *Eco* RI site on the vector. A T7→T3 orientation would therefore result in a ± 5 kb fragment

and a fragment of negligible size, ± 20 bp. Both plasmids gave such results (not shown). The selected clone was named pVCT02 (Fig. 3.1.3c).

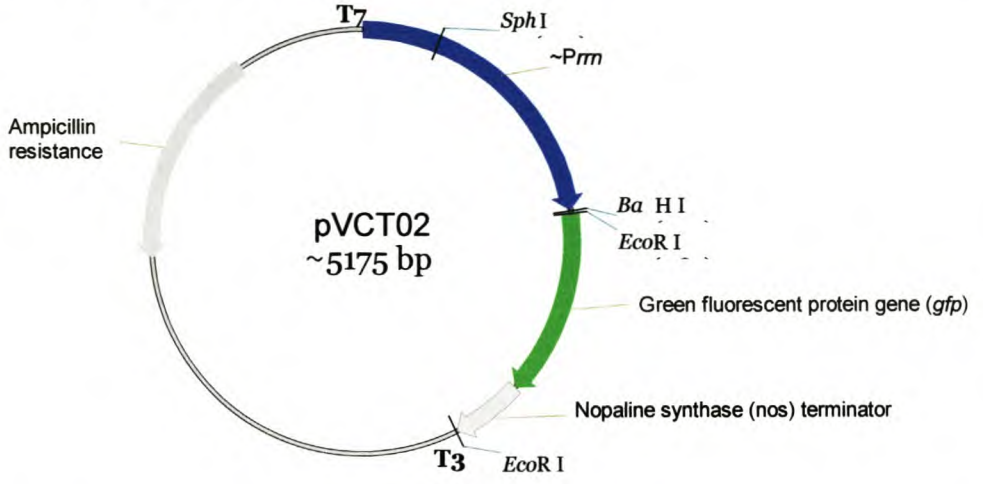


Figure 3.1.3c pVCT02: ~P*rrn* + Green fluorescent protein gene_nopaline synthase terminator (*gfp*_nos) in pT-NOT

3.1.4 Promoter shortening and cloning of a multiple cloning site into pVCT02

The promoter region in pVCT02 was reduced in size to include only the essential elements of the *rrn* promoter, reducing the region to ± 300 bp by digestion with *Sph* I and *Bam* HI. It was important to shorten this area to reduce the probability of it acting as target for homologous recombination with the native P*rrn* in the grape plastid genome (pers. comm., P Maliga, Waksman Institute, Rutgers, USA). An *Sph* I site is present ± 300 bp from the 5'-end, and a *Bam* HI site in the pT-NOT vector between the 1 166 bp region and *gfp*_nos. The digest therefore released ± 850 bp from the 3'-end of the original promoter region (Fig. 3.1.4a). This reduction was based solely on published sequence data from the NCBI. The alignments had revealed high homology illustrating the conserved nature of chloroplast genomes, a characteristic that allowed the theoretical determination of the region spanning the essential nucleotides of the promoter in the grape sequence.

A small set of unique restriction endonuclease (RE) sites were an essential addition to our vector for the ultimate incorporation of any transgene(s) of interest. The central section of the pMECA multiple cloning site was excised following digestion with *Sph* I and *Bam* HI and directionally cloned into the remaining ± 4.3 kb vector. This ± 90 bp fragment was digested from an amplification product using the vector-specific primers, T7 and T3, to ensure a high enough yield for cloning (Fig. 3.1.4b). Five colonies were PCR-screened using primers T7 and T3 to confirm the presence of the shortened promoter (P*rrn*), MCS, and *gfp*_nos regions (Fig. 3.1.4c). All 5 were positive, amplifying ± 1 600 bp fragments.

At this stage we decided to sequence the cloned elements contained in pVCT03 thus far using primers T7 and T3. The sequence generated from the T7 primer yielded the expected results; however, the reverse T3 reaction indicated an incorrect orientation for the *gfp*_nos fragment (T3→T7). This meant that the *Sal* I digests (see 3.1.3) had been misinterpreted, either due to the inactivity or absence of one of the *Sal* I sites. The *gfp*_nos fragment was then released from the current construct via *Eco* RI digestion, and religated non-directionally into the same vector (results not shown). Following colony PCR-screening using primers T7 and T3 (results not shown), the resultant ± 1 600 bp fragments from 2 colonies were used to determine the orientation of *gfp*_nos. This time, the orientation was determined by digestion with *Not* I. *Not* I sites are present on the vector flanking the combined insert (but still within the T7 and T3 sites), and between *gfp* and the nos terminator. A T7→T3 orientation therefore resulted in a ± 300 bp fragment representing the nos terminator, and a ± 1.2 kb fragment, representing P*rrn*, MCS and *gfp* (Fig. 3.1.4d Lane 3). The reverse orientation gave 2 ± 750 bp fragments (Fig. 3.1.4d Lane 2). The colony originating from the *Not* I-digested PCR fragment represented in lane 3 was thus the selected clone, named pVCT03 (Fig. 3.1.4e).

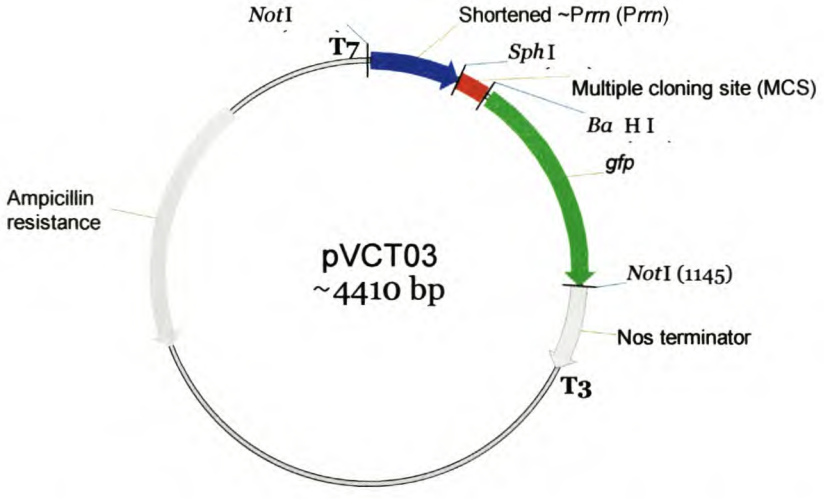


Figure 3.1.4e pVCT03: Shortened ~P*rrn* (P*rrn*) + Multiple cloning site (MCS) + *gfp*_nos in pT-NOT

3.1.5 Cloning of the spectinomycin adenyltransferase gene into pVCT03

The spectinomycin resistance gene (*aadA*, M69221) is the most widely used selectable marker used in plastid transformation due to its high specificity as a prokaryotic translational inhibitor, and its low side effects on plants (Bock, 2001). This gene was amplified from the pDG1726 plasmid using primers M69221_for_BamHI and M69221_rev_BamHI (Fig. 3.1.5a). These primers both provided *Bam* HI sites to the ends of the resultant 786 bp *aadA* PCR product. This fragment was first sub-cloned into the *Xcm* I-digested pT-NOT vector, from which a positive colony was identified through

PCR-screening using the vector-specific primers, T7 and T3. This product was then digested with *Bam* HI to facilitate cloning into the *Bam* HI site present at the 3'-end of the MCS in pVCT03 (Fig. 3.1.5b). The size difference between the PCR product in lane 2 and the digested PCR product in lane 3 is sufficient to indicate the activity of both RE sites. Following the cloning of *aadA* into the *Bam* HI-linearized pVCT03 vector, colonies were PCR-screened with the primers, M69221_rev_BamHI and T7. A fragment of ± 1 250 bp was amplified from one of the colonies, confirming both the integration and a T7→T3 orientation of the *aadA* gene (results not shown).

This intermediate construct was used as template for the amplification of the cassette including the P*rrn*, MCS, *aadA*, and *gfp*_nos elements, using vector-specific primers. The resultant PCR product was subsequently digested with *Not* I to exclude the nos terminator, and to enable transfer to the *Not* I site in the pBluescriptSK®(+/-), (2 958 bp), cloning vector MCS (Fig. 3.1.5c). A T3→T7 orientation was required in pBluescriptSK®(+/-) to facilitate further cloning, so the trnV_left_for and T7 primers were used to determine both integration and orientation, generating a fragment of ± 2 kb in 1 of the 2 colonies tested (Fig. 3.1.5d). No PCR product was amplified in colony (ii), so it was assumed that either the vector backbone had religated, or that the insert was in a T7→T3 orientation. A second PCR using the M69221_rev_BamHI and T3 primers amplified a fragment of ± 1.2 kb in colony (i), and no fragment in colony (ii), confirming the validity of the first PCR (Fig. 3.1.5d). The selected clone was named pVCT04 (Fig. 3.1.5e).

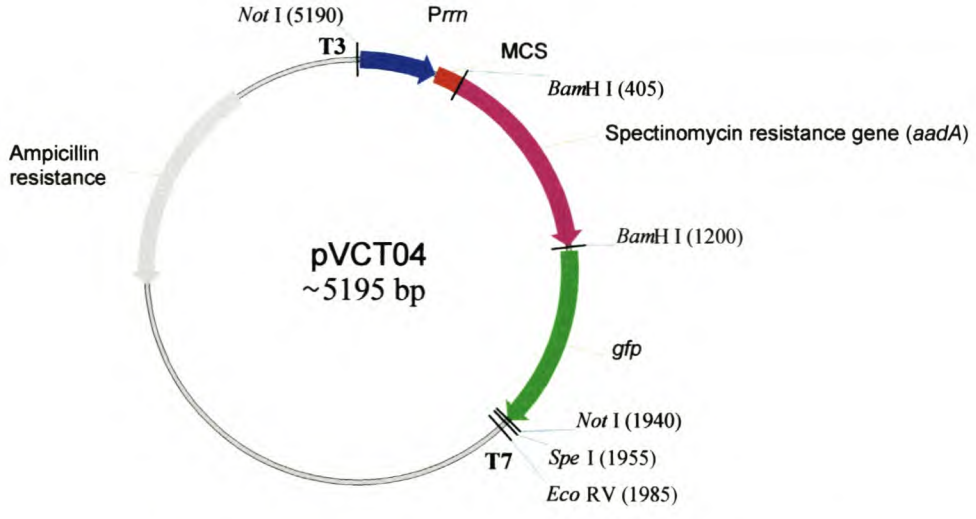


Figure 3.1.5e pVCT04: P*rrn* + MCS + Spectinomycin resistance gene (*aadA*) + *gfp* in pBluescriptSK®(+/-)

3.1.6 Cloning of the photosystem II terminator into pVCT04

The photosystem II terminator region was chosen for the stabilization of the polycistronic mRNA, transcribed by P*rrn* in our construct. Studies have shown that both the photosystem II and Rubisco large subunit 3'-untranslated regions yield stable mRNAs (Maliga, 2002), however T*psb*A appears to be more widely used in expression cassettes.

The region spanning the terminator of interest was amplified from grape ctDNA, using the degenerate primers, psbA3'_for and psbA3'_rev (results not shown). The resultant 913 bp PCR product was first cloned into the pGEM®-T Easy vector, from which 2 colonies were PCR-screened using the vector-specific primers, T7 and SP6 (Fig. 3.1.6a). Plasmid DNA from one of the positive clones was used for sequencing of the insert to aid specific primer design. A fragment of 215 bp was subsequently amplified using the T*psb*A-specific primers, psbA3'UTR_for and psbA3'UTR_rev (results not shown), designed to allow for direct cloning into the *Eco* RV site downstream of *gfp* in pVCT04. Blunt–end cloning was however unsuccessful, so we attempted an alternative method.

In the alternative approach, the T*psb*A PCR product was first sub-cloned into the pGEM®-T Easy vector, and the colonies were then PCR-screened with primers psbA3'UTR_for and T7 to determine both integration and SP6→T7 orientation (Fig. 3.1.6b). The orientation was important to enable directional cloning into pVCT04 using the *Spe* I and *Eco* RV sites, respectively. The *Spe* I site in pVCT04 is downstream of *gfp* and upstream to the *Eco* RV site. This step eliminated the *Bam* HI, *Sma* I, *Pst* I, and *Eco* RI sites present on the pBluescriptSK®(+/-) backbone, a loss that would later facilitate replacement of the *aadA* gene with the BADH gene (see 3.1.8). The SP6 – T*psb*A - T7 fragment was amplified using vector-specific primers for further manipulation (Fig. 3.1.6b), prior to final cloning into pVCT04. The PCR product was first digested with *Ksp* I (*Sac* II) (Fig. 3.1.6c), blunt-ended, and finally *Spe* I-digested (results not shown). These steps generated the (SP6) … *Spe* I – T*psb*A – *Eco* RI – *Sac* II (blunted) … (T7) fragment, now able to ligate to the *Spe* I/*Eco* RV-linearized pVCT04. The resultant colonies were screened with the psbA3'UTR_rev and T3 primers generating a ± 2.2 kb fragment, and the psbA3'UTR_for and T7 primers generating a fragment of ± 300 bp (Fig. 3.1.6d). These PCR reactions confirmed T*psb*A integration and correct orientation into pVCT04. The selected clone was named pVCT05 (Fig. 3.1.6e).

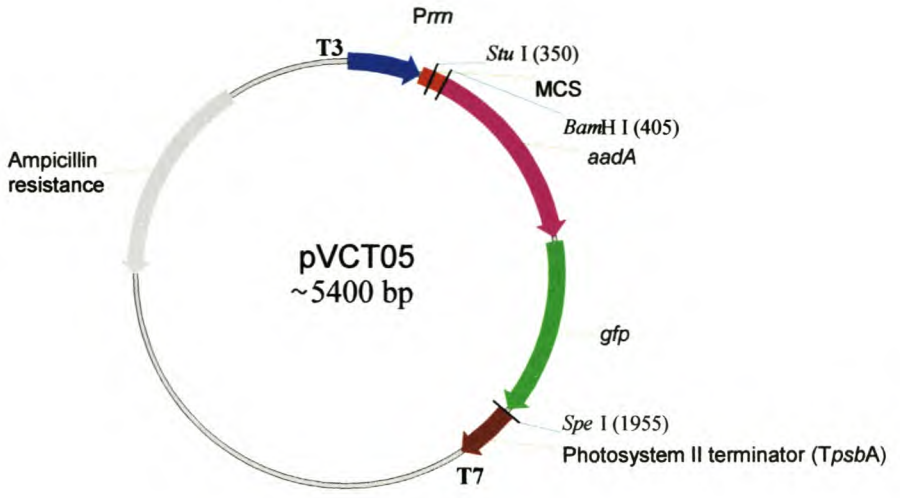


Figure 3.1.6e pVCT05: P*rrn* + MCS + *aadA* + *gfp* + Grape photosystem II terminator (T*psb*A) in pBluescriptSK®(+/-)

3.1.7 Cloning of the chloroplast ribosome binding site into pVCT05

When expressing multiple proteins from a polycistronic mRNA, it is essential to include ribosome binding sites upstream of each initiation codon to facilitate efficient translation of each open reading frame (Ruf & Kossel, 1998; Daniell *et al.*, 2002). The upstream 5'-untranslated regions (UTR) included in plastid transformation vectors have been greatly modified to optimize protein expression levels in plastids (Maliga, 2002). 5'-translation control regions (TCRs), including the 5'-UTR and amino-terminus of the source plastid gene coding region, tend to enhance protein accumulation. This seems to be related to the co-evolution of sequences surrounding the initiation codon, ensuring translation of the encoded protein (Kuroda & Maliga, 2001b). The consensus chloroplast ribosome binding site (rbs) sequence was introduced upstream of the selectable marker gene in this stage of vector construction, whereas a TCR was included upstream of the screenable marker gene (*gfp*) in the final step (see 3.1.12).

The consensus chloroplast rbs was incorporated upstream of *aadA* using a PCR-based site-directed mutagenesis protocol (Fisher & Pei, 1997). This involved the use of two specific primers in a single PCR, using the vector to be mutated as template. The T3-primed sequence of pVCT05 representing the P*rrn*, MCS, and 5'-end of *aadA*, constituted the area required for the design of primers used for the inclusion of the consensus rbs sequence, GGAGG (Daniell *et al.*, 2001). The primers were designed on opposite DNA strands with the rbs encoded at the 5'-end of the forward primer, RBS_for, directed 9 nucleotides upstream of the *aadA* initiation codon. The reverse primer, StuI_rev, annealed to the *Stu* I, *Pme* I, *Acc* I, and *Sal* I RE sites (5'→3') present in the MCS, thus excluding ± 50 bp containing the sites from *Pac* I to *Xho* I in the PCR product. This reduction in the MCS was intentional as the sites

lost were no longer unique. Both primers were phosphorylated prior to PCR to enable re-ligation of the PCR products. pVCT05 was used as template during long PCR using the *PfuTurbo*™ DNA polymerase, generating blunt-ended PCR products. Following ligation of the PCR products, wild-type template DNA was eliminated by *Dpn* I digestion which is specific for methylated adenine residues in the sequence G^{m}A↓TC. Template plasmids isolated from most common laboratory bacterial strains are methylated and therefore digested, leaving the PCR-amplified DNA intact. The *Dpn* I-digested ligation was then transformed, and plasmid DNA from 4 colonies was sequenced using the T3 primer. The sequence from the modified area was necessary to verify success. Of the 4 colonies chosen, all plasmids had correctly incorporated the rbs (bold) 9 nucleotides upstream of the *aadA* initiation codon, ATG (see below). However, DNA from one colony had not included the 5 bp directly upstream of the rbs, corresponding to the complementary nucleotides from the 5'-end of the reverse primer.

5'...CGT CGA CTG TTT AAA CGA GGC CTT A**GG AGG** CGA GGA TCC ATG AGG AGG...3'

The selected clone was named pVCT06a (Fig. 3.1.7).

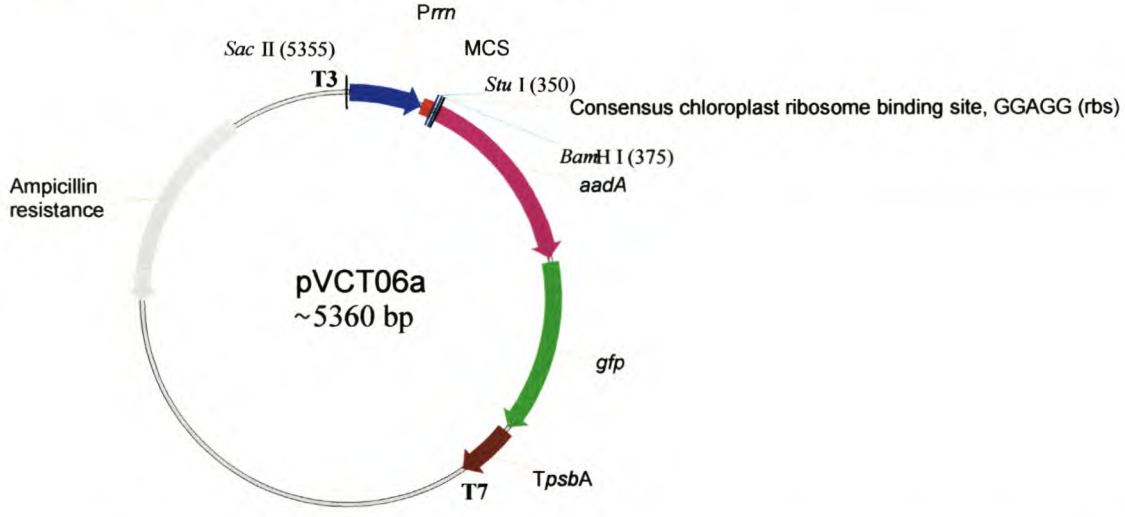


Figure 3.1.7 pVCT06a: P*rrn* + MCS + Ribosome binding site (rbs) + *aadA* + *gfp* + T*psb*A in pBluescriptSK®(+/-)

3.1.8 Cloning of the betaine aldehyde dehydrogenase gene into pVCT06a

The betaine aldehyde dehydrogenase (BADH) gene (*badh*) was used as an alternative selectable marker to the antibiotic resistance gene, *aadA*, as BADH selection may not be feasible for grapevine tissues (pers. comm., P Maliga, Waksman Institute, Rutgers, USA), and for this study was of *Amaranthus* origin (Fig. 3.1.8a, Legaria *et al.*, 1998). Naturally, the nuclear *badh* is activated in response to osmotic stress where the encoded enzyme then targeted to the chloroplast organelle, converts toxic betaine aldehyde to non-toxic glycine betaine, an osmoprotectant (Daniell *et al.*, 2001).

This gene is present in only a few plant species adapted to dry and saline environments, and is therefore a suitable selectable marker for most crop plants.

The *badh* gene was isolated from three *Amaranthus* species using reverse transcription (RT) nested PCR. Two primer sets were designed, an external and a gene-specific set, using NCBI sequences AF000132 (*ahybadh4* gene) and AF017150 (*ahybadh4* mRNA), respectively, (pers. comm., M Thomas, CSIRO Plant Industry, Adelaide, Australia). The external primers were designed to anneal to the UTR of the *badh* sequence: the forward primer, upstream of the initiation codon, and the reverse primer, downstream of the stop codon.

Initially, *Amaranthus* leaves were submerged in a salt solution to induce BADH expression. Total RNA from *A. hypochondriacus* (Fig. 3.1.8b), *A. hybridus* and *A. tricolor* species was used as template for reverse transcription using the external reverse primer, 3'UTR_rev(BADH). In the first PCR, using the external primers, 5'UTR_for(BADH) and 3'UTR_rev(BADH), a 1 609 bp fragment was amplified (results not shown), and served as template for the nested PCR that followed. The gene-specific primers, BADH_for_BamHI and BADH_rev_BamHI, amplified the 1 521 bp *badh*, and included flanking *Bam* HI sites (Fig. 3.1.8c).

The *badh* gene from *A. tricolor* was first sub-cloned into the *Xcm* I-digested pT-NOT vector. A positive clone was identified through colony PCR-screening using the vector-specific primers, T7 and T3 (Fig. 3.1.8d). Plasmid DNA was digested with *Bam* HI to release *badh* (Fig. 3.1.8e). Similarly, pVCT06a was digested with *Bam* HI, releasing the spectinomycin adenyltransferase gene, *aadA* (Fig. 3.1.8f). *Badh* was subsequently integrated into the pVCT06a vector backbone, replacing *aadA*. *Badh* integration and orientation was determined using the BADH_rev_BamHI and T3 primers, generating a fragment of ± 1.9 kb (Fig. 3.1.8g). The desired T3→T7 orientation was further confirmed using the BADH__for_BamHI and T7 primers, generating a fragment of ± 2.5 kb (Fig. 3.1.8g). The selected clone was named pVCT06b (Fig. 3.1.8h).

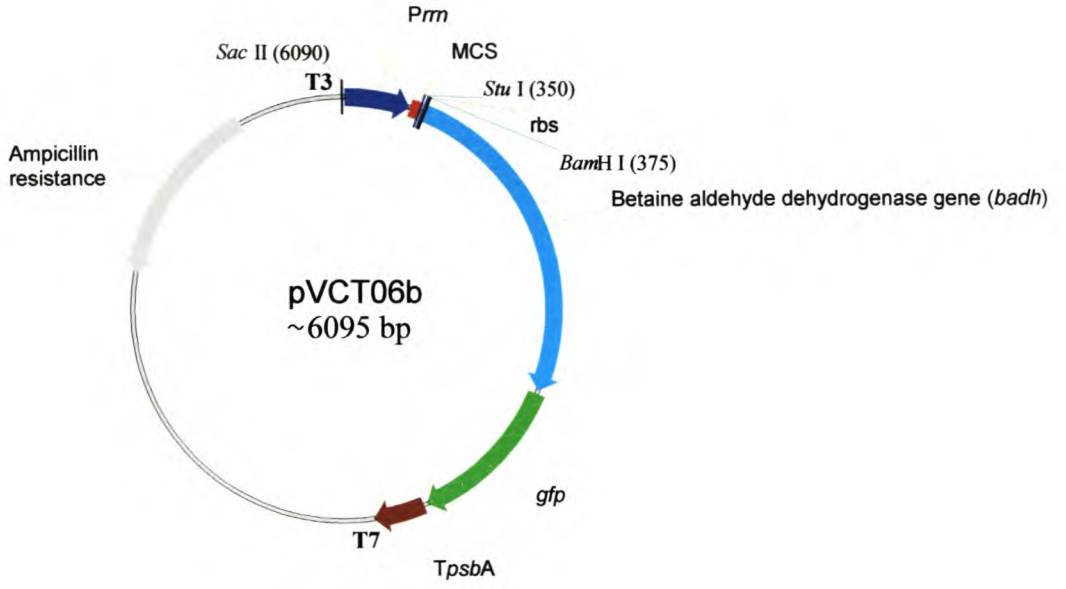


Figure 3.1.8h pVCT06b: P*rrn* + MCS + Ribosome binding site (rbs) + *badh* + *gfp* + T*psb*A in pBluescriptSK®(+/-)

The pVCT06b construct therefore contains the plant-derived selectable marker gene, *badh*, whereas pVCT06a contains the antibiotic resistance gene, *aadA*.

3.1.9 Sequencing of the grape plastid targeting region

To enable targeted insertion into the plastid genome via homologous recombination, both constructs pVCT09a and pVCT09b required at least 1 kb of plastid DNA sequence flanking the expression cassette (Zoubenko *et al.*, 1994). Insertion was directed between the divergently transcribing *rps12* and *trnV* genes so as not to disrupt endogenous gene sequence, (Zoubenko *et al.*, 1994). This target area is found in the inverted repeat region thereby enhancing potential expression levels of the transgene(s) of interest. The rps12_for(JR) and 5'16SrRNAleft_rev(JR) primers were used to amplify the region spanning the 3'-end of the *rps12* gene, *trnV*, and the 5'-end of the 16S rRNA gene. The 2 240 bp fragment amplified from grape ctDNA (Fig. 3.1.9) was used as template for progressive sequencing reactions. The first sequencing reaction used primer rps12_for(JR). The 3'-end of this sequence was then used to design the second primer, rps12_for_fillgap. Primers rps12_for_SEQ and rps12_for_trnV were similarly designed, resulting in 4 overlapping stretches of DNA that were all aligned to give a final region that reached to the next known grapevine sequence (B Rose, Dept. of Genetics, Stellenbosch University). The 3'-end of the generated consensus region aligned to the 5'-end of the 1 166 bp promoter region sequenced earlier (see 3.1.2). The total area of over 3 kb was analyzed for RE sites to facilitate cloning of the left and right flanks into the respective recipient vectors.

Accordingly, it was decided to include first the left flank into the *Sac* II site upstream of the promoter, and then the right flank into the *Hind* III and *Xho* I sites downstream of the terminator, all sites present on the pBluescriptSK®(+/-) backbone. Directional cloning of the right flank was possible due to exclusion of the MCS *Xho* I site during rbs inclusion (see 3.1.7).

3.1.10 Cloning of the left flank into pVCT06a and pVCT06b

Grape ctDNA was used as template for the amplification of the left flank (LF), a 1 134 bp fragment upstream of the *trnV* gene (Fig. 3.1.10a). The SacII_for_LF and SacII_rev_LF primers provided *Sac* II (*Ksp* I) sites to the ends of the PCR product to enable cloning into the *Sac* II site present on the pBluescript®SK(+/-) backbone, upstream of the promoter in pVCT06a and pVCT06b. The LF PCR product was first sub-cloned into the pGEM®-T Easy vector, from which a positive colony was identified through PCR-screening using the vector-specific primers, T7 and SP6 (Fig. 3.1.10b). Plasmid DNA was then digested with *Sac* II to release the LF, confirming *Sac* II activity prior to ultimate cloning (Fig. 3.1.10c). Both the pVCT06a and pVCT06b constructs were similarly digested (results not shown) and then ligated to the LF insert. LF integration and orientation was determined through amplification using the SacII_rev_LF and T3 primers, generating a ± 1.2 kb fragment in both the pVCTa- and pVCTb-derived constructs (Fig. 3.1.10d). pVCTa- and pVCTb-specific primer pairs were used to verify the initial screening using T3 with the M69221_rev_BamHI and BADH_rev_BamHI primer, respectively (Fig. 3.1.10d). The amplified products were ± 2.3 kb (pVCTa) and ± 3 kb (pVCTb) in size. The respective clones were named pVCT07a and pVCT07b (Fig. 3.1.10e and 3.1.10f).

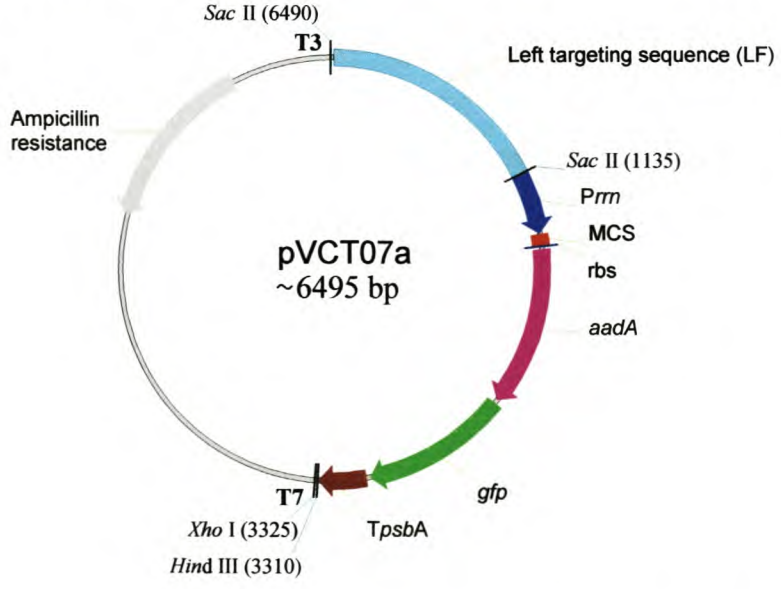


Figure 3.1.10e pVCT07a: Grape left targeting sequence (LF) + P*rrn* + MCS + rbs + *aadA* + *gfp* + T*psb*A in pBluescriptSK®(+/-)

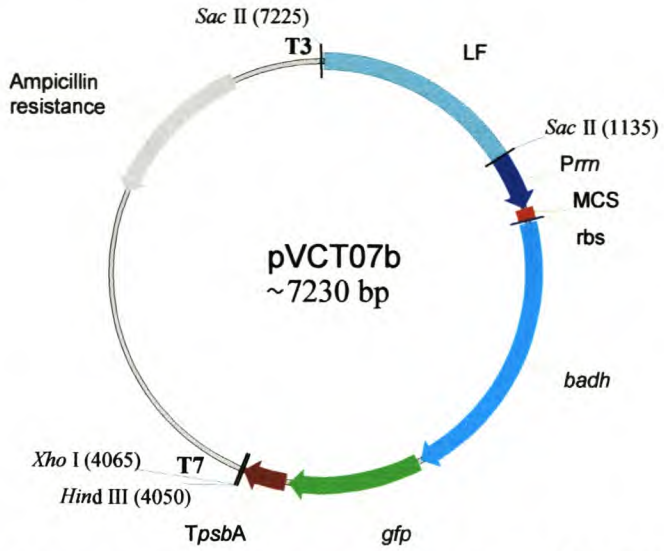


Figure 3.1.10f pVCT07b: Grape left targeting sequence (LF) + P*rrn* + MCS + rbs + *badh* + *gfp* + T*psb*A in pBluescriptSK®(+/-)

3.1.11 Cloning of the right flank into pVCT07a and pVCT07b

The initial cloning strategy employed for the incorporation of the right flank was analogous to the method followed for cloning of the left flank. Directional cloning into the *Hind* III and *Xho* I sites present on the pBluescript®SK(+/-) backbone was attempted. The primers designed for this purpose, HindIII_for_RF and XhoI_rev_RF, were used for amplification of the required fragment, and the PCR product was successfully sub-cloned into the pGEM®-T Easy vector. Cloning into the respective recipient vectors, pVCT07a and pVCT07b, however, proved unsuccessful, for reasons that remain unknown.

An alternative strategy that does not require the use of restriction endonucleases or DNA ligase (Geiser *et al.*, 2001; Chen *et al.*, 2000; Tseng *et al.*, 1999; pers. comm., M Appel, Shimoda Biotech, SA) was thus explored. This technique enables targeted, directional cloning of any PCR product into any vector, provided the sequence surrounding the point of desired insertion, is known. Restriction site-free cloning relies on the use of a pair of DNA integration primers with two portions. The 3'-portion isolates the inserts by PCR, and the 5'-portion, related to the insertion sequence, integrates the PCR products into the homologous region of the vector. The 3'-ends of the resultant PCR products, are thus generated by the 5'-portions of both primers, and are able to prime synthesis along the vector during the thermal cycling elongation step. This results in a fusion between the DNA of the PCR product and that of the vector, forming a linear recombinant molecule. In each elongation step, the original vector acts as template, generating a sense strand from the forward primer and an anti-sense strand from the reverse primer. These linear molecules form circular, recombinant plasmids during the final annealing step. No ligation reaction is necessary, as the 5'-ends of both molecules overlap,

creating single nicks flanking the original PCR product. The designed plasmid is selected following *Dpn* I digestion as *Dpn* I digests the methylated and hemimethylated DNA of the original and hybrid plasmids, respectively. The use of a *dam*$^+$ *E. coli* strain for cloning of the original vector is therefore an absolute requirement for this selection. The newly formed plasmid can then be transformed and cloned (Geiser *et al.*, 2001; Chen *et al.*, 2000; Tseng *et al.*, 1999).

This approach was employed successfully to integrate the RF between the *Hind* III and *Xho* I sites downstream of the terminator on the pBluescriptSK®(+/-) backbone. The right flank (RF) was amplified from the region downstream and adjacent to the left flank using grape ctDNA as template and the primers, Appel_RF_for and Appel_RF_rev. The nucleotide sequence of the 5'-portion of both primers was derived from the insertion region sequence from the vectors, pVCT07a and pVCT07b. The 5'-end of the forward primer included 22 nucleotides corresponding to the DNA sequence directly upstream from the point of insertion, and the reverse primer, 20 nucleotides directly downstream from the point of insertion, derived from the complementary DNA strand. The 3'-ends of the 5'-portions of the primers were 2 bp apart, thereby eliminating the *Hinc* II, *Acc* I, and *Sal* I sites. This rendered the *Sal* I site present in the MCS of the pVCTa construct unique, and thus suitable for cloning of transgenes. This did not apply to the pVCTb construct, due to a *Sal* I site present in *badh*. A 1 576 bp fragment was amplified using the Appel_RF_for and Appel_RF_rev primers (Fig. 3.1.11a).

Approximately 30–40 ng recipient plasmid (pVCT07a and pVCT07b) and 250 ng RF PCR product were used for long PCR using *PfuTurbo*™ DNA polymerase. Both ends of each strand of the PCR products anneal to their complementary strands of the vector, and the 3'-ends are then extended along the vector during the thermal cycling elongation steps. Following *Dpn* I digestion and cloning, a number of plasmids of the recipient type were still identified during colony PCR-screening (results not shown) suggesting that less should have been used. Recombinant plasmids were identified using the HindIII_RF_for and T7 primers, which amplified a fragment of ± 1.6 kb (Fig. 3.1.11b). pVCTa- and pVCTb-specific primer pairs were used to verify the initial screening using T7 with the M69221_for_BamHI and BADH_for_BamHI primers, respectively (Fig. 3.1.11b). The amplified products were ± 3.4 kb (pVCTa) and ± 4.1 kb (pVCTb) in size. The respective clones were named pVCT08a and pVCT08b (Fig. 3.1.11c and 3.1.11d).

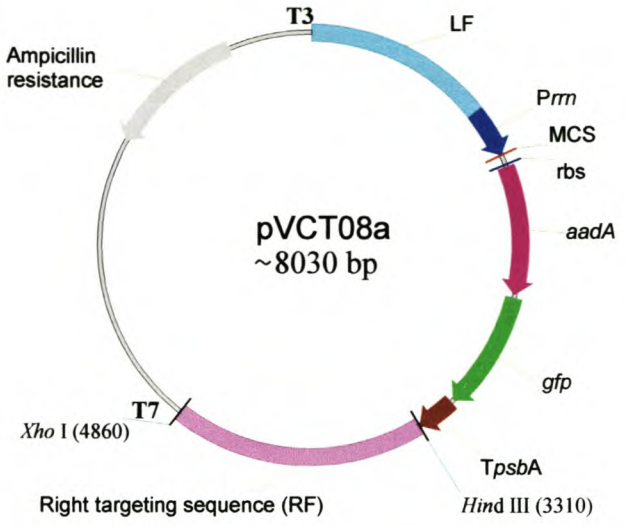


Figure 3.1.11c pVCT08a: LF + P*rrn* + MCS + rbs + *aadA* + *gfp* + T*psb*A + Grape right targeting sequence (RF) in pBluescriptSK®(+/-)

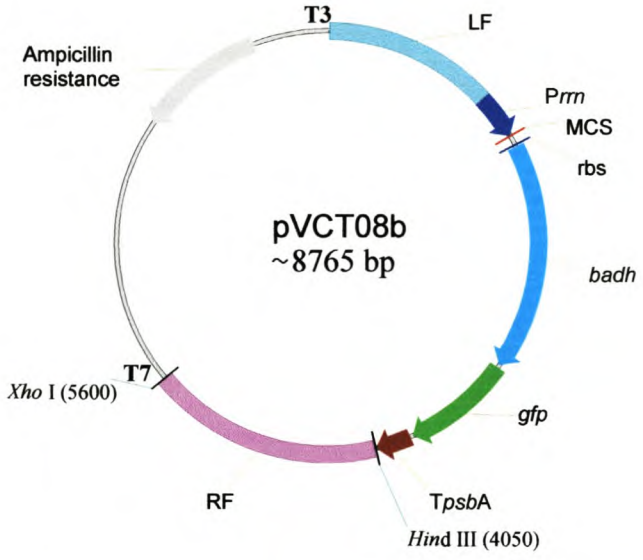


Figure 3.1.11d pVCT08b: LF + P*rrn* + MCS + rbs + *badh* + *gfp* + T*psb*A + Grape right targeting sequence (RF) in pBluescriptSK®(+/-)

3.1.12 Cloning of the Rubisco large subunit 5'-translation control region into pVCT08a and pVCT08b

In chloroplasts, post-transcriptional processes ultimately determine the level of protein expression, with mRNA abundance being a prerequisite, yet not an assurance of high-level protein accumulation (Maliga, 2001). To ensure efficient translation of *gfp*, the 5'-translation control region (TCR) of the gene encoding the large subunit of Rubisco (Ribulose Bisphosphate Carboxylase/Oxygenase), *rbcL*, was finally included into pVCT08a and pVCT08b (Kuroda & Maliga, 2001b; Maliga, 2002). The cloned TCR fragment included the 5'-untranslated region (UTR) and the first 42 nucleotides of *rbcL*,

encoding the first 14 amino acids of Rubisco, translationally fused to the second codon of *gfp*. The resultant translational fusion would thus be initiated from the first codon, ATG, of the *rbcL* sequence. The *rbcL* 5'-TCR was incorporated using a technique similar to that used for the inclusion of the right flank into pVCT07a and pVCT07b (see 3.1.11).

The sequence spanning the region from the 3'-end of the *aadA* and *badh* genes, and the 5'-region of *gfp* was first required to enable design of the 5'-portions of both primers for targeted insertion. The *gfp* sequence was obtained from a T3-primed sequencing reaction using pVCT02 as template. This sequence was used to design a reverse primer, GFP_AB_Rev, ± 350 bp downstream of the *gfp* initiation codon. The pVCT08a and pVCT08b plasmids were used as template for the respective reverse sequencing reactions. The sequence results indicated that the region separating the selectable marker genes from *gfp* was ample for the design of the 5'-portion of the forward integration primer. This meant that only one set of primers was required for the inclusion of the *rbcL* 5'-TCR into both constructs. The degenerate oligonucleotide primers, rbcL5'TCR_For_AB and rbcL5'TCR_Rev_AB, were designed to enable amplification and insertion of the *rbcL* 5'-TCR upstream of the second codon of *gfp* in pVCT08a and pVCT08b.

The *rbcL* 5'-TCR was amplified from grape ctDNA as a 140 bp fragment (Fig. 3.1.12a). Approximately 20 ng recipient plasmid (pVCT08a and pVCT08b) and 250 ng *rbcL* 5'-TCR PCR product were used for the thermal cycling elongation steps. Following *Dpn* I digestion and cloning, colonies were PCR-screened using the rbcL5'TCR_For_AB and GFP_AB_Rev primers, generating ± 450 bp fragments in positive, recombinant plasmids (Fig. 3.1.12b). This PCR indicated an ± 80% success rate (results not shown). pVCTa- and pVCTb-specific primer pairs were used to verify the initial screening using rbcL5'TCR_Rev_AB with the M69221_for_BamHI and BADH_for_BamHI primers, respectively (Fig. 3.1.12b). The amplified products were ± 900 bp (pVCTa) and ± 1 650 bp (pVCTb) in size. The respective clones were named pVCT09a and pVCT09b (Fig. 3.1.12c and 3.1.12d) and constituted the completed *Vitis* chloroplast transformation vectors.

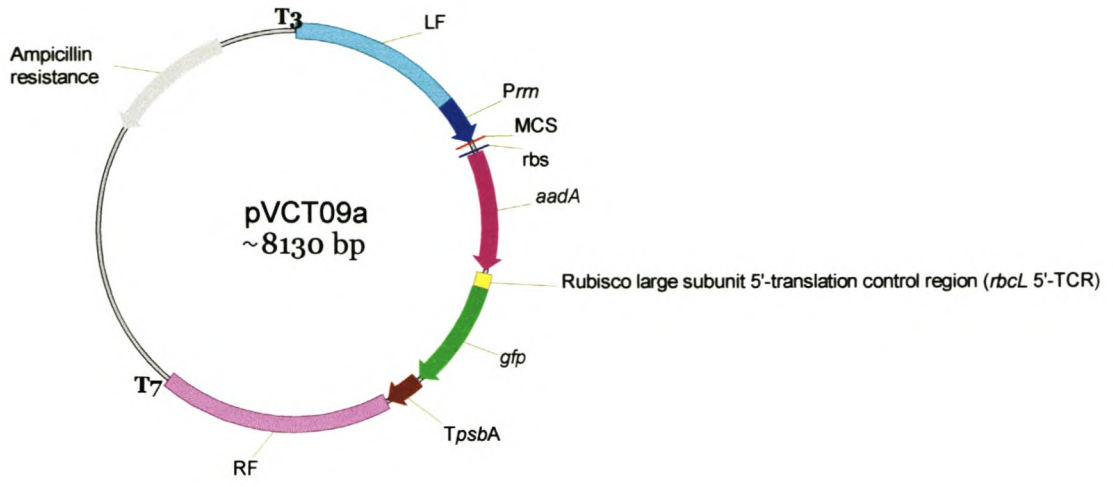


Figure 3.1.12c pVCT09a: LF + P*rrn* + MCS + rbs + *aadA* + Rubisco large subunit 5'-translation control region (*rbcL* 5'-TCR) + *gfp* + T*psb*A + RF in pBluescriptSK®(+/-)

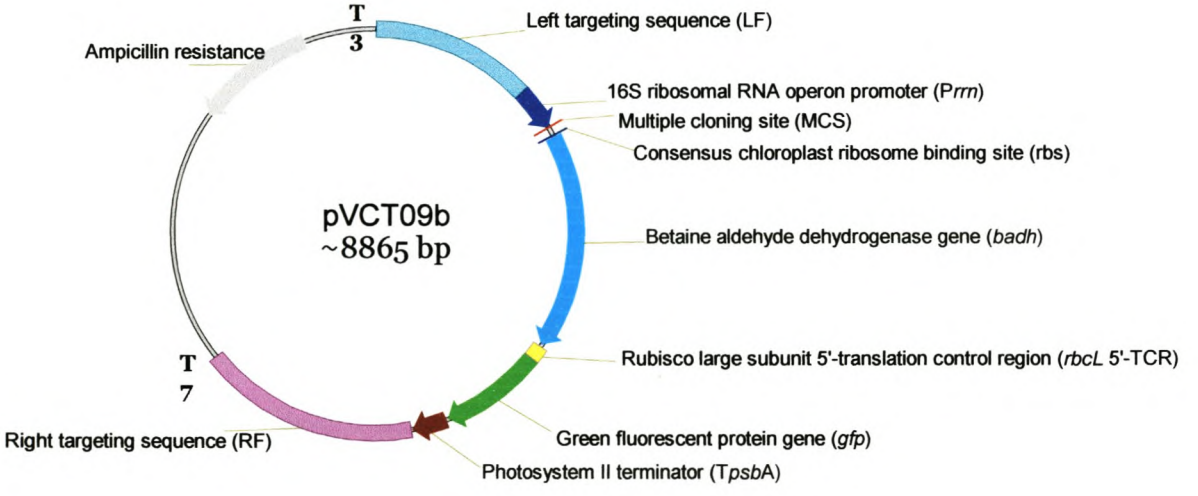


Figure 3.1.12d pVCT09b: LF + P*rrn* + MCS + rbs + *badh* + Rubisco large subunit 5'-translation control region (*rbcL* 5'-TCR) + *gfp* + T*psb*A + RF in pBluescriptSK®(+/-)

3.2 PROTEIN EXPRESSSION IN BACTERIA

3.2.1 Spectinomycin adenyltransferase expression

Expression of the spectinomycin adenyltransferase gene, *aadA*, was shown in *Escherichia coli* using the expression cassette from pVCT06a. The positive cassette (+) included the P*rrn*, MCS, rbs, *aadA*, *gfp*, and T*psb*A from pVCT06a, and was released by digestion with *Ksp* I and *Hind* III. The negative cassette (–) was similar, but excluded the promoter. This was achieved by digestion of pVCT06a with the enzymes, *Sph* I and *Hind* III. The released fragments of ± 2.1 kb (+) and ± 1.8 kb (–) in size, (Fig. 3.2.1a), were then ligated to a promoterless pMECA vector.

The pMECA MCS is flanked by the bacteriophage-derived T3 and T7 promoters. The *E. coli* RNA polymerase does not recognize these promoters, so there is virtually no transcription of downstream genes in *E. coli* (Mierendorf *et al.*, 1994). Therefore, neither promoter should have influenced the test for grape P*rrn* activity. However, due to conflicting professional opinions, it was decided to eliminate both the T7 and T3 promoters from the pMECA vector. T3 was first removed by digestion with *Eco* RI and *Pst* I, the former site being present upstream of the T3 promoter primer binding site, and the latter site, in the MCS. Both cohesive termini were blunted, then ligated, creating pMECA(-T3) (results not shown). This construct was digested with *Ksp* I and *Hind* III to facilitate cloning of the positive cassette, and *Sph* I and *Hind* III for cloning of the negative cassette (Fig. 3.2.1a). These digests resulted in the loss of the T7 promoter, as the *Hind* III site is downstream of T7, and the *Ksp* I and *Sph* I sites are present in the MCS. The respective clones were named pMECA/(+) and pMECA/(–). Resultant colonies were PCR-screened using the M69221_rev_BamHI primer with trnV_left_for and RBS_for, repectively. The former primer pair generated a fragment of ± 1.1 kb from only pMECA/(+), representing P*rrn* and *aadA*, and the latter, a ± 0.8 kb fragment from both the pMECA/(+) and pMECA/(–) clones, representing *aadA* (Fig. 3.2.1b).

Colonies harbouring the pMECA/(+) and pMECA/(–) constructs were grown in liquid cultures on variable concentrations of the antibiotic spectinomycin, and a constant concentration of the antibiotic ampicillin. The optical densities, measured at 600 nm (OD_{600}), and the corresponding spectinomycin concentrations, were recorded graphically (Fig. 3.2.1c). pMECA/(+) growth was unaffected by an antibiotic concentration of 200 µg/ml, whereas pMECA/(–) growth was greatly inhibited at a concentration of 100 µg/ml, and practically non-existent at 200 µg/ml.

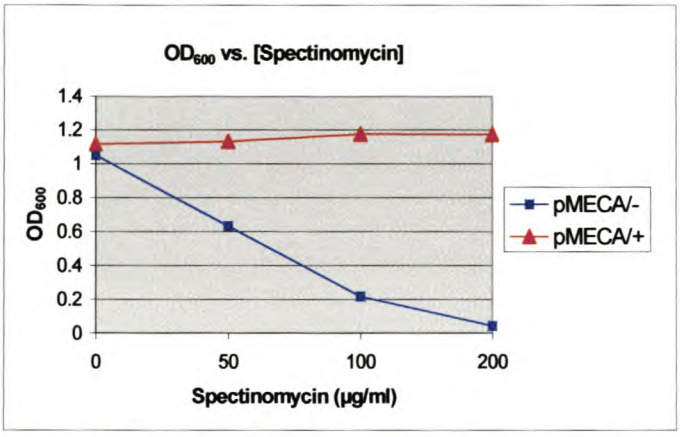


Figure 3.2.1c OD_{600} readings versus the spectinomycin concentration for pMECA/(+) and pMECA/(–). The graph illustrates active spectinomycin adenyltransferase expression in only those cells harbouring the pMECA/(+) construct.

Escherichia coli cells harbouring the pMECA/(+) construct were therefore resistant to the antibiotic, spectinomycin, and those harbouring the pMECA/(–) construct, sensitive. These results confirm grape chloroplast promoter activity and successful translation of the *aadA* gene. The encoded enzyme, aminoglycoside 3'-adenylyltransferase, catalyzes the covalent transfer of an AMP residue from ATP to aminoglycoside-type antibiotics, inactivating them. Thus, pMECA/(+) bacterial growth was unaffected by high concentrations of spectinomycin. Lack of the promoter in pMECA/(–) resulted in no transcription, and hence, no encoded enzyme, leading to sensitivity to spectinomycin, a prokaryotic translational inhibitor.

3.2.2 Green fluorescent protein expression

After inclusion of the *rbcL* 5'-TCR to yield pVCT09a and pVCT09b, transformed colonies were observed for GFP fluorescence to determine both transcription from P*rrn*, and translation via the *rbcL* 5'-TCR, of *gfp* (Brixey *et al.*, 1997; Daniell *et al.*, 2001; Kuroda & Maliga, 2001b; Maliga, 2002). *Escherichia coli* colonies harbouring the pVCT09b construct fluoresced (Fig. 3.2.2), but those harbouring the pVCT09a construct did not. Lack of fluorescence may be due to early termination of transcription in *E. coli*. However, what happens in *E. coli* may have no relevance in plastids (pers. comm., P Maliga, Waksman Institute, Rutgers, USA).

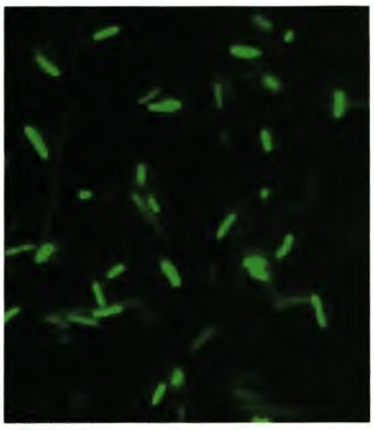


Figure 3.2.2 Illuminated *E. coli* cells harbouring the pVCT09b construct demonstrating active green fluorescent protein expression (UV 480 nm).

3.2.3 Betaine aldehyde dehydrogenase expression

The test for functional expression of the selectable marker gene, *badh*, in the pVCT09b construct, required the use of an *E. coli* mutant defective in the glycine betaine synthesis enzymes. Glycine betaine is an osmoprotectant that permits bacterial growth in high-osmolarity media (Boch *et al.*, 1996). It is synthesized in a two-step oxidation process with choline as precursor and betaine aldehyde as the intermediate. A betaine aldehyde dehydrogenase converts the betaine aldehyde to glycine betaine.

The *E. coli* genes required for glycine betaine synthesis include the choline transport gene, *betT*, the genes encoding the enzymes required for the conversion of choline to glycine betaine, *betBA*, and the regulatory gene, *betI* (Boch *et al.*, 1996). The *E. coli* strain MC4100 [Δ(*betTIBA*)] (obtained from E Bremer, Dept. of Microbiology, Philipps-University of Marburg, Germany), has a deletion in this gene cluster and is therefore unable to synthesize the osmoprotectant glycine betaine. This strain was therefore used to attempt cloning by functional complementation of the *badh* gene present in the pVCT09b expression cassette, to enable conversion of betaine aldehyde to glycine betaine. Since *betT* in MC4100 [Δ(*betTIBA*)] is deleted, this strain lacks the ability to transport exogenously provided choline or betaine aldehyde into the cell (pers. comm., E Bremer, Dept. of Microbiology, Philipps-University of Marburg, Germany). The *betT* gene was therefore included on a separate pBR322-derived plasmid, pJB004 (Boch *et al.*, 1996).

Since the pJB004 and pVCT09b plasmids both included ampicillin resistance genes and the ColE1 replicon, the pVCT09b expression cassette was cloned into the *Sma* I and *Hind* III sites in the pGL10 multiple cloning site. The pGL10 vector contains a kanamycin resistance gene and the RK2/RP4 replicon, and would therefore be compatible in its replication with pJB004. The pVCT09b expression

cassette was released by digestion with *Eco* RV and *Hind* III. Colonies were selected on kanamycin and PCR-screened using the BADH_for_BamHI and BADH_rev_BamHI primers, generating a fragment of ± 1.5 kb in size (results not shown). The selected clone was named pGL10_9bExpressionCassette (EC). Both the pGL10_9bEC and pJB004 plasmids were transformed into competent MC4100 [Δ(*betTIBA*)] cells.

The test for functional BADH expression in the pVCT09b expression cassette was carried out using the *E. coli* strain MC4100 [Δ(*betTIBA*)], including the pGL10_9bEC and pJB004 plasmids. The resultant strain was labeled MC4100/pJB004/pGL10_9bEC (SAK). The negative control was similar, however, excluded pGL10_9bEC, and was labeled MC4100/pJB004 (SA). These strains were grown in minimal media since rich media such as LB broth contains glycine betaine that could be taken up by the cells (pers. comm., E Bremer, Dept. of Microbiology, Philipps-University of Marburg, Germany). Both strains were grown in minimal media containing 200 mM NaCl and 1 mM betaine aldehyde. The ability to withstand osmotic stress through the synthesis of the osmoprotectant glycine betaine, from the precursor betaine aldehyde, would therefore suggest the presence of functional BADH. Optical densities were measured daily at 578 nanometres (OD_{578}) and recorded graphically (Fig. 3.2.3).

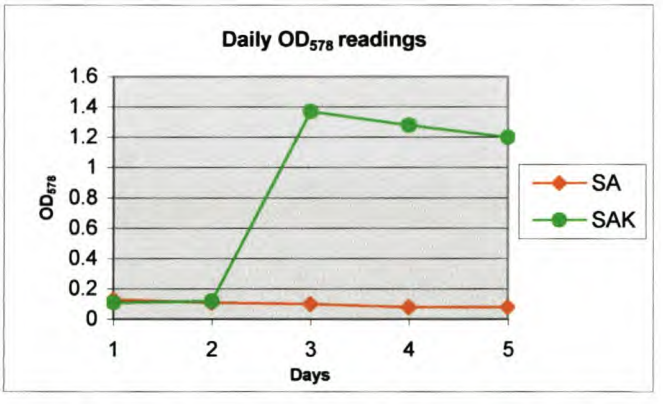


Figure 3.2.3 OD_{578} readings taken daily for the MC4100/pJB004/pGL10_9bEC (SAK) and MC4100/pJB004 (SA) strains. Both strains were grown in minimal media containing 200 mM NaCl and 1 mM betaine aldehyde. The graph illustrates increased growth of the MC4100/pJB004/pGL10_9bEC (SAK) strain after 3 days, whereas the MC4100/pJB004 (SA) strain showed no tolerance to osmotic stress during the 5 days tested.

The presence of the pJB004 plasmid in both the MC4100/pJB004 (SA) and MC4100/pJB004/pGL10_9bEC (SAK) strains, enabled uptake of the prescursor, betaine aldehyde. Betaine aldehyde accumulation in cells unable to convert it to glycine betaine is highly toxic. After five days growth, the MC4100/pJB004 (SA) strain showed no tolerance to osmotic stress, while the MC4100/pJB004/pGL10_9bEC (SAK) strain only indicated osmotic stress tolerance after three days' growth. A possible explanation for this slow response could be due to low *badh* copy number since the pVCT09b expression cassette was cloned into the pGL10 vector, a low copy number plasmid. This could have led to reduced BADH expression levels that may have resulted in delayed conversion

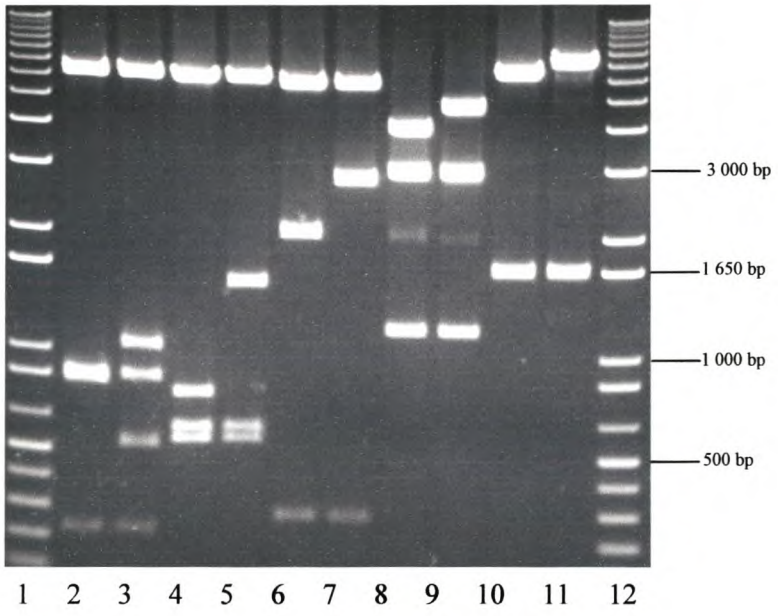
of the accumulated betaine aldehyde. Further confirmation would require cloning of the pVCT09b expression cassette into a high copy number plasmid.

3.3 SEQUENCING

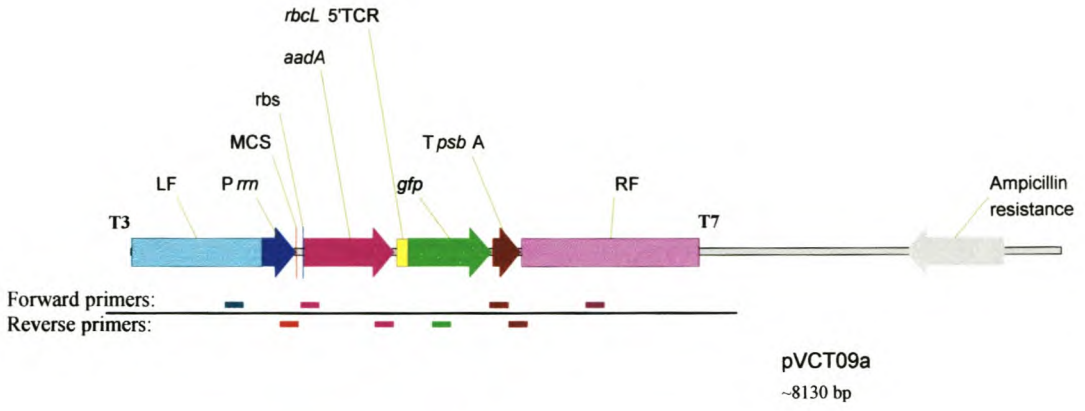
Approximately 200 ng of pVCT09a or pVCT09b plasmid DNA was digested prior to sequencing using the following restriction endonuclease combinations: *Sal* I / *Eco* RI / *Spe* I, *Bam* HI, *Not* I, *Sac* II, and *Hind* III / *Xho* I. The resultant restriction map gave the expected band profiles for both constructs.

Restriction map:

Lanes 1&12, 1 kb+ DNA ladder
Lanes 2&3, *Sal* I / *Eco* RI / *Spe* I
Lanes 4&5, *Bam* HI
Lanes 6&7, *Not* I
Lanes 8&9, *Sac* II
Lanes 10&11, *Hind* III / *Xho* I
The first lane of all digests represents pVCT09a, and the second, the pVCT09b construct.



Apart from one new primer, RF_for_seq, existing primers (T3, T7, rps12_for_trnV, M69221_for_BamHI, BADH_for_BamHI, psbA3'UTR_for, RF_for_seq, StuI_rev, M69221_rev_BamHI, BADH_rev_BamHI, GFP_AB_Rev, and psbA3'UTR_rev) were used for the complete sequencing of the cloned genetic elements in both the pVCT09a and pVCT09b constructs. A linear map of the pVCT09a construct was used to align the approximate positions of the respective primers (see below).



Forward primers:
- rps12_for_trnV
- M69221_for_BamHI
- psbA3'UTR_for
- RF_for_seq

Reverse primers:
- StuI_rev
- M69221_rev_BamHI
- GFP_AB_Rev
- psbA3'UTR_rev

The sequence results were as expected for pVCT09a, however pVCT09b had six base pairs replacing the 'A' in the start codon of *badh*. (see below: rbs – bold; *Bam* HI – underlined; 6 bp - lower case)

… **AGG AGG** CGA GGA TCC cag acg TGG CGA …

The included base pairs were likely derived from the BADH_for_BamHI primer due to incorrect oligonucleotide synthesis since the sequence from the original clone was identical. The error was corrected using a PCR-based site-directed mutagenesis protocol, analogous to that used in 3.1.7. The 5'-end of the MUTbadh_for primer annealed to the second nucleotide of *badh*, and the 5'-end of the MUTbadh_rev primer was complementary to the 'A' in the *Bam* HI site. This resulted in the removal of the following nucleotides: TCC cag acg (see below).

…**AGG AGG** CGA GGA TGG CGA…

The rbs was then directed 5 nucleotides upstream of the initiation codon, with the elimination of the *Bam* HI site.

The complete sequences of the pVCT09a and pVCT09b constructs, from the T3 to the T7 primer binding sites, are given in Appendix C (Bioedit version 5.0.9, Hall, 1999).

4. CONCLUSION

In this study, two expression vectors for the transformation of the grape chloroplast genome, namely pVCT09a and pVCT09b, were constructed. The expression cassettes included a promoter of grape chloroplast origin; a multiple cloning site for the inclusion of the desired transgene; a selectable and screenable marker with upstream ribosome binding site and translation control region respectively; a terminator, also of grape origin; and left and right flanking sequences to enable site-specific integration into the grape plastid genome. The pVCT09a construct contains an antibiotic resistance selectable marker, and the pVCT09b construct, a plant-derived selectable marker. Due to the prokaryotic origin of plastid organelles, it was possible to test the expression cassettes in *Escherichia coli*. GFP fluorescence was visualized only in those cells harbouring the pVCT09b construct, however, since the sequences controlling GFP expression are identical in both constructs, we expect this result to have no relevance to plastids (pers. comm., P Maliga, Waksman Institute, Rutgers, USA). Cells harbouring the pVCT09a construct were resistant to the antibiotic spectinomycin, confirming spectinomycin adenyltransferase expression. The expression of the selectable marker in the pVCT09b construct was tested by functional complementation using an *E. coli* strain lacking betaine aldehyde dehydrogenase activity. Initial results suggest functional BADH expression; however, the pVCT09b expression cassette should be cloned into a high copy number plasmid to attempt increased BADH expression and confirm results presented here.

The pVCT09b vector, although designed for grape, is currently being evaluated in tobacco (C Shi, Key lab of Agricultural Microbiology, Huazhong Agricultural University, China). Although the chloroplast genome is highly conserved among higher plants, the efficiency may be reduced due to the grape-specific targeting regions; however, tobacco plastid transformation is still expected to be a success using this construct (pers. comm., P Maliga, Waksman Institute, Rutgers, USA). Increased tobacco plastid transformation efficieny would be expected if grape-specific flanking sequences were replaced with tobacco-specific flanks. The pVCT09a/b vectors were designed in order to enable this change, with the release of the right flank prior to the left flank.

The multiple cloning site in the pVCT09a/b vectors is downstream and adjacent to the promoter. Since no information has yet been published on read-through transcription at an MCS (Maliga, 2002; pers. comm., P Maliga, Waksman Institute, Rutgers, USA), it was decided to include a very short MCS. This limits cloning of transgenes to a few unique RE sites. If no suitable RE sites are available, transgenes could be inserted following a protocol that does not require the use of RE or DNA ligase (Geiser *et al.*, 2001; Chen *et al.*, 2000; Tseng *et al.*, 1999). This strategy was successfully applied in

the cloning of both the right flank (see 3.1.11) and the Rubisco large subunit 5'-translation control region (see 3.1.12).

Future objectives include plastid transformation in grape using embryogenic callus as target for electroporation (pers. comm., A Perl, Agricultural Institute, Bet-Dagan, Israel). The use of non-green material will assist in visualization of the green fluorescent protein screenable marker, however, targeting the proplastids will restrict this effort to using the pVCT09a vector. The betaine aldehyde dehydrogenase selectable marker in the pVCT09b vector is specific to chloroplasts, which would require the use of green explant material (Daniell *et al.*, 2001a). To our knowledge, no protocols are currently available for the regeneration of grapevines from green explant material.

Two expression vectors for the transformation of the grape chloroplast genome have been constructed in this research. Functional expression and efficient integration of the expression cassette into the grape plastid genome is expected for both the pVCT09a and pVCT09b constructs.

APPENDIX A

Primers:

trnV_left_for:	5'-GGT GGA AGT CAT CAG TTC GAG CC-3'
16SrRNA_left_rev:	5'-AAC CAC ATG CTC CAC CGC TTG T-3'
T7:	5'-TAA TAC GAC TCA CTA TAG GG-3'
T3:	5'-ATT AAC CCT CAC TAA AGG GA-3'
M69221_for_BamHI:	5'-TTA GGA TCC ATG AGG AGG ATA TAT TTG-3'
M69221_rev_BamHI:	5'-GGC GGA TCC TTA TAA TTT TTT TAA TCT G-3'
psbA3'_for:	5'-CCA CTG CCT TRA TCC ACT TGG-3'
psbA3'_rev:	5'-GTG CTA TGC AYG GTT CCT TGG-3'
SP6:	5'-TAC GAT TTA GGT GAC ACT ATA G-3'
psbA3'UTR_for:	5'-GAT TTC CGT CTT AGT GTT TAC G-3'
psbA3'UTR_rev:	5'-GAT ACT CAA TCA TAA ACC AAC C-3'
RBS_for:	5'-GGA GGC GAG GAT CCA TGA GGA GG-3'
StuI_rev:	5'-TAA GGC CTC GTT TAA ACA GTC GAC G-3'
5'UTR_for(BADH):	5'-GTC ATC TTG CTA TAA TAT CAA CG-3'
3'UTR_rev(BADH):	5'-ATG ATG ATG GCG TCT TGC-3'
BADH_for_BamHI:	5'-TTA GGA TCC ATG GCG ATC CGT GTA CCT TCG-3'
BADH_rev_BamHI:	5'-GCT GGA TCC TCA AGG AGA CTT GTA CCA TCC-3'
rps12_for(JR):	5'-GCA GGA TCC TAG CCA TAC ACT TCA C-3'
5'16SrRNA_left_rev(JR):	5'-ATA CCG CGG AGA TTC ATA GTT GC-3'
rps12_for_fillgap:	5'-TCA CGA TCT TAT AAT AAG AAC AAG-3'
rps12_for_SEQ:	5'-TTT GCT CAT TTG GAA TCT GG-3'
rps12_for_trnV:	5'-CGT GCT TTG GTG GGT CTC-3'
SacII_for_LF:	5'-TAT CCG CGG AAT GAT ACA AGT CTT GG-3'
SacII_rev_LF:	5'-ATA CCG CGG TTA CGC AAT CGA TCG-3'
HindIII_for_RF:	5'-TGT AAG CTT GTA GCA ACG GAA CCG-3'
XhoI_rev_RF:	5'-ATA CTC GAG GTA TTA CCG CGG CTG-3'
Appel_RF_for:	5'-ATC AAG CTT ATC GAT ACC GTC GGT AGC AAC GGA ACC G-3'
Appel_RF_rev:	5'-GTA CCG GGC CCC CCC TCG AGG TAT TAC CGC GGC TG-3'
GFP_AB_Rev:	5'-GCC CTC GAA CTT CAC CTC G-3'
rbcL5'TCR_For_AB:	5'-CCC GGG CTG CAG GAA TTC GGT CGA GTA GAC CTT GTT G-3'
rbcL5'TCR_Rev_AB:	5'-CAG CTC CTC GCC CTT GCT CAC TTT RAA TCC AAC ACT TGC-3'
RF_for_seq:	5'-GTA ATG AAC TGT TGG CAC C-3'
MUTbadh_for:	5'-TGG CGA TCC GTG TAC CTT CG-3'
MUTbadh_rev:	5'-TCC TCG CCT CCT AAG GCC TCG-3'

APPENDIX B

- All gel photos
- DNA marker standards: base pairs (bp)

Grape chloroplast DNA isolation **3.1.1**

Figure 3.1.1
Lane 1, 1 kb+ DNA ladder
Lane 2, Grape ctDNA

Cloning of the 16S ribosomal RNA operon promoter **3.1.2**

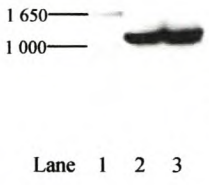


Figure 3.1.2a
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, 1 166 bp ~P*rrn*

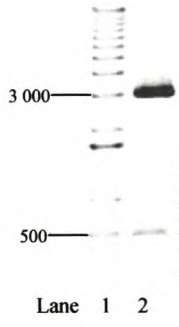


Figure 3.1.2b
Lane 1, 1 kb+ DNA ladder
Lane 2, *Xcm* I-digested pT-NOT vector, releasing the ± 500 bp stuffer fragment

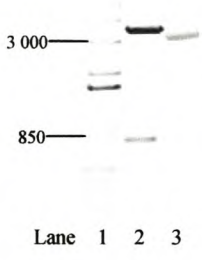


Figure 3.1.2c
Lane 1, 1 kb+ DNA ladder
Lane 2, *Ksp* I digest of pVCT01 (~P*rrn*: T7→T3 orientation) Lane 3, Undigested pVCT01

Cloning of the green fluorescent protein gene into pVCT01 **3.1.3**

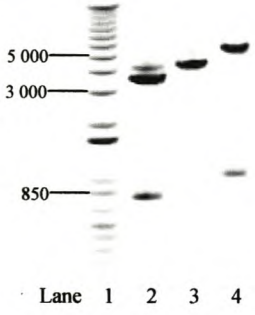


Figure 3.1.3a
Lane 1, 1 kb+ DNA ladder
Lane 2, *Ksp* I-digested pVCT01 Lane 3, *Eco* RI-digested pVCT01 Lane 4, *Eco* RI-digested pGEMUbi1-sg*fp*S65T, releasing *gfp*_nos

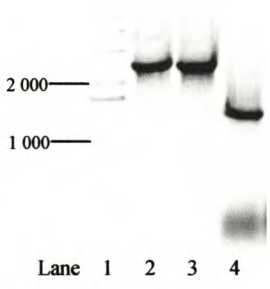


Figure 3.1.3b
T7 / T3 PCR on 3 colonies
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, Amplification of both ~P*rrn* and *gfp*_nos
Lane 4, Amplification of only ~P*rrn*

Promoter shortening and cloning of a multiple cloning site into pVCT02

3.1.4

Figure 3.1.4a
Lane 1, I kb+ DNA ladder
Lanes 2&3, *Sph* I / *Bam* HI-digested pVCT02

Figure 3.1.4b
Lane 1, I kb+ DNA ladder
Lane 2, *Sph* I / *Bam* HI-digested T7/T3 PCR product from pMECA

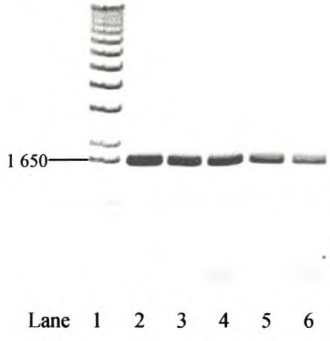


Figure 3.1.4c
Lane 1, I kb+ DNA ladder
Lanes 2-6, T7 / T3 PCR on 5 colonies representing the P*rrn*, MCS, and *gfp*_nos regions

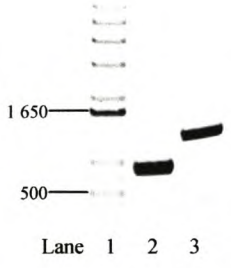


Figure 3.1.4d
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, *Not* I-digested T7/T3 PCR fragments indicating T3→T7 and T7→T3 *gfp*_nos orientations, respectively

Cloning of the spectinomycin adenyltransferase gene into pVCT03

3.1.5

Figure 3.1.5a
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, 786 bp *aadA*

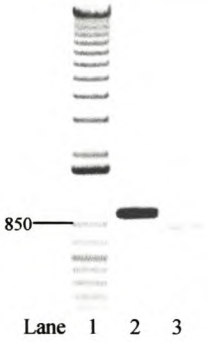


Figure 3.1.5b
Lane 1, 1 kb+ DNA ladder
Lane 2, T7/T3 *aadA* PCR product Lane 3, *Bam* HI-digested PCR product of lane 2

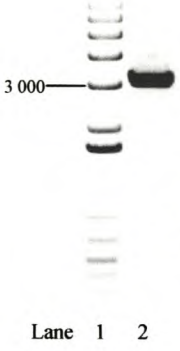


Figure 3.1.5c
Lane 1, 1 kb+ DNA ladder
Lane 2, *Not* I-digested pBluescriptSK®(+/-)

Figure 3.1.5d
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, trnV_left_for / T7 PCR on colonies (i) & (ii), respectively Lanes 4&5, M69221_rev_BamHI / T3 PCR on colonies (i) & (ii), respectively

Cloning of the photosystem II terminator into pVCT04 **3.1.6**

Figure 3.1.6a
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, T7 / SP6 colony PCR, representing the cloned *psb*A terminator region

Figure 3.1.6b
Lane 1, 1 kb+ DNA ladder
Lane 2, psbA3'UTR_for / T7 colony PCR (SP6→T7) Lanes 3&4, T7/SP6 T*psb*A PCR product

Figure 3.1.6c
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, *Sac* II-digested T7/SP6 T*psb*A PCR product

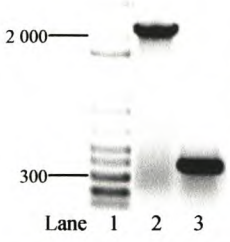


Figure 3.1.6d
Lane 1, 1 kb+ DNA ladder
Lane 2, psbA3'UTR_rev / T3 colony PCR. Lane 3, psbA3'UTR_for / T7 colony PCR

Cloning of the betaine aldehyde dehydrogenase gene into pVCT06a **3.1.8**

Figure 3.1.8a
Amaranthus plants at Welgevallen

Figure 3.1.8b
Lane 1, 1 kb+ DNA ladder
Lane 2, *A. hypochondriacus* RNA

Figure 3.1.8c
Lanes 1&2, *A. hypochondriacus* and *A. tricolor badh*, respectively (*A. hybridus badh* not shown) Lane 3, *Sty* I-digested λ DNA

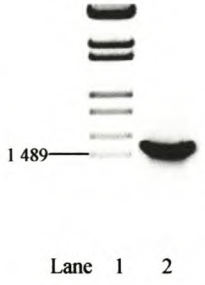


Figure 3.1.8d
Lane 1, *Sty* I-digested λ DNA Lane 2, T7 / T3 colony PCR, representing the cloned *A. tricolor badh*

Cloning of the betaine aldehyde dehydrogenase gene into pVCT06a **3.1.8**

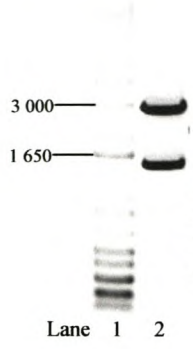


Figure 3.1.8e
Lane 1, 1 kb+ DNA ladder
Lane 2, *A. tricolor badh* released from pT-NOT via *Bam* HI digestion

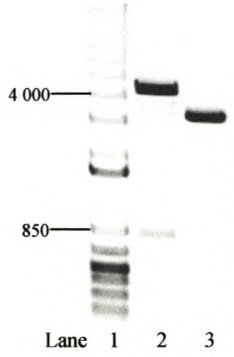


Figure 3.1.8f
Lane 1, 1 kb+ DNA ladder
Lane 2, *AadA* released from pVCT06a via *Bam* HI digestion
Lane 3, Undigested pVCT06a

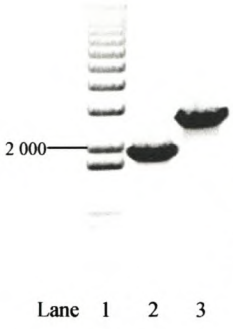


Figure 3.1.8g
Lane 1, 1 kb+ DNA ladder
Lane 2, BADH_rev_BamHI / T3 colony PCR Lane 3, BADH_for_BamHI / T7 colony PCR

Sequencing of the grape plastid targeting region **3.1.9**

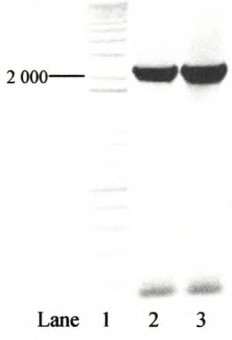


Figure 3.1.9
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, 2 240 bp PCR product

Cloning of the left flank into pVCT06a and pVCT06b **3.1.10**

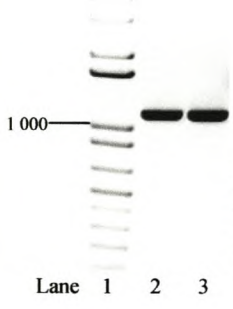


Figure 3.1.10a
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, 1 134 bp left flank (LF)

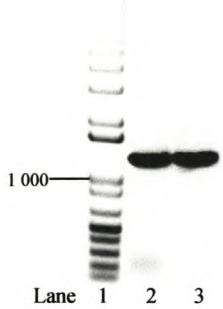


Figure 3.1.10b
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, T7 / SP6 colony PCR, representing the cloned LF

Cloning of the left flank into pVCT06a and pVCT06b **3.1.10**

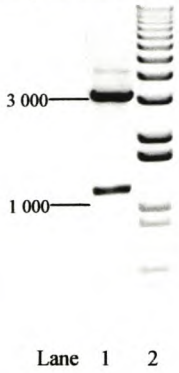


Figure 3.1.10c
Lane 1, 1 kb+ DNA ladder
Lane 2, Left flank released from pGEM®–T Easy vector via *Sac* II digestion

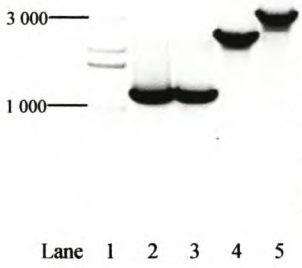


Figure 3.1.10d
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, SacII_rev_LF / T3 PCR on pVCTa- and pVCTb-colonies, respectively Lane 4, pVCTa-specific PCR Lane 5, pVCTb-specific PCR

Cloning of the right flank into pVCT07a and pVCT07b **3.1.11**

Figure 3.1.11a
Lane 1, 1 kb+ DNA ladder
Lanes 2&3, 1 576 bp right flank (RF)

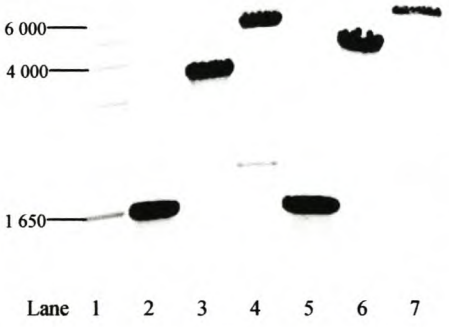


Figure 3.1.11b
Lane 1, 1 kb+ DNA ladder
Lanes 2&5, HindIII_RF_for / T7 PCR on pVCTa- and pVCTb-DNA, respectively Lane 3, pVCTa-specific PCR Lane 6, pVCTb-specific PCR Lanes 4&7, T7/T3 PCR products from pVCT08a and pVCT08b plasmids

Cloning of the Rubisco large subunit 5'-translation control region into pVCT08a and pVCT08b **3.1.12**

Figure 3.1.12a
Lane 1, 1 kb+ DNA ladder Lanes 2-4, 140 bp *rbcL* 5'-TCR

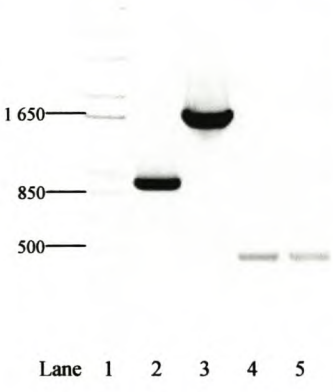


Figure 3.1.12b
Lane 1, 1 kb+ DNA ladder
Lane 2, M69221_for_BamHI / rbcL5'TCR_Rev_AB PCR on pVCTa-DNA Lane 3, BADH_for_BamHI / rbcL5'TCR_Rev_AB PCR on pVCTb-DNA Lanes 4&5, rbcL5'TCR_For_AB / GFP_AB_Rev on pVCTa- and pVCTb-specific DNA, respectively

Spectinomycin adenyltransferase expression

3.2.1

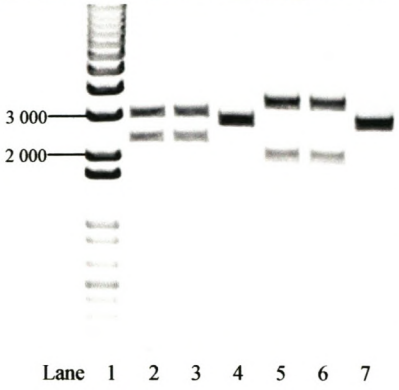


Figure 3.2.1a
Lane 1, 1 kb+ DNA ladder Lanes 2&3, *Ksp* I / *Hind* III-digested pVCT06a Lane 4, *Ksp* I / *Hind* III-digested pMECA(-T3) Lanes 5&6, *Sph* I / *Hind* III-digested pVCT06a Lane 7, *Sph* I / *Hind* III-digested pMECA(-T3)

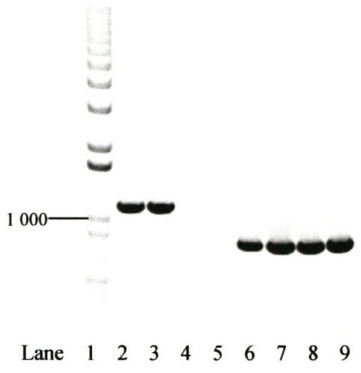


Figure 3.2.1b
Lane 1, 1 kb+ DNA ladder Lanes 2-5, M69221_rev_BamHI / trnV_left_for colony PCR on pMECA/(+), (2&3) and pMECA/(-), (4&5) Lanes 6-9, M69221_rev_BamHI / RBS_for colony PCR on pMECA/(+), (6&7) and pMECA/(-), (4&5)

APPENDIX C

pVCT09a:

```
1    TGGAGCTCCCGCGGAATGATACAAGTCTTGGATAAGAATCTACAACGCACTAGAACGCCCTTGTTGACGA 70
71   TCCTTTACTCCGACAGCGTCTAGAGTTCCTCGAACAATGTGATATCTCACACCGGGTAAATCCTTAACCC 140
141  TTCCCCCTCTTACTAAGACTACAGAATGTTCTTGTGAATTATGGCCAATACCAGGTATATAAGCAGTGAT 210
211  TTCAAATCCAGAGGTTAATCGTACTCTGGCAACTTTACGTAAGGCAGAGTTTGGTTTTTTGGGGGTGATA 280
281  GTGGAAAAGTTGACAGATAAGTCACCCTTACTGCCACTCTACAGAACCGTACATGAGATTTTCACCTCAT 350
351  ACGGCTCCTCGTTCAATTCTTTTGAAGTCATTGGATCCTTTTCCTCGTTCGAGAATCTCCTCCCCTCTTC 420
421  CACTCCGTCCCGAAGAGTAACTAGGACCAATTCAGTCACGTTTTCATGTTCCAATTGAACACTTTCGATT 490
491  TTTGATTATTCTCAAATCAAAGGAGAAGATTATTCTTTTTACCAAACATATGCGGATCAAATCACGATCT 560
561  TATAATAAGAACAAGAGATCTTTCTCGATCAATCCCTTTGCCCGAGAATCAGAAAGATCCTTTTCAAGTT 630
631  TGAATTTGCTCATTTGGAATCTGGGTTCTTCTACTTCATTTTTATTTACTTATTTATTTATTATTTTGAT 700
701  TTCCCTCTCTTTTCTTTTTTTATTCCCTTCCATCATTCCTTAAGTCCCATAGGTTTGATCCTGTAGAATC 770
771  TGACCCATTTTCTCATTGAGCGAAGGGTACGAAATAATCAGATTGATTTTTCGATCAAAGTACTATGTGA 840
841  AATCTTCGGTTTTTTCCTCTTCCTCTATCCCTATCCCATAGGTACAGCGTTTGAATCAATAGAGAACCTT 910
911  TTCTTCTGTATCTGTATGAATCGATATTATTACATTCCAATTCCTTCCCGATACCTCCCAAGGAAAATCC 980
981  CGAAGTGGATCCCAAATTGACGGGTTAGTGTGAGCTTATCCATGCGGTTATGCACTCTTCGAATAGGAAT 1050
1051 CCATTTTCTGAAAGATCCTGACTTTCGTGCTTTGGTGGGTCTCCGAGACCCTTTCGATGACCTATGTTGT 1120
1121 GTTGAAGGGATATCTATATGATCCGATCGATTGCGTAACCGCGGTGGCGGCCGCTCTAGGCCAGCCTTGG 1190
1191 TGGAAGTCATCAGTTCGAGCCTGATTATCCCTAAACCCAATGTGAGTTTTTCTATTTTGACTTGCTCCCC 1260
1261 CGCCGTGATCGTGGATAAGAGGCTCGTGGGATTGACGTGAGGGGGTAGGGATGACTATATTTCTGGGAGC 1330
1331 GAACTCCAGGCGAATATGAAGCGCATGGATACAAGTTATGCCTTGGAATGAAAGACAATTCCGAATCCGC 1400
1401 TTTGTCTACGAACAAGGAAGCTATAAGTAATGCAACTATGAATCTCATGGAGAGTTCGATCCTGGCTCAG 1470
1471 GATGAACGCTGGCGGCATGCTGCAGCCCGGGCGACGTCGACTGTTTAAACGAGGCCTTAGGAGGCGAGGA 1540
1541 TCCATGAGGAGGATATATTTGAATACATACGAACAAATTAATAAAGTGAAAAAAATACTTCGGAAACATT 1610
1611 TAAAAAATAACCTTATTGGTACTTACATGTTTGGATCAGGGAGTTGAGAGTGGACTAAAACCAAATAGTG 1680
1681 ATCTTGACTTTTTAGTCGTCGTATCTGAACCATTGACAGATCAAAGTAAAGAAATACTTATACAAAAAAT 1750
1751 TAGACCTATTTCAAAAAAAATAGGAGATAAAAGCAACTTACGATATATTGAATTAACAATTATTATTCAG 1820
1821 CAAGAAATGGTACCGTGGAATCATCCTCCCAAACAAGAATTTATTTATGGAGAATGGTTACAAGAGCTTT 1890
1891 ATGAACAAGGATACATTCCTCAGAAGGAATTAAATTCAGATTTAACCATAATGCTGTACCAAGCAAAACG 1960
1961 AAAAAATAAAAGAATATACGGAAATTATGACTTAGAGGAATTACTACCTGATATTCCATTTTCTGATGTG 2030
2031 AGAAGAGCCATTATGGATTCGTCAGAGGAATTAATAGATAATTATCAGGATGATGAAACCAACTCTATAT 2100
2101 TAACTTTATGCCGTATGATTTAACTATGGACACGGGTAAAATCATACCAAAAGATATTGCGGGAAATGCA 2170
2171 GTGGCTGAATCTTCTCCATTAGAACATAGGGAGAGAATTTTGTTAGCAGTTCGTAGTTATCTTGGAGAGA 2240
```

2241 ATATTGAATGGACTAATGAAAATGTAAATTTAACTATAAACTATTTAAATAACAGATTAAAAAAATTATA 2310

2311 AGGATCCCCCGGGCTGCAGGAATTCGGTCGAGTAGACCTTGTTTTGTTGTTGTGAAAATTCTTAATTCGT 2380

2381 GAGTTGTAGGGAGGGACTTATGTCACCACAAACAGAGACTAAAGCAAGTGTTGGATTTAAAGTGAGCAAG 2450

2451 GGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGT 2520

2521 TCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCAC 2590

2591 CGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTCACCTACGGCGTGCAGTGCTTCAGCCGC 2660

2661 TACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCA 2730

2731 CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGT 2800

2801 GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTAC 2870

2871 AACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGA 2940

2941 TCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGA 3010

3011 CGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAG 3080

3081 AAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCACGGCATGGACGAGCTGT 3150

3151 ACAAGTAAAGCGGCCGCTCTAGAACTAGTGATTGATTTCCGTCTTAGTGTTTACGAGTTTTTGAATGTAA 3220

3221 AGGAGCAATAACCAATTTCCTCTTTTATCAAGGGGGATGGTATTGCTCCTTTATTTAGTAGTCTTTTATT 3290

3291 TATCTTAGTAGTCTTTTACTTTTCTAAGTTTTTTTTATTTCTTTATTTCAACTAAAACTAAAATAGTAAT 3360

3361 AAAAAGTATTCTCATAGGTTGGTTTATGATTGAGTATCAATCGAATTCCCCTAGTGAATTCGCGGCCGCC 3430

3431 TGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCTATAGTGTCACCT 3500

3501 AAATCGTAATCAAGCTTATCGATACCGTCGGTAGCAACGGAACCGGGGAAAGTATACAGAAAAGACAGTT 3570

3571 CTTTTCTATTATATTAGTATTAGTTAGTGATCCCGGCTCAGTGAGCCCTTTCTTACGTGATGAACTGTTG 3640

3641 GCACCAGTCCTACATTTTGTCTCTGTGGACCGAGGAGAAGGGGGCTCAGCGGGAAGAGGATTGTACCATG 3710

3711 AGAGAAGCAAGGAGGTCAACCTCTTCTCTTTCAAATATACAACATGGATTCTGACAATGCAATGTAGTTG 3780

3781 GACTCTCATGTCGATCCGAATGAATCATCTTTTCAACGGAGGTAATGCCTGCTAGGTAAGAGGATAGCAA 3850

3851 GTTACAAATTCTGTCTCGGAAGGAATTTGTCCATTTTTCGGGGTCTCAAAGGGGCGTGGAAACACATAAG 3920

3921 AACTCTTGAATGGAAATGGAAAAGAGATGTAACTCCAGTTCCTTCGGAAATGGTAAGATCTTTGGCGCAA 3990

3991 GAAGAAGGGGTTGATCCGTATCATCTTGACTTGGTTCTGATTCCTCTATTTTTTTAAGAATACCGAGTCG 4060

4061 GGTTCTTCTCCTACCCGTATCGAATAGAACATGCTGAGCCAAATCTTCTTCACTGCTTGCTTTAGATCGG 4130

4131 GAAAATCGTACGGTTTTATGAAACCATGTGCTATGGCTCGAATCCGTAGTCAATCCTATTTCCGATAGGA 4200

4201 GCAGTTGACAATTGAATCCCATTTTTCCCATTATTTTCGTATCCGTAATAGTGCGAAAAGAAGGCCCGGC 4270

4271 TCCAAGTTGTTCAAGAATAGTGGCGTTGAGTTTCTCGACCCTTTGCCTTACTTAGGATTAGTCAGTTCTA 4340

4341 TTTCTCGATAGGGGCAGGGAAGGGATATAACTCAGCGGTAGAGTGTCACCTTGACGTGGTGGAAGTCATC 4410

4411 AGTTCGAGCCTGATTATCCCTAAACCCAATGTGAGTTTTTCTATTTTGACTTGCTCCCCCGCCGTGATCG 4480

4481 TGGATAAGAGGCTCGTGGGATTGACGTGAGGGGGTAGGGATGACTATATTTCTGGGAGCGAACTCCAGGC 4550

4551 GAATATGAAGCGCATGGATACAAGTTATGCCTTGGAATGAAAGACAATTCCGAATCCGCTTTGTCTACGA 4620

```
4621 ACAAGGAAGCTATAAGTAATGCAACTATGAATCTCATGGAGAGTTCGATCCTGGCTCAGGATGAACGCTG 4690
4691 GCGGCATGCTTAACACATGCAAGTCGGACGGGAAGTGGTGTTTCCAGTGGCGGACGGGTGAGTAACGCGT 4760
4761 AAGAACCTGCCCTTGGGAGGGGAACAACAGCTGGAAACGGCTGCTAATACCCCGTAGGCTGAGGAGCAAA 4830
4831 AGGAGGAATCCGCCCGAGGAGGGGCTCGCGTCTGATTAGCTAGTTGGTGAGGCAATAGCTTACCAAGGCG 4900
4901 ATGATCAGTAGCTGGTCCGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGACTCCTACGGG 4970
4971 AGGCAGCAGTGGGGAATTTTCCGCAATGGGCGAAAGCCTGACGGAGCAATGCCGCGTGGAGGTAGAAGGC 5040
5041 CCACGGGTCGTGAACTTCTTTTCCCGGAGAAGAAGCAATGACGGTATCTGGGGAATAAGCATCGGCTAAC 5110
5111 TCTGTGCCAGCAGCCGCGGTAATACCTCGAGGGGGCCCGGCCCA 5154
```

15 - 1158 Left targeting sequence
1189 - 1490 16S ribosomal RNA operon promoter
1491 - 1527 Multiple cloning site
1530 - 1534 Consensus chloroplast ribosome binding site
1544 - 2311 Spectinomycin resistance gene
2337 - 2441 Rubisco large subunit 5'-translation control region
2442 - 3158 Green fluorescent protein gene
3184 - 3398 Photosystem II terminator
3531 - 5135 Right targeting sequence

pVCT09b:

```
1    TGGAGCTCCCGCGGAATGATACAAGTCTTGGATAAGAATCTACAACGCACTAGAACGCCCTTGTTGACGA 70
71   TCCTTTACTCCGACAGCGTCTAGAGTTCCTCGAACAATGTGATATCTCACACCGGGTAAATCCTTAACCC 140
141  TTCCCCCTCTTACTAAGACTACAGAATGTTCTTGTGAATTATGGCCAATACCAGGTATATAAGCAGTGAT 210
211  TTCAAATCCAGAGGTTAATCGTACTCTGGCAACTTTACGTAAGGCAGAGTTTGGTTTTTTGGGGGTGATA 280
281  GTGGAAAAGTTGACAGATAAGTCACCCTTACTGCCACTCTACAGAACCGTACATGAGATTTTCACCTCAT 350
351  ACGGCTCCTCGTTCAATTCTTTTGAAGTCATTGGATCCTTTTCCTCGTTCGAGAATCTCCTCCCCTCTTC 420
421  CACTCCGTCCCGAAGAGTAACTAGGACCAATTCAGTCACGTTTTCATGTTCCAATTGAACACTTTCGATT 490
491  TTTGATTATTCTCAAATCAAAGGAGAAGATTATTCTTTTTACCAAACATATGCGGATCAAATCACGATCT 560
561  TATAATAAGAACAAGAGATCTTTCTCGATCAATCCCTTTGCCCGAGAATCAGAAAGATCCTTTTCAAGTT 630
631  TGAATTTGCTCATTTGGAATCTGGGTTCTTCTACTTCATTTTTATTTACTTATTTATTTATTATTTTGAT 700
701  TTCCCTCTCTTTTCTTTTTTTATTCCCTTCCATCATTCCTTAAGTCCCATAGGTTTGATCCTGTAGAATC 770
771  TGACCCATTTTCTCATTGAGCGAAGGGTACGAAATAATCAGATTGATTTTTCGATCAAAGTACTATGTGA 840
841  AATCTTCGGTTTTTTCCTCTTCCTCTATCCCTATCCCATAGGTACAGCGTTTGAATCAATAGAGAACCTT 910
911  TTCTTCTGTATCTGTATGAATCGATATTATTACATTCCAATTCCTTCCCGATACCTCCCAAGGAAAATCC 980
981  CGAAGTGGATCCCAAATTGACGGGTTAGTGTGAGCTTATCCATGCGGTTATGCACTCTTCGAATAGGAAT 1050
1051 CCATTTTCTGAAAGATCCTGACTTTCGTGCTTTGGTGGGTCTCCGAGACCCTTTCGATGACCTATGTTGT 1120
1121 GTTGAAGGGATATCTATATGATCCGATCGATTGCGTAACCGCGGTGGCGGCCGCTCTAGGCCAGCCTTGG 1190
```

```
1191 TGGAAGTCATCAGTTCGAGCCTGATTATCCCTAAACCCAATGTGAGTTTTTCTATTTTGACTTGCTCCCC 1260
1261 CGCCGTGATCGTGGATAAGAGGCTCGTGGGATTGACGTGAGGGGGTAGGGATGACTATATTTCTGGGAGC 1330
1331 GAACTCCAGGCGAATATGAAGCGCATGGATACAAGTTATGCCTTGGAATGAAAGACAATTCCGAATCCGC 1400
1401 TTTGTCTACGAACAAGGAAGCTATAAGTAATGCAACTATGAATCTCATGGAGAGTTCGATCCTGGCTCAG 1470
1471 GATGAACGCTGGCGGCATGCTGCAGCCCGGGCGACGTCGACTGTTTAAACGAGGCCTTAGGAGGCGAGGA 1540
1541 TGGCGATCCGTGTACCTTCGCGCCAGCTATTCATTGATGGAGAATGGAGGGAACCCATCAAGAAAAATCG 1610
1611 CATCCCTATCATCAATCCTTCTACTGAGGAGATCATTGGTGTTATTCCGGCTGCTACTGCTGAAGATGTG 1680
1681 GAGCTTGCAGTCGCTGCAGCTAGAAGAGCGCTTAAGAGGAACAAAGGAGAAGATTGGGCGTCTGCATCTG 1750
1751 GAGCTCATCGTGCTAAGTACCTTCGGGCCATTGCTGCTAAAATAACAGAGAAAAAAGATTATTTTGCAAA 1820
1821 ACTTGAAGCCATGGATTGTGGGAAACCACTGGATGAAGCAGCATGGGACATTGATGATGTTGCTGGATGT 1890
1891 TTTGAATATTATGCCGATCAAGCAGAAGCCCTTGATGCTAAACAAAAGGCTCCAATTGCCCTTCCTATGG 1960
1961 ACACTTTCAAATGCCATGTGCTTAAACAACCCATTGGTGTTGTTGGGTTGATTTCTCCTTGGAATTATCC 2030
2031 GCTTCTAATGGCAACATGGAAAGTTGCTCCAGCTCTTGCTGCTGGTTGCTCAGCTGTACTTAAGCCGTCT 2100
2101 GAACTGGCATCCGTAACTTGCCTAGAATTGGCTGAAGTGTGCAGAGAAGTGGGACTGCCTCCTGGCGTAT 2170
2171 TAAATATTTTAACAGGATTAGGTCCTGAAGCTGGTGGGCCGTTAGCTTGCCATCCTGATGTTGACAAGGT 2240
2241 TGCATTTACTGGGAGTACAGCTACTGGTAGCAAGGTTATGTCATCCGCTGCTCAATTGGTCAAGCCTGTT 2310
2311 ACATTAGAACTTGGAGGGAAAAGTCCTATTGTTATCTTTGAAGATGTTGACTTGGATAAAGCTGCTGAAT 2380
2381 GGACTGCTTTTGGCTGTTTTTGGACAAATGGTCAAATTTGCAGTGCAACATCGAGATTACTTGTGCATGA 2450
2451 AAGCATCGCAGCTGAATTTTTGGATAGGCTTGTAAAATGGTGCAAAAACATAAAGATCTCTGACCCGTTT 2520
2521 GAGGAAGGCTGTCGACTTGGTCCTGTTGTGAGTAAGAGTCAGTATGAAAAAGTTTTGAAGTTCATTTCAA 2590
2591 CAGCAAAGAGTGAGGGTGCAACTATTTTGTGTGGAGGTTCCCGTCCCGAGCATTTGAAGAAAGGGTATTA 2660
2661 TGTTGAACCAACAATTATAAGTGATGTCTCCACTTCCATGCAAATATGGAGGGAAGAAGTTTTCGGCCCA 2730
2731 GTCTTATGTGTCAAAACCTTTGGTTCTGAAGATGAAGCCATTGAACTGGCTAATGATACTCAGTATGGTT 2800
2801 TAGGGGCTGCTGTGTTATCGAAAGATCTTGATCGGTGTGAGAGAATAACAAAGGCATTGCAAGCTGGAAT 2870
2871 TGTGTGGGTTAACTGCTCACAACCATGCTTTTGCCAAGCTCCATGGGGAGGCACGAAGCGTAGCGGTTTT 2940
2941 GGACGTGAACTCGGGGAATGGGGTATCGAGAATTACTTGAATATCAAACAAGTGACTGAATATATTTCCG 3010
3011 ATGAACCATGGGGATGGTACAAGTCTCCTTGAGGATCCCCCGGGCTGCAGGAATTCGGTCGAGTAGACCT 3080
3081 TGTTTTGTTGTTGTGAAAATTCTTAATTCGTGAGTTGTAGGGAGGGACTTATGTCACCACAAACAGAGAC 3150
3151 TAAAGCAAGTGTTGGATTCAAAGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTC 3220
3221 GAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACG 3290
3291 GCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCAC 3360
3361 CTTCACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCC 3430
3431 GCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCG 3500
3501 CCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGA 3570
```

3571 CGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAG 3640

3641 CAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCG 3710

3711 ACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCAC 3780

3781 CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCC 3850

3851 GCCGGGATCACTCACGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCTCTAGAACTAGTGATTGATTTC 3920

3921 CGTCTTAGTGTTTACGAGTTTTTGAATGTAAAGGAGCAATAACCAATTTCCTCTTTTATCAAGGGGGATG 3990

3991 GTATTGCTCCTTTATTTAGTAGTCTTTTATTTATCTTAGTAGTCTTTTACTTTTCTAAGTTTTTTTTATT 4060

4061 TCTTTATTTCAACTAAAACTAAAATAGTAATAAAAAGTATTCTCATAGGTTGGTTTATGATTGAGTATCA 4130

4131 ATCGAATTCCCCTAGTGAATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGA 4200

4201 TGCATAGCTTGAGTATTCTATAGTGTCACCTAAATCGTAATCAAGCTTATCGATACCGTCGGTAGCAACG 4270

4271 GAACCGGGGAAAGTATACAGAAAAGACAGTTCTTTTCTATTATATTAGTATTAGTTAGTGATCCCGGCTC 4340

4341 AGTGAGCCCTTTCTTACGTGATGAACTGTTGGCACCAGTCCTACATTTTGTCTCTGTGGACCGAGGAGAA 4410

4411 GGGGGCTCAGCGGGAAGAGGATTGTACCATGAGAGAAGCAAGGAGGTCAACCTCTTCTCTTTCAAATATA 4480

4481 CAACATGGATTCTGACAATGCAATGTAGTTGGACTCTCATGTCGATCCGAATGAATCATCTTTTCAACGG 4550

4551 AGGTAATGCCTGCTAGGTAAGAGGATAGCAAGTTACAAATTCTGTCTCGGAAGGAATTTGTCCATTTTTC 4620

4621 GGGGTCTCAAAGGGGCGTGGAAACACATAAGAACTCTTGAATGGAAATGGAAAAGAGATGTAACTCCAGT 4690

4691 TCCTTCGGAAATGGTAAGATCTTTGGCGCAAGAAGAAGGGGTTGATCCGTATCATCTTGACTTGGTTCTG 4760

4761 ATTCCTCTATTTTTTTAAGAATACCGAGTCGGGTTCTTCTCCTACCCGTATCGAATAGAACATGCTGAGC 4830

4831 CAAATCTTCTTCACTGCTTGCTTTAGATCGGGAAAATCGTACGGTTTTATGAAACCATGTGCTATGGCTC 4900

4901 GAATCCGTAGTCAATCCTATTTCCGATAGGAGCAGTTGACAATTGAATCCCATTTTTCCCATTATTTTCG 4970

4971 TATCCGTAATAGTGCGAAAAGAAGGCCCGGCTCCAAGTTGTTCAAGAATAGTGGCGTTGAGTTTCTCGAC 5040

5041 CCTTTGCCTTACTTAGGATTAGTCAGTTCTATTTCTCGATAGGGGCAGGGAAGGGATATAACTCAGCGGT 5110

5111 AGAGTGTCACCTTGACGTGGTGGAAGTCATCAGTTCGAGCCTGATTATCCCTAAACCCAATGTGAGTTTT 5180

5181 TCTATTTTGACTTGCTCCCCCGCCGTGATCGTGGATAAGAGGCTCGTGGGATTGACGTGAGGGGGTAGGG 5250

5251 ATGACTATATTTCTGGGAGCGAACTCCAGGCGAATATGAAGCGCATGGATACAAGTTATGCCTTGGAATG 5320

5321 AAAGACAATTCCGAATCCGCTTTGTCTACGAACAAGGAAGCTATAAGTAATGCAACTATGAATCTCATGG 5390

5391 AGAGTTCGATCCTGGCTCAGGATGAACGCTGGCGGCATGCTTAACACATGCAAGTCGGACGGGAAGTGGT 5460

5461 GTTTCCAGTGGCGGACGGGTGAGTAACGCGTAAGAACCTGCCCTTGGGAGGGGAACAACAGCTGGAAACG 5530

5531 GCTGCTAATACCCCGTAGGCTGAGGAGCAAAAGGAGGAATCCGCCCGAGGAGGGGCTCGCGTCTGATTAG 5600

5601 CTAGTTGGTGAGGCAATAGCTTACCAAGGCGATGATCAGTAGCTGGTCCGAGAGGATGATCAGCCACACT 5670

5671 GGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATTTTCCGCAATGGGCGAAAGCCT 5740

5741 GACGGAGCAATGCCGCGTGGAGGTAGAAGGCCCACGGGTCGTGAACTTCTTTTCCCGGAGAAGAAGCAAT 5810

5811 GACGGTATCTGGGGAATAAGCATCGGCTAACTCTGTGCCAGCAGCCGCGGTAATACCTCGAGGGGGGCCC 5880

5881 GTCCCA 5886

15 - 1158 Left targeting sequence
1189 - 1490 16S ribosomal RNA operon promoter
1491 - 1527 Multiple cloning site
1530 - 1534 Consensus chloroplast ribosome binding site
1540 - 3042 Betaine aldehyde dehydrogenase gene
3068 - 3172 Rubisco large subunit 5'-translation control region
3173 - 3889 Green fluorescent protein gene
3915 - 4129 Photosystem II terminator
4262 - 5866 Right targeting sequence

5. REFERENCES

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool, J Mol Biol 215: 403-410, http://www.ncbi.nih.gov/BLAST/

Boch J, Kempf B, Schmid R & Bremer E (1996) Synthesis of the Osmoprotectant Glycine Betaine in *Bacillus subtilis*: Characterization of the *gbsAB* Genes, J Bacteriol 178;17: 5121-5129

Bock R (2001) Transgenic Plastids in Basic Research and Plant Biotechnology, J Mol Biol 312: 425-438

Bock R & Hagemann R (2000) Extranuclear Inheritance: Plastid Genetics: Manipulation of Plastid Genomes and Biotechnological Applications, Prog Bot 61: 76-90

Boynton JE, Gillham NW, Harris EH, Hosler JP, Johnson AM, Jones AR, Randolph-Anderson BL, Robertson D, Klein TM, Shark KB & Sanford JC (1988) Chloroplast Transformation in *Chlamydomonas* with High Velocity Microprojectiles, Science 240: 1534-1538

Brixey PJ, Guda C & Daniell H (1997) The chloroplast *psb*A promoter is more efficient in *Escherichia coli* than the T7 promoter for hyperexpression of a foreign protein, Biotechnol Lett 19;4: 395-399

Carrer H, Hockenberry TN, Svab Z & Maliga P (1993) Kanamycin resistance as a selectable marker for plastid transformation in tobacco, Mol Gen Genet 241: 49-56

Chen GJ, Qiu N, Karrer C, Caspers P & Page MGP (1999) Restriction Site-Free Insertion of PCR Products Directionally into Vectors, Biotechniques 28: 498-505

Corneille S, Lutz K, Svab Z & Maliga P (2001) Efficient elimination of selectable marker genes from the plastid genome by the CRE-*lox* site-specific recombination system, The Plant J 27;2: 171-178

Daniell H (1999)a The next generation of genetically engineered crops for herbicide and insect resistance: containment of gene pollution and resistant insects, Ag Biotech Net 1: 1-7

Daniell H (1999)b New tools for chloroplast genetic engineering, Nat Biotechnol 17: 855-856

Daniell H & Dhingra A (2002) Multigene engineering: dawn of an exciting new era in biotechnology, Curr Opin Biotechnol 13: 136-141

Daniell H, Khan MS & Allison L (2002) Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology, Trends Plant Sci 7;2: 84-91

Daniell H, Muthukumar B & Lee SB (2001)a Marker free transgenic plants: engineering the chloroplast genome without the use of antibiotic selection, Curr Genet 39: 109-116

Daniell H, Streatfield SJ & Wycoff K (2001)b Medical molecular farming: production of antibodies, biopharmaceaticals and edible vaccines in plants, Trends Plant Sci 6;5: 219-226

De Cosa B, Moar W, Lee S-B, Miller M & Daniell H (2001) Overexpression of the *Bt cry*2Aa2 operon in chloroplasts leads to formation of insecticidal crystals, Nat Biotechnol 19: 71-74

De Gray G, Rajasekaran K, Smith F, Sanford J & Daniell H (2001) Expression of an Antimicrobial Peptide via the Chloroplast Genome to Control Phytopathogenic Bacteria and Fungi, Plant Pysiol 127: 852-862

Elliott AR, Campbell JA, Dugdale B, Brettell RIS & Grof CPL (1999) Green-fluorescent protein facilitates rapid *in vivo* detection of genetically transformed plant cells, Plant Cell Rep 18: 707-714

Fisher CL & Pei GK (1997) Modification of a PCR-Based Site-Directed Mutagenesis Method, Biotechniques 23: 570-574

Geiser M, Cebe R, Drewello D & Shmitz R (2001) Integration of PCR Fragments at Any Specific Site within Cloning Vectors without the Use of Restriction Enzymes and DNA Ligase, Biotechniques 31: 88-92

Hager M & Bock R (2000) Enslaved bacteria as new hope for plant biotechnologists, Appl Microbiol Biotechnol 54: 302-310

Hall TA (1999) Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucl Acids Symp Ser 41: 95-98

Heifetz PB (2000) Genetic engineering of the chloroplast, Biochemie 82: 655-666

Heifetz PB & Tuttle AM (2001) Protein expression in plastids, Curr Opin Plant Biol 4: 157-161

Heiman M (1997) Webcutter 2.0, http://www.firstmarket.com/cutter/cut2.html

Hou B-K, Zhou Y-H, Wan L-H, Zhang Z-L, Shen G-F, Chen Z-H & Hu Z-M (2003) Chloroplast transformation in oilseed rape, Transgenic Res 12;1: 111-114

Inoue H, Nojima H & Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids, Gene 96: 23-28

Iocco P, Franks T & Thomas MR (2001) Genetic transformation of major wine grape cultivars of *Vitis vinifera* L., Transgenic Res 10: 105-112

Khan MS & Maliga P (1999) Fluorescent antibiotic resistance marker for tracking plastid transformation in higher plants, Nat Biotechnol 17: 910-915

Knoblauch M, Hibberd JM, Gray JC & van Bel AJE (1999) A galinstan expansion femtosyringe for microinjection of eukaryotic organelles and prokaryotes, Nat Biotechnol 17: 906-909

Kota M, Daniell H, Varma S, Garczynski SF, Gould F & Moar WJ (1999) Overexpression of the *Bacillus thuringiensis* (Bt) Cry2Aa2 protein in chloroplasts confers resistance to plants against susceptible and Bt-resistant insects, Proc Natl Acad Sci USA 96: 1840-1845

Kuroda H & Maliga P (2001)a Complementarity of the 16S rRNA penultimate stem with sequences downstream of the AUG destabilizes the plastid mRNAs, Nuc Acids Res 29;4: 970-975

Kuroda H & Maliga P (2001)b Sequences Downstream of the Translation Initiation Codon Are Important Determinants of Translation Efficiency in Chloroplasts, Plant Physiol 125: 430-436

Legaria J, Rajsbaum R, Munoz-Clares RA, Villegas-Sepulveda N, Simpson J & Iturriaga G (1998) Molecular characterization of two genes encoding betaine aldehyde dehydrogenase from amaranth. Expression in leaves under short-term exposure to osmotic stress or abscisic acid, Gene 218: 69-76

Maliga P (2001) Plastid engineering bears fruit, Nat Biotechnol 19: 826-827

Maliga P (2002) Engineering the plastid genome of higher plants, Curr Opin Plant Biol 5: 164-172

Maliga P (2003) Progress towards commercialisation of plastid transformation technology, Trends Biotechnol 21;1: 20-28

Mierendorf R, Yeager K & Novy R (1994) The pET System: Your Choice for Expression, 1: 1-3

Nagata N, Saito C, Sakai A, Kuroiwa H & Kuroiwa T (1999) The selective increase or decrease of organellar DNA in generative cells just after pollen mitosis one controls cytoplasmic inheritance, Planta 209: 53-65

O'Neill C, Horvath GV, Horvath E, Dix PJ & Medgyesy P (1993) Chloroplast transformation in plants: polyethylene glycol (PEG) treatment of protoplasts is an alternative to biolistic delivery systems, The Plant J 3;5: 729-738

Ruf S, Hermann M, Berger IJ, Carrer H & Bock R (2001) Stable genetic transformation of tomato plastids and expression of a foreign protein in fruit, Nat Biotechnol 19: 870-875

Ruf M & Kossel H (1998) Occurrence and spacing of ribosome recognition sites in mRNAs of chloroplasts from higher plants, FEBS Lett 240;1,2: 41-44

Sambrook J & Russel DW (2001) Molecular Cloning: A Laboratory Manual, 3rd edition, 1: 1.116-1.118

Sidorov VA, Kasten D, Pang S, Hajdukiewicz P, Staub JM & Nehra NS (1999) Stable chloroplast transformation in potato: use of green fluorescent protein as a plastid marker, The Plant J 19;2 209-216

Sikdar SR, Serino G, Chaudhuri S & Maliga P (1998) Plastid transformation in *Arabidopsis thaliana*, Plant Cell Rep 18: 20-24

Skarjinskaia M, Svab Z & Maliga P (2003) Plastid transformation in *Lesquerella fendleri*, an oilseed Brassicacea, Transgenic Res 12;1: 115-122

Staub JM, Garcia B, Graves J, Hajdukiewicz PTJ, Hunter P, Nehra N, Paradkar V, Schlittler M, Carrol JA, Spatola L, Ward D, Ye G & Russel DA (2000) High-yield production of a human therapeutic protein in tobacco chloroplasts, Nat Biotechnol 18: 333-338

Strefeler MS, Weeden NF & Reisch BI (1992) Inheritance of chloroplast DNA in two full-sib *Vitis* populations, Vitis 31: 183-187

Sugiura M (1992) The chloroplast genome, Plant Mol Biol 19: 149-168

Svab Z, Hajdukiewicz P & Maliga P (1990) Stable transformation of plastids in higher plants, Proc Natl Acad Sci USA 87: 8526-8530

Svab Z & Maliga P (1993) High–frequency plastid transformation in tobacco by selection for a chimeric *aadA* gene, Proc Natl Acad Sci 90: 913-917

Thomson JM & Parrot WA (1998) pMECA: a cloning plasmid with 44 unique restriction sites that allows selection of recombinants based on colony size, Biotechniques 24;6: 922-928

Tregoning JS, Nixon P, Kuroda H, Svab Z, Clare S, Bowe F, Fairweather N, Ytterberg J, van Wijk KJ, Dougan G & Maliga P (2003) Expression of tetanus toxin Fragment C in tobacco chloroplasts, Nucl Acids Res 31;4: 1174-1179

Tseng H (1999) DNA Cloning without Restriction Enzyme and Ligase, Biotechniques 27: 1240-1244

Van Bel AJE, Hibberd J, Prufer D & Knoblauch M (2001) Novel approach in plastid transformation, Curr Opin Biotechnol 12: 144-149

Wait D (2003) South African Wine Industry Statistics, http://www.sawis.co.za/SAWISPortal/uploads/SA%20Wine%20Industry%20Statistics%20nr%2027.pdf

Ye G-N, Hajdukiewicz PTJ, Broyles D, Rodriguez D, Xu CW, Nehra N & Staub JM (2001) Plastid-expressed 5-enolpyruvylshikimate-3-phosphate synthase genes provide high level glyphosate tolerance in tobacco, The Plant J 25;3: 261-270

Zhang X-H, Portis Jr AR & Widholm JM (2001) Plastid transformation of soybean suspension cultures, J Plant Biotechnol 3;1: 39-44

Zoubenko OV, Allison LA, Svab Z & Maliga P (1994) Efficient targeting of foreign genes into the tobacco plastid genome, Nucl Acids Res 22;19: 3819-3824