The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Investigation of *XvSap* promoters from the resurrection plant, *Xerophyta viscosa*

Tamaryn Lorean Ellick

Supervisor: Dr. MS Rafudeen

University of Cape Town

A thesis submitted in fulfilment of the requirements for the degree of Master of Science in the Department of Molecular and Cell Biology, Faculty of Science, University of Cape Town, South Africa
September 2012

Declaration

I know the meaning of plagiarism and declare that all work done in the document, save for which is properly acknowledged, is my own. This work has not been presented at any other university for examination or for any other purposes.

Tamaryn Lorean Ellick

University of Cape Town

Acknowledgements

My heavenly Father, You deserve all the praise and worship. Thank you for the blessings that you have bestowed upon me and giving me the ability to realise my potential.

To Dr. Revel Iyer, thank you for your thorough training and invaluable advice and mentoring throughout the duration of my research. I would like to express my sincere gratitude to my supervisor, Dr Suhail Rafudeen, for his support, understanding and guidance and to Prof Jennifer Thomson for her intellectual guidance and helpful comments.

A special thank you to Shakiera Sattar for her assistance with the generation of the promoters. This work could not have been done without you. To Bronwyn Arendze-Bailey and Rizqah Kamies, thank you for your support during my research and in various aspects of my life. To Dr. Kershini Iyer, I appreciate your guidance after all of my experimental failures. I am equally grateful to the members, both past and present, of the Plant Stress Lab at the University of Cape Town. You have all in some way gone through the motions with me. To all the other members of the Molecular and Cell Biology Department, your assistance in different ways has impacted on me and my research.

To my family and friends, I will cherish your undying support and encouragement. To my mommy, thank you for all of the prayers, always believing in me and never discouraging me even when I felt lost. You have inspired me more than you know. To my brother, Tarquin, I appreciate the humour and inspiration when I needed it most. To Jamiel, thank you for the motivational words and always availing your assistance.

I acknowledge the financial support from the National Research Foundation (DAAD), the University of Cape Town, The Maize Trust. Without this financial assistance, this research would not have been realised.

University of Cape Town

Abbreviations

The following abbreviations have been used in this thesis. The definitions apply throughout the text.

2.4-D	2.4-Dichlorophenoxy acetic acid
ABRE	Abscisic acid responsive element
bar	Bialophos resistance gene
BASTA	Bialophos
BMS	Black mexican sweetcorn
CamV	Cauliflower Mosaic Virus
cDNA	Complementary deoxyribonucleic acid
Ct	Cycle threshold
DEPC	Diethylpyrocarbonate
Dreb	Dehydration responsive element binding protein
EDTA	Ethylenediaminetretaacetate
EF-1α	Elongation factor-1α
gfp	Green fluorescent protein
gus	beta-glucuronidase
KOH	Potassium hydroxide
LB	Luria Bertani
luc	Luciferase
MS	Murashige and Skoog
NosT	Nopaline synthase terminator
qRT-PCR	Quantitative real time polymerase chain reaction
RWC	Relative water content
SWC	Soil water content

University of Cape Town

Abstract

The XvPSap1 promoter derived from *Xerophyta viscosa* has been demonstrated to be stress-inducible during dehydration in transgenic *Nicotiana tabacum*, black Mexican sweetcorn cells and *Zea mays*. To improve this promoter, for future applications in crop biotechnology, four shortened promoters, XvPSap1D, E, F and G were generated by mutagenesis. The generated promoters had *circa* 50% reduction in size and contained the 5' proximal and 3' distal regions of the XvPsap1 promoter with the internal region removed. The shortened promoters displayed no significant sequence homology to any other known plant promoter, besides XvPSap1. In addition to the shortened promoters, a newly discovered full length XvPSap2 promoter, showing a 56.41% homology with XvPSap1 was also assessed in this study. The individual promoters were cloned upstream of the luciferase reporter gene in the pTF101.1 binary vector for subsequent *Agrobacterium*-mediated transformations of *N. tabacum* and *A. thaliana* plants. The functional activities of three of the four shortened XvPSap1 (XvPSap1D, E and G) and the full length XvPSap2 promoters were assessed in response to dehydration treatments in transgenic *N. tabacum* plants. Luciferase protein activity was visually demonstrated and quantified. A trend of increased luciferase protein activity as driven by the three shortened promoters was observed during the dehydration treatment. No luciferase protein activity was visible with XvPSap2. Quantitative real time PCR was performed to confirm the observed luciferase protein activity by quantifying the luciferase mRNA. A trend of increased luciferase mRNA as driven by the three shortened promoters was observed over the dehydration treatment. Significant decreases in luciferase mRNA was noted when expression was driven by the XvPSap2 promoter. These results indicate that the XvPSap1D, E and G promoters are stress-inducible, whereas the XvPSap2 promoter is stress-repressible. The shortened promoters could therefore be used in the development of drought tolerant maize.

University of Cape Town

Contents

Declaration		i
Acknowledgements		ii
Abbreviations		iii
Abstract		iv
Table of contents		v
List of figures		ix
List of tables		xii
1.	**Literature Review**	1
1.1.	Introduction	1
1.2.	Strategies to enhance drought tolerance	2
1.2.1.	Conventional breeding	2
1.2.2.	Genetic engineering	3
1.3.	Desiccation tolerant organisms	5
1.4.	Promoters and their features	6
1.5.	Types of promoters available for genetic engineering	7
1.5.1.	Constitutive promoters	7
1.5.2.	Inducible promoters	9
1.6.	The XvPSap1 promoter	11
1.7.	Use of shortened promoters for stress tolerance	12
2.	**Research focus**	14
2.1.	Research question	14
2.2.	Research hypothesis	14
2.3.	Research aims	14
2.4.	Significance of the study	14
2.5.	Research assumptions and limitations	15
3.	**Materials and Methods**	16
3.1.	Mutagenesis of XvPSap1	16
3.2.	Generation of pBluescript::promoter::*luc*::NosT	19
3.3.	Generation of binary vector constructs	21
3.4.	*In silico* analysis of promoter sequences	23
3.5.	Transformation and screening of *Agrobacterium tumefaciens*	23
3.6.	Plant transformation	24
3.6.1.	*Agrobacterium*-mediated transformation of *Arabidopsis thaliana*	24
3.6.2.	*Agrobacterium*-mediated transformation of *N. tabacum*	26
3.6.3.	Screening of putative transgenic plants	29
3.6.4.	Dehydration of transgenic plants	30
3.6.5.	Analysis of promoter activity	32
3.6.5.1.	Live imaging of luciferase expression	32
3.6.5.2.	Luciferase expression by quantitative real-time PCR	32

University of Cape Town

4. Results 37
4.1. Mutagenesis of XvPSap1 37
4.2. Generation of pBluescript::promoter::*luc*::NosT 38
4.3. Generation of binary vector constructs 40
4.4. Bioinformatic analysis of promoters 42
4.5. Transformation of *Agrobacterium* 43
4.6. *Agrobacterium*-mediated transformation of *N. tabacum* 43
4.6.1. Transformation of *N. tabacum*, tissue culture and growth of putative positive transformants 43
4.6.2. *Agrobacterium*-mediated transformation of *A. thaliana* 45
4.7. Germination and BASTA screening of putative transgenic seed 45
4.7.1. Transgenic *N. tabacum* 45
4.7.2. Transgenic *A. thaliana* 46
4.8. Screening of putative transformants using PCR 46
4.8.1. Transgenic *N. tabacum* 46
4.8.2. Transgenic *A. thaliana* 49
4.9. Dehydration treatment of transgenic *N. tabacum* plants 51
4.9.1. Morphological changes during dehydration stress 51
4.9.2. Determination of soil water content 52
4.9.3. Determination of relative water content for each plant 53
4.10. Analysis of promoter activity in intact leaves 55
4.11. Analysis of promoter activity by quantitative real-time PCR 60
4.11.1. RNA isolation 60
4.11.2. Analysis of RNA using the Bioanalyzer software 60
4.11.3. cDNA synthesis 62
4.11.4. Primer optimisation and reference gene analysis 62
4.11.5. Quantitative real time PCR 63

5. Discussion 66
5.1. Deletion mutagenesis of XvPSap1 66
5.2. Relative water content and protein activity verification 67
5.3. Stress inducible activity 68
5.4. Stress repressible activity 69
5.5. Transcript fluctuations 69
5.6. Future work 70

Appendices 72
Appendix A: Schematic diagrams
A.1. Mutagenesis strategy to generate pBluescript::promoter construct 72
A.2. Cloning strategy to generate pBluescript::promoter::*luc*::NosT construct 73
A.3. Cloning strategy to generate binary vector constructs 74

Appendix B: Vector maps
B.1. pBluescript 75
B.2. pTF101.1 76

University of Cape Town

Appendix C: Nucleotide sequences and multiple sequence alignments
C.1. XvPSap1 77
C.2. XvPSap1D 78
C.3. XvPSap1E 78
C.4. XvPSap1F 79
C.5. XvPSap1G 79
C.6. XvPSap2 80
C.7. *luc* gene 81
C.8. nos terminator 81
C.9. Nucleotide sequence of XvPSap1E, F and G 82
C.10. Nucleotide sequence alignment of XvPSap1D, E, F, G and XvPSap2 84
C.11. Nucleotide sequence alignment of XvPSap1 and XvPSap2 89

Appendix D: Primers and accession numbers
D.1. Accession numbers of reference genes in this study 93
D.2. Primers for mutagenesis used in this study 93
D.3. Primers for screening used in this study 94
D.4. Primers for quantitative real time PCR used in this study 94

Appendix E: General protocols
E.1. PCR reaction 95
E.2. Klenow polymerase reaction 95
E.3. Blunt end ligation 96
E.4. Site specific cohesive end ligation reaction 96
E.5. PCR product purification 96
E.6. DNA purification from agarose gels 96
E.7. Plasmid DNA isolation 97
E.8. Preparation of glycerol stocks 98

Appendix F: Media and solutions
F.1. LB media 99
F.2. LB agar 99
F.3. YEP media 99
F.4. YEP agar 99
F.5. B5 vitamins 100

Appendix G: XvPSap regulatory elements
G.1. Putative regulatory elements in XvPSap2 101
G.2. Putative regulatory elements in XvPSap1G 105

Appendix H: Quantitative Real time analysis
H.1. RNA analysis 109
H.2. Melt curves, standard curves, real time runs 114

Appendix I: Wild type *N. tabacum*
I.1. Bioluminescence 118
I.2. Wild type *N. tabacum* 118

University of Cape Town

Appendix J: Supporting images

J.1. Restriction digestion of pBluescript::XvPsap1::*luc*::NosT 119
J.2. *N. tabacum* transformed with XvPSap1E displaying dwarfism 119

References 120

University of Cape Town

List of figures

Figure 3.1: Schematic illustration of the full length and shortened XvPSap1 promoters. 17

Figure 4.1: Amplification of pBluescript DNA incorporating the promoter fragments. 37

Figure 4.2: Restriction endonuclease analysis of pBluescript::promoter:: *luc*::NosT. 39

Figure 4.3: Amplification of promoter DNA cloned into pTF101.1 and transformed into competent Dh5α *E. coli* cells. 41

Figure 4.4: *Eco*RI and *Hin*dIII restriction endonuclease digests of pTF101.1::promoter::*luc*::NosT plasmid DNA. 42

Figure 4.5: Transformation and tissue culture of *N. tabacum*. **A**: Infected leaf disks on co-cultivation media. 44

Figure 4.6: Germination of transformed *N. tabacum* seeds. 45

Figure 4.7: Germination of transformed *A. thaliana* seeds. 46

Figure 4.8: Genomic DNA of putative positive *N. tabacum* plants. 47

Figure 4.9: PCR screening of putative positive plants for the presence of *luc* and XvPSap1. 48

Figure 4.10: PCR screening of putative positive plants for the presence of *luc* and XvPSap1. 50

Figure 4.11: PCR screening of putative positive plants for the presence of *luc* and XvPSap2. 51

Figure 4.12: Transgenic and wild type *N. tabacum* plants during a six day dehydration treatment. 52

Figure 4.13: Soil water content . 53

Figure 4.14: Determination of RWC for one leaf. 53

Figure 4.15: Relative water content of *N. tabacum* plants.. 54

Figure 4.16: Determination of luciferase activity in hydrated *N. tabacum* plants transformed with XvSap1D over a six day dehydration treatment. 55

University of Cape Town

Figure 4.17: Determination of luciferase activity in *N. tabacum* plants transformed with XvSap1D over a six day dehydration treatment. 56

Figure 4.18: Determination of luciferase activity in *N. tabacum* plants transformed with XvSap1E over a six day dehydration treatment. 56

Figure 4.19: Determination of luciferase activity in *N. tabacum* plants transformed with XvSap1G over a six day dehydration treatment. 57

Figure 4.20: Determination of luciferase activity in *N. tabacum* plants transformed with XvSap2 over a six day dehydration treatment. 57

Figure 4.21: Luciferase activity in transgenic and wild type *N. tabacum* plants over a six day dehydration treatment. 59

Figure 4.22: Purified RNA from transgenic *N. tabacum.* 60

Figure 4.23: Electropherogram of RNA isolated from XvPSap1E generated using the Agilent 2100 Expert software. 61

Figure 4.24: Relative stability ranking of reference genes 18S rRNA L25 ribosomal protein and *EF-1α.* 63

Figure 4.25: Relative gene expression levels of *luc* during the dehydration treatment of transgenic *N. tabacum.* 65

Figure A.1: Schematic illustration of the mutagenesis strategy to generate pBluescript::XvPSap1D. 72

Figure A.2: Generation of pBluescript::XvPSap1D::*luc*::NosT. 73

Figure A.3: Generation of binary vector construct, pTF101.1::XvPSap1D::*luc*:: NosT. 74

Figure B.1: pBluescript vector. 75

Figure B.2: pTF101.1 vector. 76

Figure C.1: Nucleotide sequence for the full length XvPSap1 promoter. 77

Figure C.2: Nucleotide sequence of the XvPSap1D promoter. 78

Figure C.3: Nucleotide sequence of the XvPSap1E promoter. 78

Figure C.4: Nucleotide sequence of the XvPSap1F promoter. 79

Figure C.5: Nucleotide sequence of the XvPSap1G promoter. 79

Figure C.6: Nucleotide sequence of the XvPSap2 promoter. 80

University of Cape Town

Figure C.7: Nucleotide sequence of the *luc* gene. 81

Figure C.8: Nucleotide sequence of NosT. 81

Figure C.9: Multiple sequence alignment of the XvPSap1D, E, F and G nucleotides generated using DNAMAN software. 82

Figure C.10: Multiple sequence alignment of the XvPSap1D, E, F, G and XvPSap2 nucleotides generated using DNAMAN software. 84

Figure C.11: Multiple sequence alignment of the XvPSap1D and XvPSap2 nucleotides generated using DNAMAN software. 89

Figure H.1A: Analysis of RNA isolated from XvPSap1D plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. 109

Figure H.1B: Analysis of RNA isolated from XvPSap1E plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. 110

Figure H.1C: Analysis of RNA isolated from XvPSap1G plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. 111

Figure H.1D: Analysis of RNA isolated from XvPSap2 plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. 112

Figure H.1E: Analysis of RNA isolated from hydrated transgenic plants over a six day dehydration period. 113

Figure H.2A: Analysis of the 18S rRNA reference gene. 114

Figure H.2B: Analysis of the elongation factor 1-α reference gene. 115

Figure H.2C: Analysis of the L25 ribosomal reference gene. 116

Figure H.2D: Analysis of the *luc* gene. 117

Figure I.1: Determination of luciferase activity in wild type *N. tabacum* plants transformed with XvSap1D over a six day dehydration treatment. 118

Figure 1.2: Untransformed wild type *N. tabacum* on rooting media supplemented with 3 mg/ml BASTA. 118

Figure J.1: Restriction endonuclease digestion of pBluescript::XvPsap1::*luc*::NosT 119

Figure J.2: *N. tabacum* transformed with XvPSap1E 119

University of Cape Town

List of Tables

Table 4.1: Efficiencies, R-values and R^2-values of the potential reference genes and *luc* gene 62

Table D.1: Accession numbers of reference genes used in this study 93

Table D.2: Primers for mutagenesis used in this study 93

Table D.3: Primer sets for screening used in this study 94

Table D.4: Primer sets for quantitative real time PCR used in this study 94

Table E.1: PCR components and final concentrations used in a standard PCR protocol 95

Table G.1: Putative regulatory elements in the XvPSap2 sequence 101

Table G.2: Putative regulatory elements in the XvPSap1G sequence 105

University of Cape Town

1. Literature Review

1.1. Introduction

After wheat and rice, maize is currently the third largest planted cereal crop worldwide (FAO, 2009). Maize is directly used for human consumption, as livestock feed or to make feed ingredients and in the production of bioethanol. In the US and EU, most maize is used as animal feed and for export to other countries (FAO, 2009). In contrast, in developing countries, particularly in sub-Saharan Africa, maize is mainly produced for human consumption where it is the staple cereal crop. Rosegrant *et al.* (2001) predict that by 2050, the demand for maize in sub-Saharan Africa will double.

Biotic and abiotic stresses are a significant constraint to crop production. Although not the focus of this research, research has shown that biotic factors such as plant pathogens, herbivorous animals and insects significantly reduced crop production (Oerke, 2006; Shiferaw *et al.*, 2011). Additionally, the influence of human populations on agricultural land has led to environmental degradation and soil nutrient depletion resulting in greater requirements for agricultural productivity per unit area (Machuka, 2001; Baldi and Jobbagy, 2012). While biotic factors and human impacts have caused considerable damage to crop productivity, Acquaah (2007) estimates that 70% of worldwide crop productivity reduction is due to abiotic stresses. Of these, drought is the most significant environmental stress (Cattivelli *et al*, 2008).

The regularity and severity of many abiotic stresses, particularly drought, has been intensified by climate change. Rainfall has become less predictable with the maize growing areas of eastern Africa experiencing a shortened rainy season and extended periods of dry spells (Segele and Lamb, 2005). This is problematic since drought has the potential to cause significant yield reductions in the major cereal crops (Lobell and Field, 2007) and is particularly harmful to maize during establishment, flowering and pollination (Otegui *et al.*, 1995; Cakir, 2004; Barnabas *et al.*, 2008).
The majority of sub-Saharan Africans rely on small-scale farming systems as their principal source of income. Nevertheless, they do not produce enough food to ensure

University of Cape Town

their families survival. In an area where people do not have the resources to purchase imported crops and food aid is often unreliable, small-holder farmers are becoming increasingly important to fulfil these demands (Livingston *et al.*, 2011; Africa's Smallholder Farmer's Group, n.d.). As these farmers rely on seasonal rain, climate is an important risk factor in their farming systems and changes in climate can increase those risks.

Drought is a major constraint to small-holder farmers and it is predicted that the occurrences of drought are likely to increase (Shanahan *et al.*, 2009). Furthermore, no naturally occurring cereal crop capable of withstanding water deficit has been identified (Zhang *et al.*, 2010). Thus strategies for increasing or maintaining agricultural productivity during dry spells must be developed so as to generate income and economic growth in risk-prone areas of Africa. The improvement of drought tolerance in staple crops is one such strategy.

1.2. Strategies to enhance drought tolerance

Different approaches have been utilised to improve drought tolerance in higher plants. The main strategies are conventional breeding, marker assisted breeding and genetic engineering. These strategies have contributed significantly to understanding and improving stress tolerance in these plants, particularly in crop plant species.

1.2.1. Conventional breeding

Conventional breeding involves a series of recurrent selection procedures aimed at maximising percentages of favourable alleles at each locus of importance in a given environment (Kling and Edmeades, 1997; Lamkey, 2002). Conventional breeding can be reduced to two essential steps. The first involves the production of a plant population that is highly favourable for specific traits that are agriculturally significant. This is achieved by identifying parent plants of the same species that show better adaptation to a particular stress, cross-pollinating the parent with another parent plant that is unable to survive the stress, but may have a different favourable trait and allowing sexual recombination to occur. A feature of sexual recombination is that the traits that differ between the parents can reassociate in the offspring that could

University of Cape Town

potentially carry improved genetic combinations. The second step involves screening the offspring for individual plants that combine the most desirable traits of both parent plants, followed by backcrossing the selected offspring with the parent that displayed the better adaptation to the stress to ensure that the offspring is most like that parent (Manshardt, 2004; McCouch, 2004).

Although conventional breeding has had successes in improvement of cereal crop tolerance to herbicides, pathogens, other biotic stresses and water-limited conditions (James, 2011), it has shortcomings. While conventional breeding is purely the normal mating process, it is manipulated through human choice of the parent plants. This approach is time consuming and labour intensive. It is difficult to alter single characteristics, and undesirable characteristics are often transferred in combination with the desirable ones. Furthermore, conventional breeding is limited to closely related species (Holmberg and Bulow, 1998; Zhang *et al*., 2000; Chinnusamy *et al*., 2005).

University of Cape Town

1.2.2. Genetic engineering

Genetic engineering is described as the process whereby the characteristics of an organism are deliberately altered by the manipulation of its genetic material via transformation of specific genes or DNA fragments to create new variations in the organism (Uzogara, 2000; Duvick, 2001). By transferring the DNA from one organism to another, new beneficial traits may be introduced to any organism.

Genetic engineering can be used as a relatively rapid and precise means of achieving biotic stress tolerance in various plants. In contrast to conventional breeding, improvements in genetic engineering have permitted the exchange of DNA between sexually incompatible plant species thus providing a wider gene pool for manipulation (Prakash, 2001). Additionally, genetic engineering allows for the transfer of the specific single or multiple genes for the desired trait to be acquired. Thus, only the minimum amount of DNA necessary is transferred and the probability of incorporating undesired genes, normally associated with conventional breeding, is reduced (Naqvi *et al*., 2010).

The first successful productions of transgenic crops were mainly engineered against biotic stresses. Transgenic maize plants were engineered to be resistant to insect pests and herbicides (Pilcher *et al.*, 1997a; Pilcher *et al.,* 1997b; Owen and Zelaya, 2005). Currently, companies such as Monsanto produce varieties of maize, soybean, canola and cotton to be insect resistant and herbicide resistant (James, 2011). However no genetically engineered drought tolerant crop species has been commercialised thus far (Xoconostle-Cazares *et al.*, 2010).

BASF have claimed that together with Monsanto, the first transgenic drought tolerant maize seeds will be made available to farmers by 2012 for planting in 2013 (James, 2011) Trials of the first drought tolerant maize under the Water Efficient Maize for Africa (WEMA) project were planted in 2010 (James, 2011). This drought tolerant maize was produced from a combination of conventional breeding and genetic engineering. The field trials are currently being conducting in Kenya, Uganda and South Africa, and it is hoped that seed will be made available to the public by 2016 (James, 2011). Transgenic rice modified for tolerance to drought is in the process of development (James, 2011). Although field trials for drought tolerant maize are in progress, development of drought tolerant crops, particularly maize, should still continue given the complex and polygenic nature of dehydration stress.

Genetic engineering does, however, have limitations. One of the greatest limitations is public acceptance of genetically engineered organisms derived from gene cloning and transformations (Singh *et al.,* 2006). Concerns raised by the public are mainly related to health, environmental, social, ethical and religious issues. To overcome this, constant developments in regulatory and motivation programs are necessary (Rashid *et al*., 2012). Another limitation is that the end product of genetic engineering could be unexpected as the expression patterns of the transferred gene can vary among populations of plants engineered with the same gene or DNA fragment (Takeda and Matsuoka, 2008). This is because each transformation event involves the random integration of the transferred gene into different sites in the plant genome (Peach and Velton, 1991; Li *et al*., 2008; Kohli *et al*., 2010). With its limitations, genetic engineering still allows for novel gene formulations to be designed and constructed. For example, this could include the use of promoter regions that target gene expression at a desired time or location in plants (Mittler and Blumwald, 2010).

University of Cape Town

The low research outputs experienced in developing drought tolerant crops could possibly be due to the limited pool of available genes that have been characterised to confer drought tolerance. However, there is possibly a large amount of untapped genes that could be involved in drought tolerance. This could potentially come from genes with unknown functions, which account for 20% to 40% of the genes in sequenced genomes (Gollery *et al.*, 2006; Horan *et al.*, 2008; Mittler and Blumwald, 2010). The challenge is to identify these genes and their promoters and to identify the best method for their delivery into the target plant to ensure that they are only functionally active under stress conditions.

1.3. Desiccation tolerant organisms

University of Cape Town

Desiccation tolerance is the ability of an organism to survive the loss of most of its cellular water for extended periods and to recover to full metabolic function upon rehydration (Gaff, 1971). It differs from drought tolerance in that drought refers to any situation in which the amount of available water is less than the amount of water required for sustaining growth and productivity or is absent completely (Deikman *et al.*, 2011). Desiccation tolerance is thus a mechanism that organisms may use to achieve drought tolerance. Desiccation tolerant organisms can be used to gain understanding of the genes and promoters shown to be uniquely upregulated in response to desiccation and thus be used to determine requirements for the bioengineering of drought tolerant plants.

Desiccation tolerance occurs in certain species of prokaryotes, such as cyanobacteria (Potts, 1994) and eukaryotes, such as nematodes, rotifers (Ricci *et al.*, 1996; Ricci, 1998) and plants (Kranner and Birtic, 2005). In plants, desiccation tolerance is common in seeds, spores and pollen but less common in vegetative tissues (Dickie and Prichard, 2002; Farrant and Moore 2011). Vegetative desiccation tolerance mainly occurs in less complex plants such as bryophytes but is rare in higher plants such as pteridophytes and angiosperms and does not occur in gymnosperms (Gaff, 1977; Porembski and Barthlott, 2000).

Resurrection plants are a group of angiosperms that possess the ability to withstand desiccation in vegetative tissues and to revive from an air-dry state upon rehydration

(Gaff, 1987). These plants can lose up to 95% of their cellular water content for extended periods but are still able to regain full metabolic function upon watering. All can recover fully from 5% cellular water, but the various species recover within different times (Gaff, 1987; Sherwin and Farrant, 1996; Mundree *et al.*, 2002). These plants have a broad geographical distribution including South America, Australia, India and South Africa (Alpert, 2005; 2006). It is estimated that there are 330 angiosperm species of resurrection plants including monocotyledonous species such as *X. viscosa* and dicotyledonous species such as *Myrothamnus flabellifolia* (Gaff, 1987; Proctor and Pence, 2002). The mechanisms of stress tolerance in these plants have been reviewed (Farrant *et al*., 2012; Farrant and Moore., 2011; Gechev *et al*., 2012).

As a result of their ability to naturally withstand prolonged periods of water deficit and based on their ecological habitats, resurrection plants have evolved structures and mechanisms to allow survival under extreme conditions. Consequently, these plants should be a source of unique genes and promoters that could be used to improve stress tolerance, particularly drought tolerance, in susceptible crop plants such as maize (Ramajulu and Bartels, 2002). A number of genes have been identified and characterised to be involved in abiotic stress tolerance in *X. viscosa* (Mundree *et al*., 2000; Mowla *et al*., 2002; Garwe *et al*., 2003; Govender *et al*., 2006). Some of these have been shown to confer drought tolerance in the model plant systems of tobacco and *Arabidopsis* (Garwe *et al*., 2006; Maredza, 2008). Furthermore, these plants should be a source of novel stress-inducible promoters (Okoth, 2009).

University of Cape Town

1.4. Promoters and their features

Eukaryotic promoters are non-coding regions of DNA that facilitate transcription of their target genes. They are generally located upstream of their target DNA and contain two main elements, the core or distal promoter and the regulatory elements or proximal promoter (Solovyev *et al*., 2010).

The core promoter commonly contains the transcription initiation site, TATA box, Initiator motif and Downstream Promoter Element (Ranish *et al*., 1999; Carey and

Smale, 2001; Smale and Kadonga, 2003). These motifs have specific functions in directing the placement of the RNA polymerase II transcription machinery during the transcription process. The events leading up to transcription eventually leads to the basal transcription machinery at the core promoter. Not only are they responsible for initiation of transcription but the core element motifs may also provide a level of transcription regulation (Butler and Kadonaga, 2002). It was initially reported that the TATA box was one of the general features of the core promoter, but increasing evidence suggests that there is no universal core promoter feature (Muller *et al*., 2007). In fact, less than 20% of characterised promoters in both yeast and human contain a TATA-box (Gross and Oelgeschlager, 2006) and less than 39% in *A. thaliana* (Bernard *et al*., 2010).

The proximal promoter contains *cis*-acting elements that are involved in the regulation of transcription directed by RNA polymerase II. They contain various recognition sites for the binding of sequence-specific DNA-binding proteins or transcription factors that may either activate or repress transcription from the core promoter (Butler and Kadonaga, 2002; Zou *et al*., 2011).

1.5. Types of promoters available for genetic engineering

Promoters can be selected that allow for either constitutive gene expression or limited expression to specific cell or tissue types or induced expression in response to chemicals or abiotic stresses. Some of these promoters are also able to function across different species (Peremarti *et al*., 2010).

1.5.1. Constitutive promoters

Constitutive promoters direct constant gene expression in almost all tissue types and are not dependant on any stimuli for functional activity. Since their activity is not dependant on any endogenous factors, constitutive promoters are active across most plant species (Park *et al*., 2010)

The most used constitutive promoters for gene expression in transgenic plants are the Cauliflower mosaic virus promoter for 35S RNA (CamV 35S) and the ubiquitin

University of Cape Town

promoters (Centre for Environmental Risk Assessment). As examples, both the CamV 35S and soybean ubiquitin promoters were able to drive high levels of expression of the green fluorescent protein (*gfp*) reporter gene with high expression levels of the green fluorescent protein observed in the leaf veins, petioles, petals, pollen, pods and seeds in mature plants and in the roots, cotyledons and plumules of younger plants (Hernandez-Garcia *et al*., 2009). In transgenic tobacco, *gfp* was highly expressed in all cell types under the control of the CamV 35S promoter (Kumar *et al*., 2011). Similar findings have been reported with the maize ubiquitin promoter. High expression levels of *gfp* were observed in transgenic plants (Takaski *et al*., 2010) and high expression levels of miraculin gene was observed in transgenic lettuce (Hirai *et al.,* 2011). Furthermore, the CamV 35S and maize ubiquitin promoters were able to express their transgenes in both monocotyledonous (Oltmanns *et al*., 2010; Ganguly *et al*., 2011) and dicotyledonous plants (Khan *et al*., 2001; Kasuga *et al*., 2004; Mann *et al*., 2011; Bahariah *et al*., 2012).

Because of the high expression levels of transgenes in most tissue types and plant species, constitutive promoters have become widely used in the generation of transgenic crop plants particularly for the expression of selectable genes and transgenes that should be constantly expressed (Christensen and Quail, 1996; Miki and McHugh; 2004). Use of the *bar* gene, which provides herbicide resistance, driven by a constitutive promoter, has been used as a selectable gene in transgenic crops including sweet potato (Yi *et al*., 2007), sugarcane (Ijaz *et al*., 2012), rice (Nakamura *et al*., 2010) and maize (Fromm *et al*., 1990; Oltmanns *et al*., 2010). Currently, most commercially available genetically engineered crop plants are modified for resistance to insecticides, herbicides and other biotic stresses. Since these characteristics should be constantly expressed, the use of a constitutive promoter to obtain these characteristics can be employed.

Constant overexpression of genes can be problematic, resulting in unwanted phenotypic characteristics in transgenic plants. These negative traits can be attributed to the constitutive expression of the transgene. Expression of genes in high amounts at stages when it is not needed is metabolically taxing to the plant (Su *et al*., 1998; Liu *et al*., 1998). For example, the CamV 35S promoter regulating constant expression of Dehydration responsive binding protein (*Dreb)1A* in transgenic

University of Cape Town

Arabidopsis (Liu *et al*., 1998), tobacco (Kasuga *et al*., 2004) and rice (Ganguly *et al*., 2011) resulted in increased tolerance to dehydration but dwarfism was observed in all of the plant species. Similarly, constitutive overexpression of the *TaDreb2* and *TaDreb3* in transgenic barley and wheat plants resulted in plants with a slower growth rate and delayed germination and flowering relative to control plants (Morran *et al*., 2011). This undesirable characteristic was not unique to use with the CamV 35S promoter. Transgenic tobacco plants, containing the genes required for trehalose synthesis driven by the constitutive AtRbcS1A promoter (Karim *et al*., 2007), bentgrass constitutively expressing the *Atbg1* gene that produces active abscisic acid, (Han *et al*., 2012) and rice containing the *OsNAC6* gene (stress responsive gene) fused to the maize ubiquitin promoter (Nakashima *et al*., 2007) all displayed growth retardation under normal conditions when compared to wild type plants. It has been reported that constitutive expression of transgenes by the CamV 35S promoter did not result in severe growth retardation (Yokotani *et al*., 2009; Quan *et al*., 2010). However, this is not always the case.

1.5.2. Inducible promoters

Developing drought tolerant transgenic plants in which the transgene is only expressed under certain conditions can reduce the genetic load on a plant. Inducible promoters regulate the expression of their transgenes in response to a stimulus. Abiotic stress-inducible promoters respond to environmental stimuli such as desiccation, light and fluctuating temperatures and therefore the expression of transgenes are only initiated when needed by the host plant (Kasuga *et al*., 2004). This should alleviate undesirable effects found with constitutive promoters.

The most characterised stress-inducible promoter for the development of drought-tolerant plants is the *rd29A* promoter. It was identified after analyses of dehydration stress in *Arabidopsis*, which revealed the presence of several stress-related genes, one being *rd29A.* A *cis*-acting element was identified in the promoter region of the *rd29A* gene and was shown to be responsible for both dehydration- and cold-induced expression (Yamaguchi-Shinozaki and Shinozaki, 1993). The use of this promoter for expression of *Dreb1a* in dicotyledonous transgenic tobacco and *Arabidopsis* (Kasuga *et al*., 1999; 2004) generated increased tolerance of the plants to dehydration while

University of Cape Town

expression of the (isopentenyltransferase) *ipt* gene in tobacco (Qui *et al*., 2011) and *dreb1a* in alfalfa (Jin *et al*., 2010) demonstrated increased tolerance to salt stress The negative effects on plant growth associated with constitutive expression were minimised. Similar results have been reported in monocotyledonous plants. Maize plants transformed with *TsCBF1* (Zhang *et al*., 2010) or *CBF3*, an abiotic stress-related gene (Al-Abed *et al*., 2007), wheat transformed with *Dreb1a* (Gao *et al*., 2009), and rice transformed with *Osdreb2a* gene (Mallikarjuna *et al*., 2011) or *dreb1* (Datta *et al*., 2012), under the regulatory control of the *rd29A* promoter displayed no undesirable phenotypic traits under normal conditions.

Other stress-inducible promoters have also led to desirable phenotypic characteristics in transgenic plants. For example, transgenic tobacco transformed with the *Atrab18* promoter fused to a trehalose synthesis gene (Karim *et al*., 2007) and the *Swpa2* promoter fused to *gus* (Kim *et al*., 2003; Bhatnagar-Mathur *et al*., 2008) displayed normal growth phenotypes when unstressed. Similarly, rice transformed with *gfp* under the control of the *Atrab21* promoter (Yi *et al*., 2010) or *Wsi18* promoter (Yi *et al*., 2011), *Arabidopsis* with *gfp* fused to the *Atmpk3* promoter (Gao *et al*., 2010) and tomatoes containing the (Abscisic acid response complex) *ABRC1* promoter (Lee *et al*., 2003) all showed normal growth under unstressed conditions and increased *gfp* transcripts during dehydration.

Although stress-inducible promoters relieve metabolic burden on plants, they do have disadvantages. In comparison to the high level transgene expression by constitutive promoters, lower expression levels are observed in transgenic plants with stress-inducible promoters (Reynolds, 1999). This was observed when the constitutive *AtRbcS1A* promoter was able to drive higher expression levels of its transgene than the stress-inducible *rab18* promoter (Karim *et al*., 2007). Similarly, in a study performed by Xiao *et al*. (2009), six stress-related genes were fused to either the constitutive *act1* promoter or stress-inducible *hva22p* promoter. In all instances, expression was higher when driven by the *act1* promoter.

Another disadvantage of inducible promoters is that they may act differently in monocotyledonous and dicotyledonous plants. This could be due in part to the fact that the molecular mechanisms involving gene expression differ between these

species or possibly the lack of suitable *cis*- and *trans*-acting regulatory proteins between species (Shaffner and Sheen, 1991; Furtado *et al*., 2008). Furthermore, some monocotyledonous promoters may cause higher transgene expression in their native species than in other monocotyledonous species. For example, the wheat *wcs120* promoter displayed higher transgene expression in transgenic wheat and barley than in rice, rye and tomatoes (Ouellet *et al*., 1998). Furtado *et al*. (2008) demonstrated that rice promoters expressed their transgenes higher than wheat and barley promoters in transgenic rice. Furtado *et al*. (2009) demonstrated that the barley *Isa* promoter did not show any activity in transgenic wheat, but displayed high gene expression in barley.

Even though different stress-inducible promoters have been identified, there is a still a shortage of efficient promoters for gene expression that display favourable characteristics in their native plants as well as in xenogenic species (Yi *et al*., 2011). Thus in finding promoters for drought tolerant maize, it would be useful and practically important to select the most useful promoter for gene expression in maize and also to determine the usefulness of these promoters in other plant species.

University of Cape Town

1.6. The XvPSap1 promoter

The *XvSap1* gene was isolated and characterised from *X. viscosa* and subsequently postulated to be one of the genes to play a role in desiccation tolerance of the plant (Garwe *et al*., 2003). Its cDNA was isolated from a cDNA library constructed from the leaves of dehydrated *X. viscosa* plants (Garwe *et al*., 2003) Investigation of this sequence specified a highly hydrophobic protein with six transmembrane regions. The deduced amino sequence revealed a 49% identity to a low temperature regulated protein from wheat and between 25% and 56% similarity to cold associated proteins in *Arabidopsis*. Analysis of gene expression indicated that the *XvSap1* gene is induced by dehydration, salt stress, high light and both high and low temperatures (Garwe *et al*., 2003). Iyer *et al.* (2007) showed that the *XvSap1* gene could function as a G protein-coupled receptor in signal transduction cascades in osmotic stress. It was thus postulated to encode an integral membrane protein, which is expressed in response to various abiotic stresses, particularly dehydration. Significantly,

transgenic *Arabidopsis* plants harbouring the *XvSap1* gene showed improved tolerance to drought (Garwe *et al.*, 2006).

The XvPSap1 promoter was isolated upstream of the *Xvsap1* gene in the *X. viscosa* genome and analysis of the sequence revealed that it displayed no significant identity to other known plant promoters. Okoth (2009) showed that it regulated the overexpression of the luciferase and green fluorescent protein reporter genes in transgenic tobacco, BMS cells and maize. These results suggest that XvPSap1 functions as a stress-inducible promoter. It is postulated that it is involved in early responses to drought as its activity peaks shortly after transgenic plants are subjected to lack of water. This promoter could therefore be used in the generation of transgenic drought tolerant plants.

More recently a second XvSap homologue was identified (Iyer, unpublished data). It was postulated that a minimum of two copies of the *Xvsap1* gene is present in *X. viscosa* (Garwe *et al.*, 2003). A putative promoter, referred to as XvPSap2, of *circa* 1800 bp was isolated. Preliminary bioinformatic analysis suggests that the XvPSap2 fragment may be functional as a promoter.

1.7. Use of shortened promoters for stress tolerance

The 2083 bp XvPSap1 promoter is larger than some of the characterized stress-inducible promoters. The *Wsi18* promoter is 1800 bp (Yi *et al.*, 2011), the *rd29A* promoter is 824 bp (Zhang *et al.,* 2005), the *swpa2* promoter is 1824 bp (Kim *et al.*, 2003), the *Atmpk3* promoter is 1016 bp (Gao *et al.*, 2010) and the *Alsap* promoter is 568 bp (Saad *et al.*, 2011). The large size of the XvPSap1 promoter may be problematic in its use in the generation of stress-tolerant crops as the bigger the transgene DNA fragment, the lower the efficiency of the transformation and stability of the desired characteristics.

In an attempt to improve on the XvPSap1 promoter and identify the shortest possible promoter fragment, which is still functionally active, two 5' deletions were generated (Okoth, 2009). The shortened promoter fragments were able to regulate expression of the reporter genes in tobacco and BMS cells under dehydration but at lower levels

University of Cape Town

than the full length promoter. This is not unusual. When 5' deletions of the *rd29A* promoter were performed, expression of *gus* decreased during dehydration when compared to the full length. It was determined that the 162 bp fragment between positions 111 and 268 contained a dehydration-responsive element (Yamaguchi-Shinozaki and Shinozaki, 1994). When a series of 5' deletions were performed on the pollen-specific promoter, *gus* expression decreased as the promoter size decreased (Swapna *et al*., 2011). However, shortening a promoter sometimes does yield positive results. For example, a shortened *Tsvp1* promoter (Sun *et al*., 2010) and the shortened *Act1* promoter (Su *et al*., 1998) were efficient in driving high levels of *gus* transcripts. Identification and analysis of the shortest promoter fragment capable of functional activity can be beneficial in the development of drought tolerant maize.

University of Cape Town

2. Research focus

2.1. Research question

This research sets out to determine whether shortened XvPSap1 (*circa* 50% reduction) and full length XvPSap2 promoters would be functionally active under dehydration stress. This evidence should indicate whether the promoters can drive expression of target genes in transgenic crop plants.

2.2. Research hypothesis

It is hypothesised that the full length XvPSap2 and shortened XvPSap1 stress-inducible promoters are functionally active under dehydration stress in *N. tabacum* and *A. thaliana*.

2.3. Research aims

The aims of this research are twofold. First, promoter truncations of the full length XvPSap1 will be generated using DNA mutagenesis. Second, the shortened XvPSap1 and full length XvPSap2 promoters will be assessed for functional activity in first generation transgenic *N. tabacum* under dehydration stress.

2.4. Significance of the study

Food security and economic growth in developing countries is to a large extent dependent on the key crop plant species. In sub-Saharan Africa, maize serves as the staple food crop and as a source of livelihood for many small-holder farmers. For these farmers, drought is one of the chief constraints to crop production and accounts for major decreases in crop yields and subsequent reductions in profits. While drought impacts all aspects of society, its effects on these farmers are direct. The development of affordable and safe drought tolerant maize will aid in serving as both a food source and profitable trade to resource poor farmers.

University of Cape Town

As yet, no minimal stress inducible promoter has been utilised in drought tolerant maize. It is hoped that a minimal XvPSap1 promoter could be used as a regulatory tool for enhancing gene expression with negligible or no negative effects to the crop itself.

2.5. Research assumptions and limitations

It is understood that although maize is the future target for the use of the promoters, it is not being used in the current research. However, the assumption made is that if the minimal XvPSap1 and full length XvPSap2 promoters are functionally active in the model plants used in this research, it would be functionally active in maize.

It is also understood that due to the nature of transgenic plant research, very few plants will be generated at the T1 stage. This may therefore limit the number of biological repeats or possibly even prevent their use. The assumption is made that trends observed in this research would provide suitable evidence towards the functionality of the promoter. However, this introduces the limitation that no absolute data will be obtained. This study therefore serves as a preliminary assessment to determine promoter activities in transgenic *N. tabacum*. In so doing, it serves to determine whether pursuing this research in a maize system would be feasible.

University of Cape Town

3. Materials and Methods

3.1. Mutagenesis of XvPSap1

The mutagenesis strategy of the XvPSap1 promoter is illustrated in Appendix A.1. The strategy is indicated for the generation of one of the shorter promoter fragments (XvPSap1D), however it is representative of the strategies for the other promoter fragments as well.

3.1.1. PCR amplification

Internal deletions were performed on the XvPSap1 promoter in the recombinant pBluescript::XvPSap1 plasmid to generate four shortened promoter fragments designated XvPSap1D, XvPSap1E, XvPSap1F and XvPSap1G (Appendix A.1). Two forward primers (Primer A and Primer B; Appendix D, Table D.2) and two reverse primers (Primer C and Primer D; Appendix D, Table D.2) were designed. Primers A and B bind to the 3'-end of XvPSap1 for amplification of 383 bp and 583 bp fragments, respectively (Fig. 3.1). Similarly, primers C and D bind to the 5'-end of XvPSap1 for amplification of 520 bp and 330 bp fragments, respectively (Fig. 3.1). The mutagenesis strategy involved the use of combinations of the respective forward and reverse primers to generate the four putative promoters in linearised pBluescript vector (Appendix B.1). Primer set B and C generated XvPSap1D, B and D generated XvPSap1E, A and C generated XvPSap1F and A and D generated XvPSap1G.

University of Cape Town

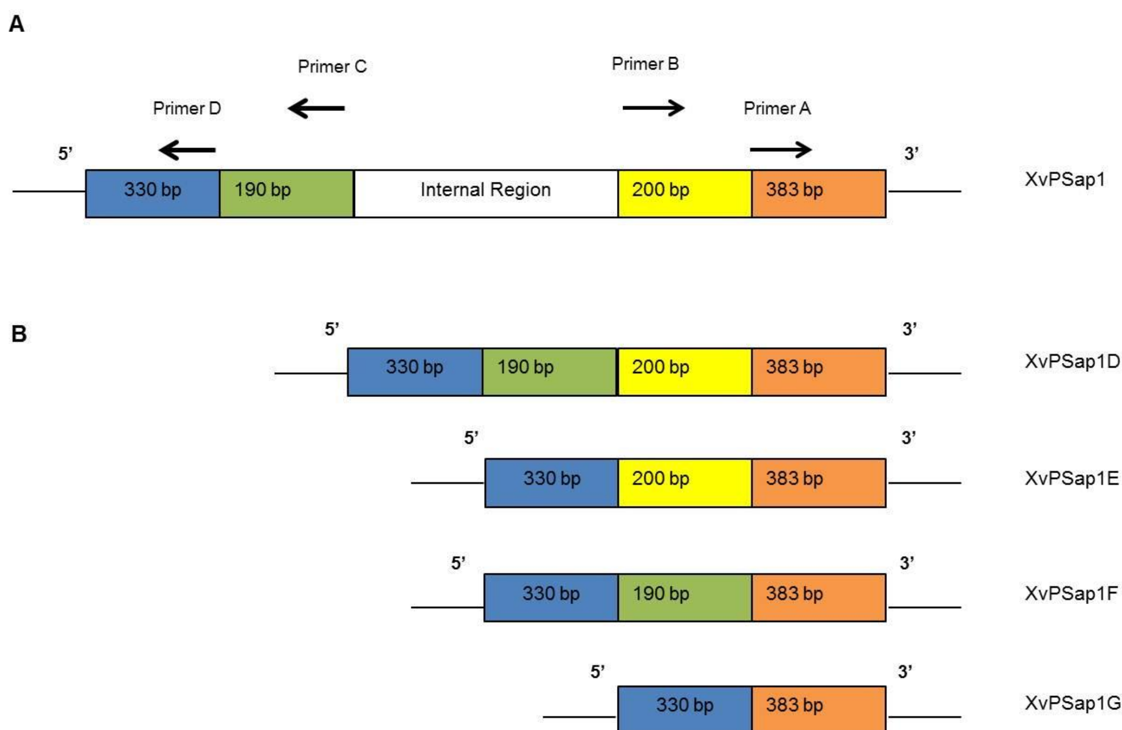


Figure 3.1: Schematic illustration of the full length and shortened XvPSap1 promoters. **A**: Full length XvPSap1. The black arrows indicate the position of binding on the XvPSap1 sequence and direction of amplification of the respective primers. **B**: The shorter promoter fragments are indicated.XvPSap1D is 1.103 kb, XvPSap1E is 0.913 kb, XvPSap1F is 0.903 kb and XvPSap1G is 0.713 kb in size.

University of Cape

For each amplification, 25 µl reaction volumes were set up with component concentrations as described in Appendix E.1. A high fidelity polymerase (Phusion High-Fidelity DNA Polymerase, Thermo Scientific, USA), which has proofreading activity was used. Amplification was carried out with the following conditions: 94ºC for 5 min; 5 cycles of 94ºC for 30 s, 52ºC for 45 s, 68ºC for 1 min; followed by 25 cycles of 94ºC for 30 s, 56ºC for 45 s, 68ºC for 1 min; and a final extension of 68ºC for 10 min. The PCR reaction was performed using a GeneAmp 9700 thermal cycler (Applied Biosystems, Singapore).

The generated PCR products were electrophoresed on a 1% ethidium bromide (EtBr) stained agarose gel. The bands of interest were excised and purified using the Wizard SV Gel Purification Kit (Promega Corporation, USA) according to the manufacturer's instructions (Appendix E.6).

3.1.2. Blunt end ligation of promoter fragments in pBluescript

The linearised pBluescript DNA was treated with Klenow Fragment exo^{-} (Fermentas, Canada), according to the manufacturer's instructions (Appendix E.2), to facilitate blunt end cloning. Standard blunt end ligation reactions were set up (as described in Appendix E.3) for ligation of the linearised pBluescript DNA. The reaction volume was made up to 10 µl, mixed well and incubated for 16 h at 4ºC.

3.1.3. Transformation of promoter constructs into *Escherichia coli* (*E. coli*)

The recombinant pBluescript plasmids (pBluescript::XvPSap1D, E, F and G) were transformed into competent *E. coli* DH5α cells. Competent cells were allowed to thaw on ice. Thereafter, 10 µl of ligation mix was added to a 100 µl aliquot of competent cells and mixed gently. This transformation mix was incubated for 10 min on ice and then heat shocked by incubation for 5 min at 37ºC followed immediately by incubation for 2 min on ice. Eight hundred microliters of Luria Bertani (LB) broth (Appendix F.1) was added to the transformed cells and incubated for 1 h at 37ºC with vigorous shaking. One hundred microliters of the transformation mix was plated on LB agar plates (Appendix F.2) supplemented with ampicillin (100 µg/ml) and incubated for 16 h at 37ºC.

University of Cape Town

3.1.4. Colony screening, plasmid extraction and screening to identify cloned promoter constructs

Colony PCR was performed to identify transformed clones using promoter specific primers (*Eco*RI-XvPsap1-F and *Bam*HI-XvPSap1-R, Appendix D, Table D.3). A 25 µl PCR amplification reaction was set up with component concentrations as described in Appendix E.1. A thermostable DNA polymerase, Supertherm Polymerase (Bertec Enterprise, Taiwan) was used for amplification. Amplification was carried out with the following conditions: 94ºC for 5 min; 30 cycles of 94ºC for 30 s, 59ºC for 45 s, 72ºC for 1 min; and a final extension of 72°C for 10 min.

Colonies that were observed to be positive by colony PCR screening were inoculated into 5 ml LB broth (Appendix F.1) supplemented with 100 µg/ml ampicillin and

incubated for 16 h at 37ºC with shaking. For each construct, plasmid DNA from 3 different clones was isolated using the Bioflux Plasmid DNA Extraction and Purification Kit (Bioer, Japan) according to the manufacturer's instructions (Appendix E.7). The purified plasmid DNA was stored at -20ºC. Plasmid DNA was sequenced and based on the sequence data, one recombinant plasmid for each promoter construct was selected for further downstream analysis.

3.2. Generation of pBluescript::promoter::*luc*::NosT

The strategy for generating the pBluescript::promoter::luc::NosT constructs are illustrated in Appendix A.2. The strategy is indicated for the generation of pBluescript::XvPSap1D::luc::NosT, however it is representative of the strategy for the other promoter constructs as well.

3.2.1. *Eco*RI and *Bam*HI double digestion of shortened promoter constructs and pBluescript::XvPSap1::*luc*::NosT

Endonuclease digestion of the pBluescript vector (Appendix B.1) containing the shorter promoters with *Eco*RI and *Bam*HI allowed for cleavage of the promoter construct from the pBluescript plasmid. Similarly, *Eco*RI and *Bam*HI double digestion of pBluescript::XvPSap1::*luc*::NosT allowed for removal of the original XvPSap1 promoter, resulting in linearised pBluescript::*luc*::NosT with *Eco*RI and *Bam*HI overhangs.

Three micrograms of each recombinant pBluescript plasmid was digested in a total volume of 60 µl using *Eco*RI and *Bam*HI (FastDigest, Fermentas, Canada). The reaction mixture contained 6 µl of 10X FastDigest buffer, 3 units FastDigest *Eco*RI and 3 units FastDigest *Bam*HI. Similarly, 4 µg of pBluescript::XvPSap1::*luc*::NosT was digested in a total volume of 80 µl using *Eco*RI and *Bam*HI (FastDigest, Fermentas, Canada). The reaction mixture contained 8 µl of 10X FastDigest buffer, 4 units FastDigest *Eco*RI and 4 units FastDigest *Bam*HI. The digestion mixtures were incubated for 1 h at 37ºC.

University of Cape Town

Digested products were electrophoresed on a 1% EtBr stained agarose gel and the desired digestion products excised and purified using the Wizard SV Gel Purification Kit (Promega Corporation, USA) according to the manufacturer's instructions (Appendix E.6).

3.2.2. Ligation of promoter DNA to pBluescript::*luc*::NosT

Site specific cohesive end ligation reactions were set up as described in Appendix E.4 for a vector to insert ratio of 1:3 using 50 ng of vector. Reaction components were mixed well and incubated for 16 h at 4ºC.

3.2.3. Transformation of recombinant promoter DNA into *E. coli*

The recombinant pBluescript plasmids (pBluescript::XvPSap1D, E, F, G::*luc*::NosT) were transformed into competent *E. coli* DH5α cells according to the protocol described in section 3.1.3. The transformed cells were plated on LB agar (Appendix F.2) supplemented with 100 µg/ml ampicillin and incubated for 16 h at 37ºC.

3.2.4. Colony screening, plasmid isolation and screening to identify pBluescript::promoter::*luc*::NosT

Colony PCR was performed to identify positively transformed clones using promoter specific primers (*Eco*RI-XvPsap1-F and *Bam*HI-XvPSap1-R, Appendix D, Table D.3). For each amplification, 25 µl reaction volumes were set up with component concentrations as described in Appendix E.1. Amplification was carried out according to the protocol described in section 3.1.4. Positive clones were inoculated into 5 ml LB broth (Appendix F.1) supplemented with 100 µg/ml ampicillin and incubated for 16 h at 37ºC with shaking. For each recombinant plasmid, 3 different clones were isolated using the Bioflux Plasmid DNA extraction and Purification Kit (Bioer, Japan) according to the manufacturer's instructions (Appendix E.7).

Recombinant plasmid DNA was assessed by endonuclease digestion. Five hundred nanograms of recombinant plasmid DNA was digested in a total volume of 20 µl using *Eco*RI and *Bam*HI (FastDigest, Fermentas, Canada). The reaction mixture

University of Cape Town

contained 2 µl of 10X FastDigest buffer, 1 unit FastDigest *Eco*RI and 1 unit FastDigest *Bam*HI. The reaction mixture was incubated for 1 h at 37ºC. Thereafter, digested products were electrophoresed on a 1% EtBr stained agarose gel. One recombinant plasmid for each promoter construct was selected for further downstream analysis. Undigested plasmid DNA was stored at -20ºC.

3.3. Generation of binary vector constructs

The cloning strategy for the generation of the binary vector constructs is illustrated in Appendix A.3. The strategy is indicated for the generation of pTF101.1::XvPSap1D::luc::NosT, however, it is a representation of the strategies for the other binary vector constructs as well.

3.3.1. Endonuclease digestion of recombinant pBluescript plasmid and pTF101.1

Digestion of pBluescript::promoter::luc::NosT constructs with *Eco*RI and *Hin*dIII allowed for cleavage of each shortened promoter cassette from the recombinant plasmid. Similarly, the *Eco*RI and *Hin*dIII restriction double digestion of pTF101.1 (Appendix B.2) resulted in linearised pTF101.1 with cohesive *Eco*RI and *Hin*dIII overhangs. pTF101.1 is a binary vector used in plant transformation protocols (Okoth, 2009).

Initially, 3 µg of each recombinant plasmid was digested in a total volume of 20 µl using *Pvu*II (Fermentas, Canada). The reaction mixture, containing 2 µl of 10X Buffer G and 1 unit *Pvu*II, was incubated for 1 h at 37ºC. The *Pvu*II digest was necessary to differentiate between the promoter cassettes and the pBluescript vector (Appendix B.1) due to their similar size (*circa* 3 kb). *Pvu*II cleaves pBluescript (Appendix B.1) to yield two fragments of *circa* 2.4 kb and a single 0.5 kb fragment. *Pvu*II does not cleave the promoter constructs. Thus, it was possible to distinguish between pBluescript and the promoter cassettes. Digested products were electrophoresed on a 1% EtBr stained agarose gel and purified using the EZ-10 Spin Column PCR Purification Kit (Bio Basic Inc, Canada) according to the manufacturer's instructions (Appendix E.5).

University of Cape Town

The *Pvu*II digested purified promoter cassettes were digested in a total volume of 50 µl using *Eco*RI and *Hin*dIII (FastDigest, Fermentas, Canada). The reaction mixture contained 5 µl of 10X FastDigest buffer, 3 units FastDigest *Eco*RI and 3 units FastDigest *Hin*dIII. The digestion mixtures were incubated for 1 h at 37ºC. Similarly, 3 µg of pTF101.1 was digested using *Eco*RI and *Hin*dIII. Digested products were electrophoresed on a 1% EtBr stained agarose gel and the desired digestion products excised and purified using the Wizard SV Gel Purification Kit (Promega Corporation, USA) according to the manufacturer's instructions (Appendix E.6).

3.3.2. Ligation of pTF101.1 and promoter cassettes

Site specific cohesive end ligation reactions were set up as described in Appendix E.4 for a vector to insert ratio of 1:3 using 50 ng of vector. Reaction components were mixed well and incubated for 16 h at 4ºC.

University of Cape Town

3.3.3. Transformation of recombinant pTF101.1 DNA into competent *E. coli*

The pTF101.1 plasmid containing the promoter::*luc*::NosT DNA fragment was transformed into competent *E. coli* DH5α cells according to the protocol described in section 3.1.3. The transformed cells were plated on LB agar (Appendix F.2) supplemented with 100 µg/ml streptomycin and incubated for 16 h at 37ºC.

3.3.4. Colony screening, plasmid extraction and screening to identify pTF101.1::promoter::*luc*::NosT

Colony PCR was performed to identify positively transformed clones using promoter specific primers (*Eco*RI-XvPsap1-F and *Bam*HI-XvPSap1-R; Appendix D, Table D.3). For each amplification, 25 µl reaction volumes were set up with component concentrations as described in Appendix A.1. Amplification was carried according to the protocol described in section 3.1.4.

Colonies observed to be positive by colony PCR screening were inoculated into 5 ml LB broth (Appendix F.1) supplemented with 100 µg/ml streptomycin and incubated for 16 h at 37ºC with shaking. For each construct, plasmid DNA from 3 different

clones was isolated using the Bioflux Plasmid DNA Extraction and Purification Kit (Bioer, Japan) according to the manufacturer's instructions (Appendix E.7).

The isolated recombinant plasmid DNA was assessed by endonuclease digestion. Five hundred nanograms of recombinant plasmid DNA was digested in a total volume of 20 µl using *Eco*RI and *Hin*dIII (FastDigest, Fermentas, Canada). The reaction mixture contained 2 µl of 10X FastDigest buffer, 1 unit FastDigest *Eco*RI and 1 unit FastDigest *Hin*dIII. The reaction mixtures were incubated for 1 h at 37ºC. Thereafter, digested products were electrophoresed on a 1% EtBr stained agarose gel. One recombinant plasmid for each promoter construct was selected for further downstream analysis. Undigested plasmid DNA was stored at -20ºC.

3.4. *In silico* analysis of promoter sequences

Following sequencing of the promoter fragments, the sequences were assessed for the presence of core and regulatory elements using the plantCARE software (Lescot *et al*., 2002).

3.5. Transformation and screening of *Agrobacterium tumefaciens*

The four pTF101.1 recombinant plasmids containing XvPSap1D,E,F,G::*luc*::NosT, were transformed into competent *A. tumefaciens* EHA101 cells. The transformation was carried according to the protocol described in section 3.1.3 with two modifications. Firstly, transformed *A. tumefaciens* cells were incubated for 6 h at 30ºC instead of 1 hr at 37ºC. Secondly, selection of transformed cells was performed on YEP agar (Appendix F.4) supplemented with 100 µg/ml spectinomycin, 30 µg/ml chloramphenicol and 50 µg/ml kanamycin.

Colony PCR was performed to identify positively transformed clones using promoter specific primers (*Eco*RI-XvPsap1-F and *Bam*HI-XvPSap1-R; Appendix D, Table D.3). For each amplification, 25 µl reaction volumes were set up with component concentrations as described in Appendix E.1. Amplification was carried according to the protocol described in section 3.1.4.

University of Cape Town

Colonies observed to be positive by colony PCR screening were inoculated into 10 ml YEP Broth (Appendix F.3) supplemented with 100 µg/ml spectinomycin, 30 µg/ml chloramphenicol and 50 µg/ml kanamycin and incubated for 16 h at 30ºC with shaking. For each construct, plasmid DNA from 3 different clones was isolated using the Bioflux Plasmid DNA Extraction and Purification kit (Bioer, Japan) according to the manufacturer's instructions (Appendix E.7).

The isolated recombinant PTF101.1 plasmid DNA was verified to contain the entire promoter cassette by endonuclease digestions with *Eco*RI and *Hin*dIII. Digestion reactions were carried out according to the protocol described in section 3.3.1. One recombinant plasmid for each promoter construct was selected for further downstream analysis. Glycerol stocks were synthesised (Appendix E.8) and stored at -80ºC.

University of Cape Town

3.6. Plant transformation

3.6.1. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* with pTF101.1 vector constructs

3.6.1.1. Sterilisation of *A. thaliana* seeds

Fifty microliters of wild type *A. thaliana* (Columbia ecotype) seed was transferred to a sterile 2 ml Eppendorf tube. One millilitre of 20% JIK (active ingredient w/v 3.5%: sodium hypochlorite) supplemented with 0.1% Tween20 was added to the seeds, briefly vortexed and incubated for 15 min at room temperature. The mixture was centrifuged for 30 s at 16000 x *g* and the JIK solution removed. The sterilised seeds were washed in 1 ml sterile water, vortexed for 30 s and centrifuged for 1 min at 16000 x *g*. The water was removed and the washes repeated 4 times. The sterilised seeds were dried between sterile filter paper and collected in a sterile 1.5 ml Eppendorf tube.

3.6.1.2. Seed germination and growth conditions

Sterilised seeds were spread on peet::vermiculite mixture in a ratio of 1:1 supplemented with 0.1 g/l Gaucho 70 WS (Bayer, South Africa) and 1.14 g/l phostrogen (Bayer, South Africa). Seeds were covered with saran wrap and vernalised by incubation for 2 days at 4ºC to minimise dormancy and improve germination rate. Sterilised vernalised seeds were transferred to the growth room with set conditions (22ºC; 16 h light, 8 h dark). Following 1 week of germination, the saran wrap was removed and seedlings watered twice a week. The first bolts were excised to encourage proliferation of secondary bolts. Plants were ready for *Agrobacterium*-mediated transformation 5 weeks post germination.

3.6.1.3. Preparation of *A. tumefaciens*

Individual *A. tumefaciens* colonies carrying the various XvPSap1 expression cassettes were inoculated into 10 ml of YEP media (Appendix F.3) supplemented with 100 µg/ml spectinomycin, 30 µg/ml chloramphenicol and 50 µg/ml kanamycin. The cultures were incubated for 16 h at 30ºC. One millilitre of the 16 h culture was thereafter inoculated into 200 ml of YEP media supplemented with the appropriate antibiotics. The cultures were incubated at 30ºC until an absorbance reading of approximately 0.8 was obtained at 600 nm. The cultures were centrifuged for 20 min at 6000 x *g* at 4ºC. The supernatant was discarded and the bacterial pellets resuspended in 5% (w/v) sucrose.

3.6.1.4. Transformation of *A. thaliana* with pTF101.1 vector constructs

Individual six week old wild type *A. thaliana* plants were transformed with pTF101::XvPSap1D, E, F and G::*luc*::NosT, using the floral dip method (Clough and Bent, 1998) with two modifications. Firstly, the inflorescences of the plants were dipped twice into the *Agrobacterium* solution for 1 min. Secondly, the plants were incubated for 24 h in the dark after which they were cultured with standard conditions as described in section 3.6.1.2 and watered twice weekly for 1 month. Plants were treated with 1.14 g/l phostrogen (Bayer, South Africa) every 2 weeks.

University of Cape Town

3.6.1.5. Seed harvesting and germination of T0 seed

Water was withheld from mature plants for 1 month and seeds collected. The T0 seed were sterilised and germinated on peet::vermiculite according to the protocols described in sections 3.6.1.1 and 3.6.1.2.

3.6.1.6. Selection of putative transformants

Two weeks after seed germination, plants were treated with 3 mg/l BASTA (Bayer, South Africa) by direct spraying onto the leaves. This was repeated every 4 days over a 2 week period. The surviving BASTA resistant plants were transferred to pots containing 0.1 g/l Gaucho 70 SW (Bayer, South Africa) treated peet::vermiculite mix.

3.6.2. *Agrobacterium*-mediated transformation of *N. tabacum* with pTF101.1 vector constructs

3.6.2.1. Sterilisation and germination of *N. tabacum* seeds

Fifty microliters of wild type *N. tabacum* (SR1 ecotype) seed was transferred to a sterile 2 ml Eppendorf. Seeds were sterilised according to the protocol described in section 3.5.2.1 and germinated on sterile potting soil mix. Plants were cultured in the growth room with set conditions (24ºC; 16 h light, 8 h dark). Plants were maintained at these conditions for 2 to 3 months and watered twice weekly. Plants were treated with 1.14 g/l phostrogen (Bayer, South Africa) every second week.

3.6.2.2. Preparation of *A. tumefaciens*

Single colonies of the transformed *A. tumefaciens* carrying the individual XvPSap1 expression cassettes were inoculated into 10 ml of YEP media (Appendix F.3) supplemented with 100 µg/ml spectinomycin, 30 µg/ml chloramphenicol and 50 µg/ml kanamycin. The cultures were incubated for 16 h at 30ºC. One ml of the 16 h culture was inoculated into 200 ml of YEP media supplemented with the appropriate antibiotics. The cultures were incubated at 30°C until an absorbance reading of approximately 0.8 at 600 nm was obtained. The cultures were centrifuged for 20 min

University of Cape Town

at 6000 x *g* at 4°C. The supernatant was discarded and the bacterial pellets resuspended in 50 ml of liquid co-cultivation media comprising hormone free MS basal salts (Highveld Biological, South Africa) supplemented with B5 vitamins (Appendix F.5), 30 g/l sucrose (Sigma-Aldrich, Germany), 0.1 mg/l α-naphthaleneacetic acid (Sigma-Aldrich, Germany), 1 mg/l 6-benzylamino purine (Sigma-Aldrich, Germany) and 100 μM/l acetosyringone (Sigma-Aldrich, Germany). The pH was adjusted to 5.4 with 1 M Potassium hydroxide (KOH).

3.6.2.3. Preparation of *N. tabacum* leaf explants

Four to six inch leaves were selected from 2 month old plants. Leaves were soaked in sterile water for 30 min and sterilised according to the protocol described in section 3.6.1.2. Sterile leaves were sliced into uniform segments of 5 mm avoiding the leaf margins and mid vein. Leaf explants were placed adaxial side up onto pre-culture media. Pre-culture was comprised of hormone free MS basal salts (Sigma-Aldrich, Germany), supplemented with B5 vitamins (Appendix F.5), 30 g/l sucrose (Sigma-Aldrich, Germany), 0.1 mg/l α-naphthaleneacetic acid (Sigma-Aldrich, Germany), 1 mg/l 6-benzylamino purine (Sigma-Aldrich, Germany), 100 μM/l acetosyringone (Sigma-Aldrich, Germany) and 8 g/l plant agar(Sigma-Aldrich, Germany). The pH was adjusted to 5.7 with 1 M KOH.

3.6.2.4. *Agrobacterium* infection of leaf disks

Leaf disks were infected for 30 min in the dark with the *Agrobacterium* inoculum containing the promoter cassettes in sterile petri dishes. The petri dish was agitated once every 10 min. Thereafter, infected leaf disks were blot dried on sterile filter paper. Negative controls infected with *A. tumefaciens* containing pTF101.1 vector (Appendix B.2) were also included.

3.6.2.5. Co-cultivation of infected explants

Each infected explant was transferred to co-cultivation medium and incubated for 3 days at 23ºC (18 h light, 6 h dark; light intensity of 140 μmol/m^2/s). The adaxial part of the leaf was kept in contact with the medium. The co-cultivation media comprised

University of Cape Town

hormone free MS basal salts (Highveld Biological, South Africa) supplemented with B5 vitamins (Appendix F.5), 30 g/l sucrose (Sigma-Aldrich, Germany), 0.1 mg/l α-naphthaleneacetic acid (Sigma-Aldrich, Germany), 1 mg/l 6-benzylamino purine (Sigma-Aldrich, Germany) and 100 μM/l acetosyringone (Sigma-Aldrich, Germany). The pH was adjusted to 5.4 with 1M KOH.

3.6.2.6. Shooting and rooting of explants

Following the 3 day co-cultivation period, leaf discs were selected on shooting medium comprising hormone free MS basal salts (Highveld Biological, South Africa) supplemented with B5 vitamins (Appendix F.5), 30 g/l sucrose (Sigma-Aldrich, Germany), 0.1 mg/l α-naphthaleneacetic acid (Sigma-Aldrich, Germany), 1 mg/l 6-benzylamino purine (Sigma-Aldrich, Germany), 10 μg/l nyastatin (Sigma-Aldrich, Germany), 250 mg/l carbenicillin (Sigma-Aldrich, Germany), and 3 mg/l BASTA (Bayer, South Africa). Leaf explants were placed under an 18 h light regime with light intensity of 140 $\mu mol/m^2/s$ at 28ºC. Putative transformants were subcultured fortnightly onto fresh media until sizable shoots were formed. BASTA resistant shoots were selected, excised and transferred to rooting media. The rooting media comprised half strength hormone free MS basal salts (Highveld Biological, South Africa) supplemented with 10 mg/l sucrose (Sigma-Aldrich, Germany), 10 μg/l nyastatin (Sigma-Aldrich, Germany), 250 mg/l carbenicillin (Sigma-Aldrich, Germany) and 3 mg/l BASTA (Bayer, South Africa).

University of Cape Town

3.6.2.7. Acclimatisation and growth of putative tobacco transformants

Putative transformants with well-established root systems were transferred to pots containing sterile potting soil and cultured with set conditions according to the protocol described in section 3.6.2.1. The plants were covered with saran wrap for 8 days to assist acclimatisation and minimise dehydration. Once acclimatised, the putative transformants were transferred to 6 inch pots containing potting soil under normal growth conditions. Mature plants were self-pollinated and seed was harvested from mature dry pods. The transformation efficiency was calculated according to the formula:

$$\text{Transformation Efficiency} = \left[\frac{\text{Number of Positive Transformants}}{\text{Number of Explants Transformed}} \right] \text{X } 100$$

3.6.2.8. Germination of putative transgenic seeds and BASTA screening of putative transformants

Putative transgenic tobacco seed was sterilised according to the protocol described in section 3.6.2.1 and germinated on MS media supplemented with 8 g/l agar (Sigma-Aldrich, Germany) and 3 mg/ml BASTA (Highveld Biological, South Africa). The pH was adjusted to 5.7 with 1 M KOH. Plants were cultured in the growth room with set conditions (24ºC; 16 h day, 8 h night). The surviving BASTA resistant plants with well-established root systems were transferred to trays containing 0.1 g/l Gaucho 70 SW (Bayer, South Africa) treated potting soil and covered with saran wrap for 1 week. Three weeks later, BASTA resistant transgenic plants were transferred to pots containing 0.1 g/l Gaucho SW treated potting soil. Plants were treated with 1.14 g/l phostrogen (Bayer, South Africa) every second week.

University of Cape Town

3.6.3. Screening of putative transgenic plants

3.6.3.1. Genomic DNA isolation of putative transgenic plants

Leaves were sampled from putative transgenic plants and were flash frozen in liquid nitrogen. Genomic DNA was extracted using the Dellaporta extraction protocol (Dellaporta *et al.*, 1983) with minor modifications. Leaf tissue was ground in liquid nitrogen using a mortar and pestle. Roughly 100 µg of ground tissue was transferred to a sterile 2 ml Eppendorf tube containing 1.4 ml of Extraction buffer (100 mM Tris-Cl pH 8, 50 mM EDTA pH 8, 500 mM NaCl and 10 mM β-mercaptoethanol) and 0.1 ml of 20% SDS and incubated for 10 min at 65ºC. Thereafter, 500 µl of 5 M potassium acetate was added and the samples shaken vigorously for 5 min followed by incubation for 20 min at 4ºC. The samples were then centrifuged for 20 min at

16000 x *g*. The supernatant was transferred to a sterile 2 ml Eppendorf tube containing 1 ml isopropanol and mixed by gentle inversion. The genomic DNA was precipitated for 24 h at -20ºC followed by centrifugation for 15 min at 16000 x *g*. The supernatant was discarded and the pellets were air dried for 10 min. The pellets were resuspended in 70 µl of Resuspension buffer (50 mM Tris-Cl pH 8, 10 mM EDTA pH 8 and 0.6 mg/ml RNase A) at room temperature. To remove insoluble debris, the samples were centrifuged for 5 min at 16000 x *g* and the supernatant transferred to a sterile 1.5 ml Eppendorf tube containing 7.5 µl of 3 M potassium acetate and 50 µl isopropanol. The samples were mixed well, incubated for 15 min at 4ºC and centrifuged for 2 min at 16000 x *g*. The supernatant was discarded and the genomic DNA pellets were washed with 1 ml of 80% absolute ethanol. The mixture was centrifuged for 2 min at 16000 x *g* and the supernatant removed. The genomic DNA pellet was redissolved in 100 µl TE (10 mM Tris-Cl, 1 mM EDTA). The quality of the extracted genomic DNA was assessed by electrophoresis on a 1% EtBr stained agarose gel.

3.6.3.2. Detection of transgene integration

The presence of the *bar* gene was determined by PCR amplification of a 421 bp DNA fragment using gene-specific primers (BarI and BarII; Appendix D, Table D.3). For each amplification, 50 µl reaction volumes were set up with component concentrations as described in Appendix E.1. Amplification was carried out with the following conditions: 94ºC for 5 min; 35 cycles of 94ºC for 30 s, 56ºC for 45 s, 72ºC for 1 min; and a final extension of 72ºC for 10 min. The PCR reaction was performed using a GeneAmp 9700 thermal cycler (Applied Biosystems, Singapore). The generated amplimers were electrophoresed on a 1% EtBr stained agarose gel.

Similarly, the presence of the promoter and *luc* gene was determined by PCR amplification of a fragment of *circa* 2 kb using a promoter-specific forward and a *luc*-specific reverse primer pair (Appendix D, Table D.3). For each amplification, 50 µl reaction volumes were set up with component concentrations as described in Appendix A.1. Amplification was carried out with the following conditions: 94ºC for 5 min; 35 cycles of 94ºC for 30 s, 54ºC for 2 min, 68ºC for 90 s; and a final extension of 72ºC for 10 min.

University of Cape Town

3.6.4. Dehydration of transgenic plants

3.6.4.1. Dehydration of transgenic tobacco plants

The dehydration treatments were carried out as described by Audran *et al* (1998) with the following modification: plants were dehydrated in soil instead of hydroponically. Six plants for each truncated promoter construct were used in the dehydration treatment. Prior to dehydration, plants were transferred to pots containing 800 g of soil and 200 ml water to allow for even drying rates. Plants were moved to Percival chambers (Percival Scientific Inc, USA) and incubated under set conditions (26ºC; 16 h day, 8 h night; 60% humidity; light intensity of 100 μmol/m^2/s) for 1 week. Dehydration stress was carried out on whole plants and achieved by withholding water for 6 days. Day 0 of the dehydration treatment represents the first day after watering ceased. Throughout the dehydration period, four stressed tobacco leaves were sampled every 24 hours. The sampled leaves were used to assay for luciferase activity (see section 3.6.5.1) and immediately frozen in liquid nitrogen and stored at -70ºC. The leaves were used for RNA isolations (see section 3.6.5.2.2).

An additional twelve transgenic and six wild type plants were included in the dehydration stress treatment. Throughout the dehydration period, four leaves (representing a single plant) were sampled every 24 hours. The sampled leaves were used to estimate the relative water content.

3.6.4.2. Determination of relative water content

The relative water content (RWC) was calculated for each sampled leaf at each time point. The fresh weight (FW) of each leaf was determined immediately after sampling. The full turgor weight (FTW) was determined after a 24 h immersion of the leaf in sterile water at room temperature. The leaves were then incubated for 48 h at 70ºC to determine the dry weight (DW). Relative water content was calculated according to the following formula:

University of Cape Town

$$RWC = \frac{(FW - DW)}{(FTW - DW)} \times 100$$

3.6.4.4. Determination of soil water content

The soil water content (SWC) of each pot containing a dehydrated plant was determined using the HH2 Moisture Meter (Delta-T Devices, United Kingdom). Each pot was probed and the soil water content was determined in triplicate.

3.6.5. Analysis of promoter activity

3.6.5.1. Live imaging of luciferase expression

Prior to sampling for RNA isolations, leaves were individually sprayed and painted with equal amounts of 5 mM luciferin (VivoGlo, Promega Corporation, USA). Luciferase activity was imaged with a 3D-luminometer consisting of a 0.5 square inch CCD camera and a field of view of 12.5 cm (Xenogen IVIS Lumina, Caliper, USA) at an exposure time of 300 s per leaf. Photon or count emission by luciferase expressing leaves was quantified using the Living Image software (Caliper, USA). The GFP assay was selected to negate any luminescence from chloroplasts.

3.6.5.2. Luciferase expression by quantitative real-time PCR

3.6.5.2.1. RNA extraction

All glassware and plastics used were double autoclaved and solutions were prepared with 0.001% diethylpyrocarbonate (DEPC) treated water. Each stored leaf was individually ground to a fine powder in liquid nitrogen with a mortar and pestle. The ground tissue was maintained at 4ºC during the extraction procedure to prevent RNA degradation. Approximately 600 mg of ground tissue was transferred to a sterile 2 ml Eppendorf tube containing 1 ml of One-Step reagent (Bio Basic Inc, Canada) and

University of Cape Town

mixed by inversion. Thereafter, samples were vortexed for 10 min at room temperature. To allow complete dissociation of the nucleoprotein complexes, the samples were incubated for 5 min at room temperature. Following incubation at room temperature, 200 µl of chloroform was added and the samples were inverted 30 times and then incubated for 3 min at room temperature. Samples were centrifuged for 15 min at 4ºC at 14000 x *g*. The upper aqueous phase was carefully transferred to a fresh sterile 2 ml Eppendorf tube containing 1 ml of isopropanol and mixed by inversion. The mixture was allowed to settle for 10 min at room temperature before the RNA was pelleted by centrifugation for 15 min at 4ºC at 14000 x *g*. The supernatant was discarded and the RNA pellet was washed with 1 ml of ice cold 75% ethanol, vortexed briefly and centrifuged for 10 min at 4ºC. The supernatant was discarded and the RNA pellet was air dried for 5 min and dissolved in 89 µl of 0.01% DEPC treated water. Dissolution was enhanced by incubating the RNA for 10 min at 55ºC. The RNA was stored at -70ºC.

3.6.5.2.2. DNAse treatment and RNA clean up

Each RNA extraction was treated with Deoxyribonuclease I (DNAse I; New England Biolabs, USA) to digest and remove any genomic DNA contamination. The DNAse I reaction mixture consisted of 89 µl of isolated RNA, 10 µl of 10X DNAse buffer and 2 units of DNAse I in a total reaction volume of 100 µl. The reaction was mixed gently and incubated for 10 min at 37ºC. The sample was purified using the GeneJet Plant RNA Miniprep Kit (Fermentas, Canada) according to the manufacturer's instructions. The purified RNA was quantified using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, USA). The RNA integrity was assessed by electrophoresis at on a 1.2% EtBr stained agarose gel. Furthermore, the quality and integrity was assessed on a RNA-6000 Nano chip using the Agilent 2100 Bioanalyzer and analysed using the Agilent 2100 Expert software (Anatech, USA).

3.6.5.2.3. cDNA synthesis

The purified RNA extracted from the four leaf samples of each *N. tabacum* plants over the dehydration period were pooled and used for cDNA synthesis. *Circa* 500 ng of RNA was used per cDNA synthesis reaction. This reaction was performed in

University of Cape Town

quadruplicate to act as technical cDNA synthesis repeats. For cDNA synthesis, M-MuLV RNase H+ reverse transcriptase (Finnzymes, Finland) was used according to the manufacturer's instructions. The cDNA synthesis reaction mix consisted of *circa* 500 ng of RNA, a final concentration of 0.05 µg oligo $(dT)_{15}$ primers, 0.45 µg random hexamers, 1X RT buffer (includes dNTPs and $MgCl_2$) and 0.04 µl M-MuLV RNase H+ reverse transcriptase (includes an RNA inhibitor) made up to a total volume of 20 µl with nuclease free water and mixed well. A ratio of 1:10 of random hexamers to oligo $(dT)_{15}$ was selected for cDNA synthesis as this method increases the sensitivity of the synthesis reaction. The PCR cycle conditions consisted of incubation at 25ºC for 5 min, 42ºC for 1 hour, 85ºC for 15 min followed by a final step at 4ºC for 2 min. An aliquot of cDNA synthesis was electrophoresed on a 1.2% agarose/EtBr gel to analyse cDNA quality. The cDNA was stored at -20ºC until further use.

3.6.5.2.4. Quantitative real-time PCR primer design

Four sets of primers for the reference genes (18S rRNA, L25 ribosomal protein and elongation factor-1α) and the gene of interest (luciferase) were designed using the Primer3 programme (Rozen and Skaletsky, 2000). The primers were assessed for secondary structure formation using the DNAMAN software (Lynnon Biosoft Copyright © 1994-2001). End point PCR and melt curve analysis were finally used to identify both specific and non-specific amplification.

3.6.5.2.5. Generation of standard curves to determine quantitative real-time PCR efficiency

Serial dilutions of pooled cDNA from the treated samples were used to generate five-point standard curves. For the 18S rRNA and luciferase standard curves, a 10-fold and 3-fold dilution series was prepared respectively. For the L25 ribosomal protein and elongation factor-1α (*ef-1α*) standard curves, a 4-fold dilution series was prepared.

Expression of the reference genes and luciferase gene was investigated by real time PCR using the Rotor Gene 6000 2 plex HRM (Corbett Life Science Research). The Kapa Sybr Fast Kit (Kapa BioSystems, South Africa) master mix containing reaction

University of Cape Town

buffer, heat activated DNA polymerase, dNTPs and a working concentration of 3mM $MgCl_2$ were used for each PCR reaction. Each PCR reaction contained a final concentration of 1X Kapa Sybr Fast, 0.04 µM gene specific primers, 2 µl of cDNA and nuclease free water to a total volume of 20 µl. Gene specific primer sets for luciferase (RT-*luc*-F2 and RT-*luc*-R2; Appendix D, Table D.4), 18S rRNA (RT-18S-F3 and RT-18S-R3; Appendix D, Table D.4), L25 ribosomal protein (RT-Rib-F3 and RT-Rib-R3; Appendix D, Table D.4) and elongation factor1-α (RT-EF-F1 and RT-EF-R1; Appendix D, Table D.4) were used in separate reactions. Each dilution point reaction was performed in quadruplicate. A no reverse transcription control (pooled RNA) as well as a no template control (NTC) was included in each real time run. Amplification was carried out with the following conditions: 95ºC for 5 min; 45 cycles of 95ºC for 5 s, 60ºC for 20 s, 72ºC for 5 s. Generation and analysis of standard curves were performed using the Rotor-Gene 6000 Series software (Corbett Life Science Research, Australia).

3.6.5.2.6. Analysis of expression stability of reference genes

Real time PCR with the primer pairs for each reference gene was used to determine the expression stability of each of the potential reference genes in dehydrated *N. tabacum* plants. The PCR reaction mix and conditions were set up as described in section 3.6.5.2.5. To minimise variations between PCR runs, all of the reactions containing one primer pair was performed in one run. The average expression levels were calculated from four technical repeats and by importing the relative standard curve into each run. Relative gene expression was determined by the amplification threshold in the exponential phase of the PCR, identifying the cycle threshold (Ct) value and comparing the Ct value to the standard curve (Muller *et al*., 2002). The stability of the potential reference genes were evaluated using both GeNorm and NormFinder (GenEx, MultiD, Sweden).

3.6.5.2.7. Relative quantification of *luc* with stable reference genes

Real time PCR reaction mixes and conditions were set up as described in section 3.6.5.2.5. To minimise variations between PCR runs, all of the reactions containing one primer pair was performed in one run. The standard curves (see section

University of Cape Town

3.6.5.2.5) were imported into each run to determine the Ct values and concentrations of the gene of interest and the reference gene. The determined values of the gene of interest were divided by that of the reference gene. The averages of the calculated values were used for relative quantification of the gene of interest. The value obtained for transcript levels on day 0 was used as a calibrator to determine whether a significant change in expression of the gene of interest occurs during dehydration. Relative quantification levels were determined using the GenEx software (MultiD, Sweden) according to the Pfaffl equation (Pfaffl, 2001). The Pfaffl equation of one sample is the ratio of the gene of interest (target) versus a calibrator sample (control) and the reference gene (reference) versus a calibrator sample (control). The amplification efficiencies (E) were calculated according to the equation: $E = 10^{[-1/slope]}$. The difference in Ct values of the target gene in the control and sample (Δ^{Ct} target) and in the reference gene in the control and sample (Δ^{Ct} reference) are considered (Pfaffl, 2001). The equation is as follows:

$$\text{Ratio} = \frac{(E_{\text{target}})^{\Delta Ct\ \text{target (control – sample)}}}{(E_{\text{reference}})^{\Delta Ct\ \text{reference (control – sample)}}}$$

University of Cape Town

4. Results

4.1. Mutagenesis of XvPSap1

Mutagenesis of the XvPSap1 promoter was performed according to the strategy outlined in Appendix A.1. The pBluescript::promoter fragments were successfully amplified, however non-specific amplification was also observed evidenced by the higher and lower molecular weight bands (Fig. 4.1). An intense band, corresponding to a fragment of *circa* 4 kb, was observed in lanes 1 and 2. This band corresponded to the expected size of the linear pBluescript::XvPSap1D, E, F and G fragments (Fig. 4.1**A**, **B**, **C** and **D**, respectively). As a negative control, control PCR reaction without template (NTC) was included. No amplification was observed in this control reaction.

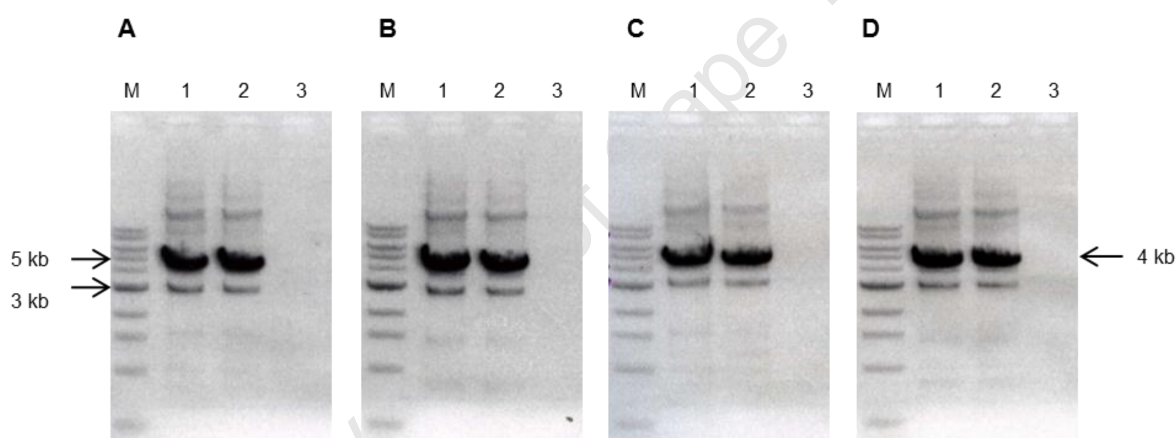


Figure 4.1: Amplification of pBluescript DNA incorporating the promoter fragments. **A**: pBluescript::XvPSap1D. **B**: pBluescript::XvPSap1E. **C**: pBluescript::XvPSap1F. **D**: pBluescript::XvPSap1G. Lane M: 1 kb DNA ladder. Lane 1 and 2: pBluescript::promoter fragments. Lane 3: NTC

University of Cape Town

The amplified DNA fragments were successfully excised, purified and treated with Klenow polymerase to facilitate blunt end ligation. Ligation mixtures were transformed into competent *E. coli* DH5α cells. Fifteen colonies each for pBluescript::XvPSap1D, E, F and G constructs were selected for screening. All of the selected colonies contained the desired promoter fragment (data not shown). Three colonies for each promoter fragment were selected for further analysis (data not

shown). Sequencing results confirmed that the four shortened promoters successfully had been ligated in pBluescript (Appendix C.2, C.3, C.4 and C.5).

4.2. Generation of pBluescript::promoter::*luc*::NosT

To facilitate later qualitative and quantitative analyses, the shortened promoters were individually cloned upstream of *luc* and NosT (Appendix A.1). To do this, the individual pBluescript vectors (Appendix B.1) containing the shortened promoters, XvPSap1D, E, F and G were successfully digested with *Eco*RI and *Bam*HI to release the respective promoter fragments. Single linear bands of 1.103 kb, 0.913 kb, 0.903 kb and 0.713 kb corresponding to the size of XvPSap1D, E, F and G, respectively and a band of 3 kb corresponding to the size of pBluescript, were observed following electrophoresis of the digestion products (data not shown). Similarly, *Eco*RI and *Bam*HI double digestion of pBluescript::XvPSap1::*luc*::NosT was successful in removing the original XvPSap1 promoter, resulting in linearised pBluescript::*luc*::NosT with *Eco*RI and *Bam*HI overhangs. A single band of *circa* 5 kb, corresponding to the expected size of pBluescript::*luc*::NosT and a band of *circa* 2 kb, corresponding to the size of XvPSap1, were observed when digestion products were electrophoresed (Appendix J.1).

University of Cape Town

The shortened fragments of interest as well as the full length XvPSap2 fragment were successfully excised, purified and cohesive end ligation reactions were set up. Ligation mixtures were successfully transformed into competent *E. coli* DH5α cells.

Fifteen colonies for each construct were selected for screening. Fourteen colonies contained XvPSap1D, E and G and fifteen colonies contained XvPSap1F. Twelve positive clones were identified for XvPSap2 (data not shown).

Three colonies for each of the promoter constructs were selected for plasmid isolation and restriction endonuclease analysis. Each isolated plasmid was subjected to *Eco*RI and *Hin*dIII digestion and thereafter electrophoresed. For all samples, a band of *circa* 3 kb was observed (Fig. 4.2). This was expected as both the vector backbone and promoter::*luc*::NosT constructs are *circa* 3 kb. An *Eco*RI and *Bam*HI restriction endonuclease reaction was also performed to differentiate between the

pBluescript vector (Appendix B.1) and the promoter::*luc*::NosT cassettes. In this instance, two distinct DNA fragments corresponding to the promoter and the pBluescript::*luc*::NosT fragments were expected. Following digestion, two distinct bands were observed (Fig. 4.2). The larger *circa* 5 kb fragment corresponded to the size of pBluescript::*luc*::NosT whereas the smaller fragments corresponded to the sizes of the various promoters.

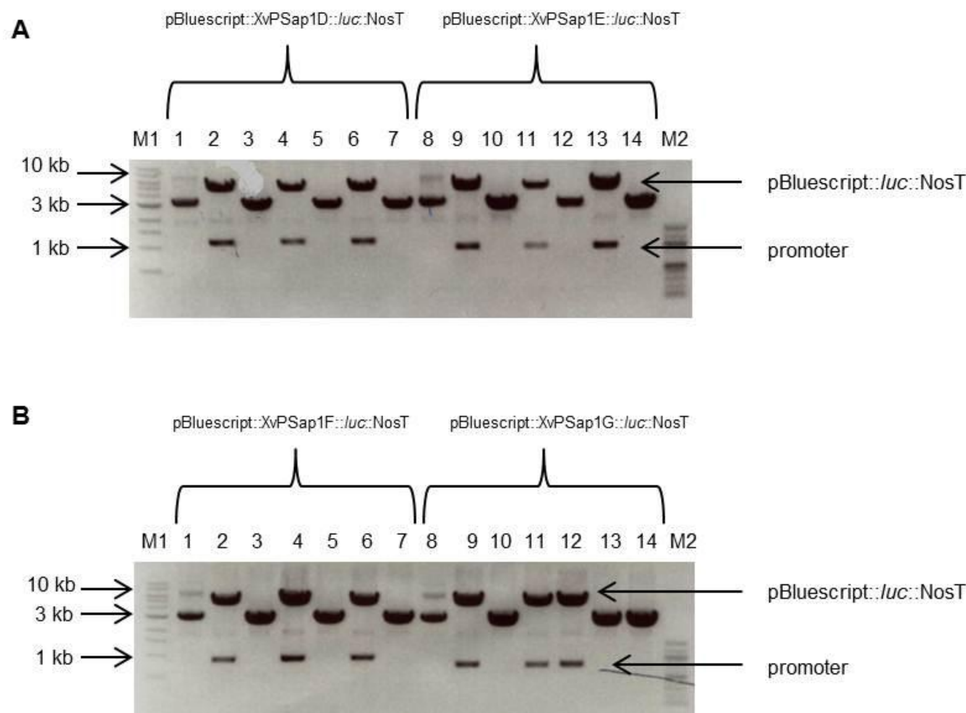


Figure 4.2: Restriction endonuclease analysis of pBluescript::promoter::*luc*::NosT. **A**: pBluescript::XvPSap1D::*luc*::NosT and pBluescript::XvPSap1E::*luc*::NosT. Lane 1: Undigested pBluescript::XvPSap1D::*luc*::NosT. Lanes 2, 4 and 6: EcoRI and BamHI restriction digests of pBluescript::XvPSap1D::*luc*::NosT. Lanes 3, 5 and 7: EcoRI and HindIII restriction digests of pBluescript::XvPSap1D::*luc*::NosT. Lane 8: Undigested pBluescript::XvPSap1E::*luc*::NosT. Lanes 9, 11 and 13: *Eco*RI and *Bam*HI restriction digests of pBluescript::XvPSap1E::*luc*::NosT. Lanes 10, 12 and 14: *Eco*RI and *Hin*dIII restriction digests of pBluescript::XvPSap1E::*luc*::NosT. **B**: pBluescript::XvPSap1F::*luc*::NosT and pBluescript::XvPSap1G::*luc*::NosT. Lane 1: Undigested pBluescript::XvPSap1F::*luc*::NosT. Lanes 2, 4 and 6: *Eco*RI and *Bam*HI restriction digests of pBluescript::XvPSap1F::*luc*::NosT. Lanes 3, 5 and 7: *Eco*RI and *Hin*dIII restriction digests of pBluescript::XvPSap1F::*luc*::NosT. Lane 8: Undigested pBluescript::XvPSap1G::*luc*::NosT. Lanes 9, 11 and 12: *Eco*RI and *Bam*HI restriction digests of pBluescript::XvPSap1G::*luc*::NosT. Lanes 10, 13 and 14: *Eco*RI and *Hin*dIII restriction digests of pBluescript::XvPSap1G::*luc*::NosT. Lanes M1 and M2: 1 kb and 100 bp DNA ladder, respectively.

4.3. Generation of binary vector constructs

To facilitate plant transformation, the promoter::*luc*::NosT constructs were cloned into pTF101.1 (Appendix B.2) as described in Appendix B.3. The pTF101.1 vector is a binary vector containing the necessary DNA sequences for integration of transgenes into the plant genomic DNA. To differentiate between the pBluescript (Appendix B.1) backbone and promoter::*luc*::NosT fragments generated by *Eco*RI and *Hin*dIII digestion, a *Pvu*II digest of the four cloned constructs was successfully performed (data not shown). The *Pvu*II endonuclease cleaves pBluescript to yield two fragments of *circa* 2.4 kb and 0.5 kb. It does not cleave the promoter cassettes of *circa* 3 kb. The digestion products were purified and subjected to *Eco*RI and *Hin*dIII digestion. This allowed for cleavage of the promoter cassettes from pBluescript. Electrophoresis of the triple digested DNA revealed the presence of three fragments of the expected size. These were the 0.5 kb and 2.4 kb fragments corresponding to the digested pBluescript vector and a fragment of *circa* 3 kb corresponding to the promoter cassettes (data not shown). The *Eco*RI and *Hin*dIII double digestion of pTF101.1 was successful and resulted in linearised pTF101.1 with *Eco*RI and *Hin*dIII overhangs situated within the T-DNA region (data not shown).

The DNA fragments corresponding to the pTF101.1 vector (Appendix B.2) and promoter cassettes were successfully excised and purified. Specific cohesive end ligation reactions were set up and used to successfully transform competent *E. coli* DH5α cells.

Ten colonies for each construct were selected for screening. Six colonies contained XvPSap1D, nine contained XvPSap1E, eight contained XvPSap1F and seven contained XvPSap1G (Fig. 4.3). Seven colonies contained XvPSap2 (data not shown).

University of Cape Town

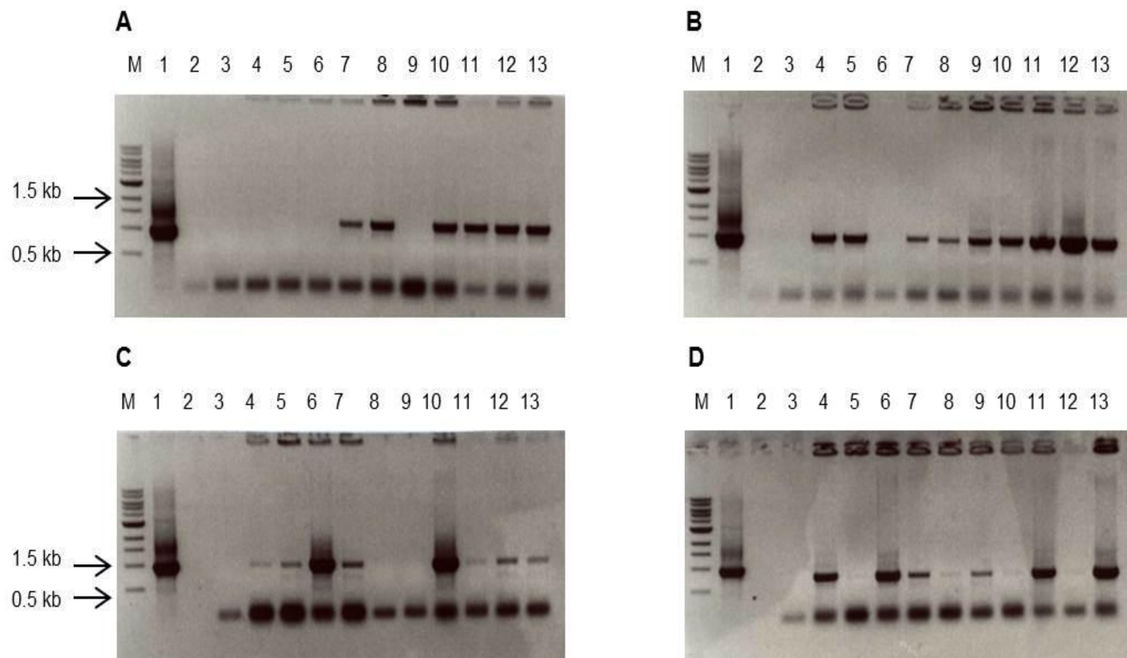


Figure 4.3: Amplification of promoter DNA cloned into pTF101.1 and transformed into competent Dh5α *E. coli* cells. **A**: XvPSap1D. **B**: XvPSap1E. **C**: XvPSap1F. **D**: XvPSap1G. Lane M: 100 bp DNA ladder. Lane 1: pBluescript::XvPSap1E::*luc*::NosT plasmid DNA (positive control). Lane 2: NTC. Lane 3: Dh5α *E. coli* (negative control). Lane 4 to 13: Selected transformed colonies.

University of Cape Town

Two colonies for each promoter were selected for plasmid isolation and *Eco*RI and *Hin*dIII restriction endonuclease analysis (Fig. 4.4). The digested pTF101.1::promoter::*luc*::NosT generated linear bands of *circa* 3 kb and 9 kb for each construct. The smaller DNA fragment was the expected size for the promoter::*luc*::NosT cassette and the larger fragment corresponded to the size of the pTF101.1 vector. Both the PCR analysis as well as digestion results indicated that each promoter::*luc*::NosT cassette had been successfully cloned into the pTF101.1 binary vector.

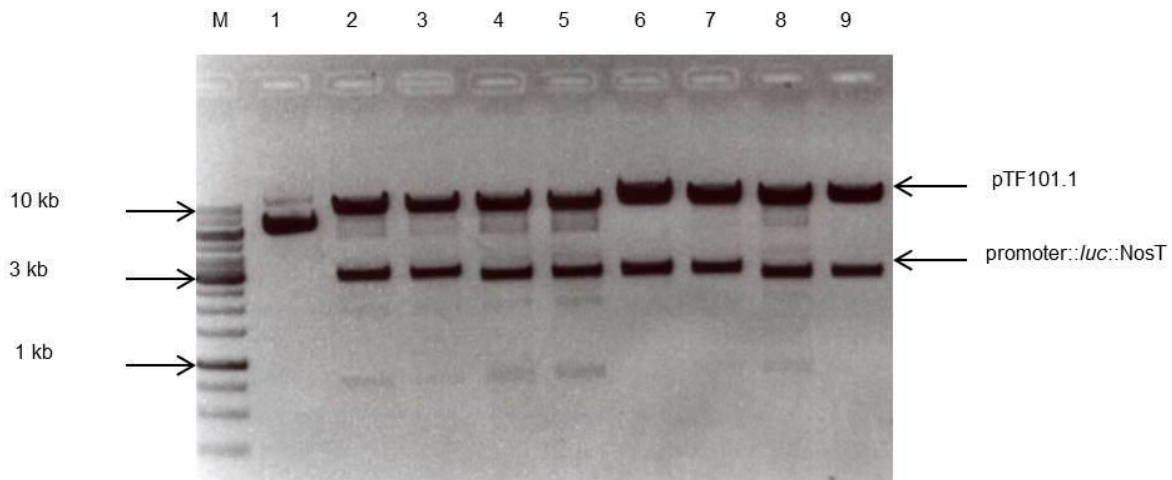


Figure 4.4: *Eco*RI and *Hin*dIII restriction endonuclease digests of pTF101.1::promoter::*luc*::NosT plasmid DNA. Lane M: 1 kb DNA ladder. Lane 1: Undigested pTF101.1::promoter::*luc*::NosT DNA. Lane 2 and 3: Digested pTF101.1::XvPSap1D::*luc*::NosT. Lane 4 and 5: pTF101.1::XvPSap1E::*luc*::NosT. Lane 6 and 7: Digested pTF101.1::XvPSap1F::*luc*::NosT. Lane 8 and 9: Digested pTF101.1::XvPSap1G::*luc*::NosT.

University of Cape Town

4.4 Bioinformatic analysis of promoters

In silico sequence analysis demonstrated that none of the promoters used in this study displayed any significant sequence homology to any other known plant promoters except to XvPSap1. The shortened XvPSap1 promoters shared 82.57% sequence homology with each other (Appendix C.9) and 49.75% homology with XvPSap2 (Appendix C.10). XvPSap2 shared 56.41% homology with XvPSap1 (Appendix C.11).

The promoter core and *cis*-acting regulatory elements were predicted by the plantCARE bioinformatics tool. Four identical TATA-boxes were predicted in each promoter within the 5'-300 bp region.

Various *cis*-acting elements were tentatively identified in the shortened XvPSap1 and unmodified XvPSap2 promoters. The *cis*-acting elements are indicated for XvPSap2 (Appendix G.1) and XvPSap1G (Appendix G.2). These identified elements were not shown for XvPSap1D and E. because, being larger in size, it is assumed that these

promoters contain the same, if not more, elements as XvPSap1G. These included elements involved in drought-, light- and low temperature- inducibility as well as in defense and biotic stress responsiveness (Lescot *et al*., 2002). *Cis*-acting elements involved in abscisic acid, salicyclic acid and gibberellin responsiveness were observed. Additionally, elements involved in meristem expression, auxin-responsiveness and circadian control were also noted (Lescot *et al*., 2002). No differences were seen in identified regulatory elements between XvPSap2 and the shortened promoters, indicating that a bioinformatic approach is not sufficient to demonstrate differences in functional activity. However, an experimental approach can be undertaken to determine differences between the promoters.

4.5. Transformation of *Agrobacterium* with pTF101.1 constructs

The pTF101.1 constructs were successfully transformed into competent *Agrobacterium* EHA101 cells. Fifteen colonies for each pTF101.1 construct were selected for screening and all contained the promoter fragment. Five colonies for each of the pTF101::promoter::*luc*::NosT constructs were selected for further analysis. The presence of the pTF101.1 constructs in EHA101 was successfully verified by *Eco*RI and *Hin*dIII digestion.

University of Cape Town

4.6. *Agrobacterium*-mediated transformation of *N. tabacum* with pTF101.1 vector constructs

4.6.1. Transformation of *N. tabacum*, tissue culture and growth of putative positive transformants

Two month old wild type *N. tabacum* plants were used for *Agrobacterium*-mediated transformation. After three weeks on shooting media, transformed leaf disks displayed minimal or no necrosis and remained green in colour. The emergence of shoots was clearly visible (Fig. 4.5**B**). When untransformed leaf disks were transferred to rooting media supplemented with 3 mg/l BASTA, total necrosis was visible with the disks exhibiting a brown colour (Appendix I.2). The transformation and tissue culture of *N. tabacum* with the promoter constructs are displayed in Fig. 4.5.

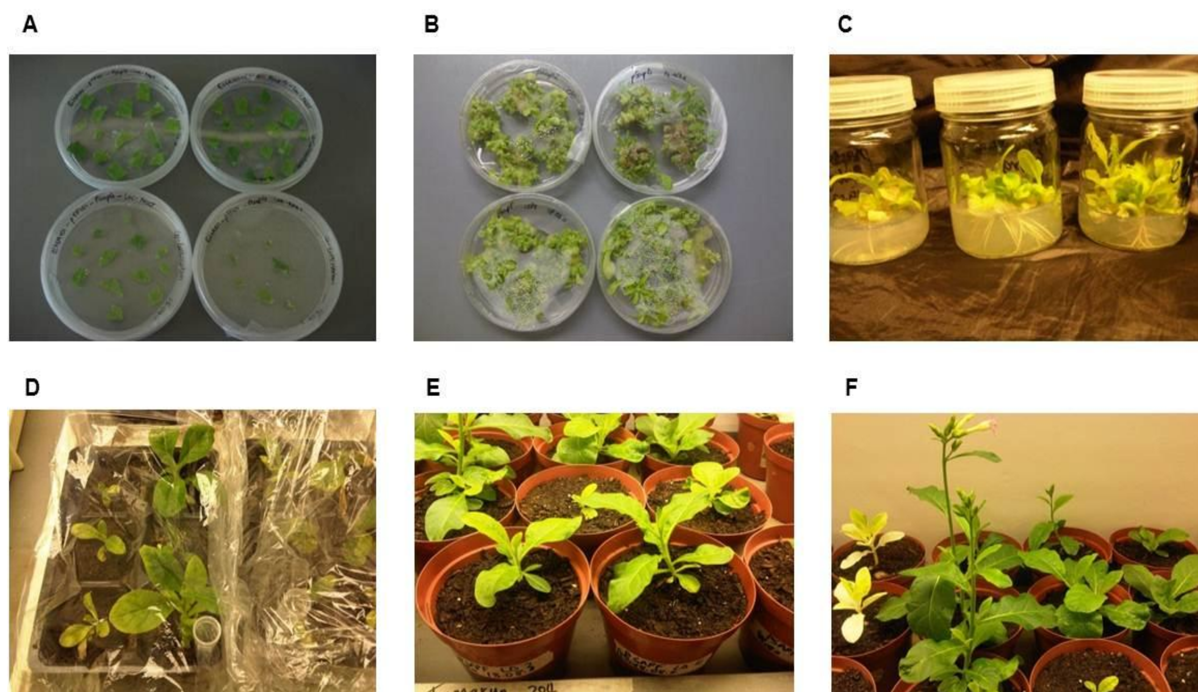


Figure 4.5: Transformation and tissue culture of *N. tabacum* with XvPSap1D. **A**: Infected leaf disks on co-cultivation media. **B**: Shoots emerging from transformed leaf discs with XvPSap1D after 3 weeks on shoot selection media with BASTA. **C**: Transformants forming roots on rooting media 5 weeks after infection. **D**: Putative transformants covered with cling wrap to minimise water loss during hardening. **E**: Putative transformants in potting soil after hardening. **F**: Flowers of putative transformants.

University of Cape Town

The putative XvPsap1D, F and G transformants did not reveal any unusual or abnormal phenotypic traits. In contrast, some of the putative XvPSap1E transformants displayed signs of dwarfism (Appendix J.2). These plants were discarded and not used for downstream experiments. The T0 seeds were collected for all plants. Overall, mature pods gave large amounts of seed. However, some XvPSap1G plants yielded pods that contained no seed. The observed absence in seed and dwarfism in the mentioned plants could be attributed to the random insertion of the promoter cassettes into the genome of the plants.

4.6.2. *Agrobacterium*-mediated transformation of *A. thaliana* with pTF101.1 vector constructs

Immature flowers of *A. thaliana* were successfully floral dipped with the transformed EHA101 *Agrobacterium* cells. For each construct, five plants were used. After six weeks, T0 seeds were collected.

4.7. Germination and BASTA screening of putative transgenic seed

4.7.1. Transgenic *N. tabacum*

Wild type seedlings displayed complete necrosis and failed to survive on the BASTA supplemented MS media. After screening the remaining plants for the presence of the promoter (see section 4.9.2), positively transformed plants were transferred to individual pots (Fig. 4.6).

Cape Town

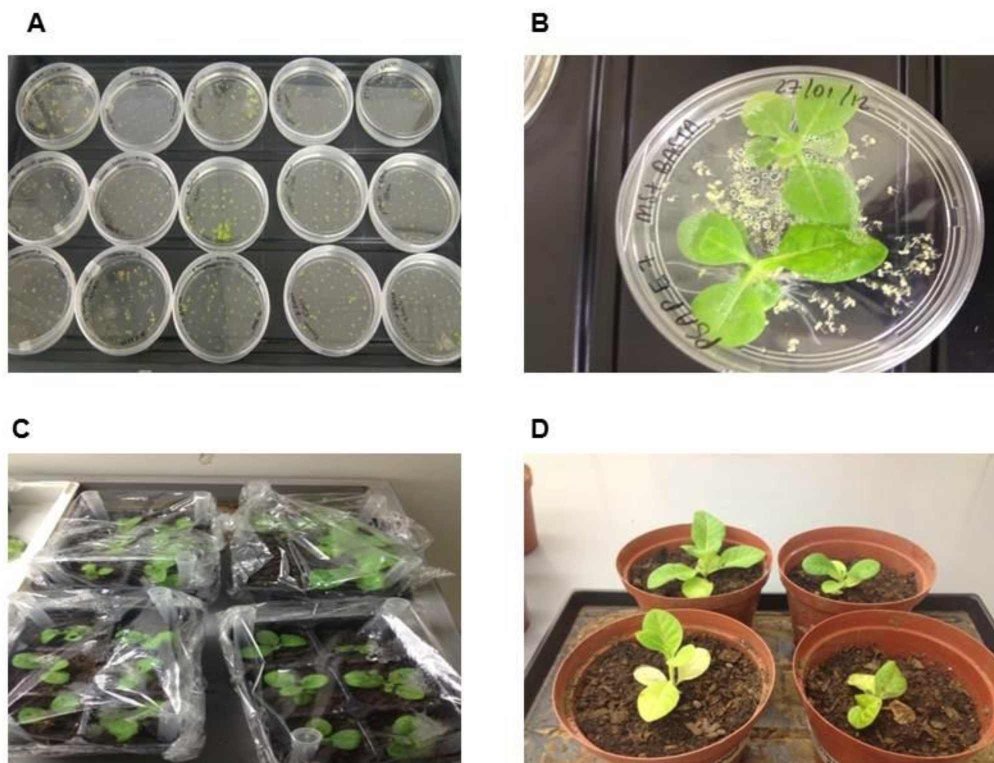


Figure 4.6: Germination of transformed *N. tabacum* seeds. **A**: Transformed tobacco seeds 4 days after plating. **B**: Transformed tobacco plants on MS media with BASTA. **C**: Hardening of putative positive plants. **D**: Putative positive plants in potting soil.

4.7.2. Transgenic *A. thaliana*

After two weeks, BASTA-resistant plants remained green in colour, while BASTA-sensitive plants turned yellow (Fig. 4.7). One week later, all of the BASTA-sensitive plants had died. The BASTA-resistant plants were transferred to trays containing new peet::vermiculite mix.

Figure 4.7: Germination of transformed *A. thaliana* seeds. **A**: Seeds 4 days post germination. **B**: BASTA treated plants. **C**: Putative positive plants surviving BASTA treatment.

University of Ca

4.8. Screening of putative transformants using PCR

4.8.1. Transgenic *N. tabacum*

For each genomic DNA isolation, an intense, high molecular weight band was visible, indicative of good, high quality DNA (Fig. 4.8).

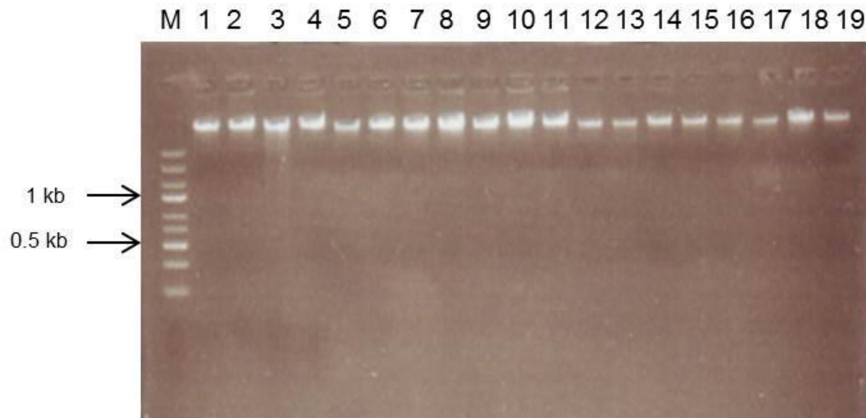


Figure 4.8: Genomic DNA of putative positive *N. tabacum* plants. Lane M: 1kb DNA ladder. Lanes 1 to 3: Isolated gDNA from 3 plants transformed with the XvPSap1D construct. Lanes 4 to 6: Isolated gDNA from 3 plants transformed with the XvPSap1E construct. Lanes 7 to 9: Isolated gDNA from 3 plants transformed with the XvPSap1F construct. Lanes 11 to 12: Isolated gDNA from 3 plants transformed with the XvPSap1G construct. Lanes 13 to 15: Isolated gDNA from 3 plants transformed with the XvPSap2 construct. Lanes 16 to 19: Isolated gDNA from 4 wild type plants.

For plants harbouring the promoter cassettes, the presence of the promoter and *luc* gene was determined by amplification of a DNA fragment of *circa* 3 kb using a promoter-specific forward and a *luc*-specific reverse primer. The *N. tabacum* plants transformed with the XvPSap1D, E, F and G constructs yielded two, one, zero and one transformation events, respectively. Since the transformation efficiency was relatively low, plants arising from a single transformation event were screened for the various promoter constructs (Fig. 4.9). In total, twenty one *N. tabacum* plants transformed with the XvPSap1D construct were identified, seven with XvPSap1E, nine with XvPSap1G. Figure 4.9 represents some of the identified positive transformants. No positive transformants were identified for XvPSap1F

Plants transformed with the XvPSap2 construct were screened for the presence of the promoter using the XvPSap2 promoter specific forward and reverse primer (Appendix D, Table D.4). In total, twelve plants with the XvPSap2 were identified. Figure 4.9D depicts four of the twelve plants.

University of Cape Town

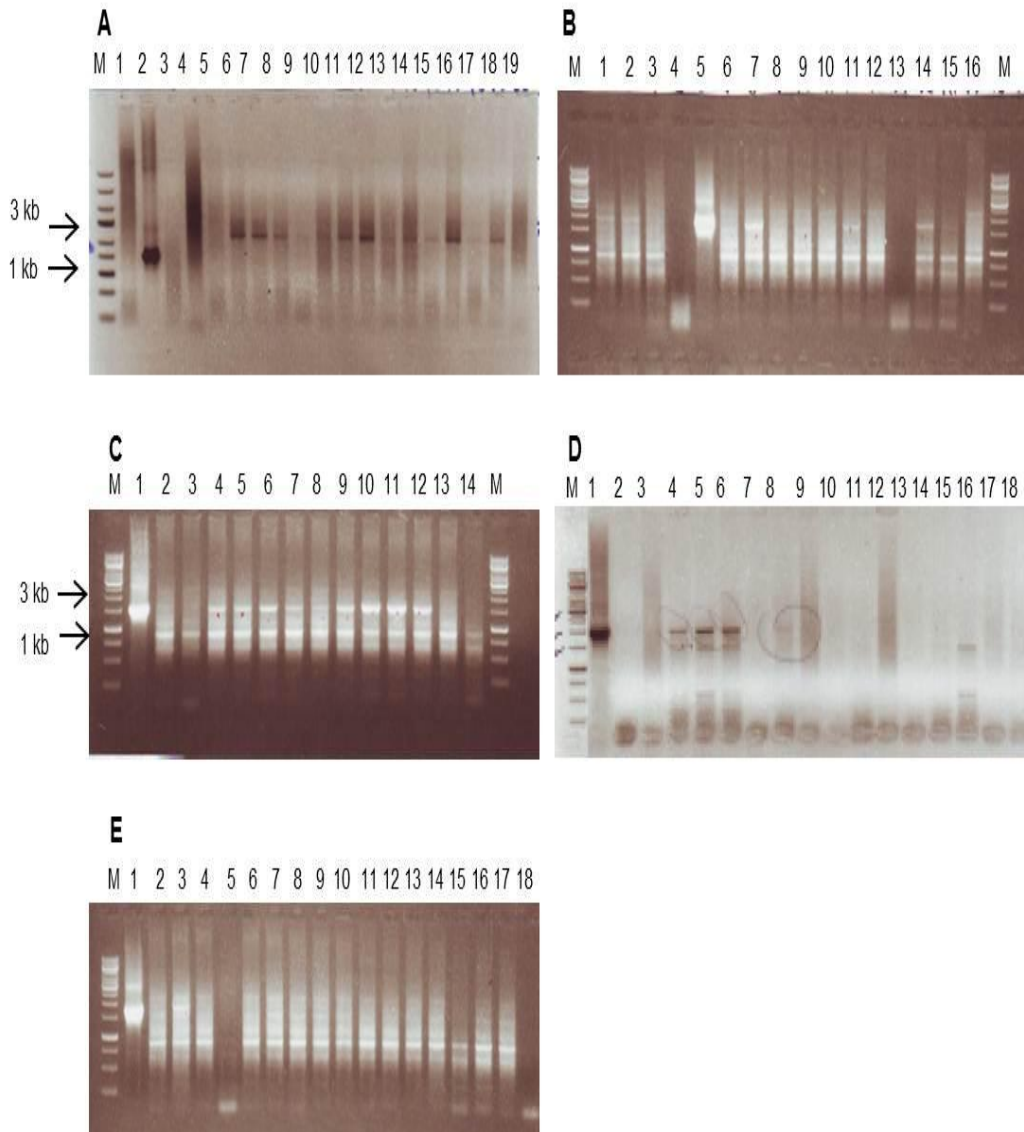


Figure 4.9: PCR screening of putative positive plants for the presence of *luc* and XvPSap1. **A**: XvPSap1D. Lane M: 100 bp DNA marker. Lane 1 and 5: Wild type tobacco genomic DNA. Lane 2: XvPSap1E construct plasmid DNA. Lane 4: NTC. Lanes 6 to 20: Genomic DNA from putative positive plants. **B**: XvPSap1E. Lane M and 20: 1 kb DNA ladder. Lane 2 and 3: Genomic DNA positive for XvPSap1D. Lane 3: Wild type tobacco genomic DNA. Lane 4: NTC. Lane 5: XvPSap1E construct plasmid DNA. Lanes 6 to 19: Genomic DNA from putative positive plants. **C**: XvPSap1 G. Lane M and 20: 1 kb DNA ladder. Lane 1: XvPSap1G construct plasmid DNA. Lane 2: Wild type tobacco DNA. Lanes 3 to 18: Genomic DNA from putative positive plants. **D**: XvPSap2. Lane M: 1 kb DNA ladder. Lane 2: XvPSap2 construct plasmid DNA. Lane 3: NTC. Lane 4: Wild type tobacco genomic DNA. Lanes 5 to 20: Genomic DNA from putative positive plants. **E**: XvPSap1F. Lane M: 1 kb ladder. Lane 1: XvPSap1F construct plasmid DNA. Lane 2: Wild type tobacco genomic DNA. Lane 3: Genomic DNA positive for XvPSap1D. Lanes 4 to 20: Genomic DNA from putative positive plants.

The transformation efficiency for each promoter construct was determined. An efficiency rate of 15%, 3.5%, 8% and 6% was calculated for XvPSap1D, E, G and XvPSap2, respectively.

4.8.2. Transgenic *A. thaliana*

Genomic DNA was successfully isolated from the leaves of T1 BASTA-resistant putative transformed *A. thaliana* plants. For each genomic DNA isolation, a high molecular weight band was visible (data not shown).

Screening of the T1 plants indicated that there were eight, three, zero, one and two positive transformants for XvPSap1D, XvPSap1E, XvPSap1F, XvPSap1G and XvPSap2, respectively. As with tobacco, no positive transformants were identified for XvPSap1F. Following screening of the T1 plants, positively transformed plants were dried and T1 seeds were collected, germinated and BASTA sprayed. Genomic DNA was isolated from the leaves of T2 plants displaying resistance to BASTA. The genomic DNA was of good quality (data not shown).

The isolated genomic DNA was screened for the presence of the promoter and *luc* gene by amplification of a DNA fragment of *circa* 2 kb using a promoter-specific forward and a *luc*-specific reverse primer. The *A. thaliana* T2 plants transformed with the XvPSap1D, E and G constructs yielded nine, two and one plants harbouring the promoter and luciferase gene, respectively (Fig. 4.10).

University of Cape Town

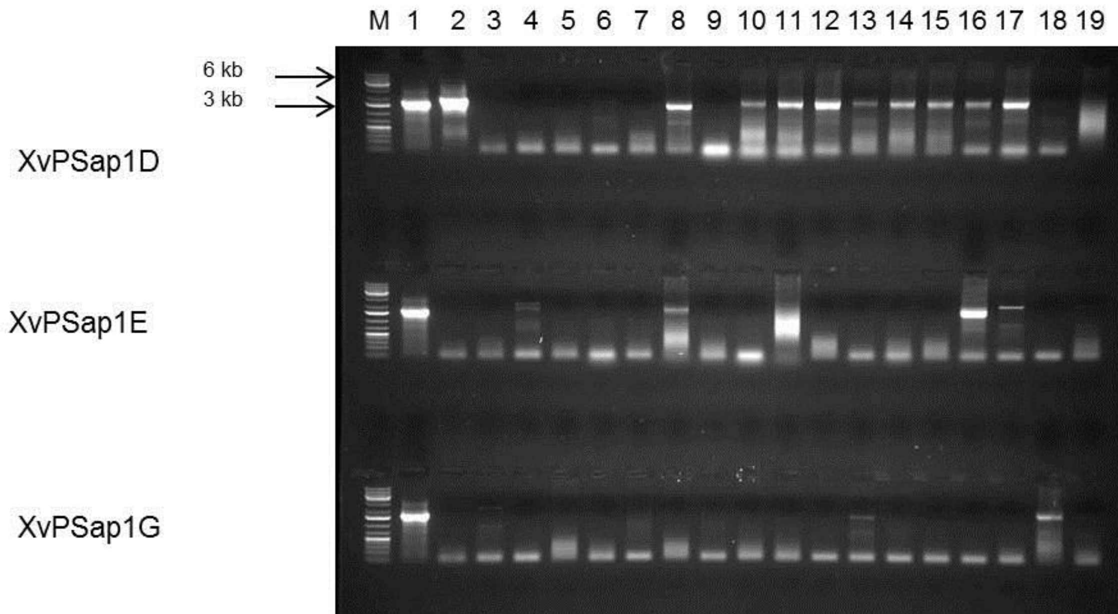


Figure 4.10: PCR screening of putative positive plants for the presence of *luc* and XvPSap1. XvPSap1D: Lane 1: pTF101.1: XvPSap1D::*luc*::NosT plasmid DNA. Lane 2: pTF101.1::XvPSap1D::*luc*::NosT genomic DNA. Lane 3: NTC. Lane 4: Wild type genomic DNA. Lane 5 to 19: Genomic DNA from putative positive plants. XvPSap1E: Lane 1: pTF101.1::XvPSap1D::*luc*::NosT plasmid DNA. Lane 2: Wild type genomic DNA. Lane 3 to 19: Genomic DNA from putative positive plants. XvPSap1G: 1 kb ladder. Lane 1: pTF101.1::XvPSap1D::*luc*::NosT plasmid DNA. Lane 2: Wild type genomic DNA. Lane 3 to 19: Genomic DNA from putative positive plants. Lane M: 1 kb DNA ladder.

University of Cape Town

Plants transformed with the XvPSap2 construct were screened for the presence of the promoter using the XvPSap2 promoter specific forward and luciferase specific reverse primer. Of the six T2 plants that were screened, only two contained the promoter and luciferase gene (Fig. 4.11).

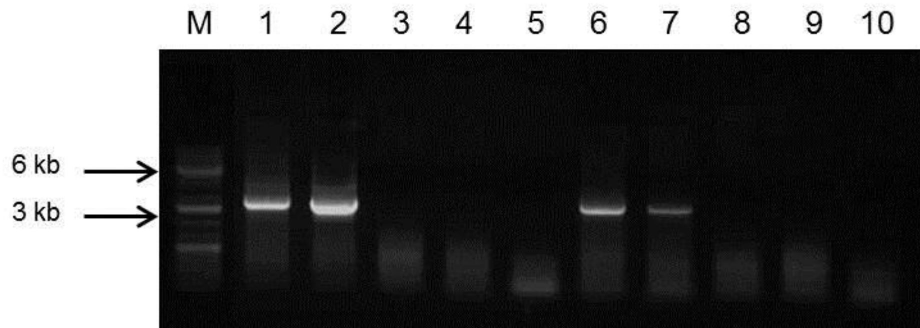


Figure 4.11: PCR screening of putative positive plants for the presence of *luc* and XvPSap2. Lane M: 1 kb ladder. Lane 1: pTF101.1::XvPSap2::*luc*::NosT plasmid DNA. Lane 2: pTF101.1::XvPSap2::*luc*::NosT genomic DNA. Lane 3: Wild type genomic DNA. Lane 4: NTC. Lanes 5 to 10: Genomic DNA from putative positive plants

University of Cape Town

An efficiency rate of 2% was achieved for XvPSap1D, whereas a 1% rate was achieved for the other promoter constructs. Since an insufficient number of positively transformed *A. thaliana* plants were obtained for XvPSap1E, G and XvPSap2 to perform meaningful downstream experiments, analysis of promoter activity was only assessed in *N. tabacum*.

4.9. Dehydration treatment of transgenic *N. tabacum* plants

To determine whether the various promoters were functionally active in response to limited water conditions, transgenic plants were subjected to a six day dehydration treatment.

4.9.1. Morphological changes during dehydration stress

Morphological changes such as changes in leaf colour and textures were observed during the dehydration stress treatment (Fig. 4.12)**.** On days 0 and 1, the leaves were green and turgid for all plants. The leaves of treated plants began folding downward on days 2 to 3 and by day 4 to 5, they were completely folded and flaccid. In contrast, the leaves of the hydrated plants remained green and turgid throughout the six day period.

There were no notable differences in the morphological changes between the leaves of the transgenic and wild type plants. Both sets of plants appeared to display similar symptoms over the course of the treatment.

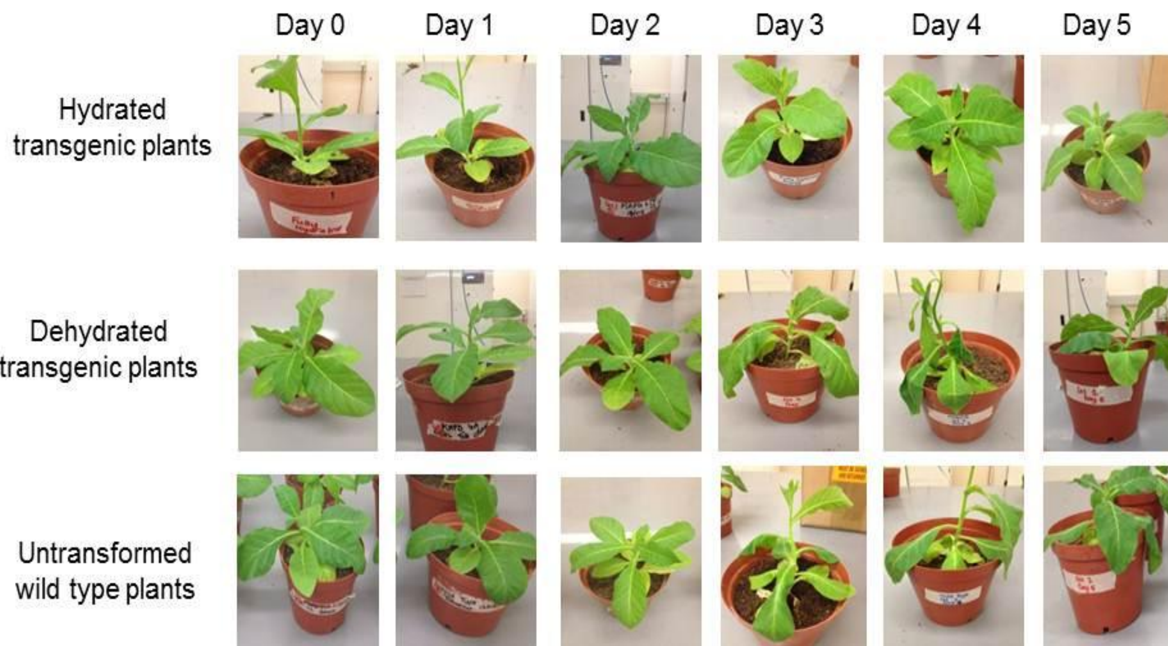


Figure 4.12: Transgenic and wild type *N. tabacum* plants during a six day dehydration treatment. The top and middle rows depict hydrated and dehydrated transgenic plants, respectively. The bottom row depicts dehydrated wild type plants. Hydrated transgenic plants were watered regularly. Transgenic plants contained XvPsap1D or G

University of Cape Town

4.9.2. Determination of soil water content (SWC)

The SWC was determined to identify whether the rate of water loss was similar across all pots. More specifically is was to determine whether the soil in pots containing the plants for RWC analysis and the pots containing plants for promoter activity analysis were not drying significantly differently. The SWC in the pots containing plants undergoing dehydration stress treatment decreased on average from 33% to 5% (Fig. 4.13). As with RWC measurements, a similar trend was observed between plants with respect to the soil drying. The SWC of the pots harbouring transgenic plants that were watered throughout the dehydration period fluctuated between 33% and 26%.

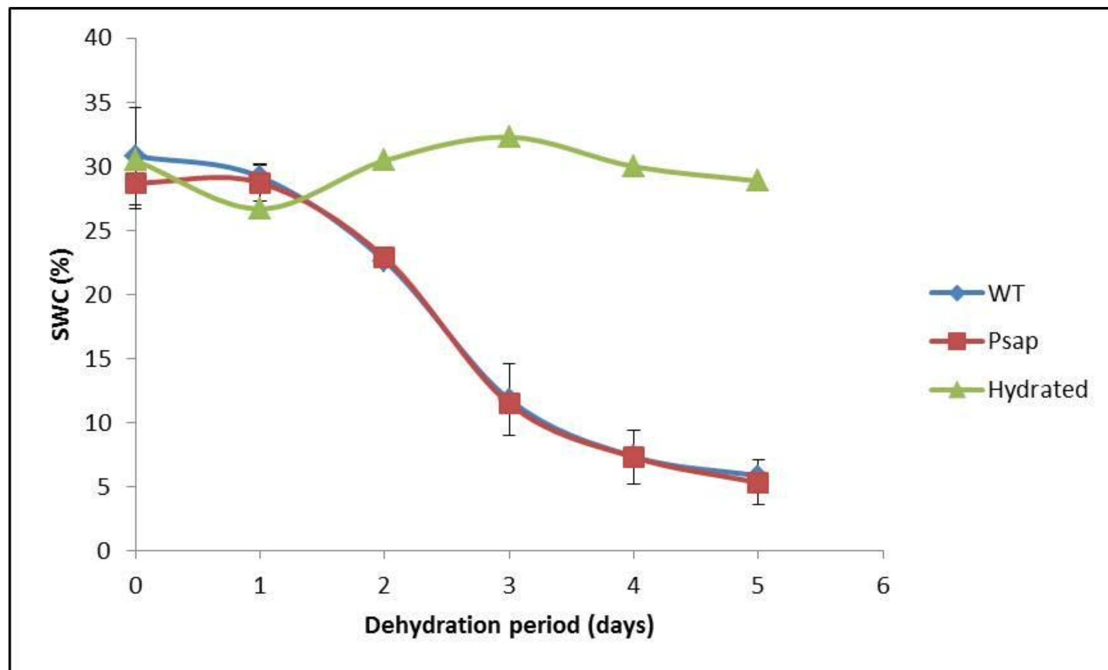


Figure 4.13: Soil water content. Each sample point indicates the average SWC calculated for each pot containing a plant. Error bars represent the standard deviation within pots.

University of Cape Town

4.9.3. Determination of relative water content (RWC) for each plant

The four leaves selected for RWC determination were situated towards the middle of the plant. This was because the leaves situated closer to the apex of the plant were smaller, young leaves while the leaves situated closer to the bottom of the plant were larger and older. RWC determination for one leaf is demonstrated in fig. 4.14.

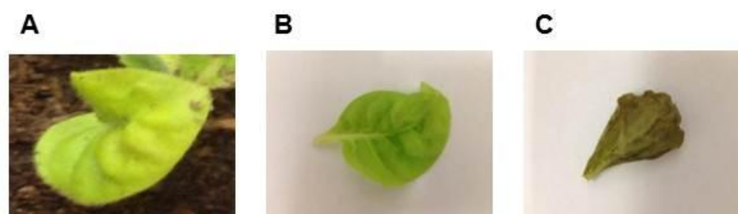


Figure 4.14: Determination of RWC for one leaf. **A**. Leaf excised from a whole plant (fresh weight). **B**. Leaf after incubation for 24 h in distilled water (full turgor weight). **C**. Leaf after incubation for 48 h at 72ºC (dry weight).

The RWC measurements obtained from both transgenic and wild type *N. tabacum* plants decreased from *circa* 95% to 53%. The rate of dehydration for each plant was similar, as demonstrated by the RWC, during the treatment (Fig. 4.15). Day 0 represents the first day after watering. On day 5 of the dehydration treatment, the transgenic plant had a lower RWC than the wild type plants. The RWC values of the hydrated transgenic plants fluctuated between 80% and 90% during the six day dehydration period.

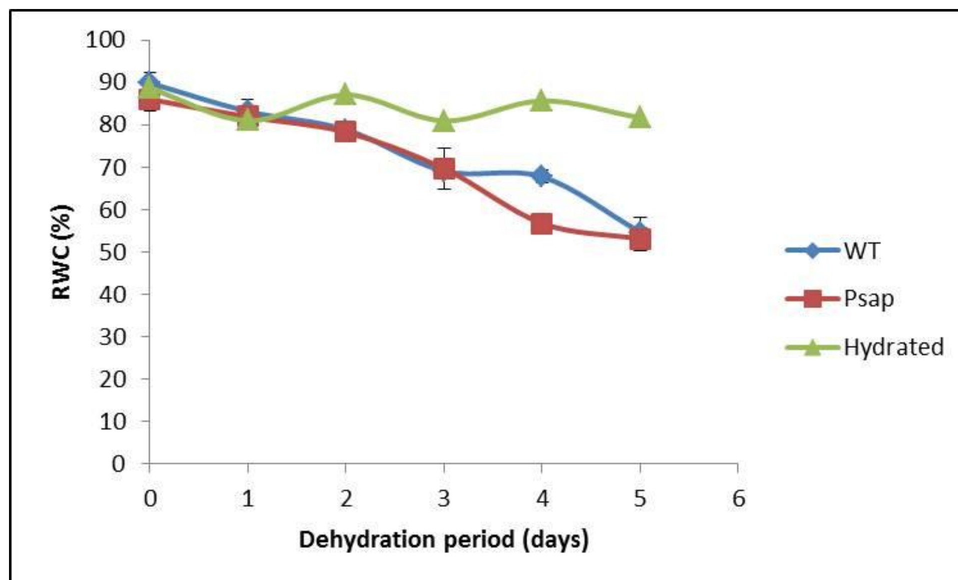


Figure 4.15: Relative water content of *N. tabacum* plants. Each sample point indicates the average % RWC calculated for **each of** the four leaves of each plant. Error bars represent the standard deviation between leaves.

4.10. Analysis of promoter activity in intact leaves

The intact leaves of each plant were assayed for bioluminescence, due to luciferase activity, over a six day dehydration treatment. One plant for each promoter construct was assayed for bioluminescence per day. As expected, no bioluminescence was detected in the wild type plants as these do not contain the luciferase gene (Appendix I.1). The expression of luciferase was visible as bioluminescence in the leaves of T1 plants transformed with XvPsap1D (Figs 4.16 and 4.17), E (Fig. 4.18) and G (4.19)This was visible as blue (minimal activity), green (medium activity) and red (high activity) colouring within each leaf. No luciferase activity was observed for plants transformed with XvPSap2 (Fig. 4.20).

Town

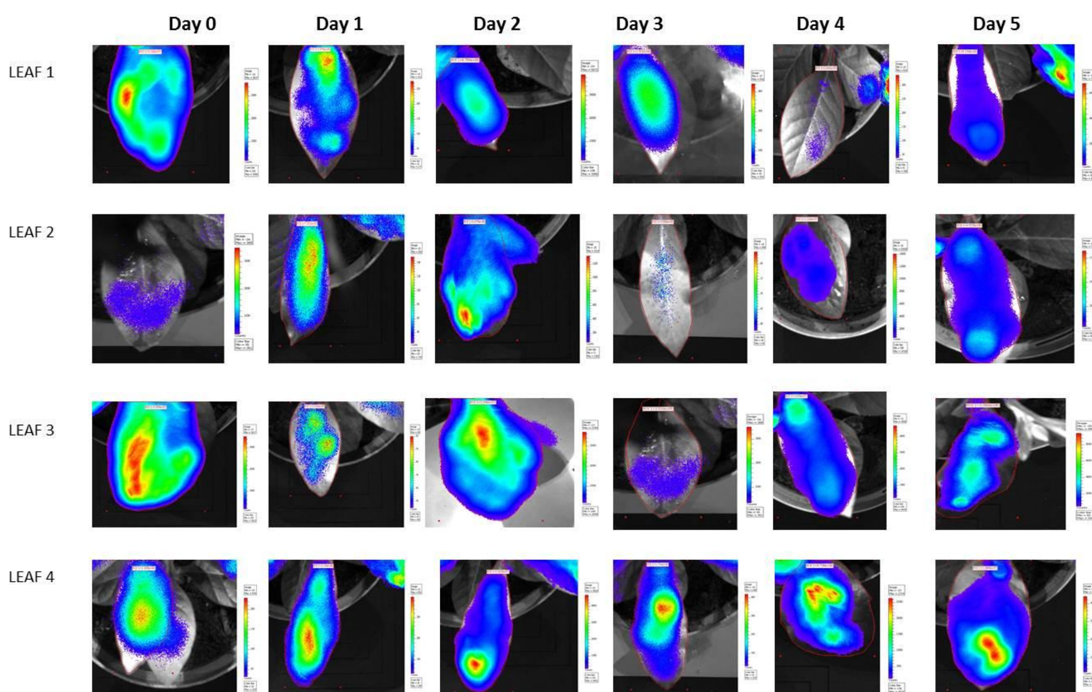


Figure 4.16: Determination of luciferase activity in hydrated *N. tabacum* plants transformed with XvPSap1D over a six day dehydration treatment.

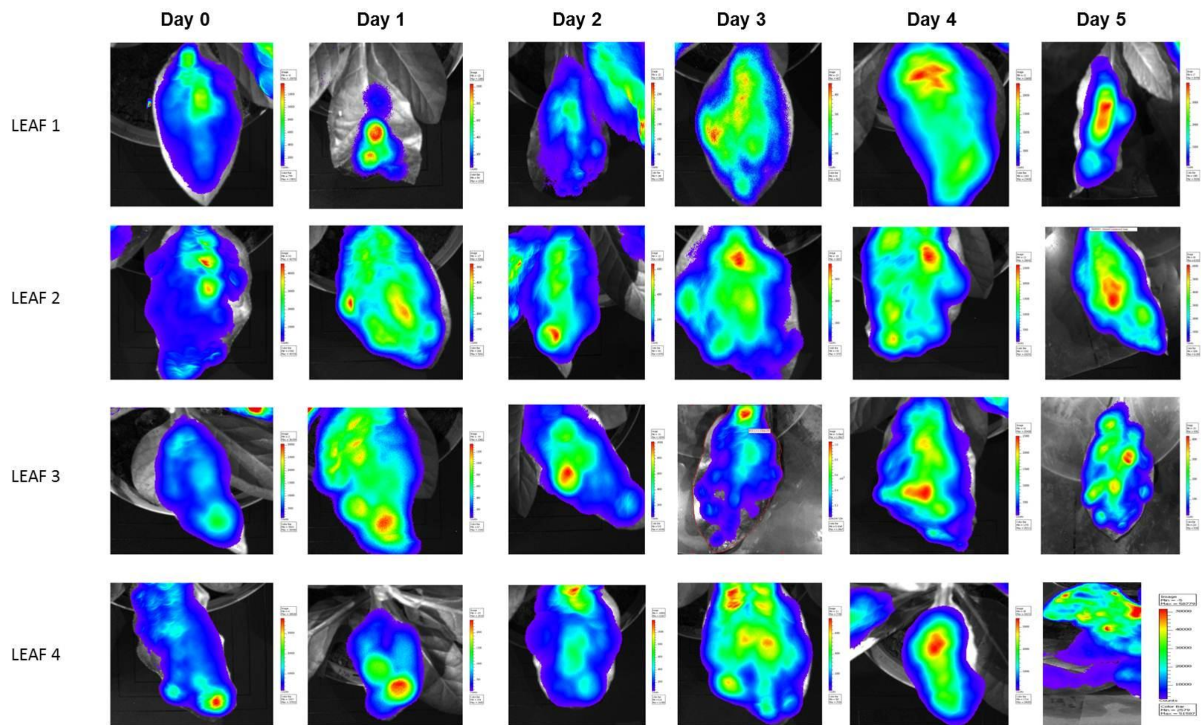


Figure 4.17: Determination of luciferase activity in *N. tabacum* plants transformed with XvPSap1D over a six day dehydration treatment.

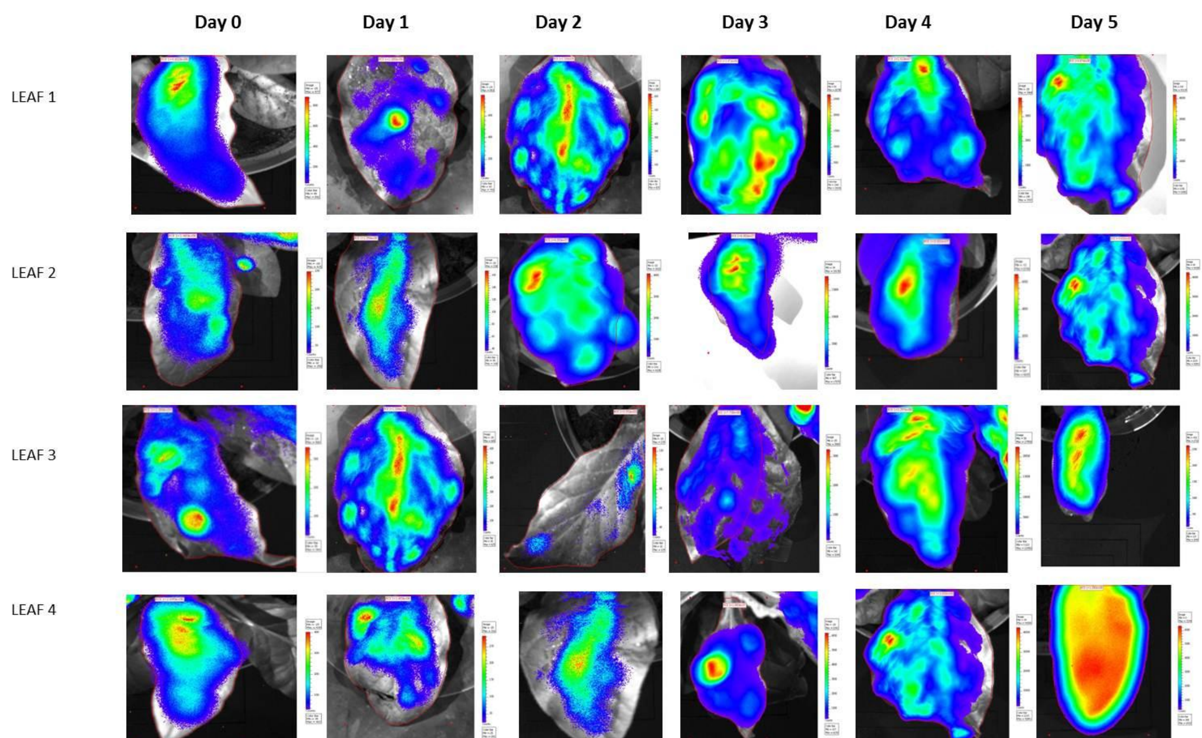


Figure 4.18: Determination of luciferase activity in *N. tabacum* plants transformed with XvPSap1E over a six day dehydration treatment.

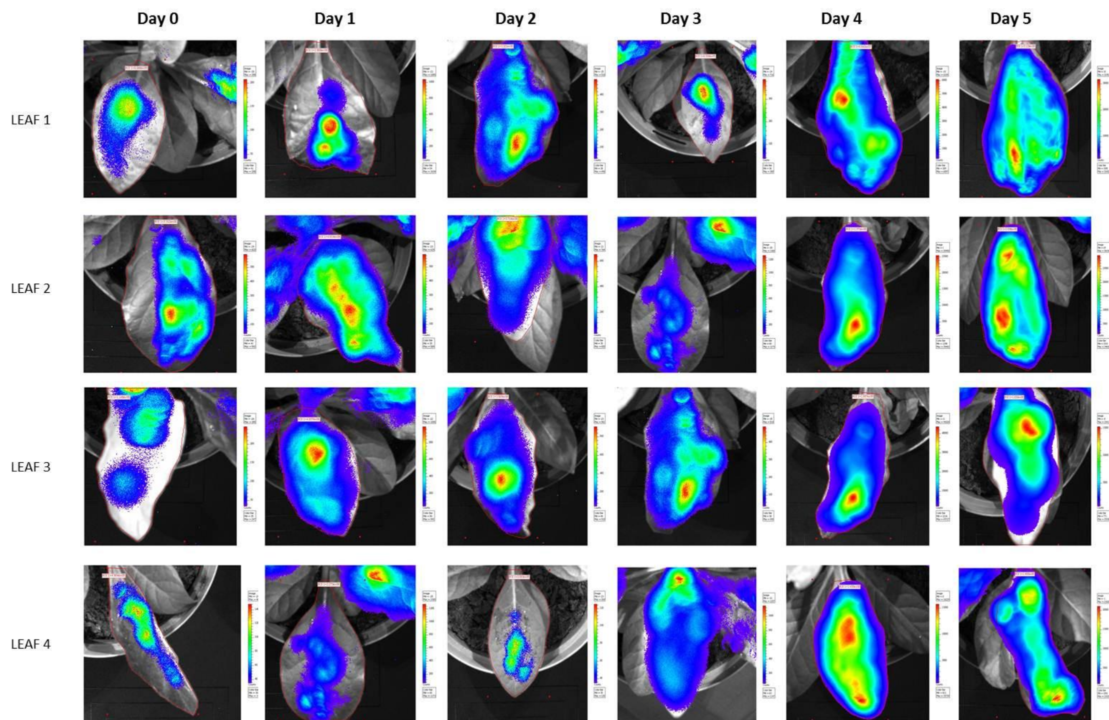


Figure 4.19: Determination of luciferase activity in *N. tabacum* plants transformed with XvPSap1G over a six day dehydration treatment.

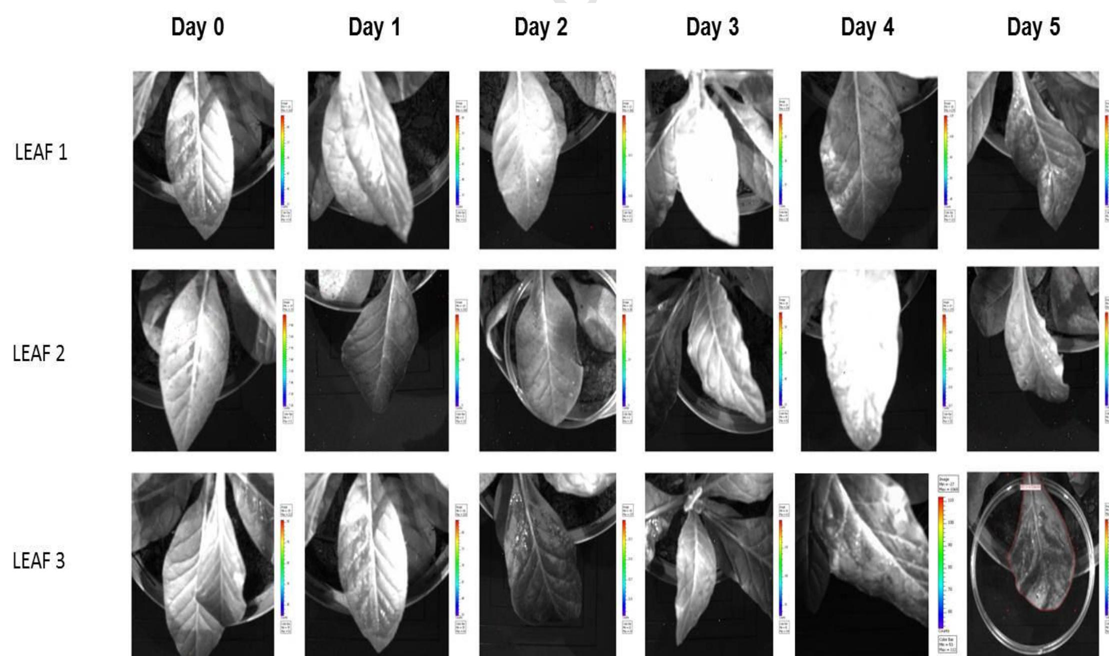


Figure 4.20: Determination of luciferase activity in *N. tabacum* plants transformed with XvPSap2 over a six day dehydration treatment.

of Cape

To quantify the levels of luciferase expression in each leaf, the region of interest (ROI) values were calculated. These values are a measure of the photons or counts that are detected due to the breakdown of luciferin to oxyluciferin. Overall, the number of counts increased over the dehydration period for plants transformed with XvPSap1D, E and G constructs (Fig. 4.21**B**, **C** and **D**, respectively). Since no biological repeats were included in the study, due to the shortage of starting material a full analysis of the data was not possible. However, the trends in promoter activity over the dehydration period were clearly established. Due to the above limitation, standard deviations could not be determined between biological repeats. Instead, standard deviations between readings across leaves in a single plant were determined.

In plants that remained hydrated over the six day period, luciferase activity remained constant (Fig. 4.21**A**). In plants transformed with XvPSap1D, E and G, there were similar levels of activity over the course of the treatment (Fig. 4.21**B**, **C** and **D**, respectively). For XvPSap1D, induction of activity began early (day 2) and was maintained over the rest of the dehydration period. For XvPSap1E and G, it began later (day 5 and 4, respectively). No induction was visible in plants transformed with XvPSap2 (Fig. 4.21, **E**). As expected, no induction was observed in wild type plants (Fig. 4.21**F**).

University of Cape Town

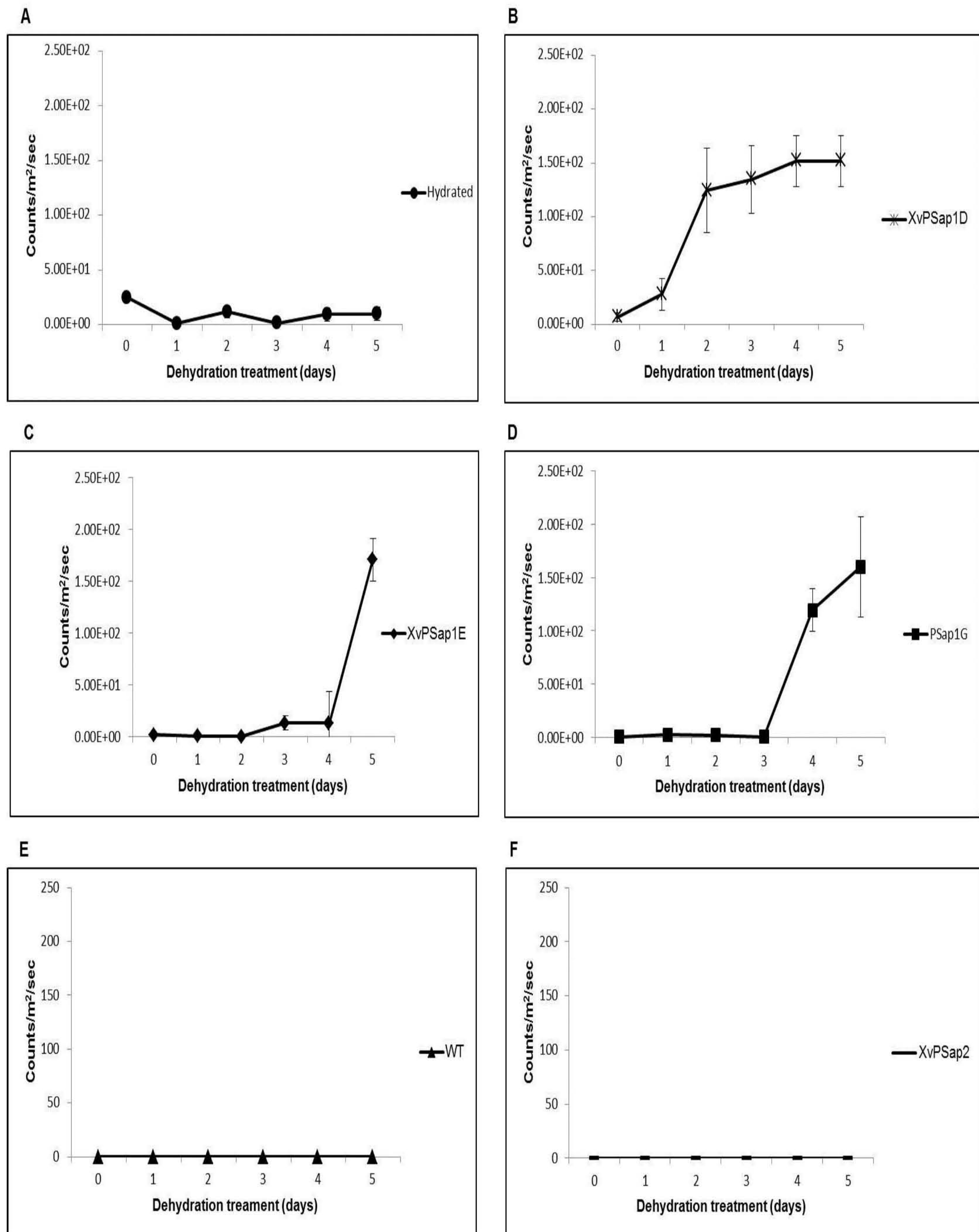


Figure 4.21: Luciferase activity in transgenic and wild type *N. tabacum* plants over a six day dehydration treatment. **A**: Hydrated transgenic plants. **B**: Transgenic plants transformed with XvPSap1D. **C**: Transgenic plants transformed with XvPSap1E. **D**: Transgenic plants transformed with XvPSap1G. **E**: Wild type plants. **F**: Transgenic plants transformed with XvPSap2. Error bars indicate the standard deviation within a single biological plant at each RWC value.

4.11. Analysis of promoter activity by quantitative real-time PCR

In order to quantify the *luc* gene expression levels, quantitative real time PCR was performed. This quantification was performed to confirm the results obtained from the promoter analysis in intact leaves.

4.11.1 RNA isolation

The concentration of RNA extracted for each isolation ranged from 89 to 543 ng/µl. Minimal to no degradation was observed when RNA was electrophoresed on a 1.2% non-denaturing agarose/EtBr gel (Fig. 4.22). The 28S and 18S ribosomal RNA (rRNA) bands were clearly visible. Bands of lower molecular weight were also observed, corresponding to the 5S rRNA and chloroplast RNA.

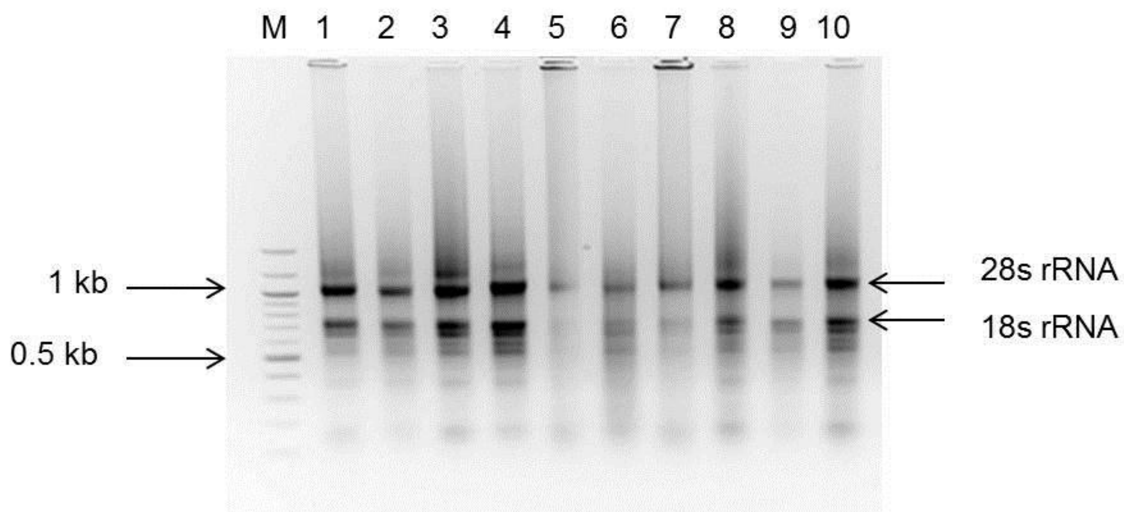


Figure 4.22: Purified RNA from transgenic *N. tabacum*. Lane M: 100 bp DNA ladder. Lanes 1 to 10: Isolated *N. tabacum* RNA.

4.11.2. Analysis of RNA using the Bioanalyzer software

The Agilent Expert software analyses RNA by assigning an RNA Integrity Number (RIN) to each sample. The RIN number is calculated based on the amount of signal found between the 5S and 18s rRNA and between the 18S and 28S rRNA bands. A RIN number of 10 indicates that the RNA is pure and that only 5S, 18S and 28S

rRNA peaks are present. However, isolation of RNA from plant tissue also includes chloroplast RNA. This decreases the RIN value, but does not imply that the RNA is of poor quality. However purified RNA with a RIN number below 5.3 is generally considered to be of poor quality and not suitable for analysis.

The Agilent 2100 Expert software generates an electropherogram with peaks corresponding to the different RNA species. The 5S, 18S and 28S rRNA bands are clearly visible as peaks (Fig. 4.23). No genomic DNA contamination is present as that would be visible as a peak downstream of the 28S rRNA peak. The presence of chloroplast RNA is visible as peaks situated between the 5S and 18S rRNA. A post 28S rRNA hump is visible, corresponding to undenatured RNA.

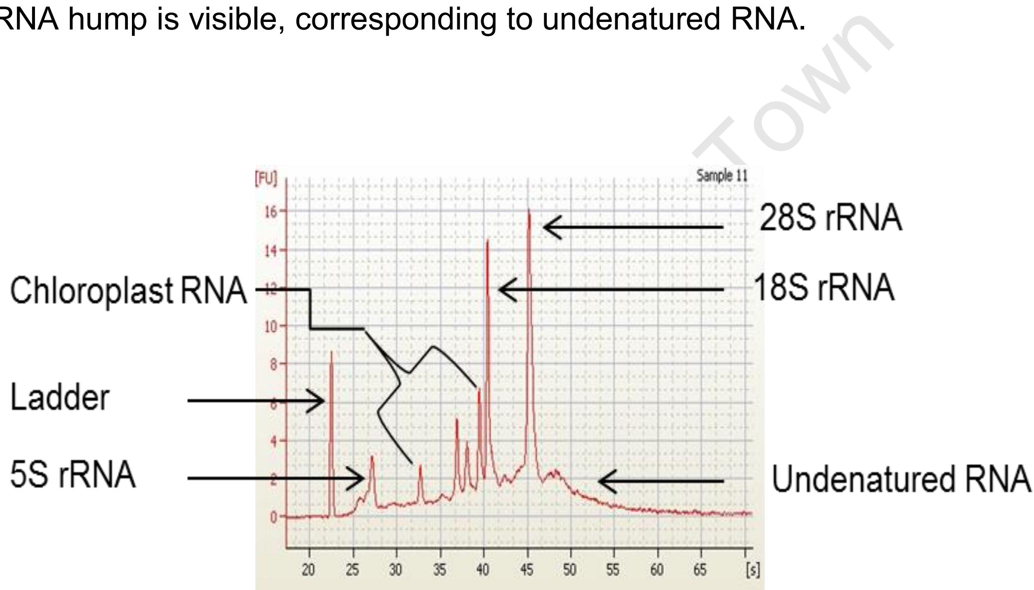


Figure 4.23: Electropherogram of RNA isolated from XvPSap1E generated using the Agilent 2100 Expert software

The RNA extracted for each isolation was analysed using the Agilent 2100 Expert software. The RIN values ranged from 4.6 to 8.8 and the concentrations from 100 ng/μl to 732 ng/μl. The electropherograms of RNA isolated from dehydration treated wild type and transgenic plants containing XvPSap1D, E, G as well as hydrated plants (Appendix H.1) displayed all of the previously mentioned RNA species. RNA of either poor quality or low concentration was discarded.

4.11.3. cDNA synthesis

The cDNA synthesis was successful as evidenced by a DNA smear following electrophoresis (data not shown).

4.11.4. Primer optimisation and reference gene analysis

A single primer set was chosen for each amplimer based on analysis of product size and whether a single peak in the melt curve had been observed. A standard curve for each potential reference gene and *luc* gene was produced (Appendix H.2).The efficiencies, R-values and R^2-values for the potential reference genes are listed in Table 4.1.

Table 4.1: Efficiencies, R-values and R^2-values of the potential reference genes and *luc* gene.

Gene	Efficiency	R-value	R^2-value
18S rRNA	0.89	0.9991	0.9982
EF-1α	0.82	0.9994	0.9988
L25 ribosomal protein	0.79	0.9992	0.9984
luc	0.99	0.9953	0.9905

University of Cape Town

Quantitative real-time PCR was performed on cDNA isolated from dehydration treated plants. The generated standard curve for each reference gene was imported and its stability of the reference gene was assessed using geNorm and NormFinder (Fig. 4.24). geNorm is based on the assumption that the expression ratio between two genes should be the same if both are stably expressed (Vandesompele *et al*., 2002). The relative stability (M) of each gene is defined as the mean pairwise variation of the gene with the other reference genes. The lower the M value, the more stable the reference gene. geNorm identified both L25 ribosomal protein and *EF-1α* with an M value of 0.037785 as being the best reference genes (Fig. 4.24**A**).

NormFinder assesses the stability of the reference genes by comparing the variation between the genes (Andersen *et al.*, 2004). The reference genes that are most stable are those with the lowest variation. Normfinder identified *EF-1α* as the most stable reference gene followed by L25 ribosomal protein (Fig. 4.24**B**). Both programmes identified 18S rRNA as the least stable of the three.

Based on both geNorm and NormFinder analysis, EF-1α and L25 ribosomal protein were chosen as reference genes. It should be noted that Schmidt and Delaney (2010) also concluded that these genes were the most suitable in a *N. tabacum* dehydration stress treatment analysis.

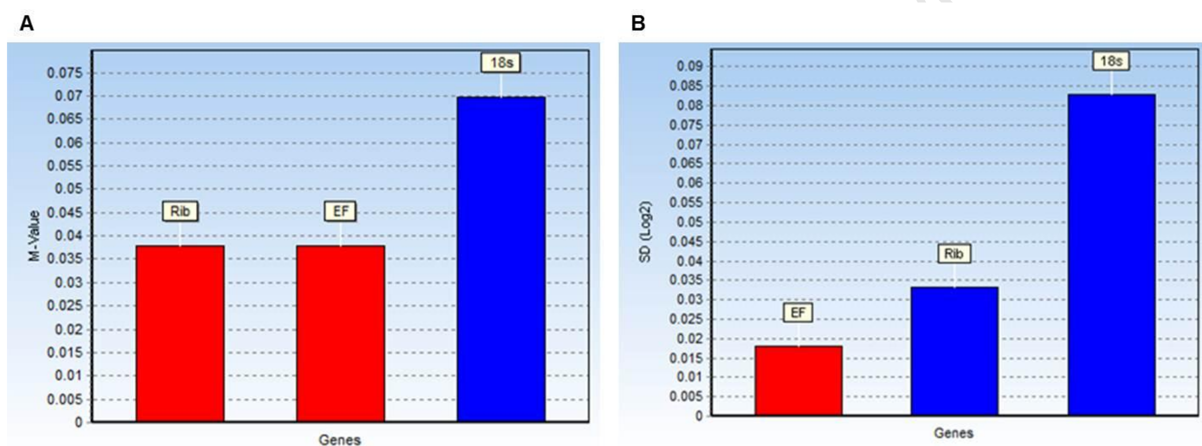


Figure 4.24: Relative stability ranking of reference genes 18S rRNA (18s), L25 ribosomal protein (Rib) and *EF-1α* (EF). **A**: geNorm. **B**: NormFinder. Most suitable reference genes are indicated by red bars.

4.11.5. qRT-PCR

The acquisition and purification of RNA is the first step in quantitative real time PCR and therefore RNA needs to be of high quality. Furthermore, the RNA should be free of genomic DNA, particularly if the target is a gene that lacks introns. Since the target gene (*luc*) does not contain any introns, an RNA control reaction was included in every real time run to demonstrate the absence of genomic DNA and subsequently confirm that any detected *luc* mRNA was not a product of genomic DNA contamination but rather of the cDNA template. In all of the real time runs, no *luc* mRNA was detected as a product of genomic DNA. The same result was obtained

for the reference gene. No products were detected in the no template control reactions as well.

As with the analysis of luciferase activity in intact leaves, promoter analysis was determined in only one plant by measuring *luc* mRNA transcripts at each time point. A consequence is that a full analysis of the data was not possible. However, qRT-PCR was used as a tool to assess the promoter activity in T1 plants and not to determine absolute activity. The overall trend of each promoter thus was assessed. As no repeats were included in the study, standard deviation between readings was determined across leaves in a single plant.

Overall, the *luc* mRNA transcripts increased over the dehydration period for plants transformed with XvPSap1D, E and G and all resulted in similar levels of induction (Fig. 4.25**B**, **C** and **D**, respectively). In plants that remained hydrated over the six day period, the *luc* mRNA transcripts remained constant (Fig. 4.25**A**). For XvPSap1E, induction began early (day 2 and 3) and was maintained over the rest of the dehydration period. For XvPSap1D, an initial decrease in *luc* mRNA transcripts was observed before induction began on day 1. Transcript levels decreased on day 3, before increasing again towards the end of the dehydration period. For XvPSap1G, induction began on day 0 and remained constant until day 2, after which a decrease in transcript levels was seen before an increase in *luc* transcript levels was noted towards the end of the dehydration period. In plants transformed with XvPSap2, the *luc* mRNA transcripts appeared to decrease steadily over the dehydration time course (Fig. 4.25**E**). This could explain why no luciferase activity was observed in the intact leaves of these plants. As expected, no induction was observed in wild type plants (Fig. 4.25**F**).

Although analysis was only determined in one biological plant at each time point, an overall trend was clearly observed. All of the shortened promoters were able to drive expression of the *luc* mRNA transcripts. Over the dehydration period, the levels of those transcripts increased to a similar level.

University of Cape Town

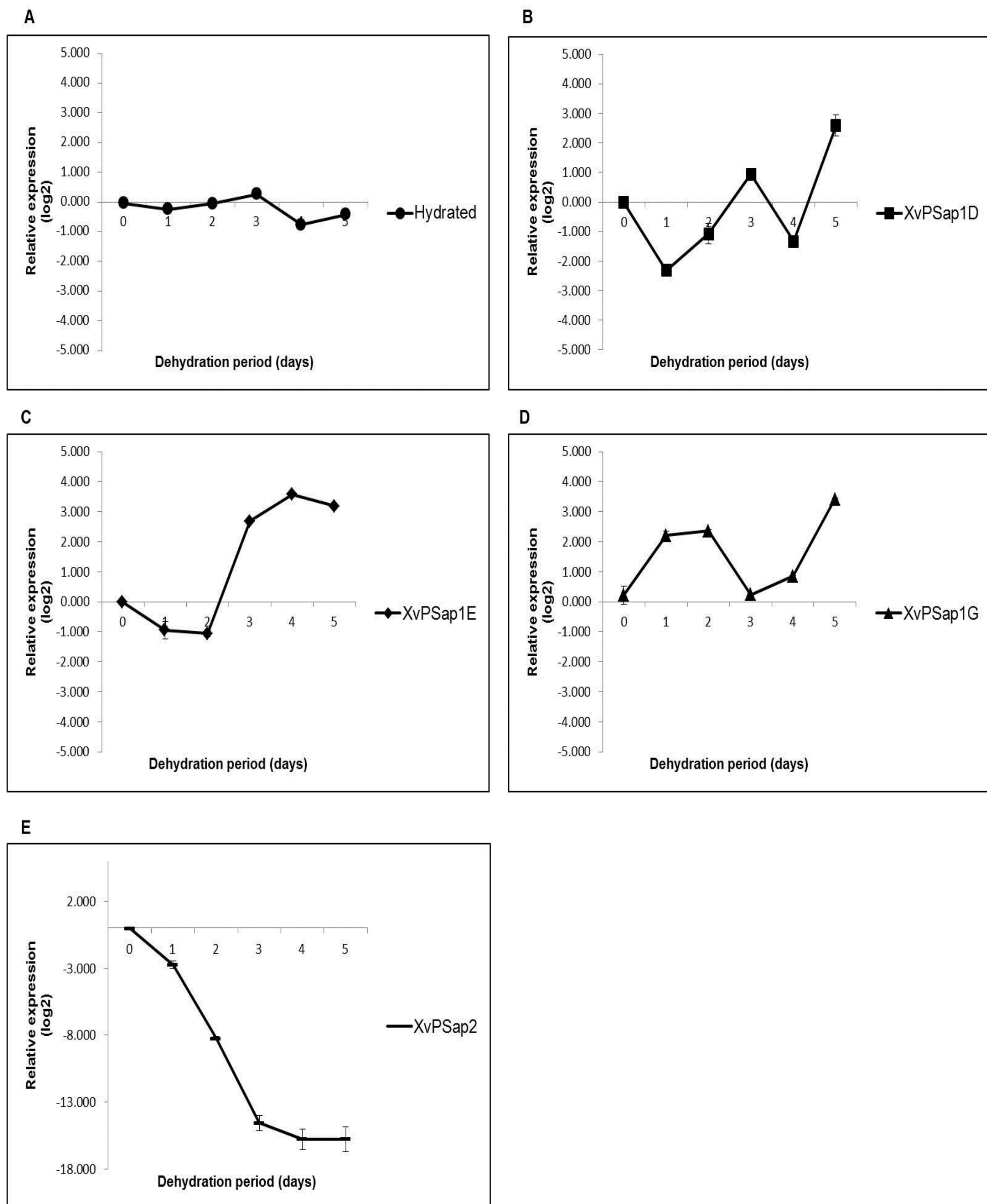


Figure 4.25: Relative gene expression levels of *luc* during the dehydration treatment of transgenic *N. tabacum*. Data points represent the mean and standard deviation of relative expression levels of *luc* compared to L25 ribosomal protein and *EF-1α* calculated using GeneX software. **A**: Hydrated transgenic plants. **B**: Transgenic plants transformed with XvPSap1D. **C**: Transgenic plants transformed with XvPSap1E. **D**: Transgenic plants transformed with XvPSap1G. **E**: Transgenic plants transformed with XvPSaP2. Error bars represent the standard deviation of four technical repeats within one biological plant at each time point.

5. Discussion

This study set out to determine whether shortened XvPSap1 (*circa* 50% reduction in size) and the full length XvPSap2 promoters would be functionally active under dehydration stress. To generate the shortened promoters, truncations on the full length XvPSap1 using DNA mutagenesis were performed resulting in four shorter promoter fragments (XvPSap1D, E, F and G). These promoters contained the 5'- and 3'-ends of the XvPSap1 promoter with the internal region removed.

The activity of the *luc* gene driven by the three shortened XvPSap1 (XvPSap1D, E and G) and unmodified XvPSap2 promoters was assessed in T1 *N. tabacum* plants over a dehydration period of six days. Promoter activity could not be determined in *A. thaliana* as insufficient positively transformed plants were obtained. The results generated in this study indicate that these promoters are active under dehydration stress in *N. tabacum*. The shortened XvPSap1 promoters displayed stress inducibility, while the XvPSap2 promoter was stress repressible. These findings answer the main objective of this study.

No positive *N. tabacum* plants containing XvPSap1F were identified. The same observation was made with transgenic *A. thaliana*. Since this occurrence was observed in two separate plant species, it suggests that the XvPSap1F promoter construct may be deleterious to the plant, possibly by driving very high expression of *luc* in a constitutive fashion. It may also be possible that the sites of integration are in a region of the genomic DNA that impairs germination or is in some other way deleterious to the plant. The former reasoning appears more plausible as the chance of the latter occurring in two separate plant species over multiple transformation events is unlikely, but not impossible.

5.1. Deletion mutagenesis of XvPSap1

This study stems from a previous study undertaken by Okoth (2009) wherein it was demonstrated that the full length XvPSap1 promoter was functionally active under dehydration stress. However, when two 5' promoter deletions were performed to

University of Cape Town

generate shorter XvPSap1 promoter fragments, they were able to drive expression of *luc* under dehydration stress but to a lesser extent than the full length promoter. Thus it was hypothesised in this study that the DNA elements required for optimal functional activity of the XvPSap1 promoter under dehydration stress might be present in the 5'-region of the promoter. Promoter deletions were conducted targeted to the internal regions of the XvPSap1 promoter, leaving the 5'- and 3'-regions intact. The 3'-region or proximal promoter was maintained as this region generally contains primary regulatory elements (Shahmuradov *et al.*, 2003). The 5'-region or distal promoter was maintained as its presence was necessary to test the hypothesis, that is, that stress inducible elements are located in this region. Four shorter promoters were consequently generated, each lacking the internal region and containing different fragment sizes at the 5' proximal and 3' distal regions to map the sequences required for functional activity of the XvPSap1 promoter.

5.2. Relative water content and protein activity verification

In this study, a trial experiment was performed prior to determination of the activities of the various XvPSap1 promoters in transgenic T1 *N. tabacum* plants in response to dehydration stress. This involved assessing promoter activity in transgenic T0 plants by visually evaluating the levels of Luc protein in the leaves of plants with the aid of the Xenogen IVIS Lumina (Caliper, USA). This yielded an unexpected result, which is that the amounts of Luc protein varied within the leaves. In one leaf the Luc protein was more concentrated at the apex, whereas in other leaves, it was concentrated towards the base or at the centre. Even more significant was that the concentrations of the Luc protein were not mirror images when the leaf was split in half. Furthermore, it was noted that the bottom, older leaves displayed higher levels of total Luc protein activity compared to the middle and younger top leaves.

The variation of Luc protein between older and younger leaves could be due to the fact that the older leaves were undergoing natural senescence, which may have an increased effect on stress gene expression. In contrast, the younger top leaves displayed the lowest levels of Luc protein activity when compared to leaves in the other parts of the plant. The variation in the localisation of the Luc protein within the leaves could be due to the fact that different parts of leaf may have different RWC

University of Cape Town

values. Possibilities also exist that gene silencing may have occurred or the nature of the gene integration within the tissue may have an effect on the expression of the Luc protein.

Based on the trial results, it was decided that using a single leaf or even half of a leaf for promoter activity analysis and for RWC determination would not generate reproducible results. Thus the dehydration treatment was performed on two sets of plants. The first set was used for RWC determination in which four leaves (positioned in the middle of the plant) were sampled, while the other set was used for promoter analysis in the four leaves. In each case, the four leaves were taken as being representative of the entire plant. This method for dehydration treatment in a tobacco species has been previously employed (Audran *et al*., 1998).

A similar trend in differing concentrations of the Luc protein in different parts of the leaf was observed in the T1 plants. This confirmed the logical basis of the selected method for dehydration, RWC and promoter activity assessment as applied in this study. In this way, unbiased representations of the RWC and promoter activities are determined in the whole plant versus unreliable determination in a portion of the plant, which could lead to a false idea of the promoter activities. A further advantage of this method is that it circumvents any possible wounding stress on the plant during the dehydration treatment when leaves are removed by excision. Excision of the leaf may be perceived as a wounding stress and this in turn might affect promoter activity (Walker-Simmons *et al*., 1984; Vian and Davies, 2006).

University of Cape Town

5.3. Stress inducible activity

The three promoters (XvPSap1D, E and G) were active in T1 *N. tabacum* plants under hydrated and dehydrated conditions. At the start of the dehydration period (days 0 and 1), when plants were considered to be hydrated with RWC values of 80 to 90%, basal levels of *luc* mRNA and protein activity were detected. From day 2 of the dehydration period, which corresponded to an RWC of *circa* 75%, differences in *luc* expression were seen in plants transformed with the various shortened XvPSap1 promoter constructs.

Although differences in *luc* expression were seen throughout the dehydration treatment, the various shortened promoters were found to induce *luc* transcription to a similar extent. Likewise, protein activity was observed to attain a similar level for all three shortened promoters towards the end of the dehydration period. This result, added to the fact that none of the T1 plants generated in this study displayed negative phenotypic characteristics and the presence of putative stress responsive *cis*-acting regulatory elements, demonstrates that the three shortened XvPSap1 promoters are stress inducible.

5.4. Stress repressible activity

In contrast to the three shorter XvPSap1 promoters, no Luc protein activity was detected in plants transformed with XvPSap2 throughout the dehydration treatment. It was initially concluded that there was no difference in Luc protein activity driven by XvPSap2. Consequently it was thought that XvPSap2 was not functionally active. However, the quantitative real time PCR results provided a different result. Here, the presence of *luc* mRNA, driven by XvPSap2, was observed at very low levels on day 0 of the dehydration. As the dehydration progressed, the levels of *luc* mRNA decreased significantly until its presence could not be measured. This decrease in transcript levels and the absence of visible Luc protein activity suggest that XvPSap2 is dehydration repressible. Interestingly, no differences were observed in the putative positive *cis*-acting regulatory elements between XvPSap2 and the shortened XvPSap1 promoters. There might however be unidentified negatively acting regulatory elements or some type of structural regulation involved.

5.5. Transcript fluctuations

Fluctuations in *luc* mRNA expression patterns in plants transformed with the three shortened promoter fragments could be attributed to the fact that the RWC and promoter activity analysis were determined on individual plants. This meant that different plants were sampled each day. As a consequence, the possibility exists that a plant may have dehydrated to a lesser or greater extent, resulting in an RWC outside of the trend. The end result would be that *luc* transcription would be influenced. Due to this limitation, such variation could be expected, although there

University of Cape Town

was minimal standard deviation between plants. This was shown by the fact that when two groups of plants were dehydrated simultaneously (wild type and transgenic), no significant differences were observed phenotypically or in the RWC values. The only significant difference was on day 4 in the RWC analysis. However, even on this day, the differences were not considerable.

5.6. Future work

Due to the time constraints, no biological replicates were incorporated in this study. In the absence of biological replicates, all that can be suggested is that biologically relevant changes either occurred or not (Karp *et al*., 2005). Furthermore, it does not allow for accurately determining absolute expression values and fold changes. Thus, in this study biologically relevant changes in the form of increased mRNA and protein were observed but absolute values for these increases could not be accurately determined. To overcome this limitation, future work should include screening of transformants to identify greater numbers of positive T1 as well as potential T2 and T3 transgenic tobacco and *Arabidopsis* plants. At the T3 stage, the plants should be largely homozygous and the promoter insertions should be stably integrated into the plant genome (James et al., 2002). At this stage, an appropriate sample size including a suitable number of biological repeats (between four and six) may be determined to suit the study and determine absolute values with a degree of confidence and statistical significance (Nakagawa and Cuthill, 2007).

Garwe. (2003) demonstrated that the *XvSap1* gene was not only upregulated in response to dehydration, but also in response to salt stress, low and high temperatures as well as high light. This suggests that the XvPSap1 promoter and consequently the shortened promoters generated in this study could be functionally active under these abiotic stresses. Their responses to various abiotic stresses could be investigated in future studies. This would be significant because, although drought is the leading cause of reductions in crop productivity, plants are significantly affected by other abiotic stresses as well (Cattivelli, 2008).

Finally, the promoter constructs should be tested in maize as it is the ultimate crop plant for which this study was intended. The time frame for generating transgenic T1

University of Cape Town

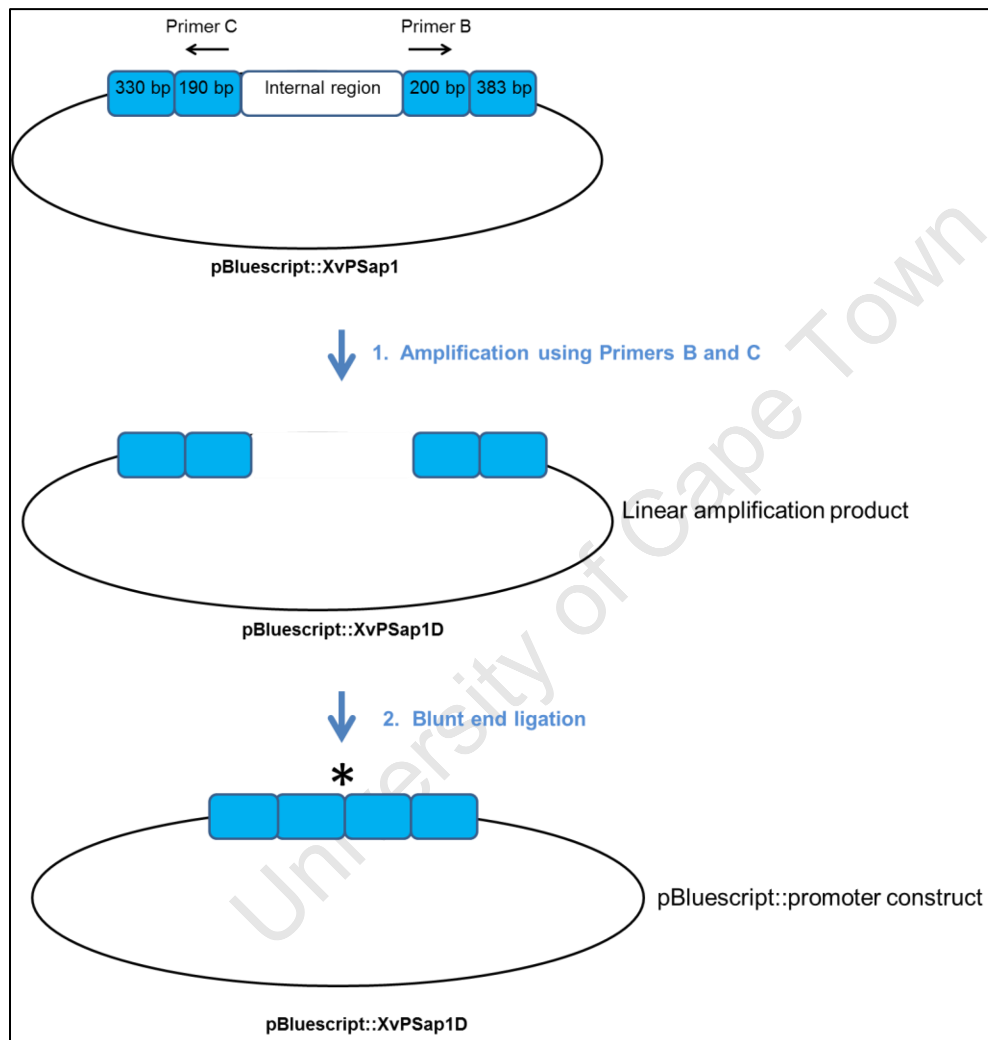
or T2 maize plants are out of the scope of the time allotted for this study. Okoth (2009) previously demonstrated that the full length XvPSap1 promoter was functionally active in both *N. tabacum* and maize, thus it is strongly suggested that the XvPSap promoters investigated in this study would be active and functional in maize. This answers the secondary objective of this study.

University of Cape Town

Appendices

Appendix A: Schematic diagrams

A.1. Mutagenesis strategy to generate pBluescript::promoter



University of Cape Town

Figure A.1: Schematic illustration of the mutagenesis strategy to generate pBluescript::XvPSap1D. **1.** Amplification using Primer B and C to generate a linear pBluescript::XvPSap1D amplification product. The internal part of the full length XvPSap1 promoter was removed. **2.** A blunt end ligation reaction of the linear amplification product was performed to generate circular pBluescript::XvPSap1D. The site of ligation is indicated with an asterisk. The same mutagenesis strategy was followed using primer set A and D, Primer set A and C and primer set B and D to generate pBluescript::XvPSap1E, XvPSap1F and XvPSap1G, respectively.

A.2. Cloning strategy to generate pBluescript::promoter::*luc*::NosT constructs

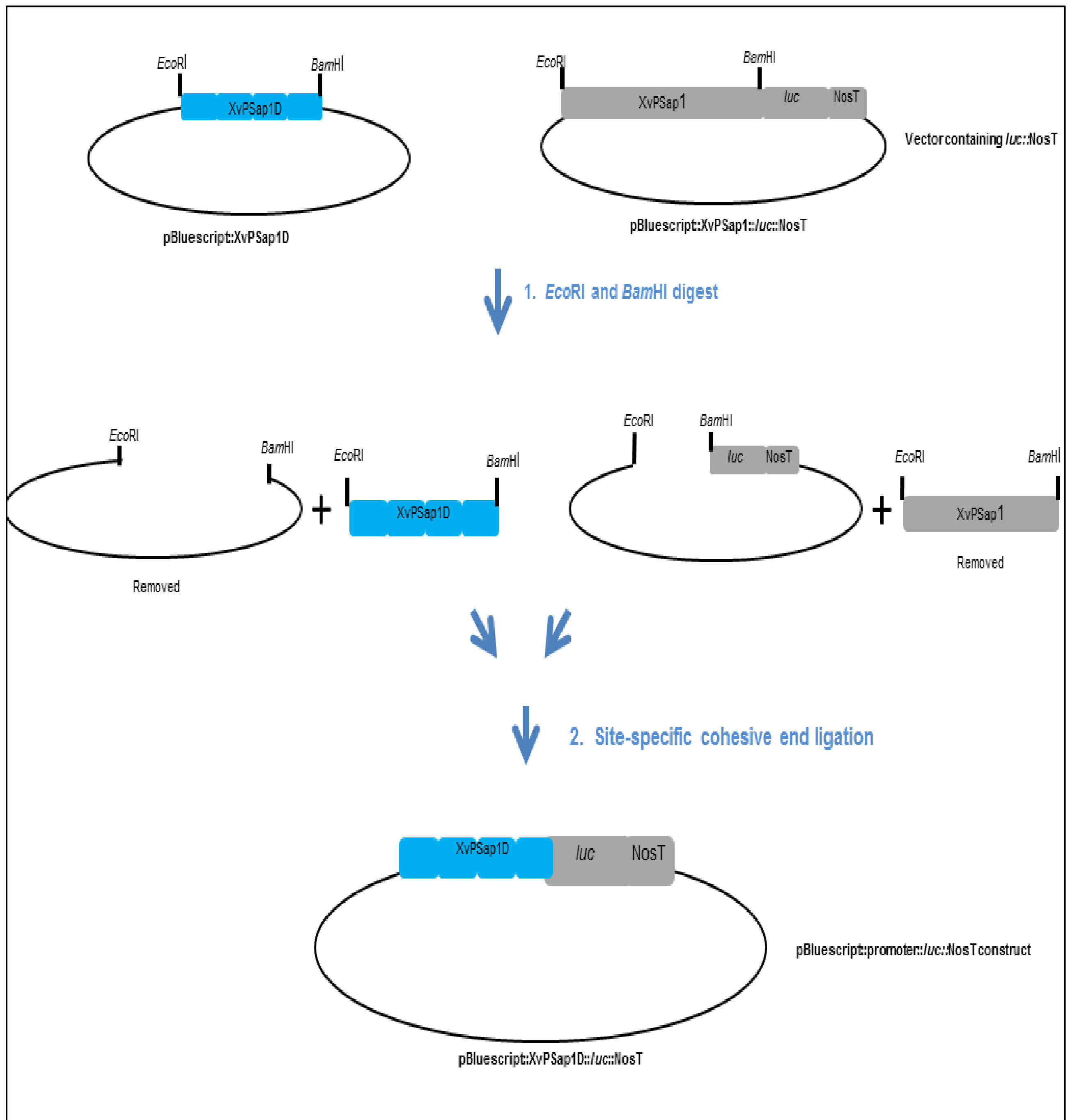


Figure A.2: Generation of pBluescript::XvPSap1D::*luc*::NosT. **1**. The pBluescript::XvPSap1D construct generated from the mutagenesis strategy was digested with *Eco*RI and *Bam*HI resulting in the removal of pBluescript from XvPSap1D. pBluescript containing the full length promoter, *luc* and NosT was also digested with *Eco*RI and *Bam*HI resulting in the removal of the full length promoter. **2**. The XvPSap1D promoter was ligated to pBluescript::*luc*::NosT resulting in the generation of the pBluescript::XvPSap1D::*luc*::NosT construct. The same cloning strategy was followed to generate all of the pBluescript::promoter::*luc*::NosT constructs.

A.3. Cloning strategy to generate the binary vector constructs

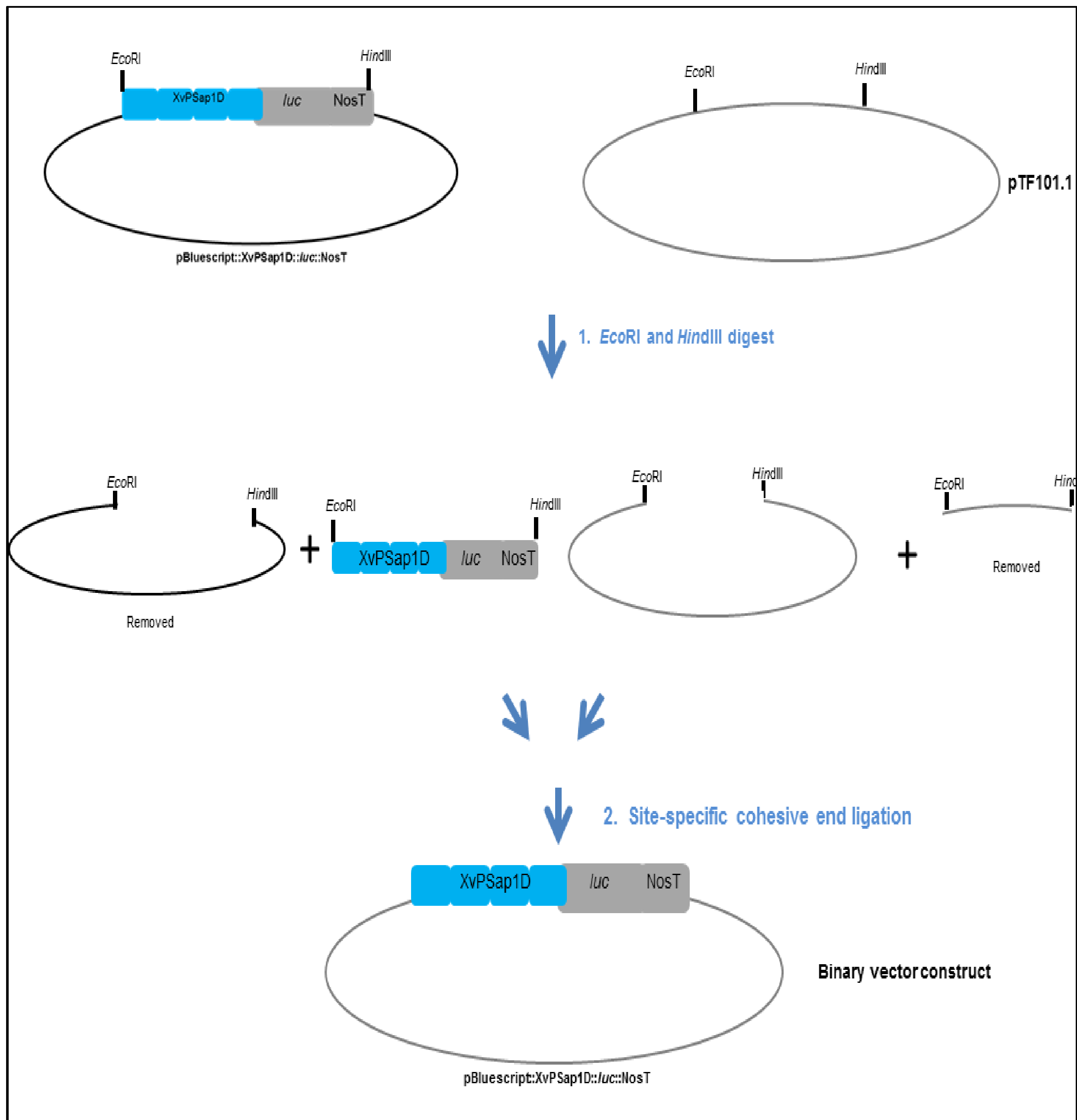


Figure A.3: Generation of binary vector construct, pTF101.1::XvPSap1D::*luc*::NosT. **1.** The pBluescript::XvPSap1D::luc::NosT construct generated from the cloning strategy in fig. A.2 was digested with *Eco*RI and HindIII resulting in the removal of pBluescript from XvPSap1D::*luc*::NosT. The binary vector, pTF101.1 was also digested with *Eco*RI and HindIII. **2.** The XvPSap1D::*luc*::NosT fragment was ligated to pTF101.1 resulting in the generation of the pTF101.1::XvPSap1D::*luc*::NosT construct. The same cloning strategy was followed to generate all of the pTF101.1::promoter::*luc*::NosT constructs.

Appendix B: Vector maps

B.1. pBluescript

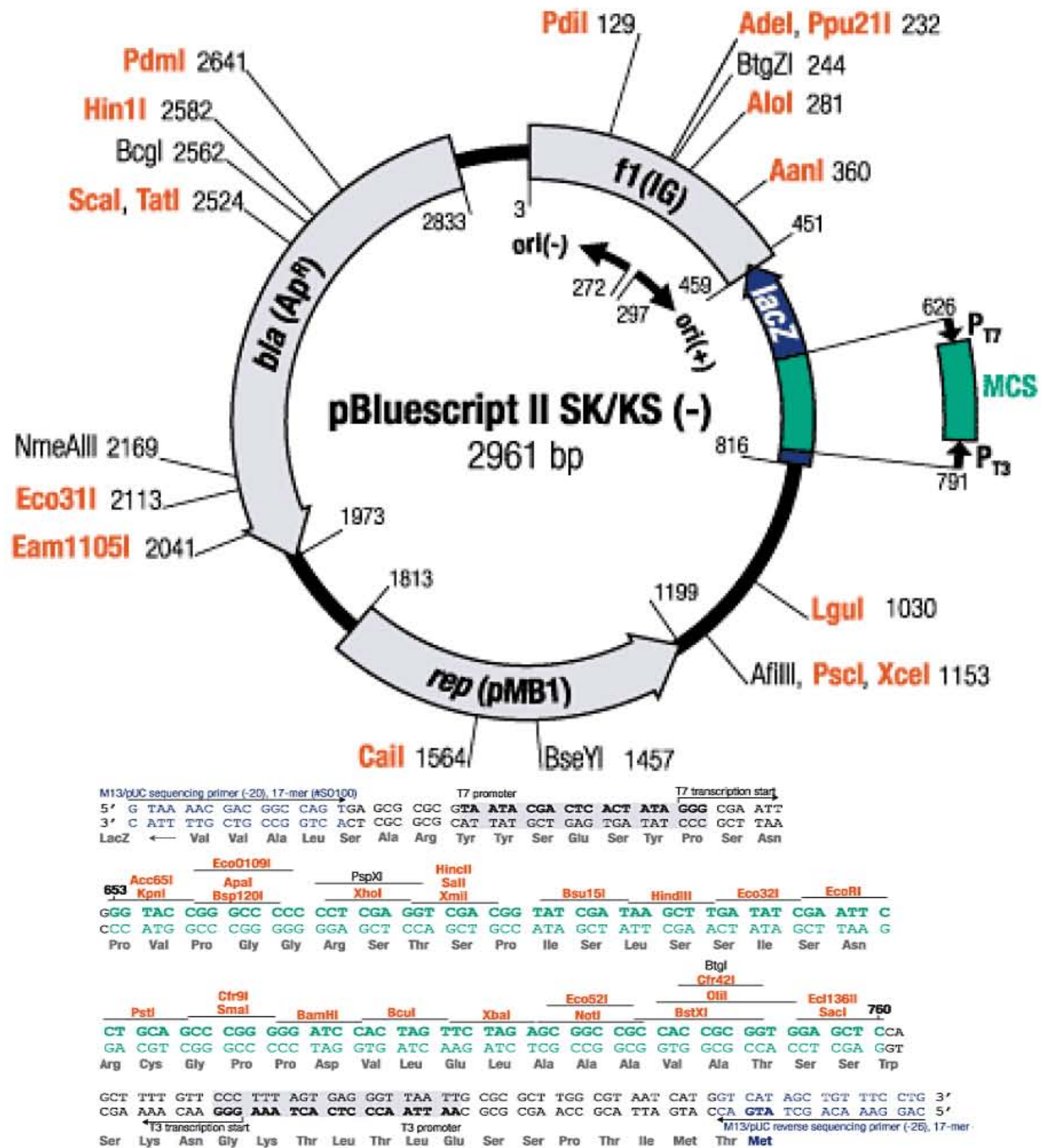


Figure B.1: pBluescript vector. The multiple cloning site is indicated below the vector map.

B.2. pTF101.1

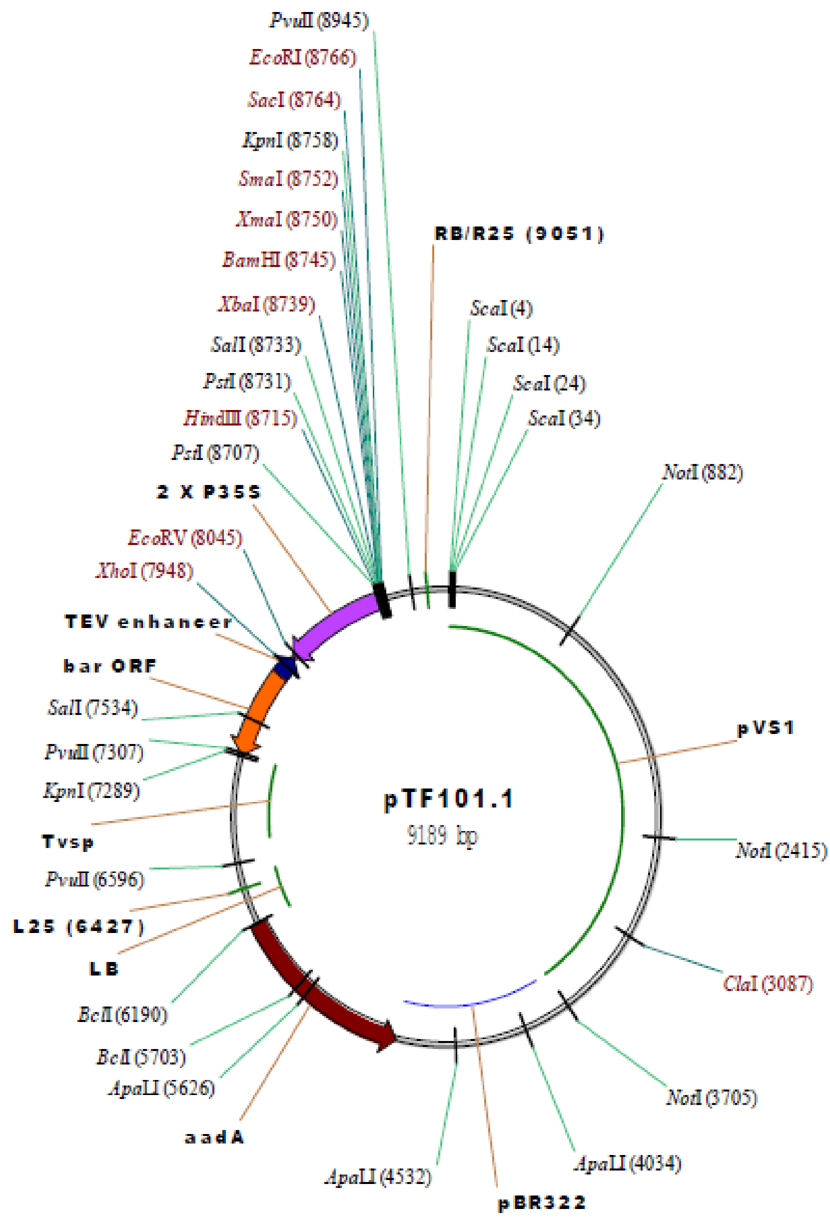


Figure B.2: pTF101.1 vector.

Appendix C: Nucleotide sequences and multiple sequence alignments

C.1. XvPSap1

```
1       ACTGTCTGGG TAGCTGGCAA TATAGAGACG TAAGAAATTC ATGGATCATC ACCCTAATTC
61      GGTCTTTCAC TCATTTTATC CTAGACCTGA CTAAAAAACT TGGTCAGAGT TTTACTTATT
121     TAAAAAAAAG AGGACTTCAT GGCATCCATG TGCAGGTACA GCTCCCAGAA AAAAAAAGCA
181     TGAAACACGA CAGGATCAAT AGCATTCGAT CTGAAACAAA AGGTTGGAGC TCAAGACTTT
241     CTCCAAAATA TTAAGATGAT CCAAAGAATT ACCCCAAGAT ATCCAACGTA TACCAATGTG
301     TATACCGAAA GTAAGAAAGT TCACGTGCAT TCTTTGATTT TTCTCCCGAG TGTTCTTTTC
361     TGAAATGAGT AAATAAGACT AGAATAAGAG CTAATGTATT TTYTTTCTAA AAAAAAGTTG
421     AATGTGGATA CAATATGATT ATACATTCAT TAGCTATTTT AAGTATATTC TATTTTTTYT
481     TTCCCCAAAA GAACACAAAT GTGTTCCGTC ACTTTCCATG GGGCAAATTA CAACTTAGGC
541     TTTATCTTAG TTGGTATGAT CTTAATTTTA TTATACTTTA AACAACTTAT CGCTAATAAT
601     TTTGTTTTGA TTTATGCGCC AATTGTAAAT ATAAATCGGA TAARTYTGAA CATTAATACT
661     TTTAGTCAGT TTCAAAAAAG AAAAAGATAA CTATGACGTT AGAGTTTGGA ATCCAGTCAA
721     ATGGAACTTA TTTTTTAGTT CATCAGAATC AACTTGATGA GATTTTTTGT ACTAGACAAT
781     CATCTTGAAT GATAGRTAGG GGACTTACCA ATCAGCCCCC CATATTTTRA AACTTTCAAC
841     GCGCCCCTCA ACATTTTTCT CTTTCAACGC GCCCYTCCAT CACTTTTCTC TTTTCATCGC
901     CCCTCCCTCA ACTTTTTGGT CGGACGGAAA TACCTCTATA TATATTTCAT CATTTCGACC
961     CCCAAAACAC CTATTTAGAA GTATTTTTTG GAAAAAATTT TAACATGAAA GTTTTAGATC
1021    TTGATGAGAT CTACAATTTT TATGTTGAAA GTTTTTCCAA AAAATACTTT TAGATAGATA
1081    TTTTGGAACT CCGAAGTGTT AATGGGTCGA CCCGCTAACT TGCGGAAATA GAAAAACATC
1141    AAGATCTACA ACTTTTATGT TGAAAGTTTT TTCAAAAAAT ACTTCTAGAT AGATATTTGA
1201    TTTGTAATTT TAATGTTGAA AGTTTTTTCA AAAAATACTT CTAGATAGAT ACTTTGGGGC
1261    TCCGGAGTGT TAACAAGTAT AGGAATATTT TTGTCTGCAA AAAATTAATT TTTCGGACAA
1321    GAGGCCGATC AGTAAGGAAT CTGGTCGGAG GGGCTGTTCG GCAATATAAG TTCAGATAGG
1381    AGAACTGATC GGATATTTTT CCTTAATTTA ATTCCAATTT GATACTATTA AGAATGAAAA
1441    CATCCTAATA ATTGTGACCA CTTTATAGCA CCACATTTAT TTTAATTTAA ATCTTTTAAA
1501    TCTTAGAATT GGACAGGGTG CTTATGATAA CAAACTTGTT CCTATCAACA ACTGCATGTT
1561    AGACAGCGCC GAATTTACAG TCCTACTGGG CGCCACTTTT CAACCCACAT CATCAAGATG
1621    AACACCACGT TATCTTCATC CGCTCCAACC ACATGGTCCA GCGCCACTGG CCAAGACCGC
1681    CAGCCAGCCA GGCCATCCAA CGTGGTGCAT TTTCTAACAC TCCACGTTCG CTGTACGGCA
1741    TTATTTCTCC AGCCAGAAAG ACCGAGACAG CGACGCTGTT GGGCGGGCCC GCGGCCTGCT
1801    CTCTCTGCTT CCCCATGAGA TTCACGGGCA TCGCTCCTCG CTCGTGCCTA CGCCCGCGCC
1861    CGCGCCCGAC CGCGCCGGTC CACGTGACGT GGCGCAGCAA TCGTTCTTAC TAGGCGCTTG
1921    CACGTGTCGT TCGCATGCGA AGCGTCCACA CTGCCAACGA CCTCCTTAAA TATCCTTGTG
1981    ATATTCGCCT TACGATCTCA CACTTCGCAC GCAAAGGCCA GTCGCAGATT TGGGTTGAAT
2041    TTGCTGCGTT TTGGCAGATT TTGAGCGAGA GATATTAGGG AAG
```

Figure C.1: Nucleotide sequence for the full length XvPSap1 promoter.

University of Cape Town

C.2. XvPSap1D

1 ACTGTCTGGG TAGCTGGCAA TATAGAGACG TAAATAATTG TCTGTAAATA GGGAGAAATT
61 CATGGATCAT CACCCTAATT CGGTCTTTCA CTCATTTTAT CATAGACCTG ACTAAAGAAC
121 TTGGTCAGAG TTTTTACTTA TTTAAAATAA AGAGGACTTC ATGGCATCCA TGTGCAGGTA
181 CAGCTCCCAG AAAAAAAAGC ATGAAACACG AGAGGATCAA TAGCATTCGA TCTGAAACAA
241 AAGGTTGCAG CTCAAGACTT TCTCCAAAAT ATTAAGATGA TCCAAAGAAT TACCCCAAGA
301 TATCCAACGT ATACCAATGT GTATACCGAA AGTAAGAAAG TTCACGTGCA TTCTTTGATT
361 TTTCTCCCGA GTGTTCTTTT CTGAAATGAG TAAATAAGAC TAGAATAAGA GCTAATGTAT
421 TTTTTTTCTA AAAAAAGTTG AATGTGGATA CAATATGATT ATACATTCAT TAGCTATTTT
481 AAGTATATTC TATTTTTTTT CCCCCCAAAA GAACACAAAT GTGTTCCGTC ACTTTCCATG
541 TCTTAGAATT GGACAGGGTG CTTATGATAC AACTTGTTCC TATCAACAAC TGCATGTTAG
601 ACAGCGCCGA ATTTACAGTC CTACTGGGCG CCACTTTTCA ACCCACATCA TCAAGATGAA
661 CACCACGTTA TCTTCATCCG CTCCAACCAC ATGGTCCAGC GCCACTGGCC AAGACCGCCA
721 GCCAGCCAGG CCATCCAACG TGGTGCATTT TCTAACACTC CACGTTCGCT GTACGGCATT
781 ATTTCTCCAG CCAGAAAGAC CGAGACAGCG ACGCTGTTGG GCGGGCCCGC GGCCTGCTCT
841 CTCTGCTTCC CCATGAGATT CACGGGCATC GCTCCTCGCT CGTGCCTACG CGACCGCGCC
901 GATCCACGTG ACGTGGCGCA GCAATCGTTC TTACTAGGCG CTTGCACGTG TCGTTCGCAT
961 GCGAAGCGTC CACACTGCCA ACGACCTCCT TAAATATCCT TGTGATATTC GCCTTACGAT
1021 CTCACACTTC GCACGCAAAG GCCAGTCGCA GATTTGGGTT GAATTTGCTG CGTTTTGGCA
1081 GATTTTGAGC GAGAGATATT AGGGAAG

Figure C.2: Nucleotide sequence of the XvPSap1D promoter.

C.3. XvPSap1E

1 ACTGTCTGGG TAGCTGGCAA TATAGAGACG TAAATAATTG TCTGTAAATA GGGAGAAATT
61 CATGGATCAT CACCCTAATT CGGTCTTTCA CTCATTTTAT CATAGACCTG ACTAAAGAAC
121 TTGGTCAGAG TTTTTACTTA TTTAAAATAA AGAGGACTTC ATGGCATCCA TGTGCAGGTA
181 CAGCTCCCAG AAAAAAAAGC ATGAAACACG AGAGGATCAA TAGCATTCGA TCTGAAACAA
241 AAGGTTGCAG CTCAAGACTT TCTCCAAAAT ATTAAGATGA TCCAAAGAAT TACCCCAAGA
301 TATCCAACGT ATACCAATGT GTATACCGAA AGTAAGAAAG TTCACGTGCA TTCTTAGAAT
361 TGGACAGGGT GCTTATGATA CAACTTGTTC CTATCAACAA CTGCATGTTA GACAGCGCCG
421 AATTTACAGT CCTACTGGGC GCCACTTTTC AACCCACATC ATCAAGATGA ACACCACGTT
481 ATCTTCATCC GCTCCAACCA CATGGTCCAG CGCCACTGGC CAAGACCGCC AGCCAGCCAG
541 GCCATCCAAC GTGGTGCATT TTCTAACACT CCACGTTCGC TGTACGGCAT TATTTCTCCA
601 GCCAGAAAGA CCGAGACAGC GACGCTGTTG GGCGGGCCCG CGGCCTGCTC TCTCTGCTTC
661 CCCATGAGAT TCACGGGCAT CGCTCCTCGC TCGTGCCTAC GCGACCGCGC CGATCCACGT
721 GACGTGGCGC AGCAATCGTT CTTACTAGGC GCTTGCACGT GTCGTTCGCA TGCGAAGCGT
781 CCACACTGCC AACGACCTCC TTAAATATCC TTGTGATATT CGCCTTACGA TCTCACACTT
841 CGCACGCAAA GGCCAGTCGC AGATTTGGGT TGAATTTGCT GCGTTTTGGC AGATTTTGAG
901 CGAGAGATAT TAGGGAAG

Figure C.3: Nucleotide sequence of the XvPSap1E promoter.

University of Cape Town

C.4. XvPSap1F

```
1       ACTGTCTGGG TAGCTGGCAA TATAGAGACG TAAATAATTG TCTGTAAATA GGGAGAAATT
61      CATGGATCAT CACCCTAATT CGGTCTTTCA CTCATTTTAT CATAGACCTG ACTAAAGAAC
121     TTGGTCAGAG TTTTTACTTA TTTAAAATAA AGAGGACTTC ATGGCATCCA TGTGCAGGTA
181     CAGCTCCCAG AAAAAAAAGC ATGAAACACG AGAGGATCAA TAGCATTCGA TCTGAAACAA
241     AAGGTTGCAG CTCAAGACTT TCTCCAAAAT ATTAAGATGA TCCAAAGAAT TACCCCAAGA
301     TATCCAACGT ATACCAATGT GTATACCGAA AGTAAGAAAG TTCACGTGCA TTCTTTGATT
361     TTTCTCCCGA GTGTTCTTTT CTGAAATGAG TAAATAAGAC TAGAATAAGA GCTAATGTAT
421     TTTTTTTCTA AAAAAAGTTG AATGTGGATA CAATATGATT ATACATTCAT TAGCTATTTT
481     AAGTATATTC TATTTTTTTT CCCCCCAAAA GAACACAAAT GTGTTCCGTC ACTTTCCATG
541     ACGTGGTGCA TTTTCTAACA CTCCACGTTC GCTGTACGGC ATTATTTCTC CAGCCAGAAA
601     GACCGAGACA GCGACGCTGT TGGGCGGGCC CGCGGCCTGC TCTCTCTGCT TCCCCATGAG
661     ATTCACGGGC ATCGCTCCTC GCTCGTGCCT ACGCGACCGC GCCGATCCAC GTGACGTGGC
721     GCAGCAATCG TTCTTACTAG GCGCTTGCAC GTGTCGTTCG CATGCGAAGC GTCCACACTG
781     CCAACGACCT CCTTAAATAT CCTTGTGATA TTCGCCTTAC GATCTCACAC TTCGCACGCA
841     AAGGCCAGTC GCAGATTTGG GTTGAATTTG CTGCGTTTTG GCAGATTTTG AGCGAGAGAT
901     ATTAGGGAAG
```

Figure C.4: Nucleotide sequence of the XvPSap1F promoter

C.5. XvPSap1G

```
1       ACTGTCTGGG TAGCTGGCAA TATAGAGACG TAAATAATTG TCTGTAAATA GGGAGAAATT
61      CATGGATCAT CACCCTAATT CGGTCTTTCA CTCATTTTAT CATAGACCTG ACTAAAGAAC
121     TTGGTCAGAG TTTTTACTTA TTTAAAATAA AGAGGACTTC ATGGCATCCA TGTGCAGGTA
181     CAGCTCCCAG AAAAAAAAGC ATGAAACACG AGAGGATCAA TAGCATTCGA TCTGAAACAA
241     AAGGTTGCAG CTCAAGACTT TCTCCAAAAT ATTAAGATGA TCCAAAGAAT TACCCCAAGA
301     TATCCAACGT ATACCAATGT GTATACCGAA AGTAAGAAAG TTCACGTGCA TACGTGGTGC
361     ATTTTCTAAC ACTCCACGTT CGCTGTACGG CATTATTTCT CCAGCCAGAA AGACCGAGAC
421     AGCGACGCTG TTGGGCGGGC CCGCGGCCTG CTCTCTCTGC TTCCCCATGA GATTCACGGG
481     CATCGCTCCT CGCTCGTGCC TACGCGACCG CGCCGATCCA CGTGACGTGG CGCAGCAATC
541     GTTCTTACTA GGCGCTTGCA CGTGTCGTTC GCATGCGAAG CGTCCACACT GCCAACGACC
601     TCCTTAAATA TCCTTGTGAT ATTCGCCTTA CGATCTCACA CTTCGCACGC AAAGGCCAGT
661     CGCAGATTTG GGTTGAATTT GCTGCGTTTT GGCAGATTTT GAGCGAGAGA TATTAGGGAA
721     G
```

Figure C.5: Nucleotide sequence of the XvPSap1G promoter

University of Cape Town

C.6. XvPSap2

```
1       TCGTACGAGA ATCGCTGTCC TCTCCAACGA GCAAGATATT CTAAAAAATT ATTTTTAAAT
61      TTATTAAGAC TTTGAAATCT CCAAAATAAG AATTTCGAAA TATTTTTTGA AAAAAATATG
121     GAATACCAAA GTTATAAATC TCATCAAAAT CTATAATATT TATACTTAAA TTTTTAAAAA
181     AATACTCCTA TATACGTATT TAGTACAATC AATTTACCAT AATATACCGA TCAAAAGTTG
241     GTTTTGACTG TCTGGGTAGC TGGCAATATA GAGACGTAAG AAATTCATGG ATCATCACCC
301     TAATTCGGTC TTTCACTCAT TTTATCCTAG ACCTGACTAA AGAACTCGGT CAGAGTTTTA
361     CTTATTTAAA AAAAAGAGGA CTTCATGGCA TCCATGTGCA GGTACAGCTC CCAGAAAAAA
421     AAAAGCATGA AACACGACAG GATCAATAGC ATTCGATCTG AAACAAAAGG TTGGAGCTCA
481     AGACTTTCTC CAAAATATTA AGATGATCCA AAGAATTACC CCAAGATATC CAACGTATAC
541     CAATGTGTAT ACCGAAAGTA AGAAAGTTCA CGTGCATTCT TTTCCCGAGT GTTCTTTTCT
601     GAAATGAGTA AATAAGACTA GAATAAGAGC TAATGTATTT TTTTTTTCTA AAAAAAAGTT
661     GAATGTGGAT ACAATATGAT TATACATTCA TTAGCTATTT TAAGTATATT CTATTTTTTT
721     TTTCCCCAAA AGAACACAAA TGTGTTCCGT CACCTTCCAT GGGGCAAATT ACAACTTAGG
781     CTTTATCTTA GTTGGTATGA TCTTAATTTT ATTATACTTT AAACAACTTA TCGCTAATAA
841     TTTTGTTTTA ATTTATGCGC CAATTGTGAA TATAAATCGG ATAAATTTGA GCATTAATAC
901     TTTTTAGTCA AAGTTTCAAA AAAGAAAAAA ATAACTATGA CGTTAGAGTT TGGAATCCAG
961     TCAAAATGGA ACTTATTTTT TTAGTTCATC ATAATCAACT TGATGAGATT TTTTGTACTA
1021    GACAATCATC TTGAATGATT AATTTAATTC CAATTTGATA CTATTAAGAA TGAAAACATC
1081    CTAAATAATA GCTGCCGTGA TCACTTTATA GCACCACATT TATCTTAATT TAAAATATTT
1141    TTAATCTTAG AATTGGACAG GGTGCTTATG ATAACAAACT TGTTCCTATC AACAGCTGCA
1201    TGTTAGACAG CGCCGAATTT ACAGTCCTAC TGGGCGCCAC TTTTCAACCC ACATCATCAA
1261    GCTGAACACC ACGTTATCTA CATCCACTCC AACCACATGG TCCAGCGGCA CTGGCCAAGA
1321    CCGCCAGCCA GCAGGCCATC CAACGTGGTG CATTTCTAAC ACTCCACGTT CGCTGTACGA
1381    CGTTATTTCT CCAGCCAGAA AGACCGGGAC AGCGACGCTG TTGGGCGGGC CCGCGGCCTG
1441    CTCTCTCTGC TTCCCCATGA GATTCACGGG CATCGCTCGT GCCTACGCCC GCGCCCGACC
1501    GCGCCGATGC ACGTGACGTG GCGCAGCAAT CGTTCTTACT AGGCGCTTGC ACGTGTCGTT
1561    CGCATGCGAA GCGTCCAGAC CGCCAACGAC CTCCTTAAAT ATCCTTGTGA TCTTCGCCTT
1621    ACGATCTCAC ACTTCGCACG CAGAGTCCAG TCGCAGATTT GGGTTGAATT TGCTGCGTTT
1681    TGGCAGATTT TGAGCGAGAG AGATTAGGGA AGATGAGGAA CGAGGGTTTT CTGAAAATGA
1741    AGACCGACGT TGGGGTCGCC GACGAGGTGA TCTCCGGAGA TCTCAAGCAG CTTGGTGACG
1801    CTGCGAAGCG GCTAGCTAAC CATGCGATCA AGC
```

Figure C.6: Nucleotide sequence of the XvPSap2 promoter

University of Cape Town

C.7 *luc* gene

```
1     ATGGTCACCG ACGCCAAAAA CATAAAGAAA GGCCCGGCGC CATTCTATCC GCTGGAAGAT
61    GGAACCGCTG GAGAGCAACT GCATAAGGCT ATGAAGAGAT ACGCCCTGGT TCCTGGAACA
121   ATTGCTTTTA CAGATGCACA TATCGAGGTG GACATCACTT ACGCTGAGTA CTTCGAAATG
181   TCCGTTCGGT TGGCAGAAGC TATGAAACGA TATGGGCTGA ATACAAATCA CAGAATCGTC
241   GTATGCAGTG AAAACTCTCT TCAATTCTTT ATGCCGGTGT TGGGCGCGTT ATTTATCGGA
301   GTTGCAGTTG CGCCCGCGAA CGACATTTAT AATGAACGTG AATTGCTCAA CAGTATGGGC
361   ATTTCGCAGC CTACCGTGGT GTTCGTTTCC AAAAAGGGGT TGCAAAAAAT TTTGAACGTG
421   CAAAAAAAGC TCCCAATCAT CCAAAAAATT ATTATCATGG ATTCTAAAAC GGATTACCAG
481   GGATTTCAGT CGATGTACAC GTTCGTCACA TCTCATCTAC CTCCCGGTTT TAATGAATAC
541   GATTTTGTGC CAGAGTCCTT CGATAGGGAC AAGACAATTG CACTGATCAT GAACTCCTCT
601   GGATCTACTG GTCTGCCTAA AGGTGTCGCT CTGCCTCATA GAACTGCCTG CGTGAGATTC
661   TCGCATGCCA GAGATCCTAT TTTTGGCAAT CAAATCATTC CGGATACTGC GATTTTAAGT
721   GTTGTTCCAT TCCATCACGG TTTTGGAATG TTTACTACAC TCGGATATTT GATATGTGGA
781   TTTCGAGTCG TCTTAATGTA TAGATTTGAA GAAGAGCTGT TTCTGAGGAG CCTTCAGGAT
841   TACAAGATTC AAAGTGCGCT GCTGGTGCCA ACCCTATTCT CCTTCTTCGC CAAAAGCACT
901   CTGATTGACA AATACGATTT ATCTAATTTA CACGAAATTG CTTCTGGTGG CGCTCCCCTC
961   TCTAAGGAAG TCGGGGAAGC GGTTGCCAAG AGGTTCCATC TGCCAGGTAT CAGGCAAGGA
1021  TATGGGCTCA CTGAGACTAC ATCAGCTATT CTGATTACAC CCGAGGGGGA TGATAAACCG
1081  GGCGCGGTCG GTAAAGTTGT TCCATTTTTT GAAGCGAAGG TTGTGGATCT GGATACCGGG
1141  AAAACGCTGG GCGTTAATCA AAGAGGCGAA CTGTGTGTGA GAGGTCCTAT GATTATGTCC
1201  GGTTATGTAA ACAATCCGGA AGCGACCAAC GCCTTGATTG ACAAGGATGG ATGGCTACAT
1261  TCTGGAGACA TAGCTTACTG GGACGAAGAC GAACACTTCT TCATCGTTGA CCGCCTGAAG
1321  TCTCTGATTA AGTACAAAGG CTATCAGGTG GCTCCCGCTG AATTGGAATC CATCTTGCTC
1381  CAACACCCCA ACATCTTCGA CGCAGGTGTC GCAGGTCTTC CCGACGATGA CGCCGGTGAA
1441  CTTCCCGCCG CCGTTGTTGT TTTGGAGCAC GGAAAGACGA TGACGGAAAA AGAGATCGTG
1501  GATTACGTCG CCAGTCAAGT AACAACCGCG AAAAAGTTGC GCGGAGGAGT TGTGTTTGTG
1561  GACGAAGTAC CGAAAGGTCT TACCGGAAAA CTCGACGCAA GAAAAATCAG AGAGATCCTC
1621  ATAAAGGCCA AGAAGGGCGG AAAGATCGCC GTGTAA
```

Figure C.7: Nucleotide sequence of the *luc* gene.

C.8. nos terminator

```
1    GAGTACGGTG GGTAGCCCGA TCGTTCAAAC ATTTGGCAAT AAAGTTTCTT AAGATTGAAT
61   CCTGTTGCCG GTCTTGCGAT GATTATCATA TAATTTCTGT TGAATTACGT TAAGCATGTA
121  ATAATTAACA TGTAATGCAT GACGTTATTT ATGAGATGGG TTTTTATGAT TAGAGTCCCG
181  CAATTATACA TTTAATACGC GATAGAAAAC AAAATATAGC GCGCAAACTA GGATAAATTA
241  TCGCGCGCGG TGTCATCTAT GTTACTAGAT CCCTAGGCTA TCTGTCACTT CATCAAAAGG
```

Figure C.8: Nucleotide sequence of NosT.

University of Cape Town

C.9. Nucleotide sequence alignment of XvPSap1D, E, F and G

```
PsapD_(B-C)   ACTGTCTGGGTAGCTGGCAATATAGAGACGTAAATAATTGTCTGTAAATAGGGAGAAATTCATGGATCAT    70
PsapE_(B-D)   ACTGTCTGGGTAGCTGGCAATATAGAGACGTAAATAATTGTCTGTAAATAGGGAGAAATTCATGGATCAT    70
PsapG_(A-D)   ACTGTCTGGGTAGCTGGCAATATAGAGACGTAAATAATTGTCTGTAAATAGGGAGAAATTCATGGATCAT    70
PsapF_(A-C)   ACTGTCTGGGTAGCTGGCAATATAGAGACGTAAATAATTGTCTGTAAATAGGGAGAAATTCATGGATCAT    70
Consensus     actgtctgggtagctggcaatatagagacgtaaataattgtctgtaaatagggagaaattcatggatcat

PsapD_(B-C)   CACCCTAATTCGGTCTTTCACTCATTTTATCATAGACCTGACTAAAGAACTTGGTCAGAGTTTTTACTTA   140
PsapE_(B-D)   CACCCTAATTCGGTCTTTCACTCATTTTATCATAGACCTGACTAAAGAACTTGGTCAGAGTTTTTACTTA   140
PsapG_(A-D)   CACCCTAATTCGGTCTTTCACTCATTTTATCATAGACCTGACTAAAGAACTTGGTCAGAGTTTTTACTTA   140
PsapF_(A-C)   CACCCTAATTCGGTCTTTCACTCATTTTATCATAGACCTGACTAAAGAACTTGGTCAGAGTTTTTACTTA   140
Consensus     caccctaattcggtctttcactcattttatcatagacctgactaaagaacttggtcagagtttttactta

PsapD_(B-C)   TTTAAAATAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAGAAAAAAAAGCATGAAACACG   210
PsapE_(B-D)   TTTAAAATAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAGAAAAAAAAGCATGAAACACG   210
PsapG_(A-D)   TTTAAAATAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAGAAAAAAAAGCATGAAACACG   210
PsapF_(A-C)   TTTAAAATAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAGAAAAAAAAGCATGAAACACG   210
Consensus     tttaaaataaagaggacttcatggcatccatgtgcaggtacagctcccagaaaaaaaagcatgaaacacg

PsapD_(B-C)   AGAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGCAGCTCAAGACTTTCTCCAAAATATTAAGATGA   280
PsapE_(B-D)   AGAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGCAGCTCAAGACTTTCTCCAAAATATTAAGATGA   280
PsapG_(A-D)   AGAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGCAGCTCAAGACTTTCTCCAAAATATTAAGATGA   280
PsapF_(A-C)   AGAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGCAGCTCAAGACTTTCTCCAAAATATTAAGATGA   280
Consensus     agaggatcaatagcattcgatctgaaacaaaaggttgcagctcaagactttctccaaaatattaagatga

PsapD_(B-C)   TCCAAAGAATTACCCCAAGATATCCAACGTATACCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCA   350
PsapE_(B-D)   TCCAAAGAATTACCCCAAGATATCCAACGTATACCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCA   350
PsapG_(A-D)   TCCAAAGAATTACCCCAAGATATCCAACGTATACCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCA   350
PsapF_(A-C)   TCCAAAGAATTACCCCAAGATATCCAACGTATACCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCA   350
Consensus     tccaaagaattaccccaagatatccaacgtataccaatgtgtataccgaaagtaagaaagttcacgtgca

PsapD_(B-C)   TTCTTTGATTTTTCTCCCGAGTGTTCTTTTCTGAAATGAGTAAATAAGACTAGAATAAGAGCTAATGTAT   420
PsapE_(B-D)   T.....................................................................   351
PsapG_(A-D)   T.....................................................................   351
PsapF_(A-C)   TTCTTTGATTTTTCTCCCGAGTGTTCTTTTCTGAAATGAGTAAATAAGACTAGAATAAGAGCTAATGTAT   420
Consensus     t

PsapD_(B-C)   TTTTTTTCTAAAAAAAAGTTGAATGTGGATACAATATGATTATACATTCATTAGCTATTTTAAGTATATTC   490
PsapE_(B-D)   ......................................................................   351
PsapG_(A-D)   ......................................................................   351
PsapF_(A-C)   TTTTTTTCTAAAAAAAAGTTGAATGTGGATACAATATGATTATACATTCATTAGCTATTTTAAGTATATTC   490
Consensus

PsapD_(B-C)   TATTTTTTTTCCCCCCAAAAGAACACAAATGTGTTCCGTCACTTTCCATGTCTTAGAATTGGACAGGGTG   560
PsapE_(B-D)   ..................................................TCTTAGAATTGGACAGGGTG   371
PsapG_(A-D)   ......................................................................   351
PsapF_(A-C)   TATTTTTTTTCCCCCCAAAAGAACACAAATGTGTTCCGTCACTTTCCATG....................   540
Consensus
```

Continued…

```
PsapD_(B-C)    CTTATGATACAACTTGTTCCTATCAACAACTGCATGTTAGACAGCGCCGAATTTACAGTCCTACTGGGCG    630
PsapE_(B-D)    CTTATGATACAACTTGTTCCTATCAACAACTGCATGTTAGACAGCGCCGAATTTACAGTCCTACTGGGCG    441
PsapG_(A-D)    ......................................................................    351
PsapF_(A-C)    ......................................................................    540
Consensus

PsapD_(B-C)    CCACTTTTCAACCCACATCATCAAGATGAACACCACGTTATCTTCATCCGCTCCAACCACATGGTCCAGC    700
PsapE_(B-D)    CCACTTTTCAACCCACATCATCAAGATGAACACCACGTTATCTTCATCCGCTCCAACCACATGGTCCAGC    511
PsapG_(A-D)    ......................................................................    351
PsapF_(A-C)    ......................................................................    540
Consensus

PsapD_(B-C)    GCCACTGGCCAAGACCGCCAGCCAGCCAGGCCATCCAACGTGGTGCATTTTCTAACACTCCACGTTCGCT    770
PsapE_(B-D)    GCCACTGGCCAAGACCGCCAGCCAGCCAGGCCATCCAACGTGGTGCATTTTCTAACACTCCACGTTCGCT    581
PsapG_(A-D)    .....................................ACGTGGTGCATTTTCTAACACTCCACGTTCGCT    384
PsapF_(A-C)    .....................................ACGTGGTGCATTTTCTAACACTCCACGTTCGCT    573
Consensus                                           acgtggtgcattttctaacactccacgttcgct

PsapD_(B-C)    GTACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCGGGCCCGCGGCCTGCTCT    840
PsapE_(B-D)    GTACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCGGGCCCGCGGCCTGCTCT    651
PsapG_(A-D)    GTACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCGGGCCCGCGGCCTGCTCT    454
PsapF_(A-C)    GTACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCGGGCCCGCGGCCTGCTCT    643
Consensus      gtacggcattatttctccagccagaaagaccgagacagcgacgctgttgggcgggcccgcggcctgctct

PsapD_(B-C)    CTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCTCGTGCCTACGCGACCGCGCCGATCCACGTG    910
PsapE_(B-D)    CTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCTCGTGCCTACGCGACCGCGCCGATCCACGTG    721
PsapG_(A-D)    CTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCTCGTGCCTACGCGACCGCGCCGATCCACGTG    524
PsapF_(A-C)    CTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCTCGTGCCTACGCGACCGCGCCGATCCACGTG    713
Consensus      ctctgcttccccatgagattcacgggcatcgctcctcgctcgtgcctacgcgaccgcgccgatccacgtg

PsapD_(B-C)    ACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGTCGTTCGCATGCGAAGCGTCCACACTGCCA    980
PsapE_(B-D)    ACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGTCGTTCGCATGCGAAGCGTCCACACTGCCA    791
PsapG_(A-D)    ACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGTCGTTCGCATGCGAAGCGTCCACACTGCCA    594
PsapF_(A-C)    ACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGTCGTTCGCATGCGAAGCGTCCACACTGCCA    783
Consensus      acgtggcgcagcaatcgttcttactaggcgcttgcacgtgtcgttcgcatgcgaagcgtccacactgcca

PsapD_(B-C)    ACGACCTCCTTAAATATCCTTGTGATATTCGCCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGCA   1050
PsapE_(B-D)    ACGACCTCCTTAAATATCCTTGTGATATTCGCCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGCA    861
PsapG_(A-D)    ACGACCTCCTTAAATATCCTTGTGATATTCGCCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGCA    664
PsapF_(A-C)    ACGACCTCCTTAAATATCCTTGTGATATTCGCCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGCA    853
Consensus      acgacctccttaaatatccttgtgatattcgccttacgatctcacacttcgcacgcaaaggccagtcgca

PsapD_(B-C)    GATTTGGGTTGAATTTGCTGCGTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAA                 1106
PsapE_(B-D)    GATTTGGGTTGAATTTGCTGCGTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAA                  917
PsapG_(A-D)    GATTTGGGTTGAATTTGCTGCGTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAA                  720
PsapF_(A-C)    GATTTGGGTTGAATTTGCTGCGTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAA                  909
Consensus      gatttgggttgaatttgctgcgttttggcagattttgagcgagagatattagggaa
```

Figure C.9: Multiple sequence alignment of the XvPSap1D, E, F and G nucleotides generated using DNAMAN software (Lynnon Biosoft). The homology levels are indicated by the shading. Black indicates 100% sequence homology across all of the sequences. Blue indicates 50% sequence homology.

C.10 Nucleotide sequence alignment of XvPSap1D, E, F, G and XvPSap2

```
XvPsap1D    ............................................................     0
XvPsap1E    ............................................................     0
XvPsap1G    ............................................................     0
XvPsap1F    ............................................................     0
XvPSap2     TCGTACGAGAATCGCTGTCCTCTCCAACGAGCAAGATATTCTAAAAAATTATTTTTAAAT    60
Consensus

XvPsap1D    ............................................................     0
XvPsap1E    ............................................................     0
XvPsap1G    ............................................................     0
XvPsap1F    ............................................................     0
XvPSap2     TTATTAAGACTTTGAAATCTCCAAAATAAGAATTTCGAAATATTTTTGAAAAAAATATG   120
Consensus

XvPsap1D    ............................................................     0
XvPsap1E    ............................................................     0
XvPsap1G    ............................................................     0
XvPsap1F    ............................................................     0
XvPSap2     GAATACCAAAGTTATAAATCTCATCAAAATCTATAATATTTATACTTAAATTTTTAAAAA   180
Consensus

XvPsap1D    .............................................ACTGTCTGGGTAGCT    15
XvPsap1E    .............................................ACTGTCTGGGTAGCT    15
XvPsap1G    .............................................ACTGTCTGGGTAGCT    15
XvPsap1F    .............................................ACTGTCTGGGTAGCT    15
XvPSap2     AATACTCCTATATACGTATTTAGTACAATCAATTTACCATAATATACCGATCAAAAGTTG   240
Consensus                                                actgtctgggtagct

XvPsap1D    GGCAATATAGAGACGTAAATAATTGTCTGTAAATAGGGAGAAATTCATGGATCATCACCC    75
XvPsap1E    GGCAATATAGAGACGTAAATAATTGTCTGTAAATAGGGAGAAATTCATGGATCATCACCC    75
XvPsap1G    GGCAATATAGAGACGTAAATAATTGTCTGTAAATAGGGAGAAATTCATGGATCATCACCC    75
XvPsap1F    GGCAATATAGAGACGTAAATAATTGTCTGTAAATAGGGAGAAATTCATGGATCATCACCC    75
XvPSap2     GTTTTGACTGTCTGGGTAGCTGGCAATATAGAGACGTAAGAAATTCATGGATCATCACCC   300
Consensus   ggcaatatagagacgtaaataattgtctgtaaatagggagaaattcatggatcatcaccc

XvPsap1D    TAATTCGGTCTTTCACTCATTTTATCATAGACCTGACTAAAGAACTTGGTCAGAGTTTTT   135
XvPsap1E    TAATTCGGTCTTTCACTCATTTTATCATAGACCTGACTAAAGAACTTGGTCAGAGTTTTT   135
XvPsap1G    TAATTCGGTCTTTCACTCATTTTATCATAGACCTGACTAAAGAACTTGGTCAGAGTTTTT   135
XvPsap1F    TAATTCGGTCTTTCACTCATTTTATCATAGACCTGACTAAAGAACTTGGTCAGAGTTTTT   135
XvPSap2     TAATTCGGTCTTTCACTCATTTTATCCTAGACCTGACTAAAGAACTCGGTCAGAG.TTTT   359
Consensus   taattcggtctttcactcattttatcatagacctgactaaagaacttggtcagagttttt

XvPsap1D    ACTTATTTAAAATAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAG..AAA   193
XvPsap1E    ACTTATTTAAAATAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAG..AAA   193
XvPsap1G    ACTTATTTAAAATAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAG..AAA   193
XvPsap1F    ACTTATTTAAAATAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAG..AAA   193
XvPSap2     ACTTATTTAAAAAAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAGAAAAA   419
Consensus   acttatttaaaataaagaggacttcatggcatccatgtgcaggtacagctcccag  aaa
```

Continued…

```
XvPsap1D    AAAAAGCATGAAACACGAGAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGCAGCTC    253
XvPsap1E    AAAAAGCATGAAACACGAGAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGCAGCTC    253
XvPsap1G    AAAAAGCATGAAACACGAGAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGCAGCTC    253
XvPsap1F    AAAAAGCATGAAACACGAGAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGCAGCTC    253
XvPSap2     AAAAAGCATGAAACACGACAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGGAGCTC    479
Consensus   aaaaagcatgaaacacgagaggatcaatagcattcgatctgaaacaaaaggttgcagctc

XvPsap1D    AAGACTTTCTCCAAAATATTAAGATGATCCAAAGAATTACCCCAAGATATCCAACGTATA    313
XvPsap1E    AAGACTTTCTCCAAAATATTAAGATGATCCAAAGAATTACCCCAAGATATCCAACGTATA    313
XvPsap1G    AAGACTTTCTCCAAAATATTAAGATGATCCAAAGAATTACCCCAAGATATCCAACGTATA    313
XvPsap1F    AAGACTTTCTCCAAAATATTAAGATGATCCAAAGAATTACCCCAAGATATCCAACGTATA    313
XvPSap2     AAGACTTTCTCCAAAATATTAAGATGATCCAAAGAATTACCCCAAGATATCCAACGTATA    539
Consensus   aagactttctccaaaatattaagatgatccaaagaattaccccaagatatccaacgtata

XvPsap1D    CCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCATTCTTTGATTTTTCTCCCGAGTG    373
XvPsap1E    CCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCAT......................    351
XvPsap1G    CCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCAT......................    351
XvPsap1F    CCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCATTCTTTGATTTTTCTCCCGAGTG    373
XvPSap2     CCAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCAT.TCTTTTCCCGAGTGTTCTTTT    598
Consensus   ccaatgtgtataccgaaagtaagaaagttcacgtgcat   tt        t      t

XvPsap1D    TTCTTTTCTGAAATGAGTAAATAAGACTAGAATAAGAGCTAATGTATTTTTTTTCTAAAA    433
XvPsap1E    ............................................................    351
XvPsap1G    ............................................................    351
XvPsap1F    TTCTTTTCTGAAATGAGTAAATAAGACTAGAATAAGAGCTAATGTATTTTTTTTCTAAAA    433
XvPSap2     CTGAAATGAGTAAATAAGACTAGAATAAGAGCTAATGTATTTTTTTTTTCTAAAAAAAAG    658
Consensus    t    t g aa  a  a     a         taa     t  t t ttt t     aaa

XvPsap1D    AAAGTTGAATGTGGATACAATATGATTATACATTCATTAGCTATTTTAAGTATATTCTAT    493
XvPsap1E    ............................................................    351
XvPsap1G    ............................................................    351
XvPsap1F    AAAGTTGAATGTGGATACAATATGATTATACATTCATTAGCTATTTTAAGTATATTCTAT    493
XvPSap2     TTGAATGTGGATACAATATGATTATACATTCATTAGCTATTTTAAGTATATTCTATTTTT    718
Consensus        tg   t  a       t    at catt   ta  t    ta  t     t t t

XvPsap1D    TTTTTTTCCCCCCAAAAGAACACAAATGTGTTCCGTCACTTTCCATG.............    540
XvPsap1E    ............................................................    351
XvPsap1G    ............................................................    351
XvPsap1F    TTTTTTTCCCCCCAAAAGAACACAAATGTGTTCCGTCACTTTCCATG.............    540
XvPSap2     TTTTTCCCCAAAAGAACACAAATGTGTTCCGTCACCTTCCATGGGGCAAATTACAACTTA    778
Consensus   ttttt  cc     aa   a a    t    tc     c  t

XvPsap1D    ............................................................    540
XvPsap1E    ............................................................    351
XvPsap1G    ............................................................    351
XvPsap1F    ............................................................    540
XvPSap2     GGCTTTATCTTAGTTGGTATGATCTTAATTTTATTATACTTTAAACAACTTATCGCTAAT    838
Consensus
```

Continued…

```
XvPsap1D   ............................................................   540
XvPsap1E   ............................................................   351
XvPsap1G   ............................................................   351
XvPsap1F   ............................................................   540
XvPSap2    AATTTTGTTTTAATTTATGCGCCAATTGTGAATATAAATCGGATAAATTTGAGCATTAAT   898
Consensus

XvPsap1D   ............................................................   540
XvPsap1E   ............................................................   351
XvPsap1G   ............................................................   351
XvPsap1F   ............................................................   540
XvPSap2    ACTTTTTAGTCAAAGTTTCAAAAAAGAAAAAAATAACTATGACGTTAGAGTTTGGAATCC   958
Consensus

XvPsap1D   ............................................................   540
XvPsap1E   ............................................................   351
XvPsap1G   ............................................................   351
XvPsap1F   ............................................................   540
XvPSap2    AGTCAAAATGGAACTTATTTTTTTAGTTCATCATAATCAACTTGATGAGATTTTTTGTAC  1018
Consensus

XvPsap1D   ............................................................   540
XvPsap1E   ............................................................   351
XvPsap1G   ............................................................   351
XvPsap1F   ............................................................   540
XvPSap2    TAGACAATCATCTTGAATGATTAATTTAATTCCAATTTGATACTATTAAGAATGAAAACA  1078
Consensus

XvPsap1D   ............................................................   540
XvPsap1E   ............................................................   351
XvPsap1G   ............................................................   351
XvPsap1F   ............................................................   540
XvPSap2    TCCTAAATAATAGCTGCCGTGATCACTTTATAGCACCACATTTATCTTAATTTAAAATAT  1138
Consensus

XvPsap1D   ......TCTTAGAATTGGACAGGGTGCTTATGATAC..AACTTGTTCCTATCAACAACTG   592
XvPsap1E   ......TCTTAGAATTGGACAGGGTGCTTATGATAC..AACTTGTTCCTATCAACAACTG   403
XvPsap1G   ............................................................   351
XvPsap1F   ............................................................   540
XvPSap2    TTTTAATCTTAGAATTGGACAGGGTGCTTATGATAACAAACTTGTTCCTATCAACAGCTG  1198
Consensus        tcttagaattggacagggtgcttatgata   aacttgttcctatcaaca ctg

XvPsap1D   CATGTTAGACAGCGCCGAATTTACAGTCCTACTGGGCGCCACTTTTCAACCCACATCATC   652
XvPsap1E   CATGTTAGACAGCGCCGAATTTACAGTCCTACTGGGCGCCACTTTTCAACCCACATCATC   463
XvPsap1G   ............................................................   351
XvPsap1F   ............................................................   540
XvPSap2    CATGTTAGACAGCGCCGAATTTACAGTCCTACTGGGCGCCACTTTTCAACCCACATCATC  1258
Consensus  catgttagacagcgccgaatttacagtcctactgggcgccacttttcaacccacatcatc
```

Continued…

```
XvPsap1D   AAGATGAACACCACGTTATCTTCATCCGCTCCAACCACATGGTCCAGCGCCACTGGCCAA    712
XvPsap1E   AAGATGAACACCACGTTATCTTCATCCGCTCCAACCACATGGTCCAGCGCCACTGGCCAA    523
XvPsap1G   ............................................................    351
XvPsap1F   ............................................................    540
XvPSap2    AAGCTGAACACCACGTTATCTACATCCACTCCAACCACATGGTCCAGCGGCACTGGCCAA   1318
Consensus  aag tgaacaccacgttatct catcc ctccaaccacatggtccagcg cactggccaa

XvPsap1D   GACCGCCAGCCAGCCAGGCCATCCAACGTGGTGCATTTTCTAACACTCCACGTTCGCTGT    772
XvPsap1E   GACCGCCAGCCAGCCAGGCCATCCAACGTGGTGCATTTTCTAACACTCCACGTTCGCTGT    583
XvPsap1G   .........................ACGTGGTGCATTTTCTAACACTCCACGTTCGCTGT    386
XvPsap1F   .........................ACGTGGTGCATTTTCTAACACTCCACGTTCGCTGT    575
XvPSap2    GACCGCCAGCCAG.CAGGCCATCCAACGTGGTGCA.TTTCTAACACTCCACGTTCGCTGT   1376
Consensus  gaccgccagccag caggccatccaacgtggtgcattttctaacactccacgttcgctgt

XvPsap1D   ACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCGGGCCCGCGG    832
XvPsap1E   ACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCGGGCCCGCGG    643
XvPsap1G   ACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCGGGCCCGCGG    446
XvPsap1F   ACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCGGGCCCGCGG    635
XvPSap2    ACGACGTTATTTCTCCAGCCAGAAAGACCGGGACAGCGACGCTGTTGGGCGGGCCCGCGG   1436
Consensus  acggcattatttctccagccagaaagaccgagacagcgacgctgttgggcgggcccgcgg

XvPsap1D   CCTGCTCTCTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCT.CGTGCCTACGC    891
XvPsap1E   CCTGCTCTCTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCT.CGTGCCTACGC    702
XvPsap1G   CCTGCTCTCTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCT.CGTGCCTACGC    505
XvPsap1F   CCTGCTCTCTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCT.CGTGCCTACGC    694
XvPSap2    CCTGCTCTCTCTGCTTCCCCATGAGATTCACGGGCATCGCTCGTGCCTACGCCCGCGCCC   1496
Consensus  cctgctctctctgcttccccatgagattcacgggcatcgctcctcgct cgtgcctacgc

XvPsap1D   GACCGCGCCGATCCACGTGACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGT    951
XvPsap1E   GACCGCGCCGATCCACGTGACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGT    762
XvPsap1G   GACCGCGCCGATCCACGTGACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGT    565
XvPsap1F   GACCGCGCCGATCCACGTGACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGT    754
XvPSap2    GACCGCGCCGATGCACGTGACGTGGCGCAGCAATCGTTCTTACTAGGCGCTTGCACGTGT   1556
Consensus  gaccgcgccgatccacgtgacgtggcgcagcaatcgttcttactaggcgcttgcacgtgt

XvPsap1D   CGTTCGCATGCGAAGCGTCCACACTGCCAACGACCTCCTTAAATATCCTTGTGATATTCG   1011
XvPsap1E   CGTTCGCATGCGAAGCGTCCACACTGCCAACGACCTCCTTAAATATCCTTGTGATATTCG    822
XvPsap1G   CGTTCGCATGCGAAGCGTCCACACTGCCAACGACCTCCTTAAATATCCTTGTGATATTCG    625
XvPsap1F   CGTTCGCATGCGAAGCGTCCACACTGCCAACGACCTCCTTAAATATCCTTGTGATATTCG    814
XvPSap2    CGTTCGCATGCGAAGCGTCCAGACCGCCAACGACCTCCTTAAATATCCTTGTGATCTTCG   1616
Consensus  cgttcgcatgcgaagcgtccacactgccaacgacctccttaaatatccttgtgatattcg
```

Continued…

```
XvPsap1D    CCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGCAGATTTGGGTTGAATTTGCTGC   1071
XvPsap1E    CCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGCAGATTTGGGTTGAATTTGCTGC    882
XvPsap1G    CCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGCAGATTTGGGTTGAATTTGCTGC    685
XvPsap1F    CCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGCAGATTTGGGTTGAATTTGCTGC    874
XvPSap2     CCTTACGATCTCACACTTCGCACGCAGAGTCCAGTCGCAGATTTGGGTTGAATTTGCTGC   1676
Consensus   ccttacgatctcacacttcgcacgcaaaggccagtcgcagatttgggttgaatttgctgc

XvPsap1D    GTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAAG........................   1107
XvPsap1E    GTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAAG........................    918
XvPsap1G    GTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAAG........................    721
XvPsap1F    GTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAAG........................    910
XvPSap2     GTTTTGGCAGATTTTGAGCGAGAGAGATTAGGGAAGATGAGGAACGAGGGTTTTCTGAAA   1736
Consensus   gttttggcagattttgagcgagagatattagggaag

XvPsap1D    ............................................................   1107
XvPsap1E    ............................................................    918
XvPsap1G    ............................................................    721
XvPsap1F    ............................................................    910
XvPSap2     ATGAAGACCGACGTTGGGGTCGCCGACGAGGTGATCTCCGGAGATCTCAAGCAGCTTGGT   1796
Consensus

XvPsap1D    ....................................                           1107
XvPsap1E    ....................................                            918
XvPsap1G    ....................................                            721
XvPsap1F    ....................................                            910
XvPSap2     GACGCTGCGAAGCGGCTAGCTAACCATGCGATCAAG                           1832
Consensus
```

Figure C.10: Multiple sequence alignment of the XvPSap1D, E, F, G and XvPSap2 nucleotides generated using DNAMAN software (Lynnon Biosoft). The homology levels are indicated by the shading. Black indicates 100% sequence homology across all of the sequences. Pink indicates 75% sequence homology. Blue indicates 50% sequence homology.

University of Cape Town

C.11. Nucleotide sequence alignment of XvPSap1 and XvPSap2

```
XvPsap1   ............................................................    0
XvPSap2   TCGTACGAGAATCGCTGTCCTCTCCAACGAGCAAGATATTCTAAAAAATTATTTTTAAAT   60
Consensus tcgtacgagaatcgctgtcctctccaacgagcaagatattctaaaaaattatttttaaat

XvPsap1   ............................................................    0
XvPSap2   TTATTAAGACTTTGAAATCTCCAAAATAAGAATTTCGAAATATTTTTTGAAAAAAATATG  120
Consensus ttattaagactttgaaatctccaaaataagaatttcgaaatattttttgaaaaaaatatg

XvPsap1   ............................................................    0
XvPSap2   GAATACCAAAGTTATAAATCTCATCAAAATCTATAATATTTATACTTAAATTTTTAAAAA  180
Consensus gaataccaaagttataaatctcatcaaaatctataatatttatacttaaatttttaaaaa

XvPsap1   ............................................................    0
XvPSap2   AATACTCCTATATACGTATTTAGTACAATCAATTTACCATAATATACCGATCAAAAGTTG  240
Consensus aatactcctatatacgtatttagtacaatcaatttaccataatataccgatcaaaagttg

XvPsap1   ......ACTGTCTGGGTAGCTGGCAATATAGAGACGTAAGAAATTCATGGATCATCACCC   54
XvPSap2   GTTTTGACTGTCTGGGTAGCTGGCAATATAGAGACGTAAGAAATTCATGGATCATCACCC  300
Consensus gttttgactgtctgggtagctggcaatatagagacgtaagaaattcatggatcatcaccc

XvPsap1   TAATTCGGTCTTTCACTCATTTTATCCTAGACCTGACTAAAAAACTTGGTCAGAGTTTTA  114
XvPSap2   TAATTCGGTCTTTCACTCATTTTATCCTAGACCTGACTAAAGAACTCGGTCAGAGTTTTA  360
Consensus taattcggtctttcactcattttatcctagacctgactaaaaaactcggtcagagtttta

XvPsap1   CTTATTTAAAAAAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAG.AAAAA  173
XvPSap2   CTTATTTAAAAAAAAGAGGACTTCATGGCATCCATGTGCAGGTACAGCTCCCAGAAAAAA  420
Consensus cttatttaaaaaaaagaggacttcatggcatccatgtgcaggtacagctcccagaaaaaa

XvPsap1   AAAAGCATGAAACACGACAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGGAGCTCA  233
XvPSap2   AAAAGCATGAAACACGACAGGATCAATAGCATTCGATCTGAAACAAAAGGTTGGAGCTCA  480
Consensus aaaagcatgaaacacgacaggatcaatagcattcgatctgaaacaaaaggttggagctca

XvPsap1   AGACTTTCTCCAAAATATTAAGATGATCCAAAGAATTACCCCAAGATATCCAACGTATAC  293
XvPSap2   AGACTTTCTCCAAAATATTAAGATGATCCAAAGAATTACCCCAAGATATCCAACGTATAC  540
Consensus agactttctccaaaatattaagatgatccaaagaattaccccaagatatccaacgtatac
```

Continued…

```
XvPsap1    CAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCATTCTTTGATTTTTCTCCCGAGTGT    353
XvPSap2    CAATGTGTATACCGAAAGTAAGAAAGTTCACGTGCATTCTTT........TCCCGAGTGT    592
Consensus  caatgtgtataccgaaagtaagaaagttcacgtgcattctttgatttttctcccgagtgt

XvPsap1    TCTTTTCTGAAATGAGTAAATAAGACTAGAATAAGAGCTAATGTA..TTTTNTTTCTAAA    411
XvPSap2    TCTTTTCTGAAATGAGTAAATAAGACTAGAATAAGAGCTAATGTATTTTTTTTTCTAAA     652
Consensus  tcttttctgaaatgagtaaataagactagaataagagctaatgtatttttttntttctaaa

XvPsap1    AAAAAGTTGAATGTGGATACAATATGATTATACATTCATTAGCTATTTTAAGTATATTCT    471
XvPSap2    AAAAAGTTGAATGTGGATACAATATGATTATACATTCATTAGCTATTTTAAGTATATTCT    712
Consensus  aaaaagttgaatgtggatacaatatgattatacattcattagctattttaagtatattct

XvPsap1    ATTTTTTNTTTCCCCAAAAGAACACAAATGTGTTCCGTCACTTTCCATGGGGCAAATTAC    531
XvPSap2    ATTTTTTTTTTCCCCAAAAGAACACAAATGTGTTCCGTCACCTTCCATGGGGCAAATTAC    772
Consensus  atttttttntttccccaaaagaacacaaatgtgttccgtcaccttccatggggcaaattac

XvPsap1    AACTTAGGCTTTATCTTAGTTGGTATGATCTTAATTTTATTATACTTTAAACAACTTATC    591
XvPSap2    AACTTAGGCTTTATCTTAGTTGGTATGATCTTAATTTTATTATACTTTAAACAACTTATC    832
Consensus  aacttaggctttatcttagttggtatgatcttaattttattatactttaaacaacttatc

XvPsap1    GCTAATAATTTTGTTTTGATTTATGCGCCAATTGTAAATATAAATCGGATAANTNTGAAC    651
XvPSap2    GCTAATAATTTTGTTTTAATTTATGCGCCAATTGTGAATATAAATCGGATAAATTTGAGC    892
Consensus  gctaataattttgttttaatttatgcgccaattgtaaatataaatcggataaatntgaac

XvPsap1    ATTAATACTTTTAGTC...AGTTTCAAAAAAGAAAAAGATAACTATGACGTTAGAGTTTG    708
XvPSap2    ATTAATACTTTTTAGTCAAAGTTTCAAAAAAGAAAAAAATAACTATGACGTTAGAGTTTG    952
Consensus  attaatacttttaagccaaagtttcaaaaaagaaaaaaataactatgacgttagagtttg

XvPsap1    GAATCCAGTC.AAATGGAACTTA.TTTTTTAGTTCATCAGAATCAACTTGATGAGATTTT    766
XvPSap2    GAATCCAGTCAAAATGGAACTTATTTTTTTAGTTCATCATAATCAACTTGATGAGATTTT   1012
Consensus  gaatccagtcaaaatggaacttatttttttagttcatcagaatcaacttgatgagatttt

XvPsap1    TTGTACTAGACAATCATCTTGAATGATAGNTAGGGGACTTACCAATCAGCCCCCCATATT    826
XvPSap2    TTGTACTAGACAATCATCTTGAATGATTA...............................   1041
Consensus  ttgtactagacaatcatcttgaatgataantaggggacttaccaatcagccccccatatt

XvPsap1    TTNAAACTTTCAACGCGCCCCTCAACATTTTTCTCTTTCAACGCGCCCNTCCATCACTTT    886
XvPSap2    ............................................................   1041
Consensus  ttnaaactttcaacgcgcccctcaacatttttctctttcaacgcgcccntccatcacttt

XvPsap1    TCTCTTTTCATCGCCCCTCCCTCAACTTTTTGGTCGGACGGAAATACCTCTATATATATT    946
XvPSap2    ............................................................   1041
Consensus  tctcttttcatcgcccctccctcaactttttggtcggacggaaatacctctatatatatt

XvPsap1    TCATCATTTCGACCCCCAAAACACCTATTTAGAAGTATTTTTTGGAAAAAATTTTAACAT   1006
XvPSap2    ............................................................   1041
Consensus  tcatcatttcgacccccaaaacacctatttagaagtattttttggaaaaaattttaacat
```

Continued…

```
XvPsap1    GAAAGTTTTAGATCTTGATGAGATCTACAATTTTTATGTTGAAAGTTTTTCCAAAAAATA   1066
XvPSap2    ............................................................   1041
Consensus  gaaagttttagatcttgatgagatctacaatttttatgttgaaagtttttccaaaaaata

XvPsap1    CTTTTAGATAGATATTTTGGAACTCCGAAGTGTTAATGGGTCGACCCGCTAACTTGCGGA   1126
XvPSap2    ............................................................   1041
Consensus  cttttagatagatattttggaactccgaagtgttaatgggtcgacccgctaacttgcgga

XvPsap1    AATAGAAAAACATCAAGATCTACAACTTTTATGTTGAAAGTTTTTTCAAAAAATACTTCT   1186
XvPSap2    ............................................................   1041
Consensus  aatagaaaaacatcaagatctacaacttttatgttgaaagttttttcaaaaaatacttct

XvPsap1    AGATAGATATTTGATTTGTAATTTTAATGTTGAAAGTTTTTTCAAAAAATACTTCTAGAT   1246
XvPSap2    ............................................................   1041
Consensus  agatagatatttgatttgtaattttaatgttgaaagttttttcaaaaaatacttctagat

XvPsap1    AGATACTTTGGGGCTCCGGAGTGTTAACAAGTATAGGAATATTTTTGTCTGCAAAAAATT   1306
XvPSap2    ............................................................   1041
Consensus  agatactttggggctccggagtgttaacaagtataggaatatttttgtctgcaaaaaatt

XvPsap1    AATTTTTCGGACAAGAGGCCGATCAGTAAGGAATCTGGTCGGAGGGGCTGTTCGGCAATA   1366
XvPSap2    ............................................................   1041
Consensus  aatttttcggacaagaggccgatcagtaaggaatctggtcggaggggctgttcggcaata

XvPsap1    TAAGTTCAGATAGGAGAACTGATCGGATATTTTTCCTTAATTTAATTCCAATTTGATACT   1426
XvPSap2    ............................ATTTAATTCCAATTTGATACTATTAAGA....   1069
Consensus  taagttcagataggagaactgatcggatatttaaccccaatttaataccaataagatact

XvPsap1    ATTAAGAATGAAAACATCCTAATAATTGTGACCACTTTATAGCACCACATTTATTTTAAT   1486
XvPSap2    ATGAAAACATCCTAAATAATAGCTGCCGTGATCACTTTATAGCACCACATTTATCTTAAT   1129
Consensus  atgaaaaagaaaaaataataacaaccgtgaccactttatagcaccacatttatcttaat

XvPsap1    TT.AAATCTTTTAAATCTTAGAATTGGACAGGGTGCTTATGATAACAAACTTGTTCCTAT   1545
XvPSap2    TTAAAATATTTTTAATCTTAGAATTGGACAGGGTGCTTATGATAACAAACTTGTTCCTAT   1189
Consensus  ttaaaatattttaaatcttagaattggacagggtgcttatgataacaaacttgttcctat

XvPsap1    CAACAACTGCATGTTAGACAGCGCCGAATTTACAGTCCTACTGGGCGCCACTTTTCAACC   1605
XvPSap2    CAACAGCTGCATGTTAGACAGCGCCGAATTTACAGTCCTACTGGGCGCCACTTTTCAACC   1249
Consensus  caacaactgcatgttagacagcgccgaatttacagtcctactgggcgccacttttcaacc

XvPsap1    CACATCATCAAGATGAACACCACGTTATCTTCATCCGCTCCAACCACATGGTCCAGCGCC   1665
XvPSap2    CACATCATCAAGCTGAACACCACGTTATCTACATCCACTCCAACCACATGGTCCAGCGGC   1309
Consensus  cacatcatcaagatgaacaccacgttatctacatccactccaaccacatggtccagcgcc

XvPsap1    ACTGGCCAAGACCGCCAGCCAGCCAGGCCATCCAACGTGGTGCATTTTCTAACACTCCAC   1725
XvPSap2    ACTGGCCAAGACCGCCAGCCAG.CAGGCCATCCAACGTGGTGCA.TTTCTAACACTCCAC   1367
Consensus  actggccaagaccgccagccagccaggccatccaacgtggtgcattttctaacactccac
```

Continued…

```
XvPsap1    GTTCGCTGTACGGCATTATTTCTCCAGCCAGAAAGACCGAGACAGCGACGCTGTTGGGCG   1785
XvPSap2    GTTCGCTGTACGACGTTATTTCTCCAGCCAGAAAGACCGGGACAGCGACGCTGTTGGGCG   1427
Consensus  gttcgctgtacgacattatttctccagccagaaagaccgagacagcgacgctgttgggcg

XvPsap1    GGCCCGCGGCCTGCTCTCTCTGCTTCCCCATGAGATTCACGGGCATCGCTCCTCGCTCGT   1845
XvPSap2    GGCCCGCGGCCTGCTCTCTCTGCTTCCCCATGAGATTCACGGGCATCGCTCGT.......   1480
Consensus  ggcccgcggcctgctctctctgcttccccatgagattcacgggcatcgctcctcgctcgt

XvPsap1    GCCTACGCCCGCGCCCGCGCCCGACCGCGCCGGTCCACGTGACGTGGCGCAGCAATCGTT   1905
XvPSap2    ......GCCTACGCCCGCGCCCGACCGCGCCGATGCACGTGACGTGGCGCAGCAATCGTT   1534
Consensus  gcctacgcccacgcccgcgcccgaccgcgccgatccacgtgacgtggcgcagcaatcgtt

XvPsap1    CTTACTAGGCGCTTGCACGTGTCGTTCGCATGCGAAGCGTCCACACTGCCAACGACCTCC   1965
XvPSap2    CTTACTAGGCGCTTGCACGTGTCGTTCGCATGCGAAGCGTCCAGACCGCCAACGACCTCC   1594
Consensus  cttactaggcgcttgcacgtgtcgttcgcatgcgaagcgtccacaccgccaacgacctcc

XvPsap1    TTAAATATCCTTGTGATATTCGCCTTACGATCTCACACTTCGCACGCAAAGGCCAGTCGC   2025
XvPSap2    TTAAATATCCTTGTGATCTTCGCCTTACGATCTCACACTTCGCACGCAGAGTCCAGTCGC   1654
Consensus  ttaaatatccttgtgatattcgccttacgatctcacacttcgcacgcaaaggccagtcgc

XvPsap1    AGATTTGGGTTGAATTTGCTGCGTTTTGGCAGATTTTGAGCGAGAGATATTAGGGAAG..   2083
XvPSap2    AGATTTGGGTTGAATTTGCTGCGTTTTGGCAGATTTTGAGCGAGAGAGATTAGGGAAGAT   1714
Consensus  agatttgggttgaatttgctgcgttttggcagattttgagcgagagagattagggaagat

XvPsap1    ............................................................   2083
XvPSap2    GAGGAACGAGGGTTTTCTGAAAATGAAGACCGACGTTGGGGTCGCCGACGAGGTGATCTC   1774
Consensus  gaggaacgagggttttctgaaaatgaagaccgacgttggggtcgccgacgaggtgatctc

XvPsap1    ..........................................................     2083
XvPSap2    CGGAGATCTCAAGCAGCTTGGTGACGCTGCGAAGCGGCTAGCTAACCATGCGATCAAG     1832
Consensus  cggagatctcaagcagcttggtgacgctgcgaagcggctagctaaccatgcgatcaag
```

Univers

Figure C.11: Multiple sequence alignment of the XvPSap1D and XvPSap2 nucleotides generated using DNAMAN software (Lynnon Biosoft). The homology levels are indicated by the shading. Black indicates 100% sequence homology across all of the sequences. Yellow indicates no homology.

Appendix D: Primers and accession numbers

Table D.1: Accession numbers of reference genes used in this study

Accession Number	Gene	Species	Reference
AJ236016	18S rRNA	*N. tabacum*	Schmidt and Delaney, 2010
AF120093	*EF-1α*	*N. tabacum*	Schmidt and Delaney, 2010
L18908	*L25* ribosomal protein	*N. tabacum*	Schmidt and Delaney, 2010

Table D.2: Primers for mutagenesis used in this study were designed using DNAMAN (Lynnon Biosoft Copyright© 1994-2001). These primers were designed in this study.

Name	Sequence	Target	Product
Primer B	5'-CAGATCTATCTTAGAATTGGACAGGGTG-3'	XvPSap1	XvPSap1D
Primer C	5'-AAGATCTCCATGGAAAGTGACGGAAC-3'	XvPSap1	
Primer B	5'-CAGATCTATCTTAGAATTGGACAGGGTG-3'	XvPSap1	XvPSap1E
Primer D	5'-AAGATCTAATGCACGTGAACTTTCTTAC-3'	XvPSap1	
Primer A	5'-CAGATCTACGTGGTGCATTTTCTAACAC-3'	XvPSap1	XvPSap1F
Primer C	5'-AAGATCTCCATGGAAAGTGACGGAAC-3'	XvPSap1	
Primer A	5'-CAGATCTACGTGGTGCATTTTCTAACAC-3'	XvPSap1	XvPSap1G
Primer D	5'-AAGATCTAATGCACGTGAACTTTCTTAC-3'	XvPSap1	

University of Cape Town

Table D.3: Primer sets for screening used in this study.

Name	Sequence	Target	Reference
*Eco*RI-XvPSap1-F	5'-GGAATTCACTGTCTGGTAGCTGG-3'	XvPSap1	(Okoth, 2009)
*Bam*HI-XvPSap1-R	5'-TCCGGATCCTCCCTAATATCTCTCGCTC-3'	XvPSap1	(Okoth, 2009)
*Eco*R1-XvPSap2-F	5'-AGAATTCATCGTACGAGAAACCCTGTC-3'	XvPSap2	(This study)
*Bam*HI-XvPSap2R-R	5'-AGGATCCGCTTGATGGCATGGTTAG-3'	XvPSap2	(This study)
*Eco*RI-XvPSap1-F	5'-GGAATTCACTGTCTGGTAGCTGG-3'	XvPSap1	(Okoth, 2009)
luc-R3	5'-AGCAGCGCACTTTGAATCTT-3'	*luc* gene	(Okoth, 2009)
*Eco*R1-XvPSap2-F	5'-AGAATTCATCGTACGAGAAACCCTGTC-3'	XvPSap2	(This study)
luc-R3	5'-AGCAGCGCACTTTGAATCTT-3'	*luc* gene	(Okoth, 2009)
BarI	5'-GGTCTGCACCATCGTCAACC-3'	*bar* gene	(Okoth, 2009)
BarII	5'-GTCATGCCAGTTCCCGTGCT-3'	*bar* gene	(Okoth, 2009)

University of Cape Town

Table D.4: Primer sets for quantitative real time PCR used in this study were designed using Primer3 (Rozen and Skaletsky, 2000).These primers were designed in this study.

Name	Sequence	Target
RT-*luc*-F2	5'-ATCCAGAAGCCACCAACGCCTTG-3'	*luc* gene
RT-*luc*-R2	5'-CGAAGATGTTGGCGTGTTGGAGC-3'	*luc* gene
RT-18S-F3	5'-GCAGCGGAAGTTTGAGGCAATAAC-3'	18S rRNA
RT-18S-R3	5'-CGACCTGATGACTCGCGCTTAC-3'	18S rRNA
RT-Rib-F3	5'-GCAGATTCAGGACAACAACACCCTTG-3'	L25 ribosomal protein
RT-Rib-R3	5'-TTGTTGGCAACGTCCGAAGCATC-3'	L25 ribosomal protein
RT-EF-F1	5'-GGTCACCAAGGCTGCTCAGAAG-3'	Elongation factor 1-α
RT-EF-R1	5'-GCCTGTCAACCACCCAGCTC-3'	Elongation factor 1-α

Appendix E: General protocols

E.1. Standard PCR reaction

The PCR amplification reactions were set up in a total of either 25 µl or 50 µl volumes. Unless otherwise stated; Supertherm Polymerase (Bertec Enterprise, Taiwan), standard buffer and dNTP mixture (Fermentas, Canada) were used.

Table E.1: PCR components and final concentrations used in a standard PCR protocol

Component	Final concentration
dH_2O	to final volume
Standard buffer	1X
Primer concentration	0.4 µM
dNTP mixture	0.2 mM
Taq polymerase	0.04 U/µl
Template DNA	5 ng/µl

E.2. Klenow reaction

DNA was treated with Klenow Fragment exo^- (Fermentas, Canada), to facilitate blunt end cloning. Two micrograms of linear DNA was added to 2 µl of 10X reaction buffer, 2.5 µl of 10mM dNTP mixture (Fermentas, Canada), 0.2 µl of the Klenow Fragment exo^- and dH_20 to a final volume of 20 µl. The reaction mixture was incubated for 15 min at 30ºC. The reaction was stopped by heat inactivation for 10 min at 75ºC.

University of Cape Town

E.3. Standard blunt end ligation

Linearised vector was religated in a reaction volume of 20 µl. Each ligation reaction contained 10 U of T4 DNA ligase and Ligase Buffer (New England Biolabs, USA) at a final concentration of 1X. The ligation reaction was mixed gently, briefly centrifuged and then incubated for 2 hours at room temperature.

E.4. Site specific cohesive end ligation reaction

Purified DNA fragments (insert DNA) were ligated to linearised vector in a reaction volume of 20 µl. Each ligation reaction contained 10 U of T4 DNA ligase and Ligase Buffer (New England Biolabs, USA) at a final concentration of 1X. The ligation reaction was mixed gently, briefly centrifuged and then incubated for 2 hours at room temperature.

E.5. PCR product purification

PCR product purification was performed using the EZ-10 Spin Column PCR Purification Kit (Bio Basic Inc, Canada). The PCR reaction mixture was transferred to a 1.5 ml Eppendorf tube and 3 volumes of Binding Buffer I was added. The mixture was then added to a column and allowed to incubate for 2 min at room temperature. The column was then centrifuged for 2 min at 10 000 x *g*. The flow through was discarded and the column was placed back into the Eppendorf tube. Five hundred microlitres of Wash Solution was added and the column centrifuged at for 2 min at 10 000 x *g*. The flow through was discarded and wash step was repeated. The column was then transferred to a clean Eppendorf tube, 30 µl of prewarmed Elution Buffer was added to the centre part of the column and incubated for 2 min at room temperature. The DNA was eluted off the column by centrifuging for 1 min at 10 000 x *g*. Purified PCR products were stored at 4ºC until further use.

E.6. DNA purification from agarose gels

Purification of DNA from agarose gels was performed using the Wizard SV Gel Purification Kit (Promega Corporation, USA). Following electrophoresis, the DNA

University of Cape Town

fragment was excised from the agarose gel using a sterile blade. The excised DNA-gel fragment was placed into a sterile Eppendorf tube and the mass was determined. For every 10 mg of excised gel, 10 µl of Membrane Binding Solution was added and incubated at 60ºC in a heating block to completely dissolve the agarose. The Eppendorf tube was vortexed every 2 to 3 min. The dissolved agarose was added to a SV minicolumn into a collection tube and incubated for 1 min at room temperature, followed by centrifugation for 1 min at 16000 x *g*. The flow through was discarded and the SV minicolumn placed back into the collection tube. Seven hundred microlitres of Membrane Wash Solution (with added ethanol) was added to the column, centrifuged for 1 min at 16000 x *g* and the flow through was discarded. Five hundred microlitres of Membrane Wash Buffer was then added to the SV minicolumn and centrifuged for 5 min at 16000 x *g*. An additional centrifugation for 1 min at 16 000 x *g* was performed and the minicolumn was then transferred to a clean, sterile 1.5 ml Eppendorf tube. Fifty microlitres of prewarmed Nuclease-Free Water was added to the centre part of the column, incubated for 1 min at 37ºC and centrifuged for 1 min at 16000 x *g* to elute the DNA. The SV minicolumn was discarded and the purified DNA was stored at 4ºC.

University of Cape Town

E.7. Plasmid DNA isolation

Plasmid DNA was isolated using the the Bioflux Plasmid DNA extraction and Purification Kit (Bioer, Japan). Five millilitres of bacterial cells were harvested by centrifugation for 1 min at room temperature at 10000 x *g*. The supernatant was discarded and pelleted bacterial cells was resuspended in 250 µl of Resuspension Buffer supplemented with 100 µg/ml RNAse A. Two hundred and fifty microlitres of Lysis Buffer was then added and the mixture was inverted gently 4 to 6 times and incubated for 4 min at room temperature. Three hundred and fifty microlitres of Neutralization Buffer was added, mixed by gentle inversion. Bacterial suspension was centrifuged for 5 min at 14000 x *g* at room temperature. Nine hundred microlitres of supernatant was added to a Spin column and centrifuged for 1 min at 6000 x *g* at room temperature. The flow through was discarded and 650 µl of Wash Buffer (with added ethanol) was added to the Spin column and centrifuged for 1 min at 12000 x *g* at room temperature. The flow through was discarded and the wash step was repeated. The flow through was discarded and the Spin column centrifuged for an

additional 1 min at 12000 x *g* at room temperature. The Spin column was transferred to a clean, sterile 1.5 ml Eppendorf tube, 50 µl of prewarmed Elution Buffer was added to the centre of the column and allowed to incubate for 5 min at room temperature. The Spin column was centrifuged for 1 min at 12000 x *g* at room temperature, discarded and the plasmid DNA stored at 4ºC.

E.8. Preparation of glycerol stocks

Following overnight culture, 200 µl of cells was added to 800 µl of autoclaved 50% glycerol and mixed gently. Glycerol stocks were at -80ºC.

University of Cape Town

Appendix F: Media and solutions

F.1 LB media

5 g Yeast Extract

10 g Tryptone

5 g NaCl

Made up to 1 litre with distilled water

Autoclave

F.2 LB agar

5 g Yeast Extract

10 g Tryptone

5 g NaCl

15 g Bacterial agar

Made up to 1 litre with distilled water

Autoclave

F.3 YEP media

5 g Yeast Extract

10 g Peptone

5 g NaCl

Made up to 1 litre with distilled water

Autoclave

F.4 YEP agar

5 g Yeast Extract

10 g Peptone

5 g NaCl

15 g Bacterial agar

Made up to 1 litre with distilled water

Autoclave

University of Cape Town

F.5. B5 vitamins

0.1 g Nicotinic acid

1 g Thiamaine-HCl

0.1 g Pyridoxine-HCl

10 g myo-inositol

Made up to 1 litre with sterile distilled water

Filter sterilise

University of Cape Town

Appendix G: XvPSap regulatory elements

G.1. Putative regulatory elements in the XvPSap2 promoter

Table G.1. Putative regulatory elements identified in the XvPSap2 sequence using the PLANTCARE software (Lescot *et al.*, 2002).

Site name	Organism	Position	Strand	Matrix Score	Sequence	Function
ABRE	Arabidopsis thaliana	343		6	CACGTG	cis-acting element involved in the abscisic acid responsiveness
ABRE	Hordeum vulgare	769	+	9	CGCACGTGTC	cis-acting element involved in the abscisic acid responsiveness
ABRE	Arabidopsis thaliana	731	+	6	CACGTG	cis-acting element involved in the abscisic acid responsiveness
ABRE	Arabidopsis thaliana	771	+	6	CACGTG	cis-acting element involved in the abscisic acid responsiveness
ABRE	Hordeum vulgare	710	-	9	CCGCGTAGGC	cis-acting element involved in the abscisic acid responsiveness
ABRE	Arabidopsis thaliana	737	+	7	ACGTGGC	cis-acting element involved in the abscisic acid responsiveness
ACE	Petroselinum crispum	561	-	9	ACTACGTTGG	cis-acting element involved in light responsiveness
ACE	Petroselinum hortense	729	-	7	ACGTGGA	cis-acting element involved in light responsiveness
ACE	Petroselinum hortense	585	-	7	ACGTGGA	cis-acting element involved in light responsiveness
CAAT-box	Arabidopsis thaliana	16	+	6	gGCAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Arabidopsis thaliana	374	-	5	CCAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	218	+	4	CAAT	common cis-acting element in promoter and enhancer

Site name	Organism	Position	Strand	Matrix Score	Sequence	Function
CAAT-box	Brassica rapa	889	-	5	CAAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Glycine max	36	-	5	CAATT	common cis-acting element in promoter and enhancer regions
CAAT-box	Petunia hybrida	802	+	7	TGCCAAC	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	315	+	4	CAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	748	+	4	CAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	18	+	4	CAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Brassica rapa	878	-	5	CAAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Arabidopsis thaliana	314	+	5	CCAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	37	-	4	CAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Glycine max	373	-	5	CAATT	common cis-acting element in promoter and enhancer regions
CAT-box	Arabidopsis thaliana	456	+	6	GCCACT	cis-acting regulatory element related to meristem expression
CAT-box	Arabidopsis thaliana	527	+	6	GCCACT	cis-acting regulatory element related to meristem expression
CATT-motif	Zea mays	223	+	6	GCATTC	part of a light responsive element
CGTCA-motif	Hordeum vulgare	735	-	5	CGTCA	cis-acting regulatory element involved in the MeJA-responsiveness

Site name	Organism	Position	Strand	Matrix Score	Sequence	Function
MBS	Arabidopsis thaliana	413	+	6	CAACTG	MYB binding site involved in drought-inducibility
O2-site	Zea mays	159	-	9	GATGACATGA	cis-acting regulatory element involved in zein metabolism regulation
O2-site	Zea mays	469	-	10	GATGATGTGG	cis-acting regulatory element involved in zein metabolism regulation
P-box	Oryza sativa	238	-	7	CCTTTTG	gibberellin-responsive element
Sp1	Oryza sativa	645	+	6	GGGCGG	light responsive element
TATA-box	Arabidopsis thaliana	19	-	9	tcTATATAtt	core promoter element around -30 of transcription start
TATA-box	Lycopersicon esculentum	143	-	5	TTTTA	core promoter element around -30 of transcription start
TATA-box	Lycopersicon esculentum	132	+	5	TTTTA	core promoter element around -30 of transcription start
TATA-box	Arabidopsis thaliana	322	+	4	TATA	core promoter element around -30 of transcription start
TATA-box	Nicotiana tabacum	41	+	11	tcTATAAATAgg	core promoter element around -30 of transcription start
TATA-box	Arabidopsis thaliana	310	+	4	TATA	core promoter element around -30 of transcription start
TATA-box	Arabidopsis thaliana	139	+	8	TATTTAAA	core promoter element around -30 of transcription start
TATA-box	Glycine max	923	-	5	TAATA	core promoter element around -30 of transcription start
TATA-box	Arabidopsis thaliana	21	+	4	TATA	core promoter element around -30 of transcription start
TATA-box	Glycine max	270	-	5	TAATA	core promoter element around -30 of transcription start
TATA-box	Lycopersicon esculentum	95	+	5	TTTTA	core promoter element around -30 of transcription start
TATA-box	Helianthus annuus	320	-	6	TATACA	core promoter element around -30 of transcription start

Site name	Organism	Position	Strand	Matrix Score	Sequence	Function
TC-rich repeats	Nicotiana tabacum	257	+	9	ATTTTCTCCA	cis-acting element involved in defense and stress responsiveness
TC-rich repeats	Nicotiana tabacum	573	+	9	ATTCTCTAAC	cis-acting element involved in defense and stress responsiveness
TCCC-motif	Spinacia oleracea	50	-	7	TCTCCCT	part of a light responsive element
TCT-motif	Arabidopsis thaliana	332	-	6	TCTTAC	part of a light responsive element
TCT-motif	Arabidopsis thaliana	755	+	6	TCTTAC	part of a light responsive element
TGA-box	Glycine max	735	+	9	TGACGTGGC	part of an auxin-responsive element
TGA-element	Brassica oleracea	776	-	6	AACGAC	auxin-responsive element
TGA-element	Brassica oleracea	806	+	6	AACGAC	auxin-responsive element
TGACG-motif	Hordeum vulgare	735	+	5	TGACG	cis-acting regulatory element involved in the MeJA-responsiveness
box II	Petroselinum hortense	729	+	9	TCCACGTGGC	part of a light responsive element
circadian	Lycopersicon esculentum	295	+	9	CAAAGATATC	cis-acting regulatory element involved in circadian control

G.2. Putative regulatory elements in the XvPSap1G promoter

Table G.2. Putative regulatory elements identified in the XvPSap1G sequence using the PLANTCARE software (Lescot *et al*., 2002).

Site	Organism	Position	Strand	Matrix score	Sequence	Function
ABRE	Arabidopsis thaliana	343	+	6	CACGTG	cis-acting element involved in the abscisic acid responsiveness
ABRE	Arabidopsis thaliana	525	+	7	ACGTGGC	cis-acting element involved in the abscisic acid responsiveness
ABRE	Hordeum vulgare	498	-	9	CCGCGTAGGC	cis-acting element involved in the abscisic acid responsiveness
ABRE	Arabidopsis thaliana	559	+	6	CACGTG	cis-acting element involved in the abscisic acid responsiveness
ABRE	Arabidopsis thaliana	351	+	6	TACGTG	cis-acting element involved in the abscisic acid responsiveness
ABRE	Hordeum vulgare	557	+	9	CGCACGTGTC	cis-acting element involved in the abscisic acid responsiveness
ABRE	Arabidopsis thaliana	519	+	6	CACGTG	cis-acting element involved in the abscisic acid responsiveness
ACE	Petroselinum hortense	373	-	7	ACGTGGA	cis-acting element involved in light responsiveness
ACE	Petroselinum hortense	517	-	7	ACGTGGA	cis-acting element involved in light responsiveness
CAAT-box	Arabidopsis thaliana	16	+	6	GGCAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	315	+	4	CAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	218	+	4	CAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Brassica rapa	677	-	5	CAAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Glycine max	36	-	5	CAATT	common cis-acting element in promoter and enhancer regions
CAAT-box	Petunia hybrida	590	+	7	TGCCAAC	common cis-acting element in promoter and enhancer regions
CAAT-box	Arabidopsis thaliana	314	+	5	CCAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	536	+	4	CAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	18	+	4	CAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Brassica rapa	666	-	5	CAAAT	common cis-acting element in promoter and enhancer regions
CAAT-box	Hordeum vulgare	37	-	4	CAAT	common cis-acting element in promoter and enhancer regions

Site Name	Organism	Position	Strand	Matrix score.	Sequence	Function
CATT-motif	Zea mays	223	+	6	GCATTC	part of a light responsive element
CGTCA-motif	Hordeum vulgare	523	-	5	CGTCA	cis-acting regulatory element involved in the MeJA-responsiveness
G-Box	Pisum sativum	343	+	6	CACGTG	cis-acting regulatory element involved in light responsiveness
G-Box	Pisum sativum	559	+	6	CACGTG	cis-acting regulatory element involved in light responsiveness
G-Box	Pisum sativum	375	+	6	CACGTT	cis-acting regulatory element involved in light responsiveness
G-Box	Antirrhinum majus	351	-	6	CACGTA	cis-acting regulatory element involved in light responsiveness
G-Box	Pisum sativum	519	+	6	CACGTG	cis-acting regulatory element involved in light responsiveness
G-box	Solanum tuberosum	168	-	7	CACATGG	cis-acting regulatory element involved in light responsiveness
G-box	Arabidopsis thaliana	519	+	6	CACGTG	cis-acting regulatory element involved in light responsiveness
G-box	Arabidopsis thaliana	517	-	9	GCCACGTGGA	cis-acting regulatory element involved in light responsiveness
G-box	Zea mays	524	-	6	CACGTC	cis-acting regulatory element involved in light responsiveness
G-box	Daucus carota	351	+	6	TACGTG	cis-acting regulatory element involved in light responsiveness
G-box	Brassica napus	523	+	8	TGACGTGG	cis-acting regulatory element involved in light responsiveness
G-box	Brassica napus	518	-	7	CACGTGG	cis-acting regulatory element involved in light responsiveness
G-box	Arabidopsis thaliana	559	+	6	CACGTG	cis-acting regulatory element involved in light responsiveness
G-box	Arabidopsis thaliana	343	+	6	CACGTG	cis-acting regulatory element involved in light responsiveness
G-box	Zea mays	375	+	6	CACGTT	cis-acting regulatory element involved in light responsiveness

Site Name	Organism	Position	Strand	Matrix score.	Sequence	Function
LTR	Hordeum vulgare	326	+	6	CCGAAA	cis-acting element involved in low-temperature responsiveness
O2-site	Zea mays	159	-	9	GATGACATGA	cis-acting regulatory element involved in zein metabolism regulation
P-box	Oryza sativa	238	-	7	CCTTTTG	gibberellin-responsive element
Sp1	Oryza sativa	433	+	6	GGGCGG	light responsive element
TATA-box	Arabidopsis thaliana	19	-	9	tcTATATAtt	core promoter element around -30 of transcription start
TATA-box	Lycopersicon esculentum	143	-	5	TTTTA	core promoter element around -30 of transcription start
TATA-box	Lycopersicon esculentum	132	+	5	TTTTA	core promoter element around -30 of transcription start
TATA-box	Arabidopsis thaliana	322	+	4	TATA	core promoter element around -30 of transcription start
TATA-box	Nicotiana tabacum	41	+	11	tcTATAAATAgg	core promoter element around -30 of transcription start
TATA-box	Arabidopsis thaliana	310	+	4	TATA	core promoter element around -30 of transcription start
TATA-box	Arabidopsis thaliana	139	+	8	TATTTAAA	core promoter element around -30 of transcription start
TATA-box	Glycine max	711	-	5	TAATA	core promoter element around -30 of transcription start
TATA-box	Arabidopsis thaliana	21	+	4	TATA	core promoter element around -30 of transcription start
TATA-box	Glycine max	270	-	5	TAATA	core promoter element around -30 of transcription start
TATA-box	Lycopersicon esculentum	95	+	5	TTTTA	core promoter element around -30 of transcription start
TATA-box	Helianthus annuus	320	-	6	TATACA	core promoter element around -30 of transcription start

Site Name	Organism	Position	Strand	Matrix score.	Sequence	Function
TC-rich repeats	Nicotiana tabacum	257	+	9	ATTTTCTCCA	cis-acting element involved in defense and stress responsiveness
TC-rich repeats	Nicotiana tabacum	361	+	9	ATTCTCTAAC	cis-acting element involved in defense and stress responsiveness
TCCC-motif	Spinacia oleracea	50	-	7	TCTCCCT	part of a light responsive element
TCT-motif	Arabidopsis thaliana	332	-	6	TCTTAC	part of a light responsive element
TCT-motif	Arabidopsis thaliana	543	+	6	TCTTAC	part of a light responsive element
TGA-box	Glycine max	523	+	9	TGACGTGGC	part of an auxin-responsive element
TGA-element	Brassica oleracea	564	-	6	AACGAC	auxin-responsive element
TGA-element	Brassica oleracea	594	+	6	AACGAC	auxin-responsive element
TGACG-motif	Hordeum vulgare	523	+	5	TGACG	cis-acting regulatory element involved in the MeJA-responsiveness
box II	Petroselinum hortense	517	+	9	TCCACGTGGC	part of a light responsive element
circadian	Lycopersicon esculentum	295	+	9	CAAAGATATC	cis-acting regulatory element involved in circadian control

Appendix H: Quantitative Real time analysis

H.1. RNA analysis

The quality and integrity of the isolated RNA was assessed on a RNA-6000 Nano chip using the Agilent 2100 Bioanalyzer and analysed using the Agilent 2100 Expert software (Anatech).

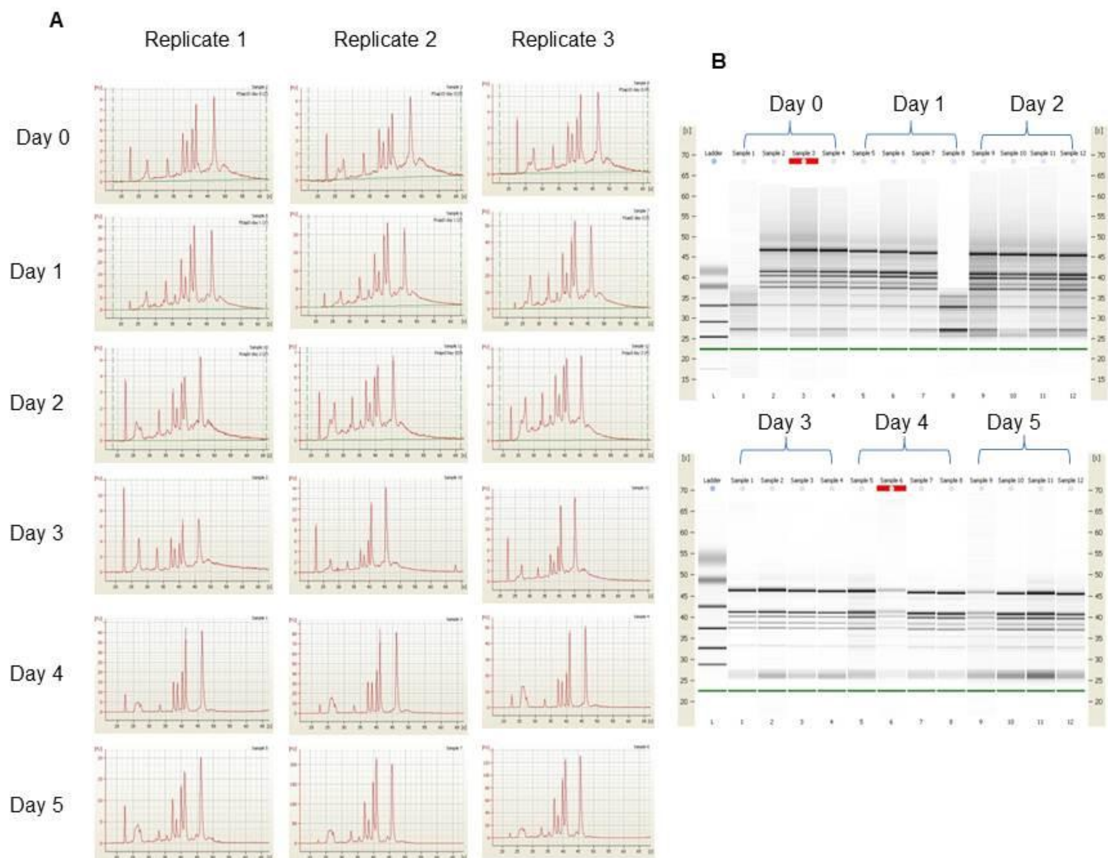


Figure H.1A: Analysis of RNA isolated from XvPSap1D plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. Four RNA isolations were performed per time point. **A**. Electropherogram indicating the three best RNA samples per time point. **B**. The resulting gel image for all RNA samples

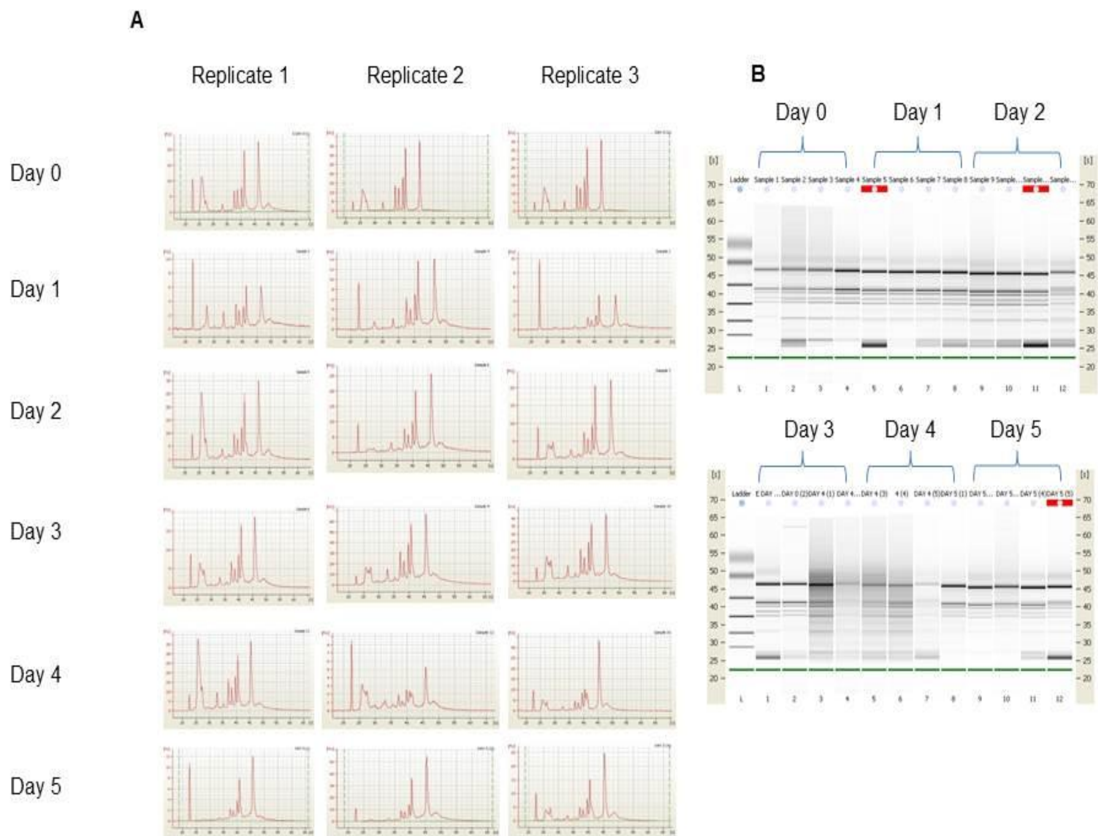


Figure H.1B: Analysis of RNA isolated from XvPSap1E plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. Four RNA isolations were performed per time point. **A**. Electropherogram indicating the three best RNA samples per time point. **B**. The resulting gel image for all RNA samples

University

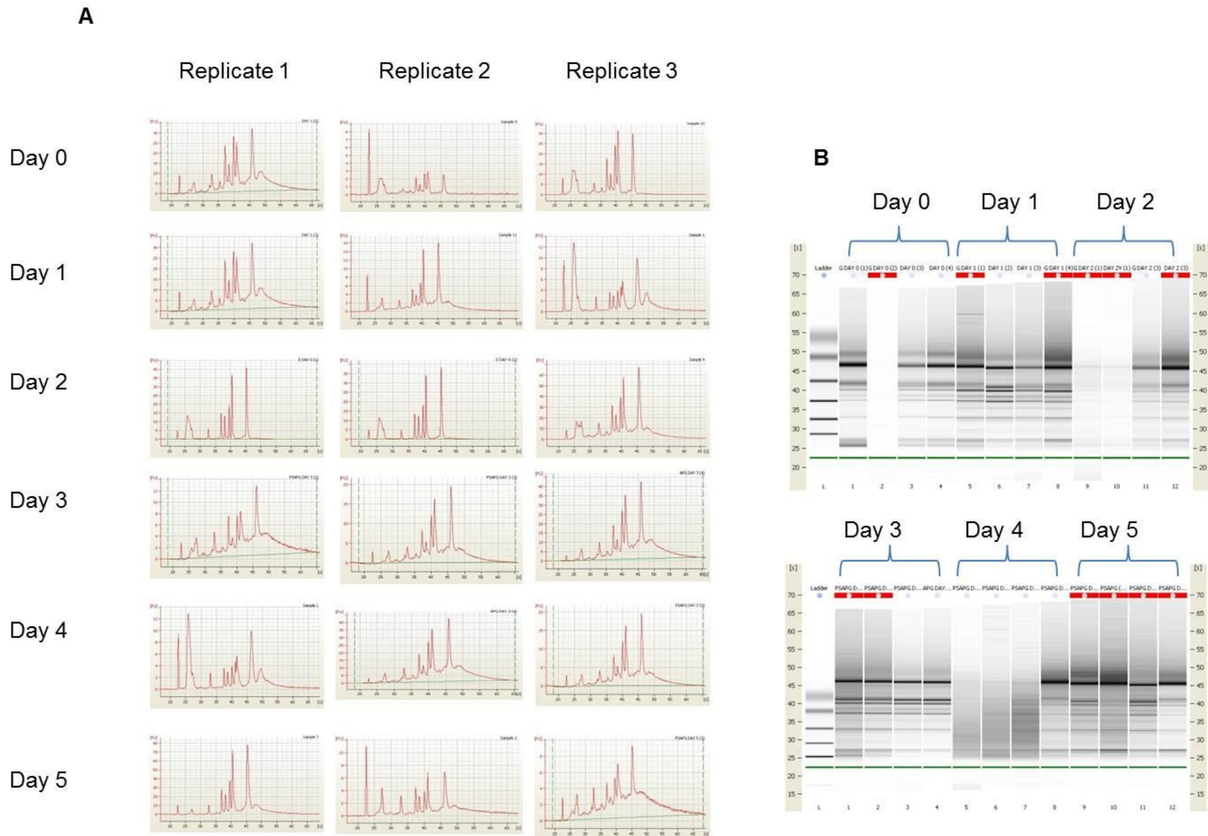


Figure H.1C: Analysis of RNA isolated from XvPSap1G plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. Four RNA isolations were performed per time point. **A**. Electropherogram indicating the three best RNA samples per time point. **B**. The resulting gel image for all RNA samples

University of Cape Town

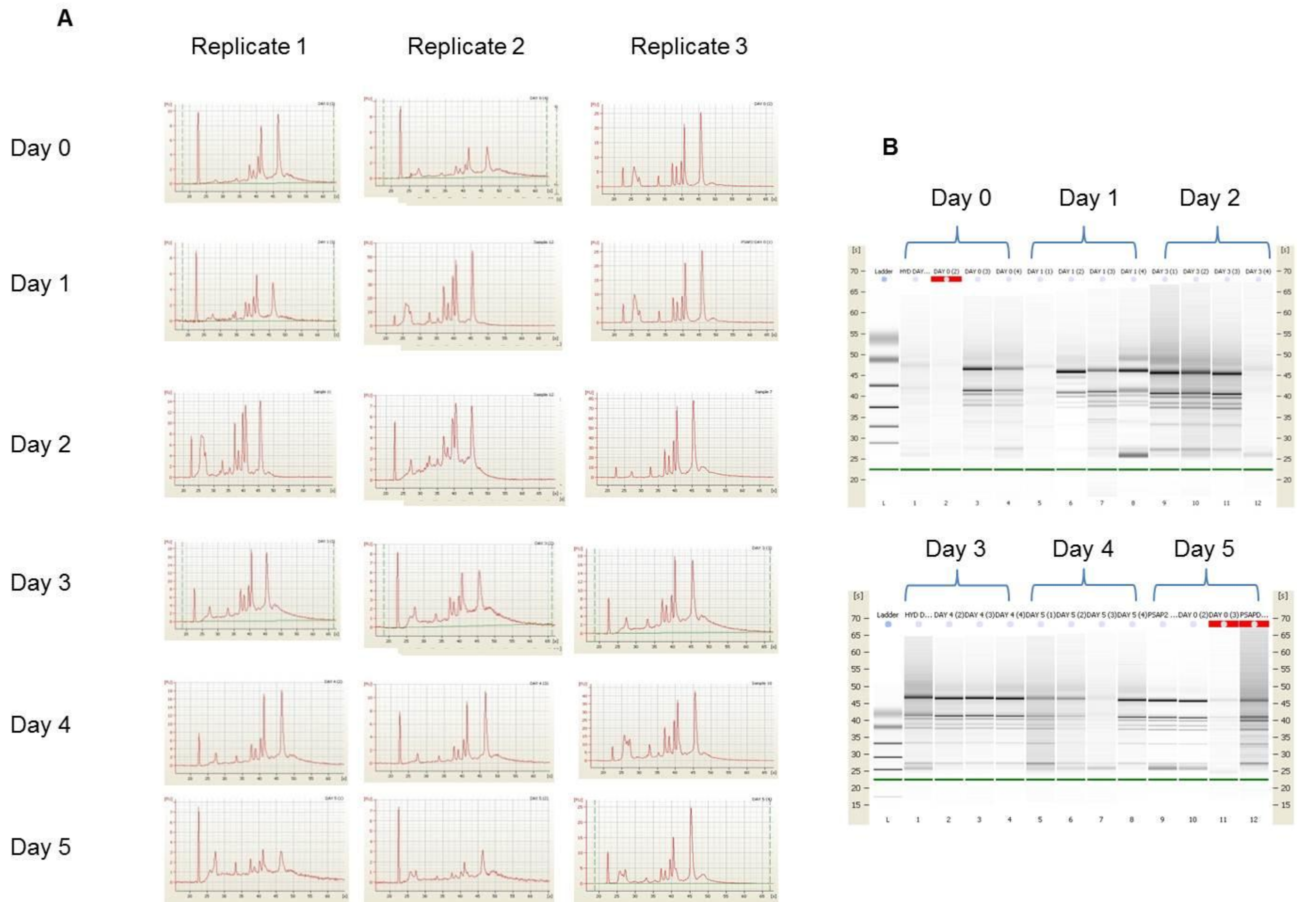


Figure H.1D: Analysis of RNA isolated from XvPSap2 plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. Four RNA isolations were performed per time point. **A**. Electropherogram indicating the three best RNA samples per time point. **B**. The resulting gel image for all RNA samples

University of Cape Town

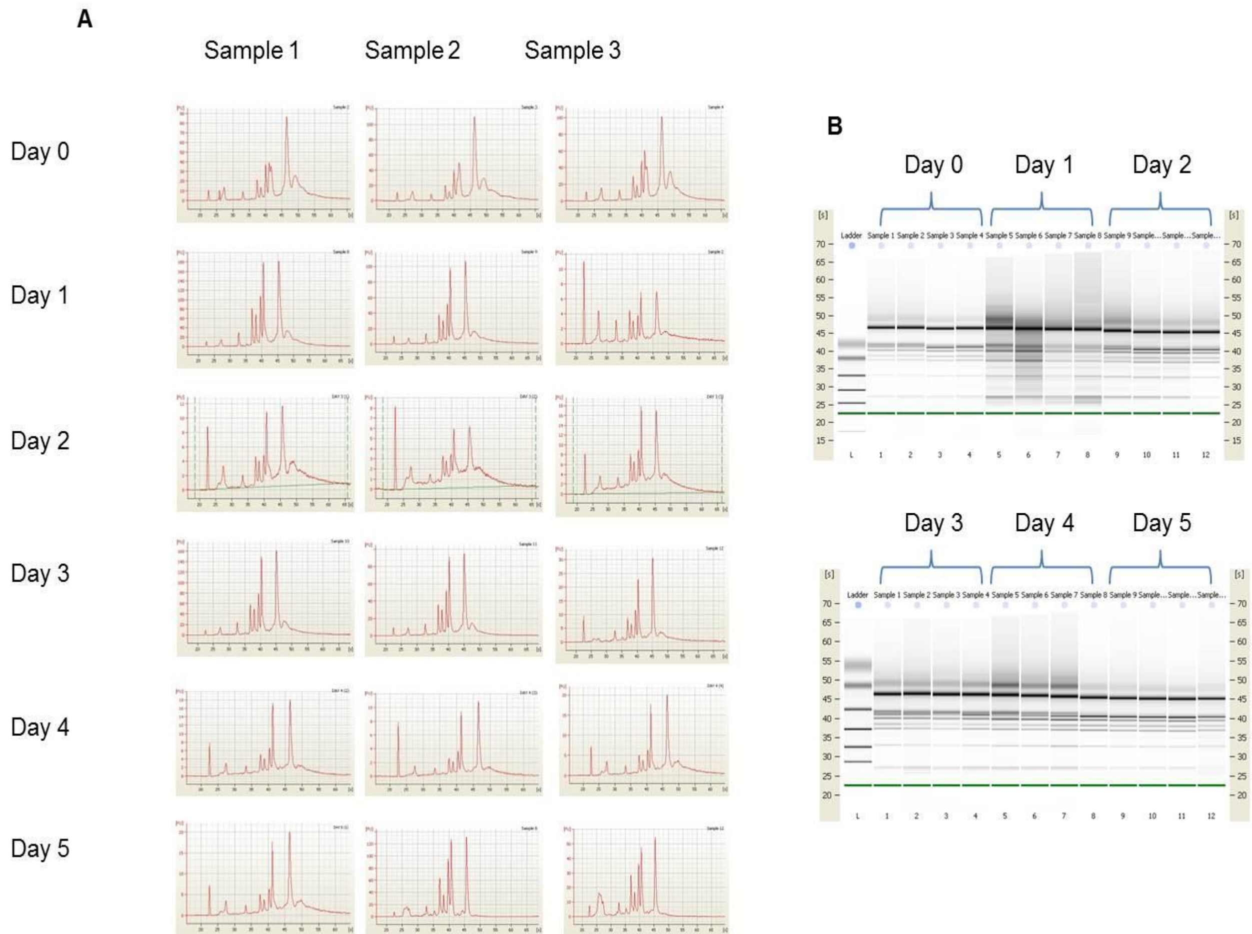


Figure H.1E: Analysis of RNA isolated from hydrated transgenic plants over a six day dehydration period. RNA was assessed using the Agilent Expert software. Four RNA isolations were performed per time point. **A**. Electropherogram indicating the three best RNA samples per time point. **B**. The resulting gel image for all RNA samples

University of

H.2. Melt curves, standard curves, real time runs

Expression of the reference genes and luciferase gene was investigated by real time PCR using the Rotor Gene 6000 2 plex HRM (Corbett Life Science Research).

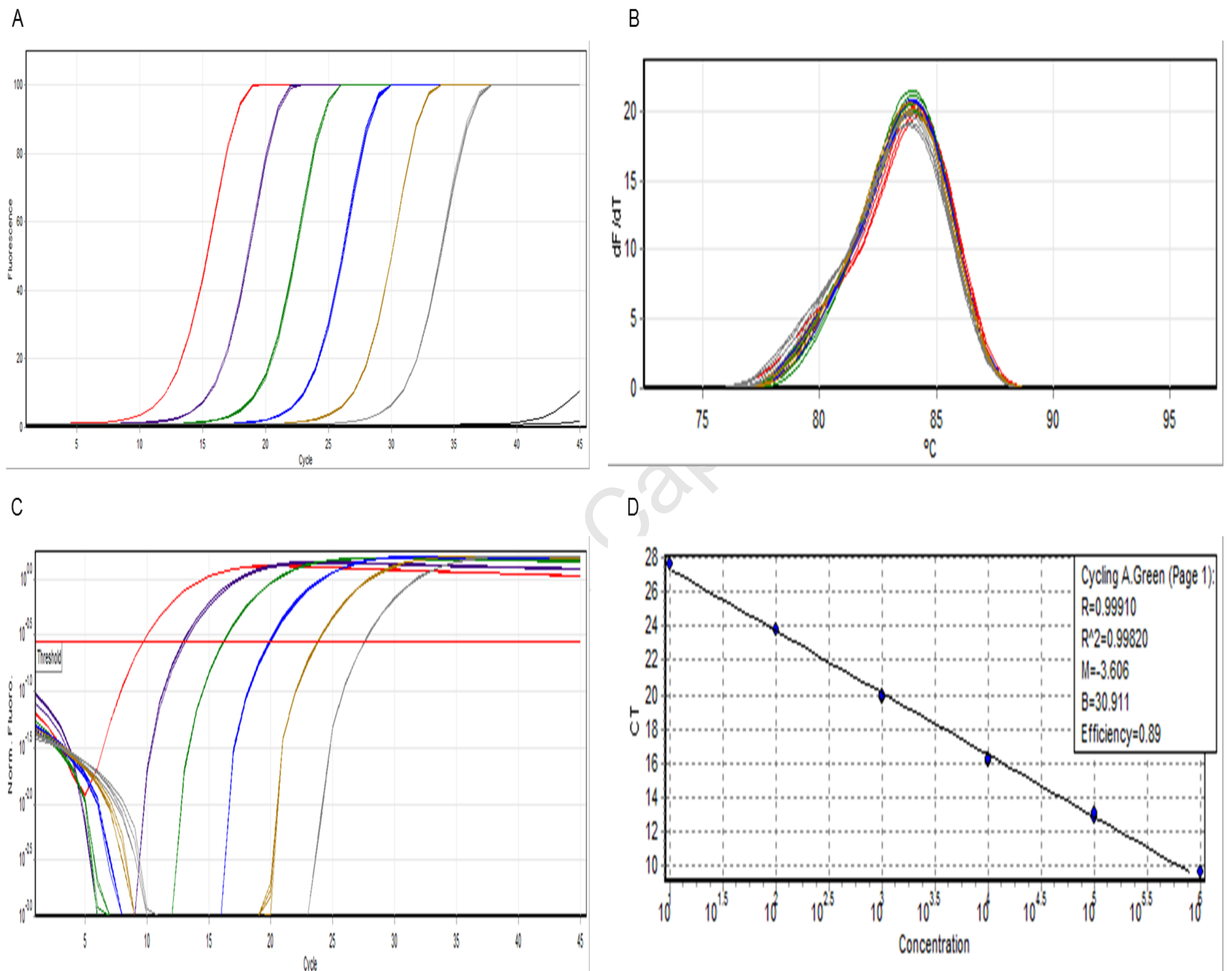


Figure H.2A: Analysis of the 18S rRNA reference gene. A: Raw data for the cycling A green. B: Melt curve analysis. C: Quantitation data for cycling A green. D: Standard curve.

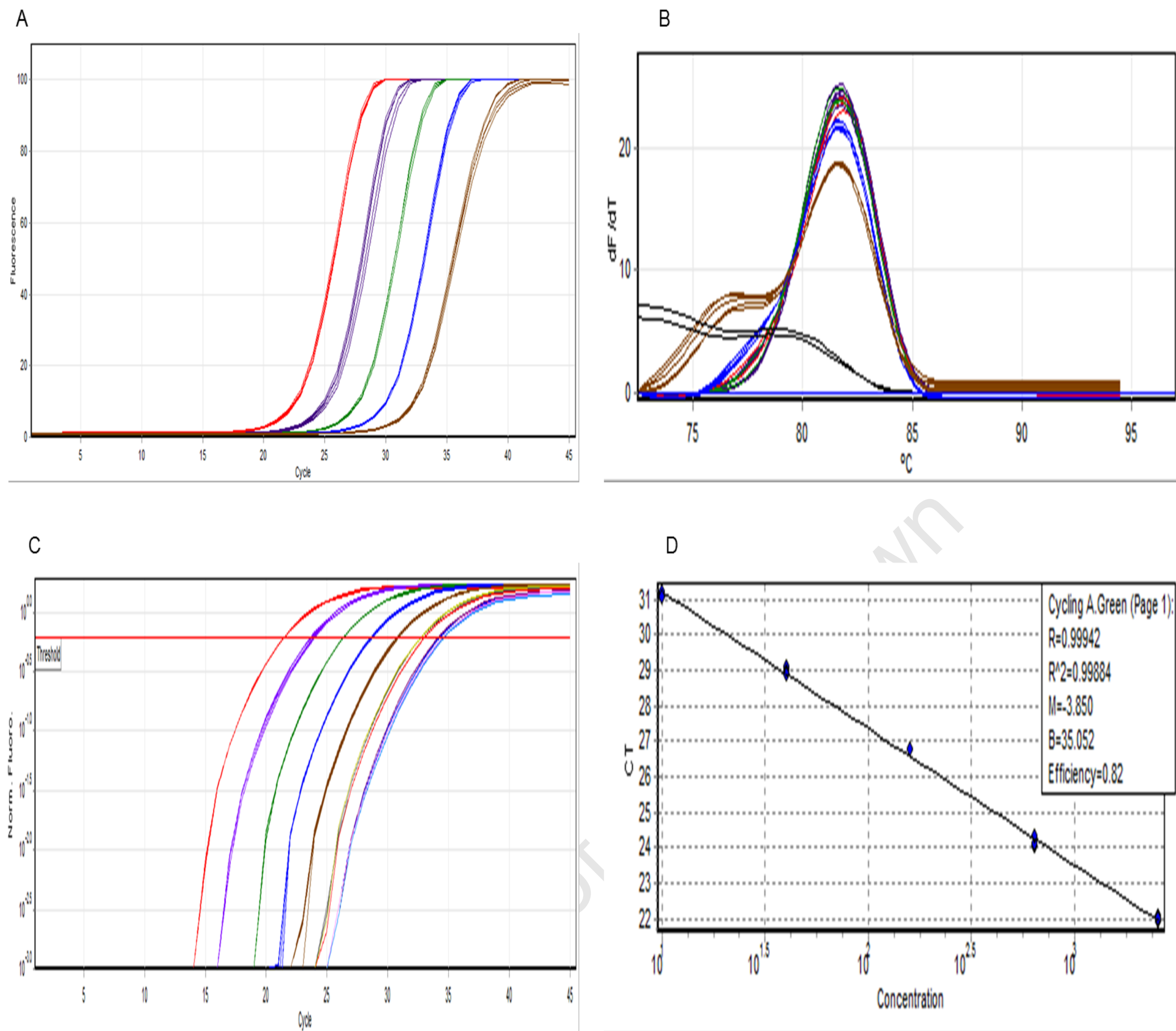


Figure H.2B: Analysis of the elongation factor 1-α reference gene. A: Raw data for the cycling A green. B: Melt curve analysis. C: Quantitation data for cycling A green. D: Standard curve.

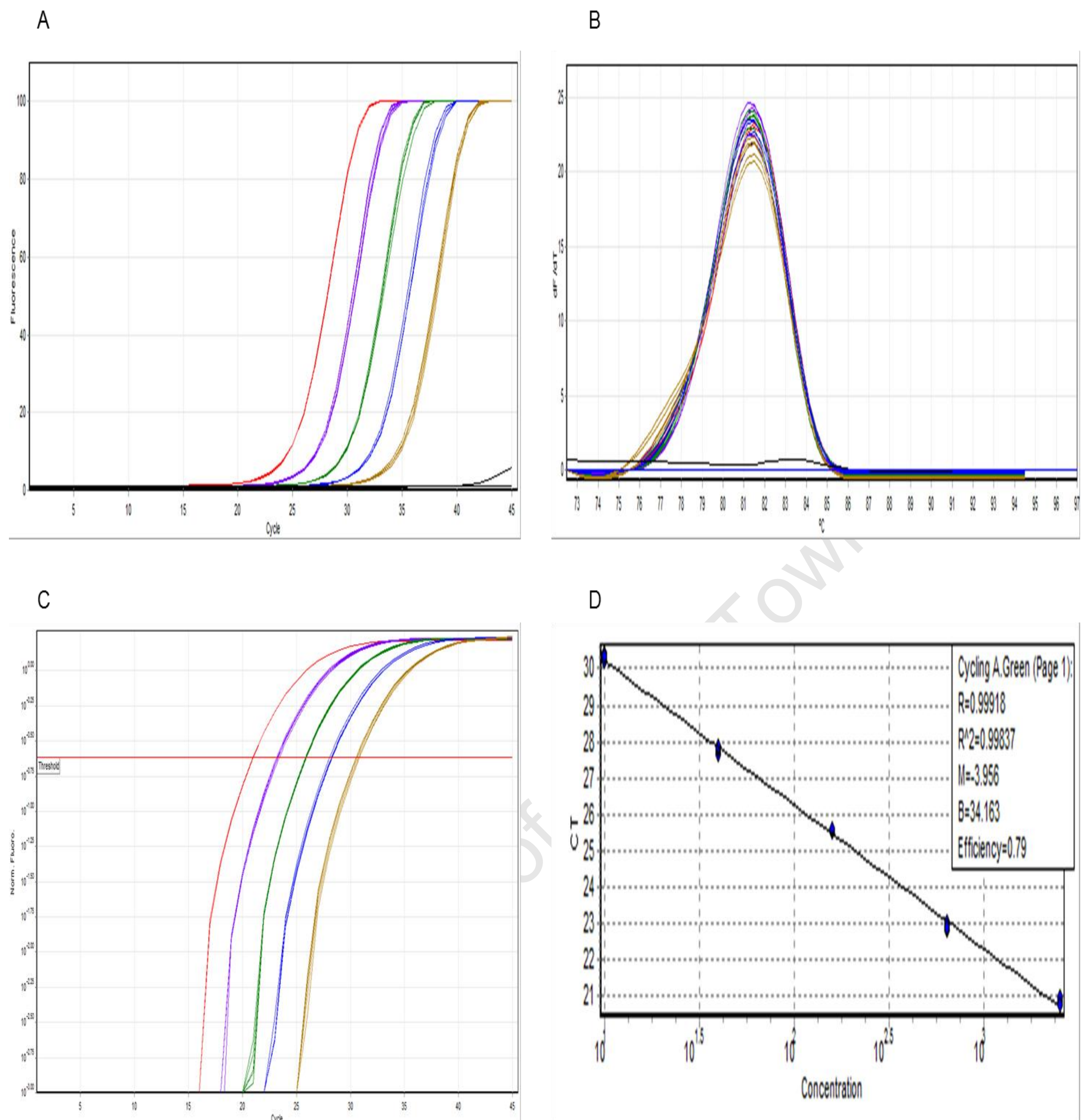


Figure H.2C: Analysis of the L25 ribosomal reference gene. A: Raw data for the cycling A green. B: Melt curve analysis. C: Quantitation data for cycling A green. D: Standard curve.

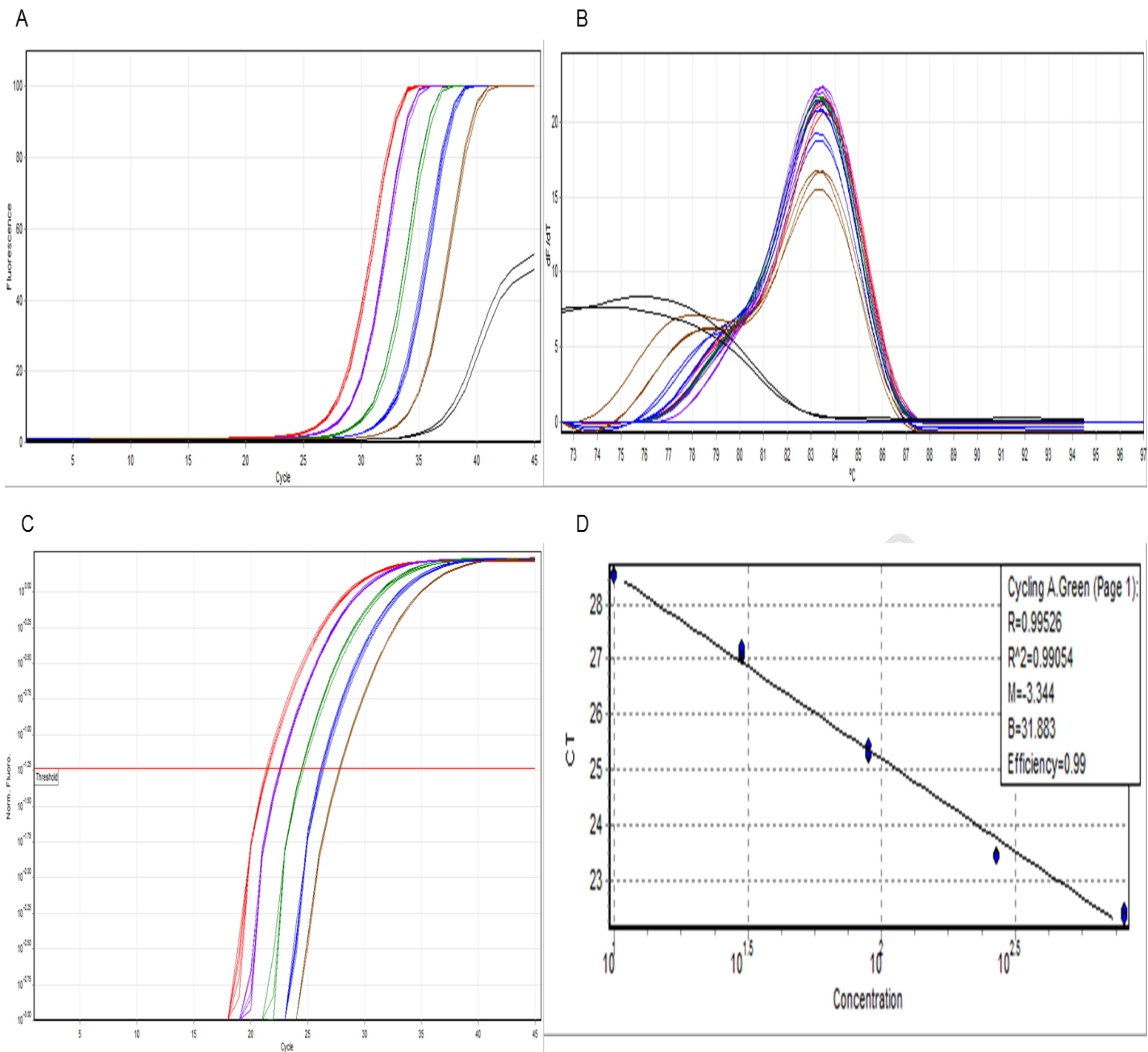


Figure H.2D: Analysis of the *luc* gene. A: Raw data for the cycling A green. B: Melt curve analysis. C: Quantitation data for cycling A green. D: Standard curve.

Appendix I: Wild type *N. tabacum*

I.1. Bioluminescence

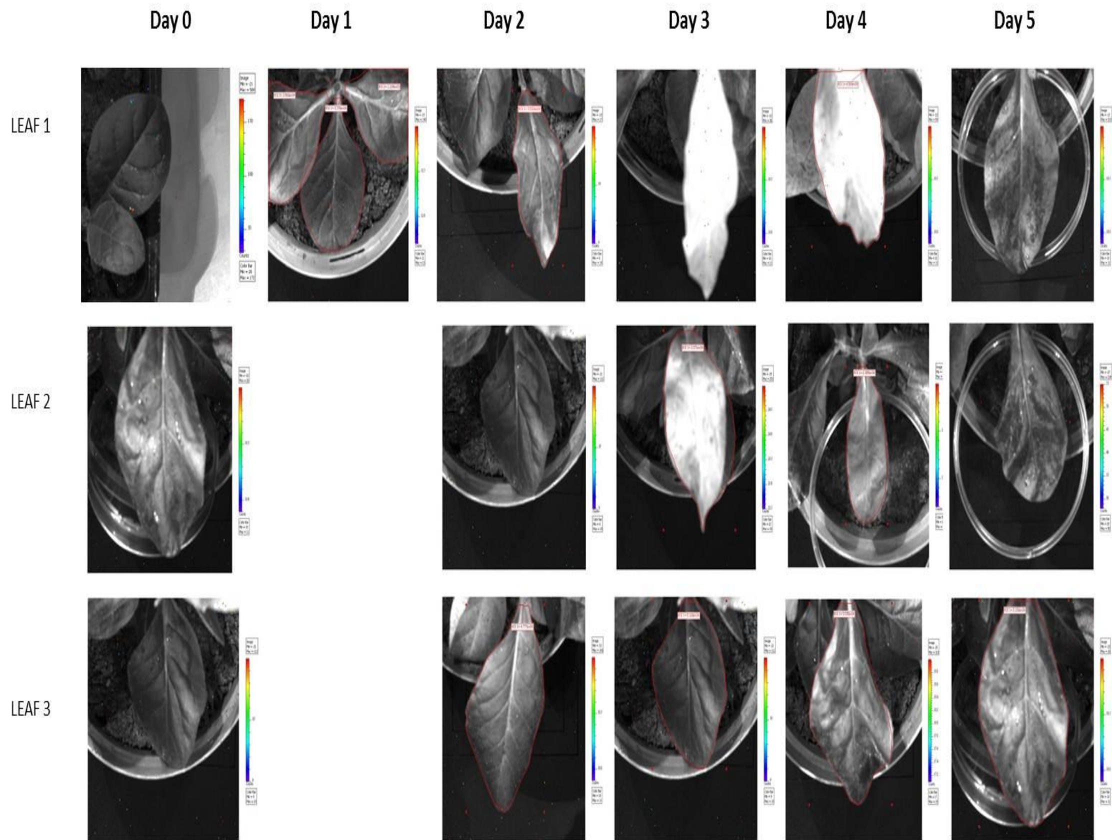


Figure I.1: Determination of luciferase activity in wild type *N. tabacum* plants transformed with XvPSap1D over a six day dehydration treatment.

1.2. Wild type *N. tabacum*

Figure I.2: Untransformed wild type *N. tabacum* on rooting media supplemented with 3 mg/ml BASTA (Bayer, South Africa).

Appendix J: Supporting images

J.1. Restriction digestion of pBluescript::XvPsap1::*luc*::NosT

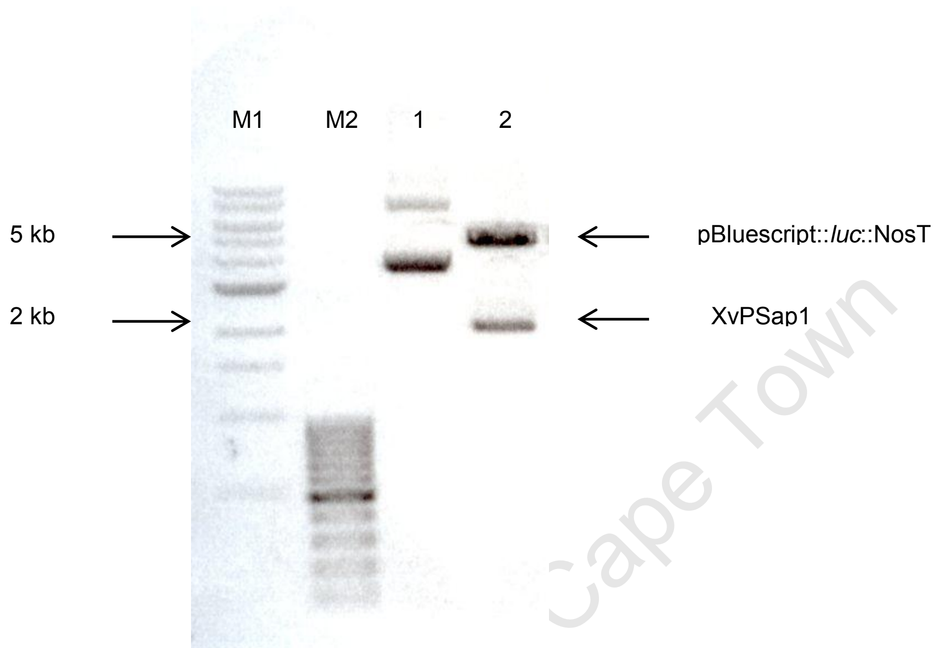


Figure J.1: Figure 4.2: Restriction endonuclease analysis of pBluescript::XvPSap1::*luc*::NosT. Lane M1 and M2: 1 kb and 100 bp DNA ladder. Lane 1: Undigested pBluescript::XvPSap1::*luc*::NosT. Lane 2: *Eco*RI and *Bam*HI restriction digest.

University of Cape Town

J.2. *N. tabacum* transformed with XvPSap1E displaying dwarfism

Figure J.2: *N. tabacum* transformed with XvPSap1E. Plants on the left represent T0 plants displaying normal growth patterns. Plants on the right represent T0 plants displaying pleiotropic characteristics.

References

Acquaah G. (2007). Importance of breeding for resistance to abiotic stress. In Principles of plant genetics and breeding. Blackwell Publishing, Malden, US. pp 385-386.

Africa's Smallholder Farmers Group (n.d.). Approaches that work for viable livelihoods. Retrieved 5 September 2012 from www.practicalaction.org/africa-smallholder-farmers-approaches.

Al-Abed D., Madasamy P., Talla R., Goldman S. and Rudrabhatla S. (2007). Genetic engineering of maize with the *Arabidopsis DREB1A/CBF3* gene using split-seed *explants. Crop Science*. **47**:2390-2402.

Alpert P. (2005). The limits and frontiers of desiccation-tolerant life*. Integrative and Comparative Biology*. **45**:485–495.

Alpert P. (2006). Constraints of tolerance: Why are desiccation-tolerant organisms so small or rare? *Journal of Experimental Biology*. **209**:1575–1584.

Andersen C.L., Jensen J.L. and Omtoft T.F. (2004). Normalization of real time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research*. **64**:5245-5250.

Audran C., Borel C., Frey A., Sotta B., Meyer C., Simonneau T. and Marion-Poll A. (1998). Expression studies of the zeaxanthin epoxidase gene in *Nicotiana plumbaginifolia. Plant Physiology.* **118**:1021–1028.

Bahariah B., Parveez G.K.A., Masani M.Y.A. and Khalid N. (2012). Construction of phosphomannose isomerase (PMI) transformation vectors and evaluation of the effectiveness of vectors in tobacco (*Nicotiana tabacum L*). *Bioinformation.* **8**(3):151-157.

University of Cape Town

Baldi G. and Jobbagy E.G. (2012). Land use in the dry subtropics: Vegetation composition and production across contrasting human contexts. *Journal of Arid Environments*. **76**:115-127.

Barnabas B., Jager K. and Feher A. (2008). The effect of drought and heat on reproductive processes in cereals. *Plant, Cell and Environment.* **31**:11-38.

Bernard V., Brunaud V. and Lecharny A. (2010). TC-motifs at the TATA-box expected position in plant genes: A novel class of motifs involved in the transcription regulation. *Biomedcentral Genomics.* **11**:166.

Bhatnagar-Mathur P., Vadez V. and Sharma K.K. (2008). Transgenic approaches for abiotic stress tolerance in plants: Retrospect and prospects. *Plant Cell Reports*. **27**: 411-424.

Butler J.E.F. and Kadonaga J.T. (2002). The RNA polymerase II core promoter: A key component in the regulation of gene expression. *Genes and Development.* **16**: 2583-2592.

Cakir R. (2004). Effect of water stress at different development stages on vegetative and reproductive growth of corn. *Field Crop Research*. **89**(1):1-6.

Carey M. and Smale S.T. (2001). Concepts and strategies: I. Promoter and the general transcription machinery. In Transcriptional regulation in eukaryotes. Cold Spring Harbor Laboratory Press, New York.

Cattivelli L., Rizza F., Badeck F., Mazzucotelli E., Mastrangelo A.M., Francia E., Mare C., Tondelli A. and Stanca A.M. (2008). Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. *Field Crops Research*. **105**:1-14.

Centre for Environmental Risk Assessment (n.d). Global status of approved genetically modified plants. Retrieved 4 July 2011 from www.cera-gmc.org.

University of Cape Town

Chinnusamy V., Jagendorf A. and Zhu J.K. (2005). Understanding and improving salt tolerance in plants. *Crop Science*. **45**:437-448.

Christensen A.H. and Quail P.H. (1996). Ubiquitin promoter-based vectors for high - level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Research*. **5**:213-218.

Clough S.J. and Bent A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant Journal*. **16**:735-743.

Datta K., Baisakh N., Ganguly M., Krishnan S., Yamaguchi-Shinozaki K. and Datta S.K. (2012). Overexpression of *Arabidopsis* and rice stress genes' inducible transcription factor confers drought and salinity tolerance to rice. *Plant Biotechnology Journal*. **10**:579–586

Deikman J., Petracek M. and Heard J.E. (2012). Drought tolerance through biotechnology: improving translation from the laboratory to farmer's fields. *Current Opinion in Biotechnology*. **23**(2):243-250

Dellaporta S.L., Wood J. and Hicks J.B. (1983). A plant DNA miniprepration: Version II. *Plant Molecular Biology*. **1**:19-21.

Dickie J.B. and Pritchard H.W. (2002). Systematic and evolutionary aspects of desiccation tolerance in seeds. In Desiccation and survival of in plants: Drying without dying. Edited by M Black and HW Pritchard. CABI Publishing, Wallingford and New York. pp 239-260.

FAO (2009). FAO Statistical Database. Food and Agricultural Organization of the United Nations. Retrieved on 5 June 2010 from www.faostat.fao.org.

Farrant J.M. and Moore J.P. (2011). Programming desiccation-tolerance: from plants to seeds to resurrection plants. *Current Opinion in Plant Biology*. **14**:340-345.

University of Cape Town

Farrant J.M., Cooper K. and Nell H. (2012). Desiccation tolerance. In Plant Stress Physiology, Edited by Sergey Shabala. CABI Publishing Wallingford, 2011.

Fromm M.E., Morrish F., Armstrong C., Williams R., Thomas J. and Klein T.M. (1990). Inheritance and expression of chimeric genes in the progeny of transgenic maize plants. *Biotechnology*. **8**:833-838.

Furtado A., Henry R.J. and Pellegrineschi A. (2009). Analysis of promoters in transgenic barley and wheat. *Plant Biotechnology Journal*. **7**(3):240-253.

Furtado A., Henry R.J. and Takaiwa F. (2008). Comparison of promoters in transgenic rice. *Plant Biotechnology Journal*. **6**:679-693.

Gaff D.F. (1987). Desiccation tolerant plants in South America. *Oecologia.* **74**:133–136.

Gaff D.F. (1977). Desiccation tolerant vascular plants in Southern Africa. *Oecologia.* **31**:95-109.

Gaff D.F. (1971). Desiccation-tolerant flowering plants in Southern Africa. *Science.* 174(**4013**):1033-1034.

Ganguly M., Roychoudhury A., Sarkar S.N., Sengupta D.N., Datta S.K. and Datta K. (2011). Inducibility of three salinity/abscisic acid-regulated promoters in transgenic rice with *gusA* reporter gene. *Plant Cell Reports.* **30**:1617-1625.

Gao F., Su Q., Fan Y. and Wang L. (2010). Expression pattern and core region analysis of *AtMPK3* promoter in response to environmental stresses. *Science China Life Science.* **53**:1315-1321.

Gao S., Chen M.C., Xia L., Xiu H., Xu Z., Li L., Zhao C., Cheng X. and Ma Y. (2009). A cotton (*Gossypium hirsutum*) DRE-binding transcription factor gene, *GhDREB*, confers tolerance to drought, high salt, and freezing stresses in transgenic wheat. *Plant Cell Reports*. **28**:301-311.

University of Cape Town

Garwe D., Thomson J.A. and Farrant J.M. (2006). *XVSAP1* from *Xerophyta viscosa* improves osmotic-, salinity- and high temperature-stress tolerance in *Arabidopsis*. *Biotechnology Journal*. **1**(10):1137-1146.

Garwe D., Thomson J.A. and Mundree S.G. (2003). Molecular characterization of *XVSAP1* gene from the resurrection plant *Xerophyta viscosa baker*. *Journal of Experimental Botany*. **54**:191-201.

Gechev T.S., Dinakar C., Benina M., Toneva V. and Bartels D. (2012). Molecular mechanisms of desiccation tolerance in resurrection plants. *Cellular and Molecular Life Sciences*. **69**:3175-3186.

Gollery M., Harper J., Cushman J., Mittler T., Girke T., Zhu J., Bailey-Serres J. and Mittler R. (2006). What makes species unique? The contribution of proteins with obscure features. *Genome Biology*. **7**:R57.

Govender K. (2006). Characterization of *XvPrx2*, a type II peroxiredoxin isolated from the resurrection plant *Xerophyta viscosa Baker*. PhD Thesis, University of Cape Town.

Gross P. and Oelgeschlager T. (2006). Core promoter-selective RNA polymerase II transcription. *Biochemical Society Symposia*. **73**:225-236.

Han Y., Cho K., Hwang O., Cho Y., Shin A., Hwang I. and Kim J. (2012). Overxpression of an *Arabidopsis β-glucosidase* gene enhances drought resistance with dwarf phenotype in creeping bentgrass. *Plant Cell Reports*. **31**:1677-1686.

Hernandez-Garcia C.M., Martinelli A.P., Bouchard R.A. and Finer J.J. (2009). A soybean (*Glycine max*) polyubiquitin promoter gives strong constitutive expression in transgenic soybean. *Plant Cell Reports*. **28**(5):837-849.

Hirai T., Shohael A.M., Kim. Y., Yano M. and Ezura H. (2011). Ubiquitin promoter-terminator cassette promotes genetically stable expression of the taste-modifying protein miraculin in transgenic lettuce. *Plant Cell Reports*. **30**:2255-2265.

University of Cape Town

Holmberg N. and Bulow L. (1998). Improving stress tolerance in plants by gene transfer. *Trends in Plant Science*. **3**:61-66.

Horan K., Jang C., Bailey-Serres J., Mittler R., Shelton C., Harper J.F., Zhu J., Cushman J.C., Gollery M. and Girke T. (2008). Annotating genes of known and unknown function by large-scale coexpression analysis. *Plant Physiology*. **147**:41-57.

Ijaz S., Anjum N., Rana I.A. and Khan I.A. (2012). Optimization of minimal inhibitory dose of selective agent (Basta) for selection of transgenics in sugarcane. *Molecular Plant Breeding*. **3**(5):50-56.

Iyer R., Mundree S.G., Rafudeen M.S. and Thomson J.A. (2007). *XvSap1*, a desiccation tolerance associated gene with potential for crop improvement. In Plant Desiccation Tolerance. Edited by Jenks M.A. and Woods A.J.. Blackwell Publishing Ltd, Oxford, UK, pp 283-296.

James C. (2011). Global status of commercialized biotech/GM crops: 2011. International Service for the Acquisition of Agri-biotech applications (ISAAA), Brief No. 43. ISAAA: Ithaca, New York.

James V.A., Avart C., Worland B., Snape J.W. and Vain P. (2002). The relationship between homozygous and hemizygous transgene expression levels over generations in populations of transgenic rice plants. *Theoretical and Applied Genetics.* **104**:553-561.

Jin T., Chang Q., Li W., Yin D., Li Z., Wang D., Liu B. and Liu L. (2010). Stress-inducible expression of *GmDREB1A* conferred salt tolerance in transgenic alfalfa. *Plant Cell, Tissue and Organ Culture*. **100**:219-227.

Karim S., Aronsson H., Ericson H., Pirhonen M., Leyman B., Welin B., Mantyla E., Palva T., Van Dijck P. and Holmstrom K. (2007). Improved drought tolerance without undesired side effects in transgenic plants producing trehalose. *Plant Molecular Biology.* **64**:371-386.

University of Cape Town

Karp N.A., Spencer M., Lindsay H., O'Dell K. and Lilley K.S. (2005). Impact of replicate types on proteomic expression analysis. *Journal of Proteome Research.* **5**:1867-1871.

Kasuga M., Liu Q., Miura S., Yamaguchi-Shinozaki K. and Shinozaki K. (1999). Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. *Nature Biotechnology*. **17**:287-291.

Kasuga M., Miura S., Shinozaki K. and Yamaguchi-Shinozaki K. (2004). A combination of the *Arabidopsis* DREB1A gene and stress-inducible *rd29A* promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer. *Plant Cell Physiology*. 45(**3**):346-350.

Khan M.A., Makhdoom R., Husnain T., Saleem M.Z., Malik K., Latif A., Altosaar I. and Riazuddin S. (2001). Expression of Bt gene in a dicot plant under promoter derived from a monocot plant. *Pakistan Journal of Biological Sciences.* 4(**12**):1518-1522.

Kim K., Kwon S., Lee H., Hur Y., Bang J. and Kwak S. (2003). A novel oxidative stress-inducible peroxidase promoter from sweetpotato: Molecular cloning and characterization in transgenic tobacco plants and cultured cells. *Plant Molecular Biology.* 51(**6**):831-838.

Kling J.G. and Edmeades G. (1997). Morphology and growth of maize, IITA/CIMMYT Research Guide 9, Ibadan, Nigeria.

Kohli A., Miro B. and Twyman R.M. (2010). Transgene integration, expression and stability in plants: Strategies for improvements. In Transgenic crop plants: Volume 1: Principles and development. Edited by Chittaranjan Kole, Charles Michler, Albert G Abbot and Tomothy C Hall. Springer Publishing. pp 201-245

Kranner I. and Birtic S. (2005). A modulating role for antioxidants in desiccation tolerance. *Integrative and Comparative Biology*. **45**:734-740.

University of Cape Town

Kumar D., Patro S., Ranjan R., Sahoo D.K., Maiti I.B. and Dey N. (2011). Development of useful recombinant promoter and its expression analysis in different plant cells using confocal laser scanning microscopy. *PLOS*. **6**:e24627.

Lamkey K. (2002). Genetically modified organisms' gene flow: A plant breeding perspective. Biotechnology, gene flow, and intellectual property rights. Agricultural Summit, Purdue University, Indianapolis.

Lee J.T., Prasad V., Yang P.T., Wu J.F., Ho D., Charng Y.Y. and Chan M.T. (2003). Expression of *Arabidopsis CBF1* regulated by an ABA/stress inducible promoter in transgenic tomato confers stress tolerance without affecting yield. *Plant, Cell and Environment*. **26**:1181–1190.

Lescot M., Dhais P., Moreau Y., De Moor B., Rouz P. and Rombauts S. (2002). PlantCARE: a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico* analysis of promoter sequences. *Nucleic Acids Research*. **30**(1):325–327.

Li J., Brunner A.M., Meilan R. and Strauss S.H. (2008). Stability of transgenes in trees: Expression of two reporter genes in poplar over three field seasons. *Tree Physiology.* **9**:299-312.

Liu Q., Kasuga M., Sakuma Y., Abe H., Miura S., Yamaguchi-Shinozaki K. and Shinozaki K. (1998). Two transcriptional factors, *DREB1* and *DREB2*, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought and low temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell.* **10**:1391–1406.

Livingston G., Schonberger S. and Delaney S. (2011). Sub-Saharan Africa: The state of smallholders in agriculture. Paper presented at the Conference on new directions for smallholder agriculture, 24-25 January 2011, Rome, International Fund for Agricultural Development.

Lobell D.B. and Field C.B. (2007). Global scale climate – crop yield relationships and the impacts of recent warming. *Environmental Research Letters*. **2**:1-7.

University of Cape Town

Machuka J. (2001). Agricultural biotechnology for Africa. African scientists and farmers must feed their own people. *Plant Physiology*. **126**:16–19.

Mallikarjuna G., Mallikarjuna K., Reddy M.K. and Kaul T. (2011). Expression of *OsDREB2A* transcription factor confers enhanced dehydration and salt stress tolerance in rice (*Oryza sativa L*). *Biotechnology Letters*. **33**:1689-1697.

Mann D.G.J., King Z.R., Lui W., Joyce B.L., Percifields R.J., Hawkins J.S., LaFayette P.R., Artelt B.J., Burris J.N., Mazarei M., Bennetzen J.L., Parrott W.A. and Stewart Jr C.N. (2011). Switchgrass (*Panicum virgatum L.*) polyubiquitin gene (PvUbi1 and PvUbi2) promoters for use in plant transformation. *Biomedcentral Biotechnology.***11**: 74.

Manshardt R. (2004). Crop improvement by conventional breeding or genetic engineering: How different are they? Honolulu (HI): University of Hawaii.

Maredza A.T. (2008). Isolation of the aldose reductase gene (*XvAld1*) from the resurrection plant *Xerophyta viscosa*, and characterisation of the gene product and transgenic plants expressing the gene. PhD thesis, University of Cape Town.

McCouch S. (2004). Diversifying selection in plant breeding. *PLOS Biology*. **2**(10):e347.

Miki B. and McHugh S. (2004). Selectable marker genes in transgenic plants: applications, alternatives and biosafety. *Journal of Biotechnology*. **107**:193-232.

Mittler R. and Blumwald E. (2010). Genetic engineering for modern agriculture: Challenges and perspectives. *Annual Review of Plant Biology.* **61**:443-462.

Morran S., Eini O., Pyvovarenko T., Parent B., Singh R., Ismagul A., Eliby S., Shirley N., Langridge P. and Lopato S. (2011). Improvement of stress tolerance of wheat and barley by modulation of expression of DREB⁄CBF factors. *Plant Biotechnology Journal.* **9**:230-249.

University of Cape Town

Mowla S.B., Thomson J.A., Farrant J.M. and Mundree S.G. (2002). A novel stress-inducible antioxidant enzyme identified from the resurrection plant *Xerophyta viscosa* Baker. *Planta*. **215**:716-726.

Muller F., Demeny M.A. and Tora L. (2007). New problems in RNA Polymerase II transcription initiation: Matching the diversity of core promoters with a variety of promoter recognition factors. *Journal of Biological Chemistry*. **282**:14685-14689.

Muller P.Y, Janovjak H, Miserez A.R. and Dobbie. Z (2002). Processing of gene expression data generated by quantitative real-time RTPCR. *Biotechniques* **32**:1 372-1379.

Mundree S.G., Baker B., Mowla S., Peters S., Marais S., Vander Willigan C., Govender K., Maredza A., Muyanga S., Farrant J.M. and Thomson J.A. (2002). Physiological and molecular insights into drought tolerance. *African Journal of Biotechnology*. **1**(2):28-38.

Mundree S.G., Whittaker A., Thomson J.A. and Farrant J.M. (2000). An aldose reductase homolog from the resurrection plant *Xerophyta viscosa* Baker. *Planta*. **211**:693-700.

Nakagawa S. and Cuthill I.C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews.* **82**:591-605.

Nakamura S., Mano S., Tanaka Y., Ohnishi M., Nakamori C., Araki M., Niwa T., Nishimura M., Kaminaka H., Nakagawa T., Sato Y. and Ishiguro S. (2010). Gateway binary vectors with the bialophos resistance gene, *bar*, as a selection marker for plant transformation. *Bioscience, Biotechnology, and Biochemistry*. **74**(6):1315-1319.

Nakashima K., Tran L.P., Van Nguyen D., Fujita M., Maruyama K., Todaka D., Ito Y., Hayashi N., Shinozaki K. and Yamaguchi-Shinkozaki K. (2007). Functional analysis of a NAC-type transcription factor, *OsNAC6* involved in abiotic and biotic stress-responsive gene expression in rice. *The Plant Journal*. **51**:617-630.

University of Cape Town

Naqvi S., Farre G., Sanahuja G., Capell T., Zhu C. and Christou P. (2010). When more is better: Multigene engineering in plants. *Trends in Plant Science*. **15**:48-56.

Oerke E.C. (2006). Crop losses to pests. *The Journal of Agricultural Science*. **144**: 31-43.

Okoth R.O. (2009). Functional analyses of the stress-inducible XvPsap1 promoter isolated from *Xerophyta viscosa*. PhD thesis, University of Cape Town.

Oltmanns H., Frame B., Lee L., Johnson S., Li B., Wang K. and Gelvin S.B. (2010). Generation of backbone-free, low transgene copy plants by launching T-DNA from the *Agrobacterium* chromosome. *Plant Physiology*. **152**:1158-1166.

Otegui M.E., Andrade F.H. and Suero E.E. (1995). Growth, water use, and kernel abortion of maize subjected to drought at silking. *Field Crops Research*. **40**:87-94.

Ouellet F., Vazquez-Tello A. and Sarhan F. (1998). The wheat wcs120 promoter is cold-inducible in both monocotyledonous and dicotyledonous species. *Federation of European Biochemical Societies Letters*. **423**:324-328.

Owen M.D.K, and Zelaya I.A. (2005). Herbicide-resistant crops and weed resistance to herbicides. *Pest Management Science*. **61**:301-311.

Park S., Yi N., Kim Y.S., Jeong M., Bang S., Choi Y. and Kim J. (2010). Analysis of five novel putative constitutive gene promoters in transgenic rice plants. *Journal of Experimental Botany*. **61**:2459-2467.

Peach C. and Velton J. (1991). Transgene expression variation (position effect) of CAT and GUS reporter genes driven by linked divergent T-DNA promoters. *Plant Molecular Biology*. **17**:49-60.

University of Cape Town

Peremarti A., Twyman R., Gomez-Galera S., Naqvi S., Farre G., Sabalza M., Miralpeix B., Dashevskaya S., Yuan D., Ramessar K., Christou P., Zhu C., Bassie L. and Capell T. (2010). Promoter diversity in multigene transformation. *Plant Molecular Biology*. **73**:363-378.

Pfaffl M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acid Research*. **29**(9):e45.

Pilcher C.D., Obrycki J.J., Rice M.E. and Lewis L.C. (1997a). Preimaginal development, survival and field abundance of insect predators on transgenic *Bacillus thuringiensis* corn. *Environmental Entomology*. **26**:446-454.

Pilcher C.D., Rice M.E., Obrycki J.J. and Lewis L.C. (1997b). Field and laboratory evaluations of transgenic *Bacillus thuringiensis* corn on secondary Lepidopteran pests (Lepidoptera: Noctuidae). *Journal of Economics and Entomology*. **90**:669–678.

Porembski S. and Barthlott W. (2000). Granitic and gneissic outcrops (inselbergs) as centers for diversity for desiccation tolerant vascular plants', *Plant Ecology*. **151**:19–28.

Potts M. (1994). Mechanisms of desiccation tolerance in cyanobacteria. *European Journal of Phycology*. **34**:319-328.

Prakash C. (2001). The genetically modified crop debate in the context of agricultural evolution. *Plant Pysiology*. **126**:8-15.

Proctor M.C.F. and Pence V.C. (2002). Vegetative tissues: bryophytes, vascular resurrection plants and vegetative propagules. In Desiccation and survival in plants: Drying without dying. Edited by Black M and Pritchard HW. Wallingford, CABI Publishing, pp 207-237.

Quan R., Hu S., Zhang Z., Zhang H., Zhang Z. and Huang R. (2010). Overexpression of an ERF transcription factor, TSRF1 improves rice drought tolerance. *Plant Biotechnology Journal*. **8**:476-488.

University of Cape Town

Qui W., Liu M., Qiao G., Jiang J., Xie L. and Zhuo R. (2011). An isopentyl transferase gene driven by the stress-inducible *rd29A* promoter improves salinity stress in transgenic tobacco. *Plant Molecular Biology Reporter*. **30**:519-528.

Ramajulu S. and Bartels D. (2002). Drought- and desiccation-induced modulation of gene expression in plants. *Plant, Cell and Environment*. **25**:141–151.

Ranish J.A., Yudkovsky N. and Hahn S. (1999). Intermediates in formation and activity of the RNA polymerase II preinitiation complex: Holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes and Development*. **13**:49-63.

Rashid B., Husnain T. and Riazuddin S. (2012). Plant genetic engineering: Problems and applications. In Crop production for agricultural improvement. Edited by Ahsraf M, Ozturk M, Ahmed MSS and Aksoy A. Springer Netherlands, pp 263-290.

Reynolds P.H.S. (1999). Inducible gene expression in plants, In Inducible gene expression. Edited by Reynolds P.H.S. CABI Publishing, Oxon, pp 9267–9283.

Ricci C. (1998). Anhydrobiotic capabilities of bdelloid rotifers. *Hydrobiologia*. **387**:1-13.

Ricci C. and Pgani M. (1996). Desiccation of *Panagrolaimus rigidus* (Nematoda): Survival, reproduction and the influence on the internal clock. *Hydrobiologia*. **347**:1-13.

Rosegrant M.W., Paisner M.S., Meijer S. and Witcover S. (2001). 2020 Global food outlook, trends, alternatives, and choices: A 2020 vision for food agriculture and the environment initiative. Technical Report, International Food Policy Research Institute.

Rozen, S. & Skaletsky, H. (2000), ‘Primer3 on the www for general users and for biologist programmers’, *Methods in Molecular Biology*. **132**, 365–386.

University of Cape Town

Saad R.B., Romdhan W.B., Zouari N., Azaza J., Miulet D., Verdiel J., Guiderdoni E. and Hassairi A. (2011). Promoter of the AlSAP gene from the halophyte grass *Aeluropus littoralis* directs developmental-regulated, stress-inducible, and organ-specific gene expression in transgenic tobacco. *Transgenic Research*. **20**:1003-1018.

Schmidt G.W. and Delaney S.K. (2010). Stable internal reference genes for normalization of real-time RT-PCR in tobacco (*Nicotiana tabacum*) during development and abiotic stress. *Molecular Genetics and Genomics*. **283**:233-241.

Segele Z.T. and Lamb P.J. (2005). Characterisation and variability of Kiremt rainy season over Ethiopia. *Meteorology and Atmospheric Physics*. **89**:153-180.

Shaffner A.R. and Sheen J. (1991). Maize rbcS promoter activity depends on sequence elements not found in dicot *rbcS* promoters. *The Plant Cell*. **3**:997-1012.

Shahmuradov I.A., Gammerman A.J., Hancock J.M., Bramley P.M. and Solovyev V.V. (2003). PlantProm: A database of plant promoter sequences. *Nucleic Acids Research.* **31**:114-117.

Shanahan T.M., Overpeck J.T., Anchukaitis K.J., Becks J.W., Cole J.E., Dettman D.L., Peck J.A., Scholz C.A. and King J.W. (2009). Atlantic forcing of persistent drought in west Africa. *Science.* **304**:377-380.

Sherwin H.W. and Farrant J.M. (1996). Differences in rehydration of three desiccation-tolerant angiosperm species. *Annuls of Botany*. **78**:703-710.

Shiferaw B., Prasanna B., Hellin J. and Banziger M. (2011). Crops that feed the world. Past successes and future challenges to the role played by maize in global food security. *Food Security.* **3**:307-327.

Singh O.V., Ghai S., Paul D. and Jain R.K. (2006). Genetically modified crops: Success, safety assessment and public concern. *Applied Microbiology and Biotechnology.* **71**:598-607.

University of Cape Town

Smale S.T. and Kadonaga J.T. (2003). The RNA polymerase II core promoter. *Annual Reviews Biochemistry*. **72**:449-479.

Solovyev V., Shahmuradov I.A. and Salamov A.A. (2010). Identification of promoter regions and regulatory sites. Computational biology of transcription factor binding. *Methods in Molecular Biology*. **674**:57-83.

Su J., Shen Q., Ho T.D. and Wu R. (1998). Dehydration stress-regulated transgene expression in stably transformed rice plants. *Plant Physiology*. **117**:913-922.

Sun Q., Gao F., Zhao L., Li K. and Zhang J. (2010). Identification of a new 130 bp *cis*-acting element in the *TsVP1* promoter involved in the salt stress response from *Thellungiella halophila. BioMed Central Plant Biology.* **10**:90.

Swapna L., Khurana R., Kumar S.V., Tyagi A.K. and Rao K.V. (2011). Pollen-specific expression of *Oryza sativa indica* pollen allergen gene (*OSIPA*) promoter in rice and *Arabidopsis* transgenic systems. *Molecular Biotechnology.* **48**:49-59.

Takaski H., Maruyama K., Kidokoro S., Ito Y., Fujita Y., Shinozaki K., Yamaguchi-Shinozaki K. and Nakashima K. (2010). The abiotic stress-responsive NAC-type transcription factor OsNAC5 regulates stress-inducible genes and stress tolerance in rice. *Molecular Genetics and Genomics.* **28**:173-183.

Takeda S. and Matsuoka M. (2008). Genetic approaches to crop improvement: responding to environmental and population changes. *Nature Reviews Genetics.* **9**:444-457.

Uzogara S.G. (2000). The impact of genetic modification of human foods in the 21st century: A review. *Biotechnology Advances*. **18:**179-206.

University of Cape Town

Vandesompele J., De Preter K., Pattyn F., Poppe B., Van Roy N., De Paepe A. and Speleman F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology.* **3**:Research0034.

Vian A. and Davies E. (2006). Two different wound signals evoke very rapid, systemic CMBP transcript accumulation in tomato. *Plant Signalling and Behaviour.* **1** (5):261-264.

Walker-Simmons M., Hollander-Czytko H., Anderson J. and Ryan C. (1984). Wound signals in plants: A systemic plant wound signal alters plasma membrane integrity. *Proceedings of the National Academy of Sciences of the United States of America.* **81**:3737-3741.

Xiao B., Chen X., Xiang C., Tang N., Zhang Q. and Xiong L. (2009). Evaluation of seven function-known candidate genes for their effects on improving drought resistance of transgenic rice under field conditions. *Molecular Plant.* **2**:73-83.

Xoconostle-Cazares B., Ramirez-Ortega F., Flores-Elenes L. and Ruiz-Medrano R. (2010). Drought tolerance in crop plants. *American Journal of Plant Physiology.* **5**:241-256.

Yamaguchi-Shinozaki K. and Shinozaki K. (1993), Characterisation of the expression of a desiccation-responsive *rd29* gene in *Arabidopsis thaliana* and analysis of its promoter in transgenic plants. *Molecular Genetics and Genomics*. **236**:331-340.

Yamaguchi-Shinozaki K. and Shinozaki K. (1994). A novel *cis*-Acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *The Plant Cell*. **6**:251-264.

Yi G., Shin Y., Choe G., Shin B., Kim Y.S. and Kim K. (2007). Production of herbicide-resistant sweet potato plants transformed with the *bar* gene. *Biotechnology Letters*. **29**:669-675.

University of Cape Town

Yi N., Kim Y.S., Jeiong M., Oh S., Jeong J.S., Park S., Jung H., Choi Y.D. and Kim J. (2010). Functional analysis of six drought-inducible promoters in transgenic rice plants throughout all stages of plant growth. *Planta*. **232**(3):743-754.

Yi N., Oh S., Kim Y.S., Jang H., Park S., Jeong J.S., Song S.I., Choi Y.D. and Kim J. (2011). Analysis of the *Wsi18*, a stress-inducible promoter that is active in the whole grain of transgenic rice. *Transgenic Research*. **20**:153-163.

Yokotani N., Ichikawa T., Kondou Y., Maeda S., Iwabichi M., Mori M., Hirochika H., Matsui M. and Oda K. (2009). Overexpression of a rice gene encoding a small C2 domain protein, OsSMCP1 increases tolerance to abiotic and biotic stresses in transgenic Arabidopsis. *Plant Molecular Biology.* **71**:391-402.

Zhang S., Li N., Gao F., Yang A. and Zhang J. (2010). Over-expression of *TsCBF1* gene confers improved drought tolerance in transgenic maize. *Molecular Breeding.* **26**:455-465.

Zhang N., Huai-Jun S.I. and Wang D. (2005). Cloning of *rd29A* gene promoter from *Arabidopsis thaliana* and its application in stress-resistance transgenic potato. *Acta Agronimca Sinica.* **31**:159-164.

Zhang J., Klueva N., Wang Z., Wu R., David-Ho T. and Nguyen H.T. (2000). Genetic engineering for abiotic stress resistance in crop plants. *In vitro Development and Biology Plant*. **36**:108–114.

Zou C., Sun K., Mackaluso J.D., Seddon A.E., Jin R. and Thomashow M.F. (2011). *Cis*-regulatory code of stress-responsive transcription in *Arabidopsis thaliana*. *Proceedings of the National Academy of the United States of America*. **188**:14992-14997.

University of Cape Town