ISOLATION OF GRAPEVINE PROMOTERS WITH SPECIAL EMPHASIS ON THE VACUOLAR PYROPHOSPHATASE

Mauritz Venter

Dissertation presented for the Degree of Doctor of Philosophy (Plant Biotechnology) at the University of Stellenbosch

Supervisor: Prof. Frederik C. Botha

Co-Supervisor: Dr. Bernard Potier

April 2004

Declaration

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

January 2004

For My Parents

ABSTRACT

Understanding the complex nature of grapevine molecular biology is of great importance for viticulturists. Progress in the elucidation of key events on a genetic level could provide further insight into the underlying cues responsible for the precise control of physiological and metabolic changes during a specific condition such as fruit development. The use and analysis of molecular 'tools', such as promoters controlling the site and level of gene activity, could assist in the understanding of grapevine biology and serve as a platform for the future design and development of recombinant DNA protocols and strategies for *Vitis vinifera* L.

A high-throughput gene expression system, cDNA-AFLPs, was successfully used to analyse large-scale transcriptional activity during berry ripening. Candidate cDNA fragments were selected on the basis of desired expression patterns and/or known gene function for subsequent promoter isolation. From three candidate cDNAs selected, the promoter of a gene encoding vacuolar pyrophosphatase (V-PPase) was isolated for computational and comparative analyses. Promoter activity was evaluated on a transient level using the green fluorescent protein (GFP) reporter gene. Comparative integration has allowed for putative correlation of *cis*-elements, acting as receptors within promoter regions, to regulate V-PPase gene expression in response to development, environmental stress and tissue-specificity.

In this study, integration of genetic data have advanced the understanding and transcriptional role of a key enzyme (V-PPase) during grape ripening. Although never a replacement for experimental verification, this integrative strategy of combining gene expression profiles with bioinformatics and regulatory data will greatly assist in further elucidation of various other key components and regulatory cues associated with grapevine molecular biology. This study has allowed us to use molecular tools that could assist in gaining further insight into genetic complexities and could serve as a platform for a more refined genetic manipulation strategy in *Vitis vinifera* L.

OPSOMMING

Begrip van die komplekse aard van wingerd molekulêre biologie is van groot belang vir wingerdkundiges. Vooruitgang in die begrip van belangrike gebeurtenisse op ń genetiese vlak behoort verdere insig in die onderliggende instruksies vir die noukeurige beheer van fisiologiese en metaboliese veranderinge tydens ń spesifieke kondisie soos vrug rypwording te bevorder. Die gebruik en analise van molekulêre 'instrumente' soos promoters, wat die posisie en vlak van geen aktiwiteit beheer, kan bydra tot n beter begrip van wingerd biologie en sodoende dien as ń platform vir die toekomstige ontwerp en ontwikkeling van rekombinante DNS (deoksi-ribonukleiensuur) protokolle en strategieë vir *Vitis vinifera* L.

ń Hoë-kapasiteit geen uitdrukkings sisteem, nl. kDNS-AFLPs (komplementêre deoksi-ribonukleiensuur-geamplifiseerde fragment lengte polimorfisme), is suksesvol gebruik vir die analise van grootskaalse transkripsionele aktiwiteit tydens druif rypwording. Kandidaat kDNS fragmente is geselekteer, gebaseer op verlangde uitdrukkings-patrone en/of bekende geen funksie vir daaropvolgende promoter isolering. Van drie geselekteerde kandidaat kDNS fragmente, is die promoter van ń geen wat vakuolêre pirofosfatase (V-PPase) kodeer geïsoleer vir rekenaar- en vergelykende analise. Promoter aktiwiteit is op ń nie-stabiele vlak deur die gebruik van ń groen-fluoresserende proteien (GFP) verklikker geen geëvalueer. Vergelykende integrering het dit moontlik gemaak om veronderstelde korrelasies van *cis*-elemente, wat as reseptore binne ń promoter area dien, en die regulering van V-PPase geen uitdrukking, in reaksie tot ontwikkeling, omgewings stres en weefsel-spesifisiteit, te maak.

Tydens hierdie studie, het die integrering van genetiese data gehelp om die transkripsionele rol van ń belangrike ensiem (V-PPase) tydens druif rypwording beter te verstaan. Alhoewel dit nooit ń plaasvervanger vir eksperimentele bewyse sal wees nie, kan hierdie gëintegreerde strategie, wat die kombinasie van geen-uitdrukkingsprofiele met bioinformatika en regulatoriese data behels, grootliks bydra om verskeie ander belangrike komponente en regulatoriese

aanwysings geassosieërd met wingerd molekulêre biologie te ontrafel. Hierdie studie het verdere insig in genetiese kompleksiteite verleen, en kan nou dien as ń platform vir ń meer presiese genetiese manipulering strategie in *Vitis vinifera* L.

ACKNOWLEDGEMENTS

This study would not have been possible without Divine intervention

My sincere gratitude goes to my supervisor and mentor, Prof. Frederik C. Botha, for his unwavering support and guidance (scientific and personal) throughout this study period.

I would like to thank my co-supervisor, Dr. Bernard Potier for his continuous encouragement, useful suggestions and scientific guidance.

I wish to express my sincere thanks to Anita Burger for her advice, patience and support. It was a privilege to work with her.

I thank Hennie and Dr. Sarita Groenewald for the lab discussions, advice and willingness to help.

Many thanks for the moral support, technical assistance, coffee sessions, humorous moments and friendship from the Staff & Students of the IPB.

I thank Winetech, the National Research Foundation and the Department of Trade and Industry for financial support.

My heartfelt thanks to my Parents for their endless support, love and for believing in me.

Contents

List of Figures and Tables

Figures

Tables

List of Abbreviations

^{o}C	degrees Celsius
µl	microlitre
µg	microgram
%	percent
α-^{32}P	Alpha Phosphorus 32 isotope
γ-^{33}P	Gamma Phosphorus 33 isotope
ABA	abscisic acid
ABRE	abscisic acid responsive element
ADH	alcohol dehydrogenase
AFLP	amplified fragment length polymorphism
ATP	adenosine 5'-triphosphate
AuxRE	auxin responsive element
BLAST	basic local alignment search tool
bp	base pair
BSV	banana streak badnavirus
CaMV	cauliflower mosaic virus
CBF	C-repeat binding factor
cDNA	complementary deoxyribonucleic acid
ChIP	chromatin immunoprecipitation
Ci	Curie
cpm	counts per minute

cv	cultivar
dCTP	deoxycytidine 5'-triphosphate
DD-PCR	differential display PCR
DIG	digoxygenin
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide 5’-triphosphate
DPE	down-stream promoter element
DPF	Dragon Promoter Finder
dUTP	deoxyuridine 5’-triphosphate
EDTA	ethylenediaminetetraacetate
EphB	epoxide hydrolase
EPD	eukaryotic promoter database
ERE	ethylene responsive element
EST	expressed sequence tag
EtBr	ethidium bromide
EtOH	ethanol
g	gram
GA	gibberellin
GARE	gibberellin responsive element
GFP	green fluorescent protein
Grip	grape ripening induced protein
GTF	general transcription factor
GUS	β-glucuronidase
h	hour

HCl	hydrochloric acid
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high-performance liquid chromatography
HSE	heat shock element
HSF	heat shock transcription factor
kb	kilobase
LR-iPCR	long range inverse PCR
LTRE	low temperature responsive element
LUC	luciferase
M	molar
MeJA	methyl jasmonate
mg	milligram
$MgCl_2$	magnesium chloride
min	minute
ml	millilitre
mM	millimolar
mol	mole
mRNA	messenger RNA
MS	Murashige and Skoog
m/v	mass per volume
MW	molecular weight
NaAc	sodium acetate
NaCl	sodium chloride
NaOH	sodium Hydroxide

ng	nanogram
nm	nanometer
NNPP	neural network promoter prediction
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pH	Acidity
PIC	pre-initiation complex
PLACE	plant *cis*-acting regulatory DNA elements database
PlantCARE	plant *cis*-acting regulatory element database
PPi	pyrophosphate
PVPP	polyvinylpoly-pirolidone
RNA	ribonucleic acid
rRNA	ribosomal RNA
SA	salicylic acid
SAGE	serial analysis of gene expression
SCPD	S*accharomyces cerevisiae* promoter database
SD	standard deviation of a series
SDS	sodium dodecyl sulfate
sec	seconds
SSC	standard saline citrate

TBP	TATA-box binding protein
T-DNA	transfer DNA
TE	Tris-EDTA
TF	transcription factor
TFCC	transcription-factor-centric clustering
TRANSFAC	transcription factor database
Tris	tris(hydroxymethyl)-aminomethane
TSS	transcription start site
U	units
Ubi	Ubiquitin
UTR	untranslated region
UV	ultraviolet
v/v	volume per volume
V-ATPase	H^+-transporting/vacuolar adenosine triphosphatase
V-PPase	H^+-translocating/vacuolar inorganic pyrophosphatase
W	watts

Chapter 1

Introduction

The grape is one of the world's most economically valuable crops, mostly used for wine but also important as table grapes and dried fruit. During the late 1980's and early nineties wine industries realised that conventional breeding and selection programs for cultivar improvement were inefficient. Progress in plant genomics has introduced renewed interest in grape genetic improvement. Consequently, several industries and institutes have restructured their research programmes to include a focus on recombinant DNA technology for the preservation and enhancement of viticultural resources around the world. Recent years have witnessed spectacular progress in (1) the understanding of various grapevine traits such as fruit-ripening, metabolism, disease and stress tolerance, wine aroma, (2) DNA-marker assisted breeding as well as characterization of cultivar and rootstock biodiversity, and (3) manipulation of grapevine through embryogenic transformation. (for review see Coombe 1992, Davies and Boss 2000, Vivier and Pretorius 2000, Robinson and Davies 2000, Roubelakis-Angelakis 2001).

Successful genetic manipulation of grapevine would require useful genes to confer specific traits, suitable promoters to control tissue-specificity and expression of gene activity and an efficient grapevine transformation and regeneration system. (Kikkert *et al*. 2001, Iocco *et al*. 2001). Genetic enhancement of grapevine has yielded promising results in the areas of disease and stress responsiveness (Colova-Tsolova *et al*. 2001). The genetic structure of grapevine is relatively simple, however, the complexity of a multitude of mechanisms (although well elucidated individually) associated with grapevine biology together with the variability between different cultivars are still poorly understood (Davies and Boss 2000, Kikkert *et al*. 2001). Therefore, it could be that the improvement of grapevine will depend on accomplishing an integrative approach between modern genomics, genetic manipulation and conventional

breeding strategies. This potential could, however, only be realised when there is a better understanding of the complex genetic, physiological and metabolic characteristics of the species (Vivier and Pretorius 2000, Kikkert *et al*. 2001, Colova-Tsolova *et al*. 2001).

Fruit ripening entails a diverse range of complex and developmentally orchestrated events and research efforts have mostly been conducted on model-organisms such as *Arabidopsis*, tomato and strawberry (Giovannoni 2001). In *Vitis vinifera* L., there are three important, clearly distinguishable metabolic stages during the development of grape berries: 1) During the first phase, organic acids dominate and the reducing sugars and sucrose are present in low and approximate equal proportions. 2) The second phase is a period of slow overall growth during which the embryos develop rapidly. 3) The third phase is characterised by dramatic changes in the appearance of the berries, the rapid accumulation of reducing sugars and the concomitant disappearance of organic acids. Inception of the third stage is referred to as *véraison* (Coombe 1992).

Metabolic control during grape berry development influences quality traits such as sugar-organic acid ratios, cell wall metabolism associated with berry softening and anthocyanin compositions that determine colour development, but is not well understood (Coombe 1992). It is therefore conceivable that various metabolic reactions could be targeted in a genetic manipulation program aimed at improving fruit quality. A vast array of differentially expressed genes/ESTs (expressed sequence tags) has been identified in various physiological, environmental, metabolic and ripening related cues during different stages of grape berry development. However, large-scale identification of cDNAs could only be conducted where the construction of cDNA libraries was a prerequisite (Sparvoli *et al*. 1994; Davies and Robinson 2000; Terrier *et al*. 2001, Nunan *et al*. 2001). Although other (viral and/or plant) promoters have been used, currently grapevine transformation relies, almost exclusively, on the use of constitutive promoters such as the CaMV 35S (Iocco *et al*. 2001, Li *et al*. 2001). The number of

grapevine regulatory sequences available for public domain is limited, although it is known that in the private sector, grapevine promoters have been isolated and patented.

The overall aim of this study was twofold. Firstly, to increase the molecular genetic resources available for grapevine genetic manipulation. Secondly, to specifically target promoter elements that could be utilized for the modification and further understanding of the fruit-ripening process. Due to the seasonal growth cycle and long regeneration time of grapevine during transformation (Davies and Boss 2000), it was realized before the onset of this project that stable transformed plants bearing fruit was not realistic for the evaluation of molecular 'tools' such as promoter elements and transgenes. Instead it was decided that this project would concentrate on: 1) the application of a high-throughput gene expression technology for large-scale identification of genes differentially expressed during berry ripening and 2) promoter isolation, integration and putative correlation of regulatory data with differential gene expression patterns of a candidate gene implicated as an important target to study the control of berry ripening.

The major objective of this study was to identify and apply the molecular 'tools' used in conjunction with a candidate gene-approach to enrich the current knowledge and preliminary understanding of grapevine biology. In plant species such as grape, genomic sequence is limited to EST and DNA marker data. Therefore, an integrative approach (discussed in chapter 2) where the combination of a high throughput gene expression system with bioinformatics and promoter data could be used to form a more accurate understanding and putative prediction of complex regulatory events i.e. during berry development. The current status of plant genomics especially in the areas of transcription, promoters (isolation and evaluation strategies), bioinformatics and combinatorial regulatory network analysis are discussed in chapter 2.

A high-throughput gene expression system, cDNA-AFLP technology, was applied for the first time in grapevine to identify differentially expressed genes during berry ripening (see chapter

3). This system has allowed for the rapid identification of a large amount of grape cDNAs to be selected on the basis of gene expression profile and putative function for promoter isolation.

The integrative approach, described in chapter 2, was implemented and evaluated in grapevine in this study. For this purpose the *Vitis vinifera* L. vacuolar pyrophosphatase promoter was isolated and an integrative analysis of putative regulatory elements and transcriptional activity during stress and ripening was conducted in an attempt to advance the current knowledge of this proton pump and its role in relation to grapevine development (see chapter 4).

References

Colova-Tsolova V, Perl A, Krastanova S, Tsvetkov I, Atanassov A. 2001. Genetically engineered grape for disease and stress tolerance. In: Roubelakis-Angelakis KA, ed. Molecular biology and biotechnology of the grapevine. Dordrecht: Kluwer Academic Publishers. p 411-432.

Coombe BG. 1992. Research on development and ripening of the grape berry. Am.J.Enol.Vitic. 43(1): 101-110.

Davies C, Boss PK. 2000. The use of molecular techniques to study and manipulate the grapevine: Why and how? Australian Journal of Grape and Wine Research 6: 159-167.

Davies C, Robinson SP. 2000. Differential screening indicates a dramatic change in mRNA profiles during grape berry ripening. Cloning and characterization of cDNAs encoding putative cell wall and stress response proteins. Plant Physiol 122: 803-812.

Giovannoni J. 2001. Molecular biology of fruit maturation and ripening. Annu Rev Plant Physiol Plant Mol Biol 52:725-749.

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

Kikkert JR, Thomas MR, Reisch BI. 2001. Grapevine genetic engineering. In: Roubelakis-Angelakis KA, ed. Molecular biology & biotechnology of the grapevine. Dordrecht: Kluwer Academic Publishers. p 393-410.

Li Z, Jayasankar S, Gray DJ. 2001. Expression of a bifunctiol green fluorescent protein (GFP) fusion marker under the control of three constitutive promoters and enhanced derivatives in transgenic grape (*Vitis vinifera*). Plant Sci 160: 877-887.

Nunan KJ, Davies C, Robinson SP, Fincher GB. 2001. Expression patterns of cell wall-modifying enzymes during grape berry development. Planta 214: 257-264.

Robinson SP, Davies C. 2000. Molecular biology of grape berry ripening. Australian Journal of Grape and Wine Research 6: 175-188.

Roubelakis-Angelakis KA. 2001. Molecular biology and biotechnology of the grapevine. 1st ed. Dordrecht: Kluwer academic publishers. 474 p.

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C. 1994. Cloning and molecular analysis of structural genes involved in the flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol Biol 24: 743-755.

Terrier N, Ageorges A, Abbal P, Romieu C. 2001. Generation of ESTs from grape berry at various developmental stages. J Plant Physiol 158: 1575-1583.

Vivier MA, Pretorius IS. 2000. Genetic improvement of grapevine: Tailoring grape varietiesfor the third millennium - A review. S Afr J Enol Vitic 21: 5-26.

Chapter 2

Promoter analysis and transcription profiling: Integration of genetic data enhances understanding of gene expression

---

Mauritz Venter and Frederik C. Botha (2004) *Physiologia Plantarum* 120 (1): 74-83.

2.1 Abstract

It is increasingly evident that transcription control might be conserved among organisms. For this reason, genome sequencing and gene expression profiling methods, which have yielded a plethora of data in different organisms, may be applied in species where genomic sequence is limited to mostly expression array and EST data. The identification of transcription factors and promoters associated with gene expression profiles and ESTs could therefore contribute to elucidate and predict complex regulatory events in plants.

2.2 Introduction

Genome sequencing of complex and multi-cellular eukaryotes such as *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster, Homo sapiens* and a draft of *Oryza sativa*, opened new opportunities for identifying the genetic mechanisms and networks that control gene expression in response to different cues. Complete determination of a species' genome is not only valuable for gaining insight into the complexity of a particular organism, but can also serve as a model for other organisms. One potential application may be the study of complex gene regulatory processes such as transcription. The combination of recent computational, high-throughput gene expression and comparative sequencing approaches proved to be powerful tools for mapping, predicting and deciphering gene regulatory functions in eukaryotes, where yeast *S. cerevisiae* has been used extensively as a model (for review see Wyrick and Young 2002).

Completion of two plant genome sequences, a dicotyledonous plant, *A. thaliana* (The *Arabidopsis* Genome Initiative, 2000) and more recently sequence drafts of two subspecies from the monocotyledous member of the graminae, *Oryza sativa* (Yu *et al*. 2002, Goff *et al*. 2002), have paved the way for comprehensive functional characterisation of genes, transcription factors (TFs) and promoters in plants. Genome sequencing initiatives for other plants, not necessarily of commercial or medical interest, could allow for genomic comparisons across a variety of plant species that would subsequently be invaluable to gain further insights into plant evolution, phylogeny and genomic organisation (Pryer *et al*. 2002). Although there are no other plant species for which complete genomic maps exist, a vast array of ESTs (Expressed Sequence Tags) does exist as well as some regulatory sequences. This could allow genomic comparisons with other plants, as well as other multi-cellular eukaryotes. This discussion focuses on current knowledge in the analysis of transcription factors and promoter elements in combination with expression profile data. Such analysis may lead to a better understanding of the regulatory mechanisms involved in gene expression in agriculturally important species.

2.3 High throughput gene expression analysis

The capability to extrapolate biological data from genome-wide or multi-gene expression patterns became feasible after the development of high throughput gene expression technologies such as cDNA microarrays, DD-(Differential display) PCR, cDNA-AFLP (Amplified Fragment Length Polymorfism) and SAGE (Serial analysis of gene expression). The major advantages and drawbacks of several transcript profiling technologies, specifically used in plants, have been compared extensively on the basis of sensitivity, reproducibility, labor intensity, cost, possibility to integrate genetic data and measurement of expression level (see reviews, Kuhn 2001, Donson *et al*. 2002, Breyne and Zabeau 2001). The principles, preparation and applications of specifically cDNA-AFLPs (Bachem *et al*. 1998) and cDNA-microarrays (Aharoni and Vorst 2001) have been described for functional plant genomics and both methods constitute an important advance for genome annotation and identification of promoter elements (Donson *et al*. 2002).

Scientific progress in the fields of medicine, agriculture and an overall better understanding of our and other organisms' biological make-up, will advance significantly when we associate the keywords 'integration' and 'combination' with scientific progress. While gene expression and transcript imaging technologies are unable on their own, to reveal information regarding the underlying genetic control of genes expressed during certain conditions, they can be combined with promoter/motif data to provide a framework for the deciphering and possible integration of coordinated gene activity.

The combination of gene expression profiles with other technologies is certainly not new, other groups have recently used methods such as chromatin immunoprecipitation assays or 'ChIPs' for the rapid identification of target promoters by probing a human transcription factor E2F-antibody to CpG island microarrays instead of cDNA microarrays (Weinmann *et al*. 2002).

With this approach they could successfully identify a large set of known promoters as well as promoter regions of unknown mRNAs. Although not combined with transcript profiles, the ChIP approach was successfully used on *Arabidopsis* embryonic culture tissue to isolate target genes regulated by a MADS-domain protein (AGL15) that preferentially accumulates in embryos (Wang *et al*. 2002). The ChIP-technology is relatively new for plant systems and will be useful when used in conjunction with microarrays to elucidate regulatory networks in plants (Wang *et al*. 2002).

The combination of immunoprecipitated cross-linked protein-DNA complexes with DNA microarrays have been exploited by Simon *et al*. (2001) in yeast and Ren *et al*. (2002) in human. Simon *et al*. (2001) showed how known cell cycle transcription factors regulate global gene expression during different stages of the yeast cell cycle thus enabling them to construct a regulatory map of the transcriptional machinery directly associated with the cell cycle. The results of Ren *et al*. (2002) identified both known and novel genes, regulated by the human E2F transcription factor family, required for cell cycle progression. From these studies it was evident that microarray gene expression data used in conjunction with immunoprecipitated cross-linked protein-DNA complexes are a powerful approach to identify the direct connection of regulatory motifs with coordinated gene expression events during e.g the cell cycle. These studies underscore the importance of combining gene expression profiles with other technologies and databanks for the prediction, understanding and management of biological data in plant systems.

2.4 Plant promoters

The basic plant promoter consists of a core promoter region and upstream *cis*-elements a.k.a. TF-binding sites. Core promoter architecture and orchestrated pre-initiation complex (PIC) assembly are basically universal in all eukaryotes where the core promoter is the minimal promoter region necessary to direct gene transcription. Transcriptional initiation of protein

coding genes by RNA polymerase II involves the stepwise assembly of general transcription factors (GTFs) to the core promoter, usually containing the TATA-box and/or an initiator element, to form a stable PIC (Roeder 1996). There are however promoters with no TATA-boxes or initiator elements and it has been demonstrated in *Drosophila* that some TATA-less promoters have substitutes known as down-stream promoter elements (DPEs) (Burke and Kadonaga 1996). Knowledge of TATA-less promoters in plants is limited but it was recently found that the majority of TATA-less promoters were common to genes associated with photosynthesis (Nakamura *et al*. 2002).

Gain-of function experiments have shed light on a vast array of *cis*-elements that are associated with specific expression patterns in response to constitutive, developmental, tissue-specific, hormonal and environmental regulation. The nature of specific *cis*-elements acting as activators, repressors, enhancers and chromatin modifiers is crucial for the combinatorial transcriptional regulation in plants. Although not discussed in this review, it is increasingly evident that the interaction of a multitude of different overlapping *cis*-elements and TFs are responsible for diverse gene expression events.

A number of promoters responsive to specifically environmental and hormonal stimuli have been characterized. Promoters induced by environmental cues include light-, heat-, cold-, anaerobic stress (hypoxia)-, dehydration- and elicitor-responsive promoters and are usually associated with plant defence and survival. Hormonal responsive promoters are induced by auxin, abscisic acid (ABA), gibberellic acid (GA), ethylene, salicylic acid (SA) and/or methyl jasmonate (MeJA) and play an important role during growth, development, ripening, flowering and certain stress responses. Genes expressed during seed development code for storage proteins and have distinct promoters with conserved motifs required for seed-specific expression (for a comprehensive review of the structure of plant promoters see Guilfoyle 1997).

Various strategies have been used for promoter isolation. These approaches are based upon the isolation/cloning of flanking genomic DNA regions, especially in non-model plants where only cDNAs exist. The three basic methods include genomic DNA library screening (Yoshida and Shinmyo 2000, Trindade *et al*. 2003), inverse PCR (Ochman *et al*. 1988, Chabouté *et al*. 2000) and promoter trapping utilising a mobile element (T-DNA or transposon) containing a reporter gene (Meissner *et al*. 2000). The advantage of fully sequenced model plants such as *Arabidopsis* facilitates for easier promoter identification and comparative analysis. Seki *et al*. (2002) identified promoter regions by comparing data of the 5'-ends of isolated full-length cDNAs, from *Arabidopsis* plants, to genomic sequences of *Arabidopsis* and constructed a promoter database using data from a plant transcription factor database, PLACE (Higo *et al*. 1999).

Stable or transient promoter activity can be assessed when the genomic 5'-untranslated region (UTR) is fused to a reporter gene such as luciferase (LUC; Miller *et al*. 1992), β-glucuronidase (GUS; Jefferson *et al*. 1987) (Figure 2.1A) or green fluorescent protein (GFP; Elliott *et al*. 1999) (Figure 2.1B) and transferred into plant cells by *Agrobacterium*-mediated infection or particle bombardement that allows detection by fluorescence (LUC, GFP) or activity staining *in situ* (GUS) (see Yoshida and Shinmyo 2000 and references therein).

The cauliflower mosaic virus (CaMV) 35S promoter is the workhorse-promoter for transgenic analysis in plants mostly used as a control for comparison to test other promoter elements (Holtorf *et al*. 1995) or to confer constitutive and high level expression of specific genes (Kay *et al*. 1987, Yoshida and Shinmyo 2000). Today, however, the model promoters are divided for optimal gene expression in *dicots* and *monocots.* Although previously utilised in most plant species, the CaMV 35S promoter is more efficient in dicot-plants than in grass-like monocots. With the use of the maize ubiquitin (Ubi-1) promoter (Christensen *et al*. 1992), a more than 10-fold higher level of expression in comparison to the CaMV 35S promoter was possible. With the focus on combinatorial control and understanding of transcription in plants, it was Benfey

and Chua (1990) that initially used the CaMV 35S promoter to investigate synergistic interactions of the promoters' subdomains. This early study revealed how specific *cis*-element (subdomain) combinations of the CaMV 35S promoter conferred different β-glucuronidase (GUS) expression patterns between tobacco and petunia (Benfey and Chua 1990). Since 1990, numerous other promoters (viral and plant) have been isolated and used for high-level transgene expression. Schenk *et al*. (2001) evaluated promoters from the banana streak badnavirus (BSV) in transgenic tobacco, banana and sugarcane plants. The BSV-promoters conferred high-level transgene expression that was similar to activities of the maize ubiquitin (Ubi-1) and CaMV 35S promoters. Nevertheless, the CaMV 35S promoter has paved the way for analysing the integration of regulatory elements with gene expression patterns. Used as a control, to confer foreign gene expression or as a platform for the initiation of transcription in synthetic promoters, the promoter of choice is from the CaMV 35S.

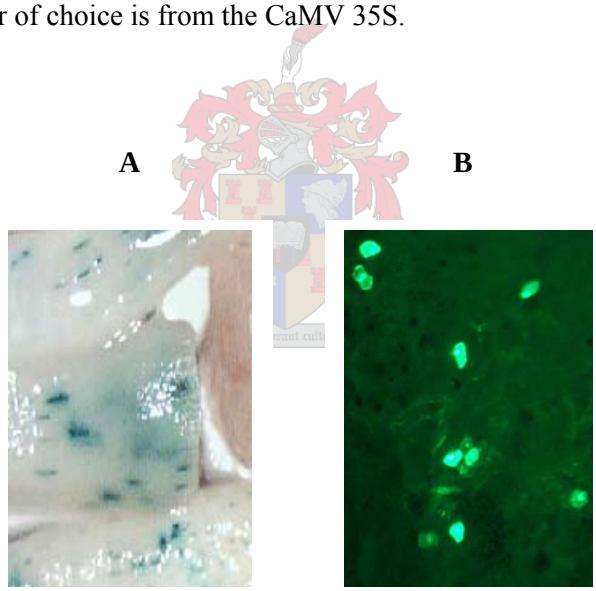


Figure 2.1 Evaluating promoter activity in transgenic plant tissue: A) Ubi1-promoter driven β-glucuronidase (GUS) expression (blue) in young sugarcane stalks and B) CaMV 35S-promoter driven green fluorescent protein (GFP) expression in young green pepper.

2.5 Plant transcription factors

Besides the general transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH) necessary to initiate transcription, other TFs associated with protein-protein and DNA-protein interactions have been analysed in higher eukaryotes. Binding of the transcription factor IID (TFIID), containing a TATA-box binding protein (TBP), to the TATA-box of the core promoter plays an essential part in the transcription-initiation complex (see reviews Novina and Roy 1996, Singh 1998, Martinez 2002). Functional significance of specific transcription initiation interactions such as the TBP-TFIIB interaction, has been evaluated in humans and yeast but in plants it was shown that transcriptional activity of the cauliflower mosaic virus 35S and maize ubiquitin promoters does not always necessitate strong TBP-TFIIB interactions (Pan *et al*. 2000).

Plant transcription factor (TF) studies have shed light on complex and diverse aspects of plant development, tissue-specific gene regulation and response to environmental stimuli. Interactions of TFs with combinations of *cis*-motifs within promoter regions account for the specificity of gene expression and allow the elucidation of combinatorial control in plants (Singh 1998, Zhou 1999). These regulatory functions are mostly poorly understood. However numerous labour-intensive efforts have revealed specific transcriptional regulators such as the Dof genes (Yanagisawa 2002) and the WRKY ('worky') proteins (Eulgem *et al*. 2000) to be plant specific and/or functionally related.

Apart from common TF-functions, studies revealed that certain TFs in *Arabidopsis*, responsible for floral meristem development, could move between cells and remain active (Sessions *et al*. 2000). It has recently been shown that transcription factors could be used to modulate plant metabolism for the production and engineering of plant pharmaceutical compounds such as alkaloids and flavonoids. These studies show great promise for the industrial production of nutraceuticals and other therapeutic active molecules (Gantet and Memelink 2002). The

functional specificity of different plant transcription factors will however not be discussed. For comprehensive reviews of plant transcription factors and TF-binding sites within plant promoters see Schwechheimer *et al*. (1998), Eulgem *et al*. (2000), Yanagisawa (2002) and Guilfoyle (1997).

We focus on transcriptome profiling, specifically in *Arabidopsis*, for genome-wide analysis of plant transcription factors. Major transcription factor gene families and their functions have been characterised in *Arabidopsis* (Riechmann and Ratcliffe 2000). Some of these TFs and their functions (in brackets) include; Myb (signal transduction), AP2/EREBP (ABA and ethylene response), bZIP (seed-storage gene expression), MADS (flower development), WRKY (defence response), ARF-Aux/IAA (Auxin response) and Dof (endosperm-specific expression). Comparative studies of fully sequenced organisms *S. cerevisiae*, *C. elegans* and *D. melanogaster* to *A. thaliana* revealed certain TFs (WRKY, Dof, ARF-Aux/IAA) to be plant-specific or present in all species (bZIP, MADS). However, certain TFs such as the MADS-box genes, associated with flower development, revealed to be functionally redundant. It is envisaged that extensive microarray analysis will prove to be exceptionally useful for genome-wide transcriptome analysis to obtain a holistic view of transcriptional regulation in plants (Riechmann and Ratcliffe 2000, Riechmann *et al*. 2000).

Consigning to the fully sequenced plant model *Arabidopsis,* Chen *et al*. (2002) have extrapolated a considerable amount of data from 402 possible stress-related genes encoding known or putative transcription factors with the use of microarrays. Potential functions of these genes were annotated by evaluating their expression profiles in different developmental stages during more than 80 stress conditions and from different tissue samples (stems, roots, leaves etc). These results were then analysed to construct a 2-D transcription matrix of genes vs. treatments or developmental stages and tissues. Further evaluation and confirmation of these results are recommended to overcome microarray technical complications (Chen *et al*. 2002), nevertheless, under diverse stress and developmental-stage specific conditions, the

multifunctional role of transcription factors could be observed. It was revealed that 74 TF-genes responsive to pathogen infection had additional responses to plant hormones, showing that the transcription factor genes play a multifunctional role. From clusters of genes induced by specific stresses (pathogen and cold stress), conserved promoter *cis*-motifs were identified (Chen *et al*. 2002). Recently transcriptome profiling in *Arabidopsis* was used to identify cold responsive genes activated from multiple cold tolerance pathways. Apart from the orchestrated activation of cold responsive genes from the CBF (C-repeat-binding factor) regulon, other genes were transiently expressed, suggesting independent and overlapping (during development) cold responsiveness (Fowler and Thomashow 2002). The tissue specific nature of TF gene expression is poorly understood, thus suggesting that some TFs have general 'house-keeping' functions as well as specific cues in different plant organs depending on the occurrence frequency or combination of TF-elements and binding sites within promoters.

2.6 Bioinformatic tools and database assistance

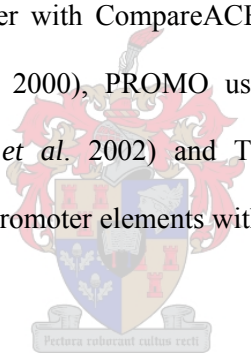
The accumulation of plant biological data, specifically of plant promoter/TF data, and subsequent database construction can be utilised for more accurate comparisons and analysis of plant regulatory elements. Several software tools exist, either for deposition and organisation of biological data and/or for the prediction of putative functions or the structure of complex molecules. Experimental characterisations of regulatory promoter elements are usually conducted by sequential deletion of promoter fragments and promoter (with gain-of-function *cis* element) activities are assessed on transient and stable level in transgenic plants (as recently described by Cho and Cosgrove 2002, Koch *et al*. 2001). However, here we focus on computational tools for promoter and regulatory element prediction, identification of TF-binding sites relating to gene expression profiles as well as the use of transcription factor databases (specifically for plants).

At this stage only three major databases exist for the identification of TF-binding sites and *cis*-elements in plant promoters (Hehl and Wingender 2001). They are PLACE (Higo *et al*. 1999), PlantCARE (Rombauts *et al*. 1999) and TRANSFAC (Wingender *et al*, 2000; references in Hehl and Wingender 2001). The eukaryotic promoter database (EPD) is a collection of currently 1402 Polymerase II promoters from multicellular species of which 198 are from plants (see references in Périer *et al*. 2000). These databases are cross-referenced and contain regulatory information from model organisms such as *Arabidopsis thaliana* (Seki *et al*. 2002) and *Saccharomyces cerevisiae* (SCPD: *Saccharomyces cerevisiae* promoter database; Park *et al*. 2002) to allow for comparative sequence and gene expression profile analysis. Numerous software links such as GETools; to integrate *Escherichia coli* microarray data with regulatory elements (Huerta *et al*. 2002) and GenEST; linking cDNA-AFLP transcript profiles with EST data either to select or to discard ESTs on the basis of expression profiles (Qin *et al*. 2001) are available. These integrative software tools allow for comparison, analyses and prediction of new entries to several experimentally validated levels of data e.g. gene expression profile or *cis*-elements in the promoter.

Eukaryotic polymerase II promoter and TF-binding site identification systems are constructed and used with the implementation of specific or combinations of strategies. These strategies include frequency analyses of repetitive and/or similar sequences (expected frequency versus observed frequency), alignment methods, neural network, clustering and association algorithms and database mining. Motif clustering approaches are developed on the basis of finding motif similarities by comparing sequences. Various sequence alignment tools exist, but the most widely used and well-established method is the BLAST (Basic local alignment search tool; Altschul *et al*. 1990) method. The limitation is however that BLAST only allows alignment of two sequences, therefore systems like CLUSTAL W and CLUSTAL X (Aiyar 2000) that allow for alignment of multiple sequences, have been designed. Various detection algorithms such as Dragon Promoter Finder (DPF; Bajic *et al*. 2002), PromoterInspector (Scherf *et al*. 2000) and Promoter2.0 (Knudsen 1999) are designed to discriminate between core-promoter elements

(such as the TATA-box and transcription start site:TSS), exons and introns. However, because of the sensitivity of noise; 1) presence of promoter sequences not containing the desired motif or 2) intergenic sequences that are not biologically functional, the algorithms are updated and improved continuously (Ohler and Niemann 2001, Thijs *et al*. 2001).

The important promoter detection methods, based on maximum likelihood estimation, are Gibbs-sampling (Lawrence *et al*. 1993) and expectation maximization (MEME; Bailey and Elkan 1995). Gibbs-sampling and MEME are known as probabilistic methods that perform a local optimisation on motifs in sequences to be unknown (hidden motif in noisy background sequence) in order to identify the most conserved sequence (Thijs *et al*. 2001, Ohler and Niemann 2001). Other methods used for the identification of transcriptional regulatory elements include AlignACE (together with CompareACE and ScanACE) based on a Gibbs-sampling algorithm (Hughes *et al*. 2000), PROMO using sequence information from the database TRANSFAC (Messeguer *et al*. 2002) and Transcription-factor-centric clustering (TFCC) designed to link regulatory promoter elements with the binding TFs *in silico* (Zhu *et al*. 2002).



Although not complete, the databases serve as invaluable tools for the transcriptional prediction and elucidation of new-entry sequences with known and/or unknown functions. Ultimately however, complete databases would be useful for the higher accurate detection and synthetic construction of plant promoters (Figure 2.4) designed *in silico* for desired gene expression patterns (Hehl and Wingender 2001).

2.7 Regulatory network analysis

Recent comprehensive regulatory network studies on *Saccharomyces cerevisiae* and *Escherichia coli* have led to the identification and network 'building block' assembly of a vast array of transcriptional regulators and interconnected motifs (Lee *et al*. 2002; Milo *et al*. 2002;

Shen-Orr *et al*. 2002). Complete genome sequence data on these 'relative simple' model organisms have allowed for the experimental design of genome-wide binding/location analysis. Genome-wide binding analysis (e.g. using ChIPs) was integrated with microarray expression profile data to identify target genes and promoters in vivo (Lee *et al*. 2002). These motif network studies are invaluable to gain a holistic view of the complete transcriptional regulatory network of organisms (Shen-Orr *et al*. 2002, Milo *et al*. 2002), emphasizing the importance of integrative experimental designs and strategies.

Experimental data on yeast allowed Pilpel *et al*. (2001) to correlate occurrences of transcription factor (TF) binding site combinations with specific gene expression patterns. Computational and statistical methods were utilized to generate TF/motif synergy maps that demonstrated how specific motif combinations were associated with specific gene expression patterns during certain events such as sporulation or stress conditions. It is evident that the synergestic motif approach of Pilpel *et al*. (2001) could allow for the identification of different combinations of similar TF-motifs associated with different specific expression profiles. The strategy of Pilpel *et al*. (2001) led to the identification of specific co-clustered motifs associated with stress response (heat shock and treatment with DNA-damaging agents) and the cell cycle (G_1-, G_2-phase and sporulation). Motif synergy maps were constructed showing motif relationships associated with specific conditions (e.g. heat shock, cell-cycle, sporulation and DNA-damage). Figure 2.2 represents a simplified motif synergy map where different stress conditions are represented by coloured lines. Motif 1, known as the common motif, is present in the promoters of genes expressed during all the stress conditions. The synergy map (Figure 2.2) shows how the common motif (motif 1) is connected (connection represent stress condition) in combination with other motifs (motifs 2,3 and 4) within the promoters of genes associated with specific stress conditions. Synergistic *cis*-motif pairs 1 and 2 are present in promoters of genes associated with cold- and heat shock, motif 1 and 3 are associated with heat shock and DNA-damage and motif 1 and 4 are associated with cell-cycle, sporulation and DNA-damage (Figure

2.2). According to these maps, motifs associated with similar conditions e.g. cell cycle, clustered together and synergized with each other.

These synergistic links (analysis of co-occurred motifs) were visualised with so-called "combinograms" providing a global view of the motif combinations. During different stages of the cell cycle or different stress responses, a specific/common motif was present in a cluster (in combination with other motifs). The nature of the identified common motif was predicted/speculated i.e. if the motif was necessary to invoke a specific expression or if the motif was the major determinant of the expression pattern. Specific motif function (activate, repress, chromatin modifier) could not be determined, however, based on promoter similarities of different species (e.g. human and mouse) Pilpel *et al*. (2001) suggested that their strategy could be harnessed to predict expression profiles of genes with no experimentally validated data available.

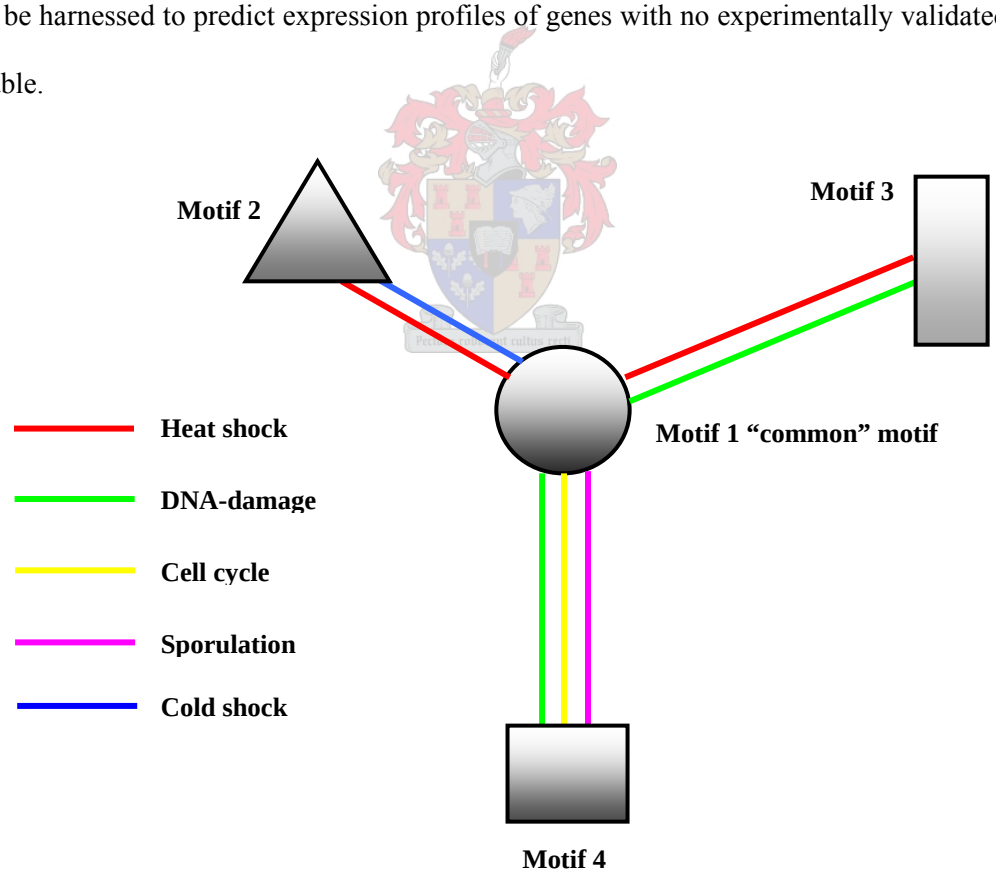


Figure 2.2 Simplified representation of a motif synergy map, illustrating interplay of motifs during different stress conditions (represented by coloured lines connecting motifs).

Another study (Fessele *et al*. 2002) similarly accentuated the strategy of Pilpel *et al*. (2001) by evaluating the functional organisation of different TF-binding sites in the human RANTES/CCL5 promoter in five different tissues under stimulated and unstimulated conditions. The RANTES/CCL5 promoter was characterised and a model of the promoter was constructed on the basis of previous experimental analyses. This information was used to subsequently construct organisational submodels of the RANTES/CCL5 promoter for five different cell-types allowing *in silico* visualisation of TF-binding sites/motifs and overlapping binding of transcriptional factors (experimentally confirmed) in different combinations. From all these studies, only the organisational structure of the promoter sequence in its steady state can be observed. In all organisms, where the function of many genes has been identified, we need to ask why were those genes expressed during those specific conditions in a tissue-specific manner. An important part of the answers lies within the promoter area of those genes, thus associating a specific condition and gene expression pattern with the regulatory organisation that allows for the elucidation of why and how that gene was switched on during a specific time and place. Specifically in plants, this approach will not only help us elucidate the regulatory nature of genes expressed e.g. during development or flowering, but also to accurately address what elements acting as receptors are induced, within the promoter, to drive gene expression under a specific condition.

A simplified approach, utilizing the strategies of Pilpel *et al*. (2001) and Fessele *et al*. (2002), for the identification, understanding and subsequent prediction of a promoter motif combination associated with a gene expression in plants is presented in figure 2.3. Although other motifs are associated with specific expression patterns, a certain combination is associated with a specific condition. Comparative analysis of experimentally confirmed plant database results could allow us to verify a specific regulatory context to a specific gene expression pattern and subsequently to the same expression pattern from variety of plant species because of the conserved nature of plant (and other) transcription factor binding sites. As shown by Pilpel *et al*. (2001), the presence or absence of other motifs within a known combination allows for genes to be

expressed during other conditions either with or without expression during one specific condition. Utilising the approach of Pilpel *et al*. (2001), in plants (specifically non-model systems), it could be possible to integrate plant gene expression profiles with promoter information to form a more accurate understanding and prediction of gene expression when compared to other combinations of promoter TFs under the same condition. With this strategy, although not a replacement for experimental verification, it would be possible to predict the gene expression pattern of previously unknown EST sequences in plant species where the whole-genome approach is not feasible either because of the vast genome size of a specie itself, priority/importance of the crop or even different genome sizes of the same specie e.g. sugarcane. Whole genome sequencing has laid the groundwork for large-scale investigations pushing us into the 'post-genomic' era with the relative 'new' frontier of *proteomics* to be explored. For the monocot and dicot plants where this 'prerequisite' of genome sequencing have not been met, the combination of a high throughput transcript profile system (cDNA-AFLPs or –microarrays) with bioinformatics, genomic comparisons and TF/promoter data is a powerful approach for the regulatory functional analysis of known and unknown genes.

2.8 Transcriptional similarity assists in the identification of unknown genes

Comprehensive studies by Klok *et al*. (2002) and Moseyko *et al*. (2002), utilizing a similar approach to Pilpel *et al*. (2001), revealed the identification of common promoter motifs to be involved in similar gene expression profiles during low oxygen (Klok *et al*. 2002) and early gravitropic (Moseyko *et al*. 2002) responses. Both groups used microarray technology to identify known and discover new genes expressed in *Arabidopsis.*

Gravitropic-response microarray profiles (Moseyko *et al*. 2002) revealed genes regulated by gravitropic responses as well as to mechanical perturbations (during reorientation of plants). Genes with specific expression profiles were clustered together, analysed and divided into functional categories of which the majority of gravity-regulated genes were involved in

oxidative burst and plant defence (Moseyko *et al*. 2002). AlignACE software (Hughes *et al*. 2000) was used for promoter analyses that led to the identification of common sets of known and unknown motifs when compared to TF-binding site information available from the TRANSFAC database (Hehl and Wingender 2001).

Klok *et al*. (2002) compared identified motifs, within promoters from clustered groups of genes, to the TRANSFAC (Hehl and Wingender 2001) and PlantCARE (Rombauts *et al*. 1999) databases. Genes that responded to low-oxygen stress conditions were clustered in groups according to their differential expression profile when monitored at specific time points. Common *cis*-regulatory motifs within promoter regions belonging to genes in specific clusters were identified, emphasizing the coordinated (or competitive) role of regulatory motifs associated with a specific gene expression profile under certain stress responses. It was apparent that previously well-characterised promoters, used as references, are invaluable for the identification of common sets of motifs associated with a specific gene expression profile. Klok *et al*. (2002) subjected a cluster of genes with similar expression (microarray confirmed by real-Time PCR) profiles to the expression pattern of the alcohol-dehydrogenase gene (*ADH1*) to search common *cis*-elements. The promoter of the *ADH1* gene is well characterised and served as a reference promoter for comparative motif analysis.

From the results obtained by Klok *et al*. (2002) and Moseyko *et al*. (2002), coupling microarray data with promoter structure in plants, it was evident that promoters of genes (or cluster of genes) with a specific expression profile contained similar motifs. We illustrate this strategy (simplified in Figure 2.3) starting from: high-throughput gene expression profiles during certain condition, identification of promoter elements of ESTs with unknown function, identification of conserved motifs present in all the promoters of clustered ESTs, comparison to well characterized 'reference' promoter of gene with similar expression profile, putative functional prediction of ESTs expressed during certain condition. Although not a replacement for

experimental verification, we suggest that a characterized promoter and/or combination of motifs, as reference, could assist in the identification of unknown ESTs/genes.

2.9 Future prospects

As a consequence of these studies and information gained, we can envisage a more optimistic *in silico* 'cut and paste' approach for the construction of synthetic promoters (Figure 2.4). Synthetic promoters to be induced in transient plant expression systems are not new and previous attempts have successfully implemented synthetic promoters to confer desired transgene expression (Puente *et al*. 1996, Yoshida and Shinmyo 2000). However, the uncertainty of these attempts arose when the functional retention of different *cis*-elements, removed from their local promoter regions, was questioned (Rushton *et al*. 2002). The extensive study by Rushton *et al*. (2002) demonstrated the construction of synthetic pathogen-induced promoters using previously studied *cis*-elements such as the W-boxes (Eulgem *et al*. 2000) and other elements associated with pathogen-inducibility. Expression patterns of promoters were evaluated during various interactions with different pathogens, number and spacing of *cis*-elements, responses to wounding and combinations of defined *cis*-acting elements. The minimal promoter region (TATA-box region for binding to the 'PIC' pre-initiation complex) of the CaMV 35S was used to initiate transcription after induction of specific *cis*-elements (similar representation in Figure 2.4). Spacing of elements from each other as well as from the TATA-box revealed only changes in promoter strength but omissible effects on the inducibility could be observed. From all these results, Rushton *et al*. (2002) emphasized the importance of previous experimental verification as supported by Hehl and Wingender (2001) utilising database assistance for promoter analysis. The synthetic promoter work described by Rushton *et al*. (2002) is promising and could be used as a replacement of conventional 'wild-type' promoters. Today as we approach the near completion of genome sequencing initiatives (at least for model organisms), our predictions to understand gene regulatory events must await experimental confirmation.

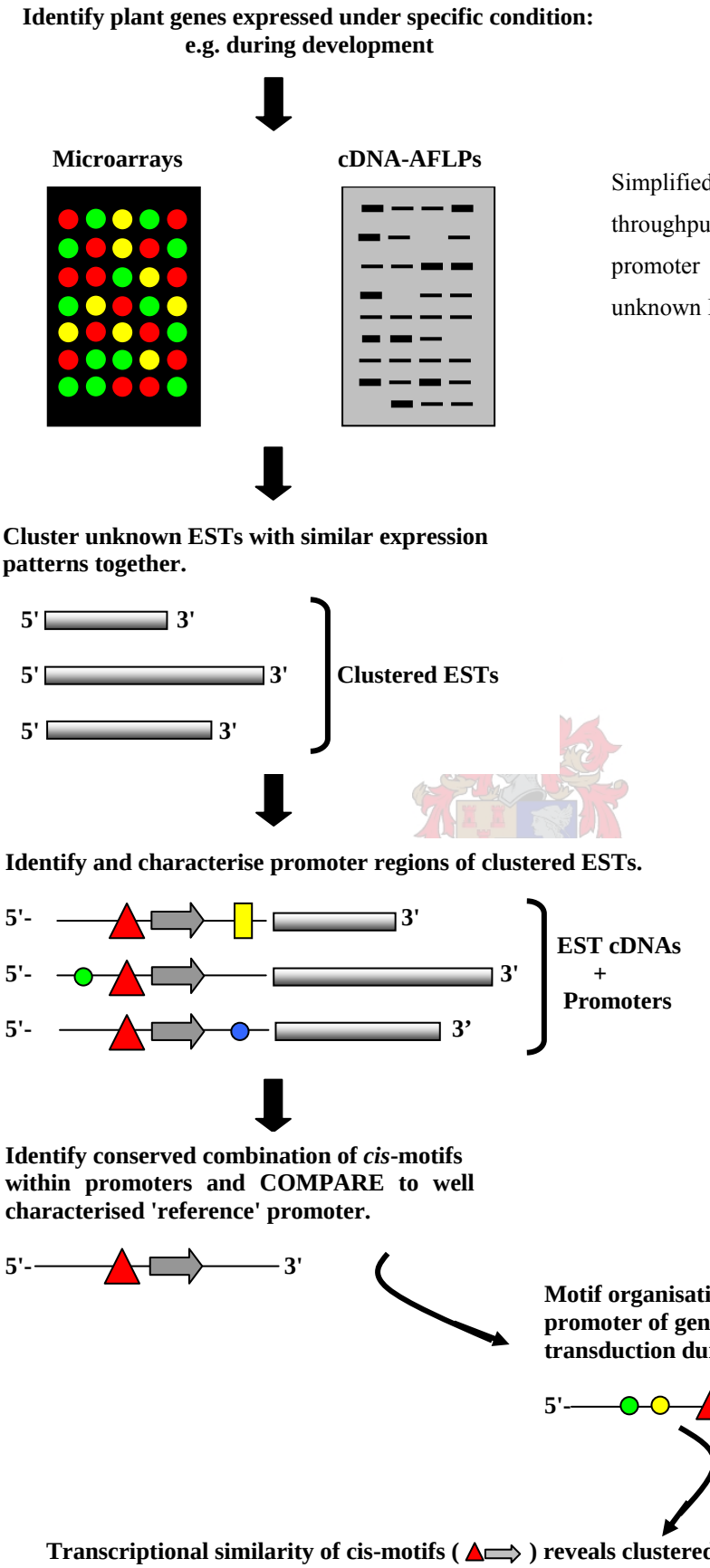


Figure 2.3

Simplified representation of integrating high-throughput gene expression profiles with promoter data for the identification of unknown ESTs.

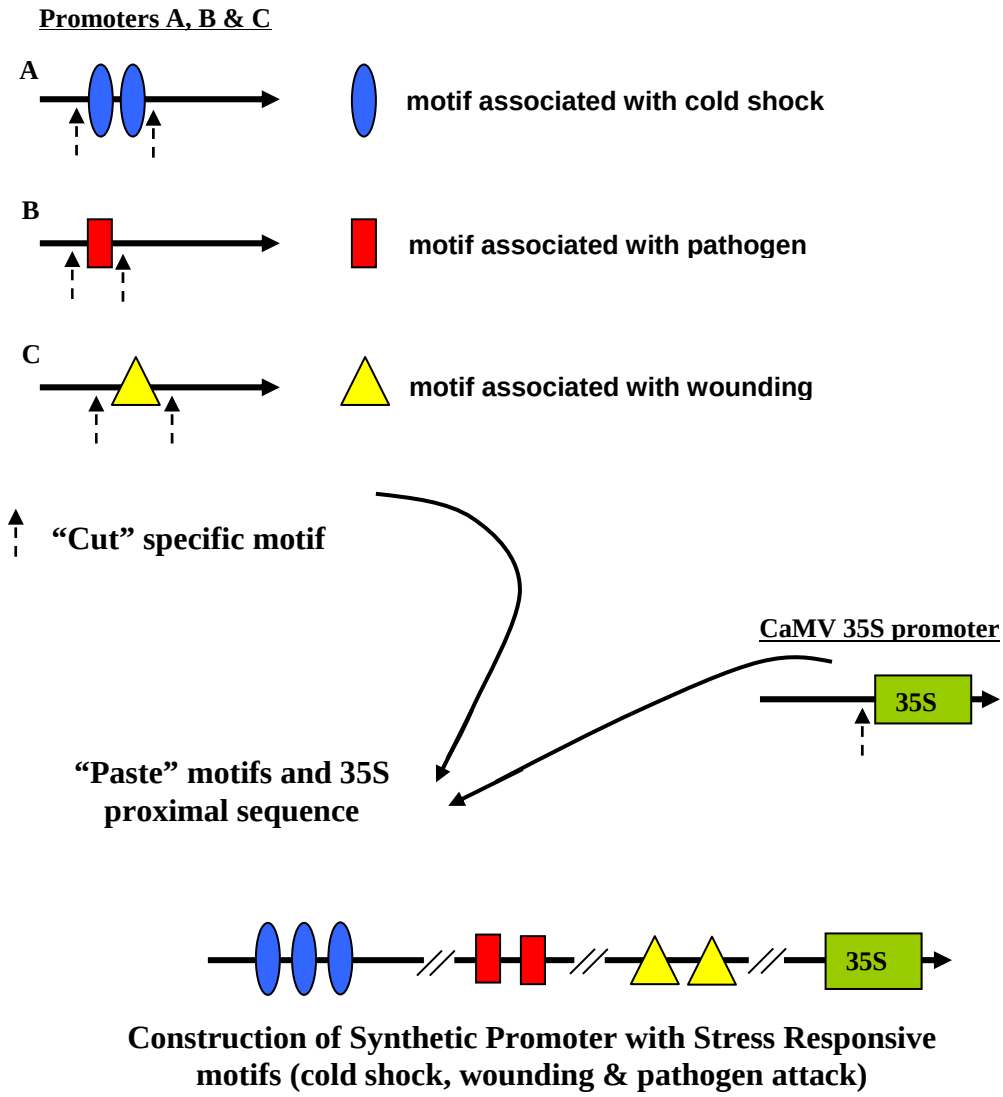


Figure 2.4 The 'Cut and Paste' approach of different TF-binding sites for the construction of a synthetic promoter to be induced by cold shock, wounding and pathogen attack.

Acknowledgments

The authors thank Jennie Van der Merwe and Sue Bosch for helpful comments on the manuscript.

References

Aharoni A, Vorst O (2002) DNA microarrays for functional plant genomics. Plant Mol Biol 48: 99-118

Aiyar A (2000) The use of CLUSTAL W and CLUSTAL X for multiple sequence alignment. Methods Mol Biol 132: 221-241

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Nucleic Acids Research 215: 403-410

Bachem CWB, Oomen RJFJ, Visser RGF (1998) Transcript imaging with cDNA-AFLP: A step-by-step protocol. Plant Molecular Biology Reporter 16: 157-173

Bailey TL, Elkan C (1995) Unsupervised learning of multiple motifs in biopolymers using Expectation Maximization. Mach Learn 21: 51-80

Bajic VB, Seah SH, Chong A, Zhang G, Koh JL, Brusic V (2002) Dragon Promoter Finder: recognition of vertebrate RNA polymerase II promoters. Bioinformatics 18: 198-199

Benfrey PN, Chua N-H (1990) The Cauliflower Mosaic Virus 35S promoter: Combinatorial regulation of transcription in plants. Science 25: 959-966

Breyne P, Zabeau M (2001) Genome-wide expression analysis of plant cell cycle modulated genes. Curr Opin Plant Biol 4: 136-142

Burke TW, Kadonaga JT (1996) Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev. 10: 711-724

Chaboute ME, Clement B, Sekine M, Philipps G, Chaubet-Gigot N (2000) Cell cycle regulation of the tobacco ribonucleotide reductase small subunit gene is mediated by E2F-like elements. Plant Cell 12: 1987-2000

Chen W, Provart NJ, Glazebrook J, Katagiri F, Chang H-S, Eulgem T, Mauch F, Luan S, Zou G, Whitham SA, Budworth PR, Tao Y, Xie Z, Chen X, Lam S, Kreps JA, Harper JF, Si-Amour A, Mauch-Mani B, Heinlein M, Kobayashi K, Hohn, T, Dangl JL, Wang X, Zhu T (2002) Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses. The Plant Cell 14: 559-574

Cho HT, Cosgrove DJ (2002) Regulation of root hair initiation and expansin gene expression in Arabidopsis. Plant Cell 14: 3237-3253

Christensen AH, Sharrock RA, Quail PH (1992) Maize polyubiquitin genes: structure,

thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mol.Biol 18: 675-689

Donson J, Fang Y, Espiritu-Santo G, Xing W, Salazar A, Miyamoto S, Armendarez V, Volkmuth W (2002) Comprehensive gene expression analysis by transcript profiling. Plant Mol Biol 48: 75-97

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

Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Trends Plant Sci 5: 199-205

Fessele S, Maier H, Zischek C, Nelson PJ, Werner T (2002) Regulatory context is a crucial part of gene function. Trends Gen 18: 60-63

Fowler S, Tomashow MF (2002) Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold responsive pathway. Plant Cell 14: 1675-1690

Gantet P, Memelink J (2002) Transcription factors: tools to engineer the production of pharmacologically active plant metabolites. Trends Pharmacol Sci 23: 563-569

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R,

Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296: 92-100

Guilfoyle TJ (1997) The structure of plant gene promoters. In: Setlow JK (ed) Genetic Engineering. Plenum Press, New York

Hehl R, Wingender E (2001) Database-assisted promoter analysis. Trends Plant Sci 6: 251-255

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res 27: 297-300

Holtorf S, Apel K, Bohlmann H (1995) Comparison of different constitutive and inducible promoters for the transgenes in Arabidopsis thaliana. Plant Mol Biol 29: 637-646

Huerta AM, Glasner JD, Gutiérrez-Ríos RM, Blattner FR, Collado-Vides J (2002) GETools: gene expression tool for analysis of transcriptome experiments in *E.coli*. Trends Gen 18: 217-218

Hughes JD, Estep PW, Tavazoie S, Church GM (2000) Computational identification of cis-regulatory elements associated with groups of functionally related genes in Saccharomyces cerevisiae. J Mol Biol 296: 1205-1214

Jefferson RA, Kavanagh TA, Caboche M (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901

Kay R, Chan A, Daly M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236: 1299-1302

Klok EJ, Wilson IW, Wilson D, Chapman SC, Ewing RM, Somerville SC, Peacock WJ, Dolferus R, Dennis ES (2002) Expression profile analysis of the low-oxygen response in *Arabidopsis* root cultures. Plant Cell 14: 2481-2494

Koch MA, Weisshaar B, Kroymann J, Haubold B, Mithell-Olds T (2001) Comparative genomics and regulaory evolution: Conservation and function of the *Chs* and *Apetala3* promoters. Mol Biol Evol 18: 1882-1891

Knudsen S (1999) Promoter2.0: for the recognition of Pol II promoter. Bioinformatics 15: 356-361

Kuhn E (2001) From library screening to microarray technology: strategies to determine gene expression profiles and to identify differentially regulated genes in plants. Ann Bot 87: 139-155

Lawrence CE, Altschul SF, Boguski MS, Liu JS, Neuwald AF, Wootton JC (1993) Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. Science 262: 208-214

Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM , Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren

B, Wyrick JJ, Tagne JB, Volkert TL, Fraenkel E, Gifford DK, Young RA (2002) Transcriptional regulatory networks in Saccharomyces cerevisiae. Science 298: 799-804

Martinez E (2002) Multi-protein complexes in eukaryotic gene transcription. Plant Mol.Biol. 50: 925-947

Meissner R, Chague V, Zhu Q, Emmanuel E, Elkind Y, Levy AA (2000) Technical advance: a high throughput system for transposon tagging and promoter trapping in tomato. Plant J 22: 265-274

Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM (2002) PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics 18: 333-334

Miller AJ, Short SR, Hiratsuka K, Chua N-H, Kay SA (1992) Firefly luciferase as a reporter of regulated gene expression in higher plants. Plant Mol Biol Rep 10: 345-361

Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U (2002) Network motifs: simple building blocks of complex networks. Science 298: 824-827

Moseyko N, Zhu T, Chang H-S, Wang X, Feldman LJ (2002) Transcription profiling of the early gravitropic response in Arabidopsis using high-density oligonucleotide probe microarrays. Plant Physiol 130: 720-728

Nakamura M, Tsunoda T, Obokata J (2002) Photosynthesis nuclear genes generally lack

TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. Plant J. 29: 1-10

Novina CD, Roy AL (1996) Core promoters and transcriptional control. Trends Gen 12: 351- 355

Ochman H, Gerber AS, Hartl DL (1988) Genetic applications of an inverse polymerase chain reaction. Genetics 120: 621-623

Ohler U, Niemann H (2001) Identification and analysis of eukaryotic promoters: recent computational approaches. Trends Genet 17: 56-60

Pan S, Czarnecka-Verner E, Gurley WB (2000) Role of the TATA binding protein-transcription factor IIB interaction in supporting basal and activated transcription in plant cells. Plant Cell 12: 125-136

Park PJ, Butte AJ, Kohane IS (2002) Comparing expression profiles of genes with similar promoter regions. Bioinformatics 18: 1576-1584

Périer RC, Praz V, Junier T, Bonnard C, Bucher P (2000) The Eukaryotic Promoter Database (EPD): recent developments. Nucleic Acids Res 28: 302-303

Pilpel Y, Sudarsanam P, Church GM (2001) Identifying regulatory networks by combinatorial analysis of promoter elements. Nat Gen 29: 153-159

Puente P, Wei N, Deng XW (1996) Combinatorial interplay of promoter elements constitutes the minimal determinants for light and developmental control of gene expression in Arabidopsis. EMBO J 15: 3732-3743

Pryer KM, Schneider H, Zimmer EA, Banks JA (2002) Deciding among green plants for whole genome studies. Trends Plant Sci 7: 550-554

Qin L, Prins P, Jones JT, Popeijus H, Smant G, Bakker J, Helder J (2001) GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. Nucleic Acids Res 29: 1616-1622

Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290: 2105-2110

Riechmann JL, Ratcliffe OJ (2000) A genomic perspective on plant transcription factors. Curr Opin Plant Biol 3: 423-434

Ren B, Cam H, Takahashi Y, Volkert T, Terragni J, Young RA, Dynlacht BD (2002) E2F integrates cell cycle progression with DNA repair, replication, and G_2/M checkpoints. Gen Dev 16: 245-256

Roeder,RG (1996) The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem Sci 21: 327-335

Rombauts S, Déhais P, Van Montagu M, Rouzé P (1999) PlantCARE, a plant cis-acting regulatory element database. Nucleic Acids Res 27: 295-297

Rushton PJ, Reinstädler A, Lipka V, Lippok B, Somssich E (2002) Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen- and wound-induced signaling. The Plant Cell 14: 749-762

Schenk PM, Remans T, Sági L, Elliott AR, Dietzgen RG, Swennen R, Ebert PR, Grof CPL, Manners JM (2001) Promoters for pregenomic RNA of banana streak badnavirus are active for transgene expression in monocot and dicot plants. Plant Mol Biol. 47: 399-412

Scherf M, Klingenhoff A, Werner T (2000) Highly specific localization of promoter regions in large genomic sequences by PromoterInspector: a novel context analysis approach. J Mol Biol 297: 599-606

Schwechheimer C, Zourelidou M, Bevan MW (1998) Plant transcription factor studies. Annu.Rev.Plant Physiol Plant Mol.Biol. 49, 127-150

Seki M, Narusaka M, Kamiya A, Ishida J, Satou M, Sakurai T, Nakajima M, Enju A, Akiyama K, Oono Y, Muramatsu M, Hayashizaki Y, Kawai J, Carninci P , Itoh M, Ishii Y, Arakawa T, Shibata K, Shinagawa A, Shinozaki K (2002) Functional annotation of a full-length Arabidopsis cDNA collection. Science 296: 141-145

Sessions A, Yanofsky MF, Weigel D (2000) Cell-cell signaling and movement by the floral transcription factors LEAFY and APETALA1. Science 289: 779-782

Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet 31: 64-68

Simon I, Barnett J, Hannett N, Harbison CT, Rinaldi NJ, Volkert TL, Wyrick JJ, Zeitlinger J, Gifford DK, Jaakkola TS, Young RA (2001) Serial regulation of transcriptional regulators in the yeast cell cycle. Cell 106: 697-708

Singh KB (1998) Transcriptional regulation in plants: the importance of combinatorial control. Plant Physiol 118: 1111-1120

The *Arabidopsis* Genome Initiative, AGI (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796–815

Thijs G, Lescot M, Marchal K, Rombauts S, De Moor B, Rouze P, Moreau Y (2001) A higher-order background model improves the detection of promoter regulatory elements by Gibbs sampling. Bioinformatics 17: 1113-1122

Trindade LM, Horvath B, Bachem C, Jacobsen E, Visser RGF (2003) Isolation and functional characterization of a stolon specific promoter from potato (*Solanum tuberosum* L.). Gene 303: 77-87

Wang H, Tang W, Zhu C, Perry SE (2002) A chromatin immunoprecipitation (ChIP) approach to isolate genes regulated by AGL15, a MADS domain protein that preferentially accumulates in embryos. Plant J 32: 831-843

Weinmann AS, Yan PS, Oberley MJ, Huang TH-M, Farnham PJ (2002) Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. Gen Dev 16: 235-244

Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Pruss M, Reuter I, Schacherer F (2000) TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res 28: 316-319

Wyrick JJ, Young RA (2002) Deciphering gene expression regulatory networks. Curr. Opin. Genet Dev 12: 130-136

Yanagisawa S (2002) The Dof family of plant transcription factors. Trends Plant Sci 7: 555-560

Yoshida K, Shinmyo A (2000) Transgene expression systems in plant, a natural bioreactor. J Bioscience and Bioengineering 90: 353-362

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296: 79-92

Zhou D-X (1999) Regulatory mechanisms of plant gene transcription by GT-elements and GT-factors. Trends Plant Sci 4: 210-214

Zhu Z, Pilpel Y, Church GM (2002) Computational identification of transcription factor binding sites via a transcription-factor-centric clustering (TFCC) algorithm. J Mol Biol 318: 71-81

Chapter 3

Molecular analysis of fruit ripening: The identification of differentially expressed sequences in *Vitis vinifera* using cDNA-AFLP technology

---

Mauritz Venter, Anita L. Burger and Frederik C. Botha (2001) *Vitis* 40 (4): 191-196.

3.1 Abstract

Differential gene expression patterns were studied during the ripening process of grape (*Vitis vinifera* cv. Chardonnay) berries. Thirty *Pst* I + *Mse* I primer combinations were used to generate 213 fragments that appeared to be differentially expressed of which 94% were successfully re-amplified. Reverse northern dot-blot analysis indicated that 35% of the fragments had similar gene expression profiles to cDNA-AFLPs regarding developmental-stage specificity. Northern blot analyses confirmed the tissue and/or developmental stage specific expression of three of these cDNA fragments. This work illustrates that developmentally regulated sequences can be identified from grape berry tissue using cDNA-AFLP technology.

Key words: Differential expression, cDNA-AFLP, berry ripening, Vitis vinifera

3.2 Introduction

The understanding of the regulation of gene expression during fruit development has important agricultural implications. Fruit-specific genes can be used as molecular tools to modify the fruit ripening process (Edwards and Coruzzi 1990). Methods such as differential screening have been used to identify and isolate differentially expressed sequences/genes in grape berries during ripening (Davies and Robinson 2000). Other methods used for the identification of differentially expressed sequences include random sampling, subtraction cloning and differential display. All these methods are invaluable tools to select differentially expressed sequences but some of them suffer several drawbacks including the fact that these methods are labor intensive and time consuming (Sagerström *et al*. 1997). cDNA-AFLP technology largely overcomes these limitations, produces more reliable results than differential display (Habu *et al*. 1997) and is a broadly applicable technique for the identification of developmentally regulated genes (Bachem *et al*. 1996). However, the suitability of this methodology has been evaluated in a very limited number of plant species.

With respect to grape *Vitis vinifera* L., which is a non-climacteric fruit, extensive research on fruit ripening has been conducted to identify certain biochemical and physiological changes during the developing process (Coombe 1992). Major changes in several characteristics (shape, size, colour and metabolic changes) occur during fruit development and ripening which eventually have an effect on taste and quality (Archer 1981). However, success in biotechnological applications will only be possible if a better understanding is gained in the biochemical control and gene expression patterns in the grape berries. The implementation of an effective gene manipulation strategy is dependent on the isolation and characterizing of genes that are specifically expressed in grape berry tissue. In this article we report on the isolation of differentially expressed fragments. Collecting the data indicates that the cDNA-AFLP technique allows for the rapid identification of differentially expressed genes during grape berry ripening.

3.3 Materials and methods

Plant Material

Grape berries (*Vitis vinifera* L. cv. Chardonnay) were collected and classified into six stages of development. The first stage is 26 days after anthesis, there after 5 stages until the berries reached maturity (105 days after anthesis). Berries were deseeded, crushed in liquid nitrogen and stored at –80 °C until use.

Sugar and organic acid extraction

Fifty mg of frozen material was transferred to 2 ml eppendorf tubes and suspended in 1.5 ml 80% (v/v) EtOH containing 100 mM Hepes (pH 7.5) and 20 mM $MgCl_2$. The suspension was incubated at 70 ^{o}C for 14 h and the insoluble material removed through centrifugation.

HPLC analysis

Sugars and organic acids were prepared for HPLC as previously described (Whittaker and Botha 1997). All analyses were conducted on a Shimadzu SCL-10AVP HPLC system. Sugars were separated over a 20 min period on a Supelco™ LC-NH_2 column using with 80 % (v/v) acetonitryl as the mobile phase and a flow rate of 1.2 ml.min^{-1}. Sugars were quantified with differential refractometry (Shimadzo RID-10A). Organic acids were separated over a 15 min period on an Aminex ion exclusion HPX-87H column with 0.02 M H_2SO_4 as the mobile phase and at a flow rate of 0.6 ml.min^{-1}. Organic acids were quantified by UV spectrometry at 210 nm (Shimadzo SPD-10AVP UV/Vis).

RNA isolation and cDNA synthesis

Total RNA was extracted from 4 g of ground, frozen berry material with a modified Na-perchlorate method (Rezaian and Krake 1987). The extraction buffer contained 5 M sodium-perchlorate; 1 M Tris-HCl (pH 8.3); 10 % (m/v) SDS; 20 % (m/v) PEG 6000; 10 % (m/v) PVPP

and 1 % (v/v) β-mercaptoethanol. RNA was quantified fluorometrically (BIO-TEK® Instruments Inc., Winooski, Vermont, USA) and quality was visualised in ethidium bromide-stained 2 % (m/v) agarose gels.

Five μg total RNA, from each tissue sample, was used for first strand cDNA synthesis (Superscript™II, GibcoBRL Life Technologies Inc., Gaithersberg, MD, USA) followed by second strand cDNA synthesis (Universal Riboclone® cDNA synthesis system, Promega Corporation, Madison, USA).

cDNA-AFLP analysis

All AFLP-associated procedures were carried out according to a modified method (Vos *et al*. 1995).

Double strand cDNA templates were digested with 2.5 U of both *Mse* I and *Pst* I restriction enzymes at 37 °C overnight. Non-phosphorylated adaptor sequences were ligated to the restriction fragments at 20 °C overnight. The restriction-ligation products were subjected to 30 cycles of pre-amplification (94 °C denaturation, 30 sec; 56 °C annealing, 1 min; 72 °C polymerization, 1 min) using primers with no selective nucleotides to obtain a sufficient amount of template. The pre-amplification products were diluted 1:10 with 1xTE (10 mM Tris pH 8.0; 0.1 mM EDTA) and visualized in ethidium bromide-stained 1.5 % (m/v) agarose gels with expected sizes ranging from 100 bp to 1000 bp. The *Pst* I forward primer was radioactively labelled using 0.5 μCi γ^{33}P-ATP. Selective amplification was performed with 30 combinations of *Pst* I primer (5'-GACTGCGTACATGCAG+N-3') and *Mse* I primer (5'-GATGAGTCCTGAGTAA+N-3') extensions where 'N' represents two or three selective nucleotides (Table 3.1). Thirty-five cycles of amplification (12 cycles: 94 °C denaturation, 30 sec; 65 °C annealing, 30 sec; 72 °C polymerization; 1 min then 23 cycles: 94 °C denaturation, 30 sec; 56 °C annealing, 30 sec; 72 °C polymerization, 1 min) were carried out where the

annealing temperature was lowered gradually from 65 °C to 56 °C at which efficient primer binding occurs. Thermocycling was started at 65 °C annealing temperature for optimal primer selectivity.

Table 3.1 Total *Pst* I and *Mse* I primers used in combinations with two or three selective nucleotides. ***** *Pst* I-primer extensions +GT, +CT and +GTA in combinations with *Mse* I-primer extensions +CAA, +CAC and +CAG which generated the highest amount of polymorphic fragments.

***Pst* I primer extension/s**	***Mse* I primer extension/s**
1) GT *	**CAA, CAC, CAG**
2) CT *	**CAA, CAC, CAG**
3) GTA	**TG, CAT, CTG**
4) TTT	**TG, CAT, CTG**
5) TTT, GTA, GT	**CA**
6) GTA *	**CAA, CAC, CAG**
7) GA	**CAA, CAC, CAG**
8) GC	**CAA, CAC, CAG**
9) GT	**TG, CAT, CTG**
10) TTT	**CAA, CAC, CAG**

Amplified products were heated at 95 °C for 5 min after addition of an equal amount of formamide dye (98 % (v/v) formamide, 10 mM EDTA pH 8.0 and 1 mg/ml each of bromophenol blue and xylene cyanol) and immediately chilled on ice. Fragments were

separated in 5 % (m/v) denaturing polyacrylamide gels and all gels were developed at 80 Watts for about 1 h 40 min. Gels were dried on to Whatman 3M paper on a slab gel dryer (Biorad Laboratories Inc., Hercules, CA, USA).

Radioactively labelled cDNA fragments were visualized on BioMAX MR film (Eastman Kodak Company, Rochester, New York) after exposure times ranging between 18 h and 72 h. Fragments that appeared to be selectively expressed were excised from the dried gels. cDNA was recovered from each band after heat treatment of 95 °C in 30 µl distilled water for 10 min. Fragments were re-amplified using the same selective primers and PCR conditions as used in the initial pre-amplification procedures and all re-amplified cDNA fragments were visualized in ethidium bromide-stained 2 % (m/v) agarose gels.

Reverse northern dot-blot analysis

A total of 192 fragments were selected and 2 µl of each re-amplified cDNA product were dot blotted on a nylon membrane (Boehringer Mannheim, GmbH Mannheim, Germany) in a 2x96 well PCR-plate configuration. Seven identical membranes were prepared. cDNA were denatured (1.5 M NaCl; 0.5 M NaOH), neutralized (1 M Tris pH 7.4; 1.5 M NaCl), rinsed in 2xSSC (0.15 M NaCl; 0.015 Tri-sodium citrate pH 6.8; citric acid) and UV cross-linked before hybridization began.

Single strand cDNA probes were prepared from 5 µg total RNA of each berry ripening stage and leaf sample using reverse transcriptase and a equimolar mix of primer 5'-AGTCTGCAGT$_{12}$-N-3', with 'N' representing A, C or G respectively (Superscript™II, Gibco BRL Life Technologies Inc., Gaithersberg, MD, USA). Modifications regarding 10 µCi ^{32}P-dCTP incorporation were made in our laboratory. Equal counts (1.5x10^{7} cpm/ml) of cDNA probes were used to probe the membranes. Hybridization was visualized by autoradiography. Dot intensities of sequences were analyzed using the AlphaImager™2000 documentation and

analysis system (Alpha Innotech Corporation, San Leandro, USA). Fragments were selected according to fruit-specificity and/or abundant expression. Promising fragments based on differential expression were cloned using pGEM®-T Easy Vectors (Promega Corporation, Madison, USA) to be used for Northern analysis.

Sequence analysis

Selected cDNA clones were sequenced (ABI PRISM™ dye terminator cycle sequencing) using the ready reaction kit with AmpliTaq® DNA polymerase (The Perkin Elmer Corporation, Norwalk, USA). The cDNA sequences were edited to discard the vector/linker and primer sequences.

Northern blot analysis

Northern blot membranes were prepared using total RNA (visualised in ethidium bromide-stained 1 % (m/v) agarose gels) from grape ripening stages 1 to 6, young grapevine leaf and root (10 μg/track). RNA was transferred to a positively charged nylon membrane (Boehringer Mannheim, GmbH Mannheim, Germany) by upward capillary blotting (Sambrook *et al*. 1989) using 10xSSC (standard saline citrate). The RNA was UV cross-linked and all hybridization (using ULTRAhyb™ ultrasensitive hybridization buffer) and washing procedures were carried out as described by the manufacturer (Ambion, Austin, USA). For preparation of probes, the re-amplified fragment of interest was radioactively labelled using 25 μCi [α-^{32}P] dCTP by four cycles re-amplification PCR using the same conditions as used in the initial pre-amplification procedures of this study. Hybridization was visualized using the Cyclone™ Storage Phosphor System (Packard Instrument Co., Inc., Meriden, USA).

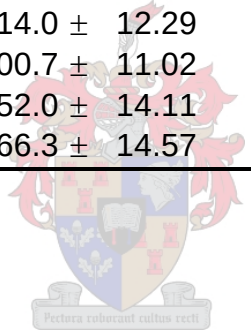
3.4 Results and discussion

Stages of fruit development

It is evident that the berries collected at stages 1 and 2 are typical of grape berries prior to *véraison*, i.e. they contain similar levels of reducing sugars and sucrose and high acid levels (table 3.2). Stage 3 is characterised by a very rapid increase in both glucose and fructose and a decrease in malate levels. Stage 6 represents fully ripened fruit where total sugars are now in excess of 15% of the total fresh mass and the acid levels are low.

Table 3.2 Sucrose, glucose, fructose and malate levels in the grape berries isolated at different stages during development and ripening. Each value is the average ± SD of three extractions.

Stage	Sucrose	Glucose	Fructose	Malate
	mol g-1 fresh mass			
1	21.0 ± 2.65	21.3 ± 2.08	23.3 ± 2.08	286.7 ± 30.55
2	19.3 ± 1.53	25.7 ± 1.53	24.0 ± 2.65	328.3 ± 16.07
3	17.3 ± 1.53	214.0 ± 12.29	211.3 ± 8.74	195.0 ± 13.23
4	19.0 ± 3.61	300.7 ± 11.02	330.0 ± 26.46	158.3 ± 18.93
5	18.7 ± 3.51	352.0 ± 14.11	355.0 ± 21.79	108.3 ± 10.41
6	25.3 ± 1.53	366.3 ± 14.57	392.7 ± 22.48	63.3 ± 18.93



cDNA-AFLP analysis

Total RNA was extracted from grape berry tissue and a decline of RNA yield (24 μg/g to 2.75 μg/g fresh weight) could be observed over the period from early to late berry ripening stages. cDNA pre-amplification products ranged in size from 200 bp to 1 kb. For cDNA-AFLP *Pst* I and *Mse* I in combinations according to different selective nucleotide extensions on the primers were used. Although 6-bp restriction enzyme recognition sites would be present in only a minimal fraction of cDNA species (Habu *et al*. 1997), we retrieved 213 putative polymorphic bands with the primer combinations used for this study.

cDNA-AFLP reproducibility was examined by comparing reaction products that were derived from two sets of independent samples of total RNA, prepared from early and late developmental stages. Two different primer combinations, *Pst* I +CT with *Mse* I +CAT and +CTG gave

identical band-patterns (Figure 3.1). A total of 213 polymorphic fragments were isolated after visual analysis of cDNA-AFLP profiles using 30 *Pst* I + *Mse* I primer combinations. *Pst* I-primer extensions +GT, +CT and +GTA in combinations with *Mse* I-primer extensions +CAA, +CAC and +CAG generated the highest amount of polymorphic fragments. Stage-dependant expression as well as the gradual increase or decrease of gene expression intensities were observed (Figure 3.2). cDNA-AFLP analysis conducted over a period across six stages of berry development, verified the presence or absence of bands at different ripening stages. This analysis was repeated in leaf material from the same cultivar. Ninety-four percent of all the fragments excised, could successfully be re-amplified (Figure 3.3).

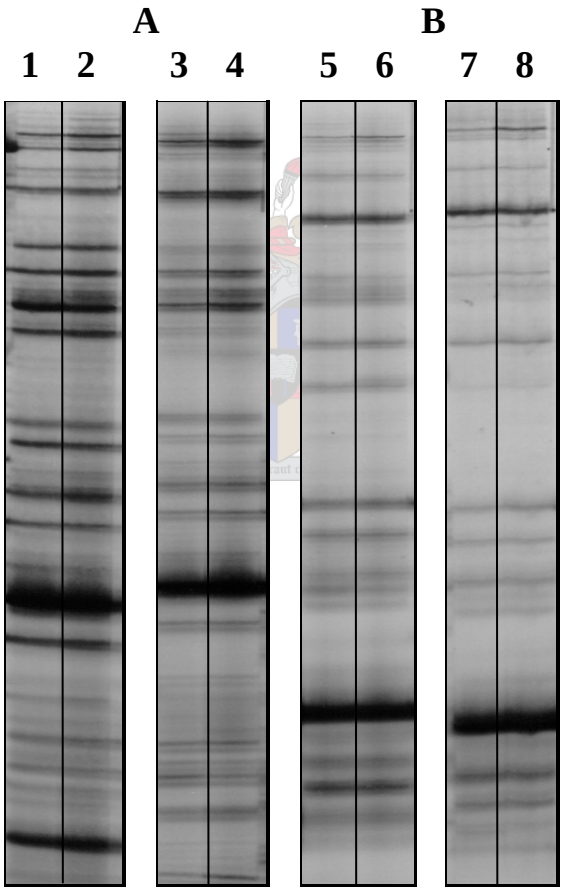


Figure 3.1 Four sections of an autoradiograph with primer combination *Pst* I +CT with *Mse* I +CAT (**A)** and *Pst* I +CT with *Mse* I +CTG (**B)**. Reproducibility examined with two independent total RNA samples of stage 1 (lanes 1 & 2 and lanes 5 & 6) and two independent total RNA samples of stage 5 (lanes 3 & 4 and lanes 7 & 8).

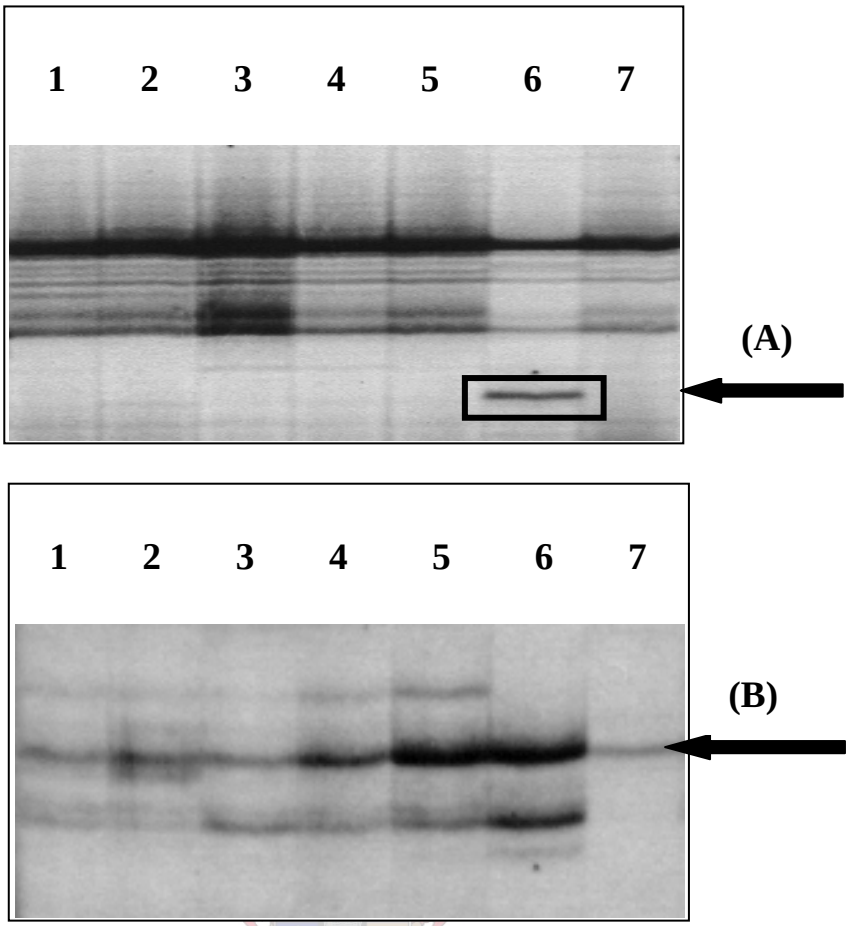


Figure 3.2 Autoradiograph sections with primer combination *Pst* I+*G*T and *Mse* I+CAC showing (**A)** stage-specific expression and primer combination *Pst* I+CT and *Mse* I+CAA showing **(B)** gradual change of expression levels from early to late stages of grape development. Lanes 1 to 6 represent cDNA from ripening stages 1 to 6 and lane 7 is cDNA from young, field grown leaf material.

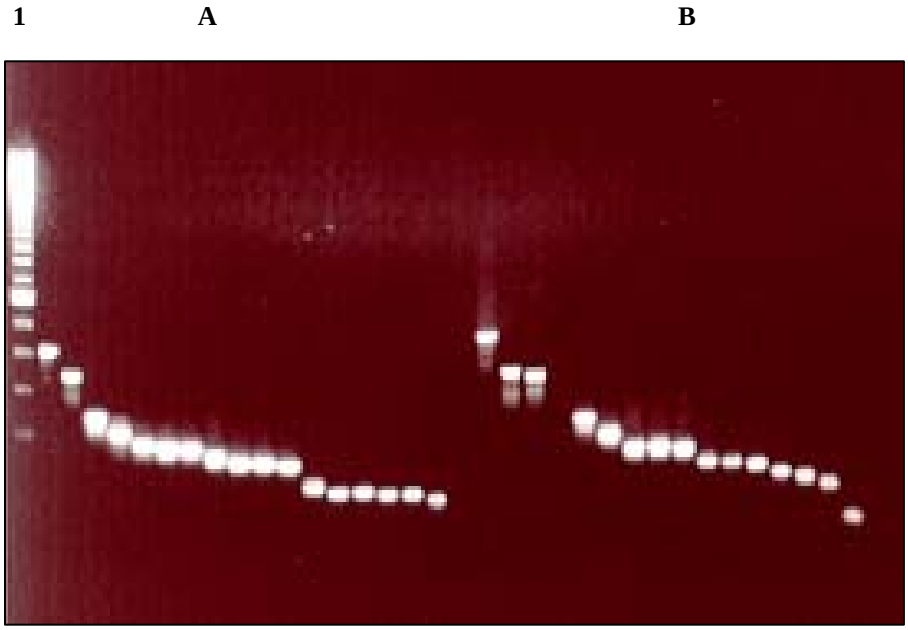


Figure 3.3 Fragments excised and re-amplified with primer combination **(A)** *Pst* I+GTA and *Mse* I+CAC and **(B)** *Pst* I+GTA and *Mse* I+CAA. CDNA fragments range in sizes from 100 bp to 500 bp. Lane 1 is the 100 bp DNA ladder (Promega Corporation, Madison, USA).

Reverse northern and differential expression analysis

The reverse northern dot-blot technique is an effective method to test the feasibility of differential screening (Zhang *et al.* 1996). Dot-blot results confirmed changes in gene expression patterns over a period from early to late berry development (Figure 3.4).

Reverse northern dot-blot analysis revealed the presence of the 98 most abundantly expressed fragments of which 58 were only expressed in the berry and not leaf, and designated as fruit-specific during this study. Expression levels of 10 randomly chosen fragments were examined (Figure 3.5). Analysis over a period of early to late berry development (stage 1 to stage 6) revealed that 60 % (Figure 3.5 fragments A, B, C, E, H, and I) of these fragments showed differential gene expression patterns similar to cDNA-AFLP profiles. Collectively this data

indicate that approximately 35 % of the fragments identified during the initial cDNA-AFLP analysis are truly differentially expressed. Most of the fragments analyzed (Figure 3.5 fragments A, B, C, D, E, F, H and I) showed an increase of expression levels from early to late development with the highest expression occurring in the late stages of ripening. Some fragments (Figure 3.5 fragments G and J) revealed high expression levels in both early and late berry development with lower expression between stages 1 and 6.

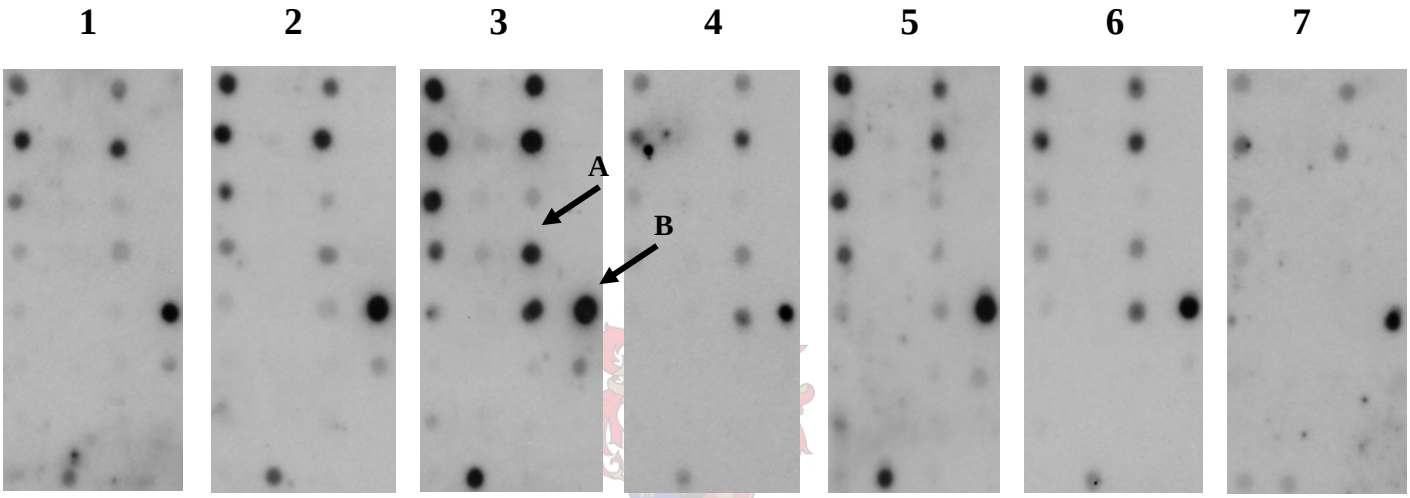


Figure 3.4 Identical sections of original reverse northern blot membranes with examples of **(A)** stage-specific and **(B)** constitutive gene expression during berry ripening. Sections 1 to 6 represent membranes probed with cDNA from ripening stages 1 to 6 and section 7 probed with cDNA from young leaf material.

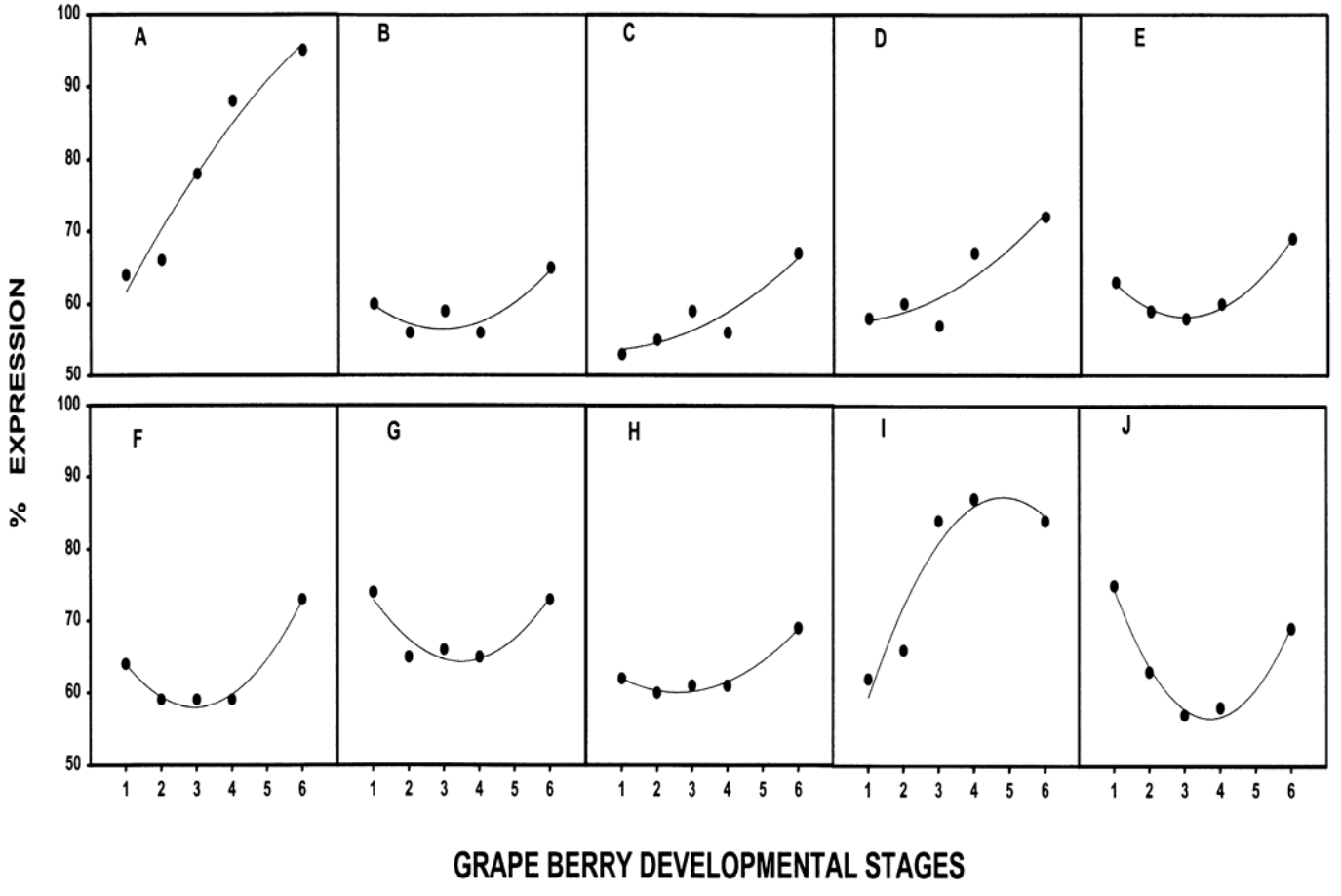


Figure 3.5 Gene expression patterns, only visible in the berry, of 10 randomly selected sequences (A to J) from early to late berry development (stage 1 to stage 6).

Northern blot and sequence analysis

Ten PCR fragments, abundantly differentially expressed as visualized with reverse northern analysis, and irrespective of their tissue and/or stage specificity, were selected and successfully cloned (Figure 3.6). Fragments excised from dried polyacrylamide gels usually contain more templates than the desired one, therefore two clones of each PCR product were isolated and successfully re-amplified. A total of twenty cloned fragments, designated as M1.A/M1.B to M10.A/M10.B (Figure 3.6), were sequenced and analyzed (Table 3.3). Sequence-search (Altschul *et al.* 1990) results revealed two of the clones, M2.B and M8.A, to be homologues to

known grape ripening related “Grip” genes (Davies and Robinson 2000) and fragment M10.A had homology to a H+-pyrophosphatase gene from *Vitis vinifera* (AF192308.1). These three cDNA clones, M2.B, M8.A and M10.A, which had sequence similarity to grape-genes were used as probes for northern blot analysis (Figure 3.7). Fragment M10.A showed constitutive expression during berry ripening but not present in leaf RNA. Fragments M8.A and M2.B were fruit-specific and showed gradual increase and stage specific expression patterns during ripening.

It has been shown in another study (Singh and Cheah 2000), using the differential display technique (Liang and Pardee 1992), that the lack of homology to known sequences in *Genbank* could be due to the fact that the cloned cDNAs were only partial length (200-600 bp). However, in this study we obtained partial cDNAs (62-315 bp) of which only 4 sequences showed no homology to known plant sequences. Other cDNAs had homology to genes from *Vitis vinifera* and to known DNA sequences, mostly from *Arabidopsis thaliana.* Sequences M1 and M7.B (Table 3.3) had similarity to expressed but as yet unidentified proteins. Analyses have also revealed sequence similarity to a putative epoxide hydrolase EphB gene from *Bradyrhizobium japonicum* which could be stress-induced and to a 19S ribosomal RNA gene from *Rafflesa pricei* (Table 3.3).

Table 3.3 Putative sequence identities of clones M1A/B to M10.A/B. The clones of five selected PCR products M1, M4, M5, M6 and M9 were identical. *** Fragments of which gene expression profiles were evaluated during berry maturation (Figure 3.7)**

Clone nr.	Length (bp)	Sequence similarity (accession no.)
M1.A=M1.B	296	*Arabidopsis thaliana* unknown protein (MBK23.13/AT5g41600) mRNA (AY035169.1)
M2.A	62	*Arabidopsis thaliana* chromosomeII section 208 of 255 of the complete sequence (AC005499.2)
M2.B*	237	mRNA for putative proline-rich cell wall from *Vitis vinifera* (AJ237982.1) Similarity to genes 'Grip' 3, 4, 13 and 15
M3.A	291	No significant similarity to plant sequences
M3.B	308	No significant similarity to plant sequences
M4.A=M4.B	179	*Arabidopsis thaliana* DNA, chromosome 5, BAC clone F21E1 (AL391716.1)
M5.A=M5.B	315	*Bradyrhizobium japonicum* putative epoxide hydrolase EphB (ephB), putative stress-induced protein Ohr (U33833.2)
M6.A=M6.B	193	*Arabidopsis thaliana* genomic DNA, chromosome 5, P1 clone:MNB8 (AB018116.1)
M7.A	173	*Arabidopsis thaliana* chromosome 1 BAC T22A15 genomic sequence (AC021666.5)
M7.B	170	*Arabidopsis thaliana* unknown protein (F12M12_190/AT3g46220) mRNA (AY034937.1)
M8.A*	292	mRNA for putative ripening-related protein from *Vitis vinifera* (AJ237986.1) Similarity to gene 'Grip' 31
M8.B	272	No significant similarity to plant sequences
M9.A=M9.B	128	No significant similarity to plant sequences
M10.A*	289	mRNA for H+-pyrophosphatase from *Vitis vinifera* (AF192308.1)
M10.B	289	*Rafflesia pricei* 19S ribosomal RNA gene, mitochondrial gene for mitochondrial rRNA (U96694.1)

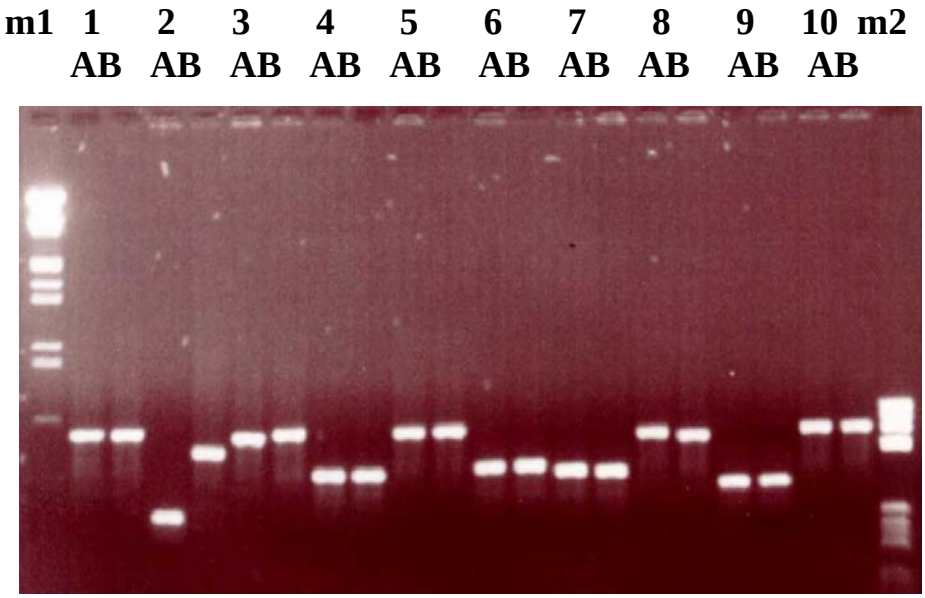


Figure 3.6 Numbers 1 to 10 represent the 10 most abundantly expressed PCR fragments isolated after dot-blot analysis. Two amplified clones (A and B) of each PCR product can be visualized. Lanes m1 and m2 are the molecular weight markers III and V respectively (Boehringer Mannheim, GmbH Mannheim, Germany).

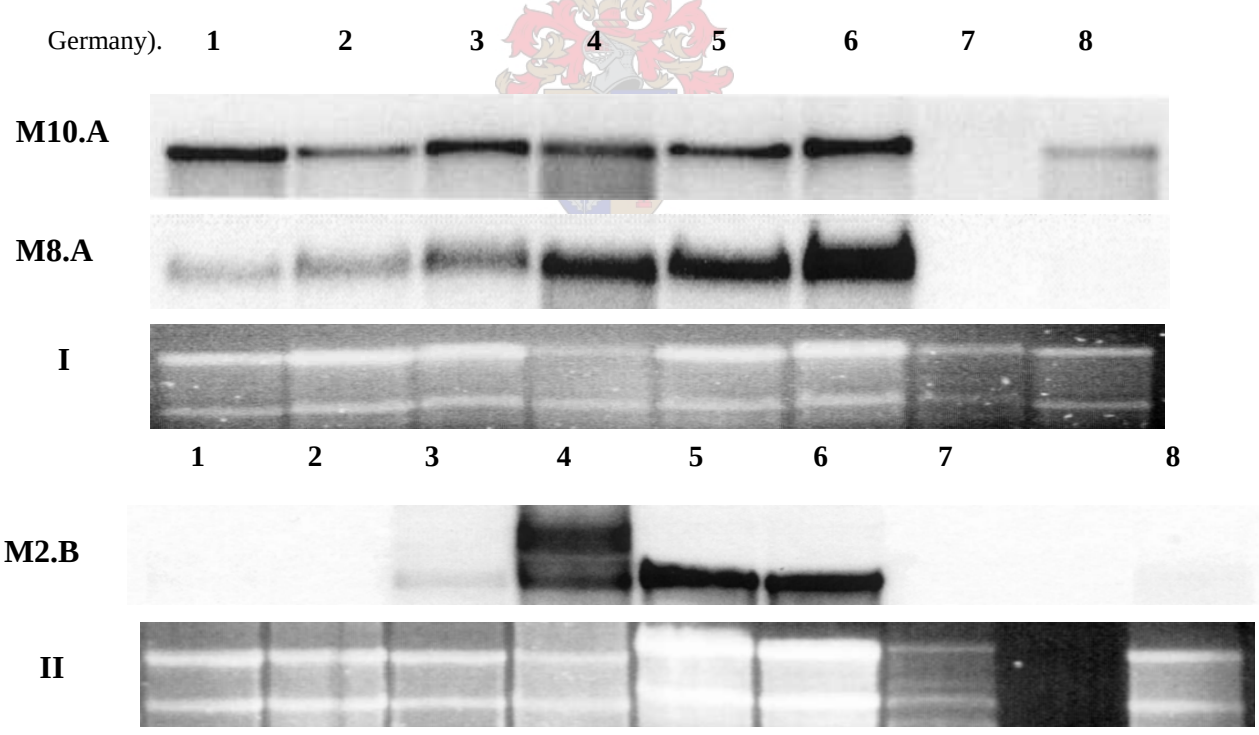


Figure 3.7 Northern blot results of selected, tissue-specific fragments (marked by asterisks in table 3.3) expressed constitutively **M10.A** and differentially **M2.B** & **M8.A**. Panels **I** and **II** shows approximately equal amounts of intact total RNA with lanes 1 to 6 representing RNA from ripening stages 1 to 6 and lanes 7 to 8 represent RNA from young leaf and root material respectively.

3.5 Conclusion

The reverse northern dot-blot technique confirmed differential expression similarity of isolated fragments to cDNA-AFLPs. Both methods used in conjunction, proved to be powerful and effective tools to identify and screen large quantities of polymorphic bands in grapevine. Northern blot results confirmed tissue and/or stage specific expression verifying the authenticity of the selected differentially expressed sequences. As mentioned before, a putative, polymorphic fragment initially excised from the dried gel, could be a mixed template and this can be considered as a major drawback. Therefore, cloning of excised fragments is a prerequisite for final evaluation and analysis. Nevertheless, results obtained during northern blot and sequence analysis suggests that the cDNA-AFLP method is a fast and reliable technique for identifying differentially expressed genes in grapevine.

Acknowledgements

The authors would like to thank Winetech, the National Research Foundation and the department of Trade and Industry for financial support. We also acknowledge the contributions of Dr. Bernard Potier and Cobus Olivier.

References

Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J.; 1990: Basic local alignment search tool. J. Mol. Biol. **215**, 403-410.

Archer, E.; 1981: Rypwording en Oesmetodes. Wingerdbou in Suid-Afrika. Eds: Burger, J.; Deist, J.; Maskew Miller. Cape Town. 1st Edition. Chapter 17: 463-476. (Translation: Archer, E.; 1981: Ripening and Harvesting methods. Viticulture in South Africa.)

Bachem, C.W.B.; Van der Hoeven, R.S.; De Bruijn, S.M.; Vreugdenhil, D.; Zabeau, M.; Visser, R.G.F.; 1996: Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. The Plant J. ***9*(5)**, 745-753.

Coombe, B.G.; 1992: Research on development and ripening of the grape berry. Am. J. Enol. Vitic. **43**, 101-110.

Davies, C.; Robinson, S.P.; 2000: Differential screening indicates a dramatic change in mRNA profiles during grape berry ripening. Cloning and characterization of cDNAs encoding putative cell wall and stress response proteins. Plant. Phys. **122**, 803-812.

Edwards, J.W.; Coruzzi, G.M.; 1990: Cell specific gene expression in plants. Ann. Rev. Gen. **24**, 275-303.

Habu, Y.; Fukada-Tanaka, S.; Hisatomi, Y.; Iida, S.; 1997: Amplified restriction

fragment length polymorphism-based mRNA fingerprinting using a single restriction enzyme that recognizes a 4-bp sequence. Biochem. Biophys. Res. Commun. **234(2)**, 516-521.

Liang, P.; Pardee, A.B.; 1992: Differential display of eukaryotic messenger RNA by means of polymerase chain reaction. Science **257**, 967-971.

Rezaian, M.A.; Krake, L.R.; 1987: Nucleic acid extraction and virus detection in grapevine. J. Vir. Methods. **17**, 277-285.

Sambrook, J.; Fritsch, E.F.; Maniatis, T.; 1989: Molecular cloning: A laboratory manual, 2nd Edition. Cold Spring Harbor Laboratory Press, New York.

Sagerström, C.G.; Sun, B.I.; Sive, H.L.; 1997: Subtractive cloning: Past, present and future. Ann. Rev. Biochem. **66**, 751-783.

Singh, R.; Cheah, S-C.; 2000: Differential gene expression during flowering in the oil palm (*Elaeis guineensis*). Plant Cell Reports. **19**, 804-809.

Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; Van de Lee, T.; Hornes, M.; Frijters, A.; Pot, J.; Peleman, J.; Kuiper, M.; Zabeau, M.; 1995: AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. **23(21)**, 4407-4414.

Whittaker, A.; Botha, F.C.; 1997: Carbon partitioning during sucrose accumulation in sugercane internodal tissue. Plant Physiol. **115**, 1651-1659.

Zhang, H.; Zhang, R.; Liang, P.; 1996: Differential screening of gene expression difference enriched by differential display. Nucleic Acids Res. **24(12)**: 2454-2455.

Chapter 4

The vacuolar H^+-pyrophosphatase of *Vitis vinifera* L.: Gene expression and promoter analysis

(Manuscript in preparation)

Mauritz Venter, Jacobus P. Zwiegelaar, Bernard Potier and Frederik C. Botha

4.1 Abstract

Different expression patterns for the vacuolar H+-pyrophosphatase (V-PPase EC 3.6.1.1.) of *Vitis vinifera* L. have been reported (Terrier *et al*. 2001; Venter *et al*. 2001). RNA gel blots revealed differential V-PPase gene expression patterns during fruit ripening. Southern blot and sequence analyses as well as RNA *in situ* hybridisation experiments were conducted to gain further insight into the genomic organisation as well as the cellular gene expression patterns of one cloned grapevine V-PPase gene. The expression of this particular V-PPase gene was induced by cold shock and drought stress. The promoter region of this gene was isolated, fused to a GFP reporter gene and promoter activity was confirmed in green pepper (*Capsicum* spp.) exocarp tissue. The V-PPase promoter sequence was analysed using plant *cis*-motif databases and promoter alignment software. *In silico* promoter characterisation was used to assist in the preliminary understanding of V-PPase gene expression patterns during berry ripening and under specific environmental stress conditions.

Key words

vacuolar pyrophosphatase, fruit ripening, promoter, Vitis vinifera, gene expression

4.2 Introduction

The two major electrogenic proton pumps in the plant vacuolar membrane are H^+-transporting adenosine triphosphatase (V-ATPase) and H^+-translocating inorganic pyrophosphatase (V-PPase) (Rea and Poole 1993). The pumps, primarily responsible for vacuolar acidification (Maeshima 2000) and generation of a proton gradient to drive transport (Ratajczak 2000), maintain the critical functions of the vacuole under diverse stress conditions (Davies 1997). The V-PPase proton pump utilizes inorganic pyrophosphate (PP_i) instead of ATP to drive proton movement (Taiz and Zeiger 1998) and could, therefore play an important role especially under energy limiting conditions.

V-PPase is present in plants and bacteria. Although a single isoform of V-PPase was reported from higher plants such as *Vigna radiata* (Nakanishi and Maeshima 1998) and *Hordeum vulgare* (Tanaka *et al*. 1993), multiple copies of the same gene have been identified in other plants (Drozdowicz and Rea 2001, Kim *et al*. 1994, Lerchl *et al*. 1995, Sakakibara *et al*. 1996, Suzuki *et al*. 1999). V-PPase activity is present in the fruit of tomato (Milner *et al*. 1995) grape (Terrier *et al*. 1998 and 2001) and pear (Suzuki *et al*. 1999), but apparently absent in lemon (Müller *et al*. 1996). The V-PPase activity decreases during fruit development in tomato (Milner *et al*. 1995) and increases in the grape berry (Terrier *et al*. 2001). In the pear fruit there is no correlation between the V-PPase transcript abundance and protein level (Suzuki *et al*. 1999). V-PPase activity is induced during chilling and anoxia (Carystinos *et al*. 1995), as well as during salt stress (Ballesteros *et al*. 1996).

The importance of solute transport in and out of the vacuole, specifically affecting sugar-acid metabolism in the grape berry, as well as the influence of different stress conditions have led researchers to characterize the expression patterns and promoter elements of specific target genes in order to study the control of grape berry development. Different V-PPase gene expression profiles during grape berry ripening have been reported (Terrier *et al*. 2001, Venter

et al. 2001). This strongly suggests that different V-PPase transcriptional activity during the same condition (i.e. development) could be indicative of more than one isoform of V-PPase and that these genes are regulated by different promoter elements.

Here we report that the V-PPase described by Venter *et al*. (2001) is different from the gene previously described (Terrier *et al*. 2001) in grape berries. In addition we show that different promoter elements identified *in silico*, could assist to gain further insight on sub-cellular and environmental stress responsive gene expression patterns of the vacuolar pyrophosphatase pump in *Vitis vinifera* L.

4.3 Materials and methods

Plant material and sugar/organic acid determination

Grape berries (*Vitis vinifera* L. cv. Chardonnay) were collected and classified into eight stages of development. The first stage is 14 days after anthesis, there after 7 stages (with 2 week intervals) until the berries reached maturity (112 days after anthesis). Berries were deseeded, crushed in liquid nitrogen and stored at –80 °C until use. To determine the exact developmental stage and the time of onset of *véraison*, sugars were extracted from the grape berries, and analysed according to the protocol described in Famiani *et al*. (2000). The concentrations of the soluble sugars and malate in the berries were determined and plotted against the ripening stage to enable comparison of the data to that of previous studies and other growing seasons. For stress induction treatments grape berries (*Vitis vinifera* L. cv. Chardonnay) were collected 4 weeks after anthesis which coincides with the early developmental stage prior to *véraison*. To confirm herbaceous stage, a representative sample of collected berries was crushed in liquid nitrogen and sugar and organic acid analyses were conducted as described by Famiani *et al*. (2000).

DNA extraction and Southern blot analysis

Genomic DNA was isolated from grapevine leaves based on the method described by Steenkamp *et al*. (1994). Grapevine genomic DNA (8 µg) was digested using 15 U of restriction enzymes at 37 ^{o}C overnight. Digested DNA was precipitated with 1/20 volume of 3 M NaAc and 2 volumes of 95 % (v/v) EtOH overnight. DNA was collected, resuspended in distilled water, electrophorised and visualised in a 2 % (m/v) ethidium bromide stained agarose gel. DNA was transferred to a positively charged nylon membrane (Boehringer Mannheim, GmbH Mannheim, Germany) by upward capillary blotting (Sambrook *et al*. 1989) using 10xSSC (standard saline citrate). The DNA was UV cross-linked and all hybridisation (using Rapid-hyb buffer) and washing procedures were carried out as described by the manufacturer (Amersham Pharmacia Biotech Inc, USA). The V-PPase cDNA fragment M10.A of 289 bp (AJ430532, Venter *et al*.2001), cloned in a pGEM®-T Easy Vector (Promega Corporation, Madison, USA) was used as a probe. The cDNA was radioactively labelled using 25 µCi [α-^{32}P] dCTP by eight cycles of amplification (94 °C denaturation, 30 sec; 56 °C annealing, 1 min; 72 °C polymerization, 1 min) using primers T7 (5'-AATACGACTCACTATAGG-3') and SP6 (ATTTAGGTGACACTATAG-3'). Hybridisation was visualized using the Cyclone™ Storage Phosphor System (Packard Instrument Co., Inc., Meriden, USA).

In situ hybridisation

Berries from two, 8 and 14 weeks post flowering (stages 1, 4 and 7) were used. The roots and leaves were collected from grape plants cultured *in vitro* on medium containing Murashige and Skoog (1962) (MS) basal medium with 3 % (m/v) sucrose and 0.2 % (m/v) Gelrite under a 16 h photoperiod at 22 °C. Tissue was fixed in paraformaldehyde, dehydrated in a graded ethanol series, cleared in xylene and infiltrated with Paraplast X-tra (Sigma-Aldrich Chemie, GmbH Steinheim, Germany). *In situ* hybridisation experiments were carried out on 10 µm dewaxed sections pretreated with 0.125 mg.ml^{-1} pronase (Sigma-Aldrich Chemie, GmbH Steinheim, Germany) in 0.05 M Tris-HCl (pH 7.5), 0.005 M EDTA for 10 min at room temperature, 0.2 %

(m/v) glycine in PBS for 2 min and 1 % (v/v) acetic anhydride in 0.1 M triethanolamine (pH 8) for 10 min at room temperature and dehydrated with a graded ethanol series. Hybridisation was carried out using DIG-labeled RNA probes. M10.A (AJ430532) sense and antisense probes were generated by *in vitro* transcription of linearized template DNAs in the presence of DIG-labeled dUTP as described by the manufacturer (Boehringer Mannheim, GmbH Mannheim, Germany). The probe was diluted in hybridisation buffer (Sigma-Aldrich Chemie, GmbH Steinheim, Germany) to a final concentration of 200 $ng.ml^{-1}$. After overnight hybridisation and washing at 42 °C in 2x SSC 50 % formamide, the slides were treated with 1 % blocking reagent (Roche Molecular Biochemicals, GmbH Mannheim, Germany) in 100 mM Tris-HCl (pH 7.5) 150 mM NaCl for 1h at room temperature and incubated for 1h at room temperature with the antibody (Anti-DIG Fab fragments, Roche Molecular Biochemicals, GmbH Mannheim, Germany) diluted 1:3000 in the blocking solution. After washing the antibody was detected by incubating the slides in the dark with 150 $\mu g.ml^{-1}$ Nitroblue tetrazolium, 75 $\mu g.ml^{-1}$ 5-Bromo-4-chloro-3-indolylphosphate, 10 % polyvinyl alcohol (MW=70 000-100 000), 100 mM Tris-HCl (pH 9.5) 100 mM NaCl, 50 mM $MgCl_2$ for up to 3 days. The slides were mounted and visualized with a Nikon Eclipse E400 and photographed with a Nikon Coolpix 990.

Stress induction experiments

Stress treatments were conducted for 48 h. Berries were placed in petri dishes on a filter-paper and wetted with 15 ml distilled water. Berries were subjected to heat stress at a constant 37 ^{o}C. For cold shock treatment, berries were kept at 25 ^{o}C for 12 h then transferred to 4 ^{o}C for 12 h and the cycle was repeated once again. For hypoxia induction, berries were completely submerged in degassed distilled water, in airtight bottles. The bottles were filled so that there was no airspace. Submerged berries were kept between 25 ^{o}C and 27 ^{o}C. For osmotic stress treatments berries were transversely sliced in half and placed in petri dishes at 37 ^{o}C on filter paper saturated with 10 ml 0.2 M, 0.4 M and 0.8 M Mannitol respectively. Control treatments were berries, placed in petri dishes on a filter-paper, wetted with 15 ml distilled water and kept between 25 ^{o}C and 27 ^{o}C for 48h. All treatments were repeated once.

RNA gel blot analysis

Total RNA was extracted from 4g of ground, frozen berry material (after stress treatments) with a modified Na-perchlorate method (Rezaian and Krake 1987). The extraction buffer contained 5 M sodium-perchlorate; 1 M Tris-HCl (pH 8.3); 10 % (m/v) SDS; 20 % (m/v) PEG 6000; 10 % (m/v) PVPP and 1 % (v/v) β-mercaptoethanol. RNA was quantified spectrophotometrically and quality was visualised in ethidium bromide-stained 1.2 % (m/v) agarose gels. Gel blot membranes were prepared using total RNA from grape berries (5 μg/track). RNA was transferred to a positively charged nylon membrane (Roche Diagnostics, GmbH Mannheim, Germany) by upward capillary blotting (Sambrook *et al*, 1989) using 10xSSC (standard saline citrate). The RNA was UV cross-linked and all hybridisation (using Rapid-hyb buffer) and washing procedures were carried out as described by the manufacturer (Amersham Pharmacia Biotech Inc, USA). For preparation of the probe, the re-amplified cDNA fragment (M10.A; AJ430532) was radioactively labelled using 25 μCi [α-^{32}P] dCTP by four cycles re-amplification PCR using conditions as conducted by Venter *et al*. 2001. Hybridisation was visualized using the Cyclone™ Storage Phosphor System (Packard Instrument Co., Inc., Meriden, USA).

Isolation, cloning and sequencing of the Vitis V-PPase promoter

Long-range inverse PCR amplification (LR-iPCR) was conducted by designing inverse M10.A-forward (5'-CATGGAAGGCACTGCCAAGC-3') and M10.A-reverse (5'-TAGAGCTGCAGACCCAATTGC-3' towards the 5'-UTR) primers from the M10.A cDNA fragment (AJ430532). Three μg of genomic DNA were digested with 6 bp-cutter restriction enzymes in a total volume of 30 μl overnight. Digested DNA (600 ng) was resuspended in a ligation mix containing 9 U T4 DNA ligase, 40 μl ligase buffer (Promega Corporation, Madison, USA) and made up in a final volume of 400 μl. Reaction was initiated at 16 ^{o}C and allowed to proceed overnight. Circularised DNA was phenol/chloroform (1:1) extracted,

precipitated and resuspended in 40 µl distilled water. Digestion and recircularisation of genomic DNA were visualised in ethidium bromide-stained 0.8 % (m/v) agarose gels. LR-iPCR reactions were performed with 150 ng recircularised genomic DNA in a final volume of 50 µl containing 1 µl Elongase®enzyme mix, 10µl total volume of Buffer A and B (Invitrogen Corporation), 0.2 mM of each dNTP and 0.2 mM of each primer (M10.A forward and M10.A reverse). PCR samples were denatured by 94 °C for 30 sec followed by 35 cycles of amplification (94 °C denaturation, 1 min; 50 °C annealing, 1 min; 68 °C polymerization, 10 min). A grapevine genomic library was constructed using the lambda FIX®II/*Xho*I partial fill-in vector kit (Stratagene, LaJolla, CA). The library was used in parallel with LR-iPCR to obtain a grapevine V-PPase genomic clone. Approximately 224000 recombinant plaques were screened with M10.A (AJ430532). The cDNA fragment was [α-^{32}P] dCTP-labelled by eight cycles of amplification (94 °C denaturation, 30 sec; 56 °C annealing, 1 min; 72 °C polymerization, 1 min) using primers (forward: 5'-GACTGCGTACATGCAGGA-3') and (reverse: 5'-GATGAGTCCTGAGTAACAC-3') that were initially used as cDNA-AFLP primers for fragment re-amplification (Venter *et al*. 2001). First and second round screenings were conducted using standard procedures as described by Sambrook *et al*. (1989) and all hybridisation (using Rapid-hyb buffer) and washing procedures were carried out as described by the manufacturer (Amersham Pharmacia Biotech Inc, USA). Hybridisation was visualized using the Cyclone™ Storage Phosphor System (Packard Instrument Co., Inc., Meriden, USA). Putative positive clones were identified, punched out of plaques and stored in SM buffer (100 mM NaCl; 8 mM $MgSO_4.7H_2O$; 1 M Tris-HCl pH 7.5; 2 % (v/v) Gelatin). Chloroform was added to kill the bacteria, phage/SM buffer suspension was mixed and stored at 4 °C to be used as PCR template (10 µl) for positive clone amplification. PCR amplification of promoter regions included a predenaturation at 94 °C for 30 sec followed by 35 cycles of amplification (94 °C denaturation, 30 sec; 56 °C annealing, 1 min; 72 °C polymerization, 1 min) using primers M10.A-reverse and vector primer T7. All PCR reactions were carried out in a Perkin-Elmer GeneAmp® Thermocycler 9700. The 5'-flanking region of the V-PPase gene was cloned

in a pGEM®-T Easy Vector (Promega Corporation, Madison, USA) and sequenced (ABI PRISM® 3100 genetic analyser) using the ready reaction kit with AmpliTaq® DNA polymerase (The Perkin Elmer Corporation, Norwalk, USA). The genomic DNA sequence was edited to discard the vector/linker and primer sequences. To identify preliminary proximal promoter regions (TATA-box and putative transcriptional start site), the genomic clone sequence was analysed by promoter prediction database software, Neural Network Promoter Prediction (NNPP, http://www.fruitfly.org/seq tools/promoter.html) (Reese and Eeckman 1995) as well as a plant transcription factor database software, PlantCARE (Rombauts *et al*. 1999).

Determination of promoter activity

Promoter-reporter gene constructs were generated using the pGEM-T vector (Promega Corporation, Madison, USA) with a modified green fluorescent protein (GFP)-reporter gene (S65T) as described in Elliot *et al*. (1999) and the 3'-non-coding region of the nopaline synthase (*nos*) gene inserted. The following fusions were prepared for microprojectile bombardment: pGFP + promoter region of the grapevine V-PPase gene, pGFP + 35S promoter of the cauliflower mosaic virus (CaMV) as a positive control and pGFP + *Hordeum vulgare* HVP1 gene (AB032839) as a promoterless negative control. Bombardments of tungsten particles alone were used as an additional negative control. Microprojectile bombardments were carried out as described by Bower *et al*. (1996). Young and mature green peppers (*Capsicum* spp.), which are non-climacteric fruits, were used as target tissue.

Sequence analyses

All sequence similarity search and alignment analyses were conducted using BLAST (Altschul *et al*. 1990 and 1997), CLUSTAL W (Higgins *et al*. 1994) and DNASIS®MAX software (Hitachi Software Engineering Co., Ltd., Japan). Two major plant transcription factor databases, PLACE (Higo *et al*. 1998) and PlantCARE (Rombauts *et al*. 1999) were used to identify putative transcription factor binding sites in all promoter sequences analysed. Conserved noncoding sequences (CNS) in V-PPase promoters from three different plant species

were analysed using VISTA (visualisation tool for alignment; Mayor *et al*. 2000). Abbreviated names and database accession numbers (in brackets) of all nucleotide sequences analysed during this study are as follows: VVPP1 (AF257777), M10A clone (AJ430532), vpp2 (AJ557256), HVP1 (AB032839), PVP3 (AB097115), TVP31 (X83729), AVP3 (M81892), OVP1 (D45383), OVP2 (D45384), VPP promoter (AJ544719), AVP3 promoter (AB015138), OVP2 promoter (AB012766) and CaMV 35S promoter (E05206).

4.4 Results and Discussion

Genomic organisation and analysis of the M10A V-PPase transcript

Investigations by Terrier *et al*. (2001) and Venter *et al*. (2001) revealed different RNA expression patterns of the V-PPase transcript during berry ripening. These gene expression patterns could be due to diverse environmental conditions and/or cultivar differences as observed with genes involved in the anthocyanin biosynthesis (Boss *et al*. 1996). Additionally this could indicate different V-PPase isoforms in grapevine with different expression patterns. Therefore hybridisation with the same radioactive-labelled cDNA fragment M10.A, which yielded constitutive fruit-specific expression during RNA gel blot analysis (Venter *et al*. 2001), was carried out for Southern blot analysis (Figure 4.1). None of the restriction sites, used for Southern blot analysis, were detected in the sequence of the 289 bp M10.A cDNA probe. Two to three bands could be identified in genomic DNA digested with *Hind* III and *Xba* I and three to four bands were detected in DNA digested with *Eco*R I and *Eco*R V. Genomic organisation indeed suggests that more than one V-PPase gene is present in the grapevine genome (Figure 4.1).

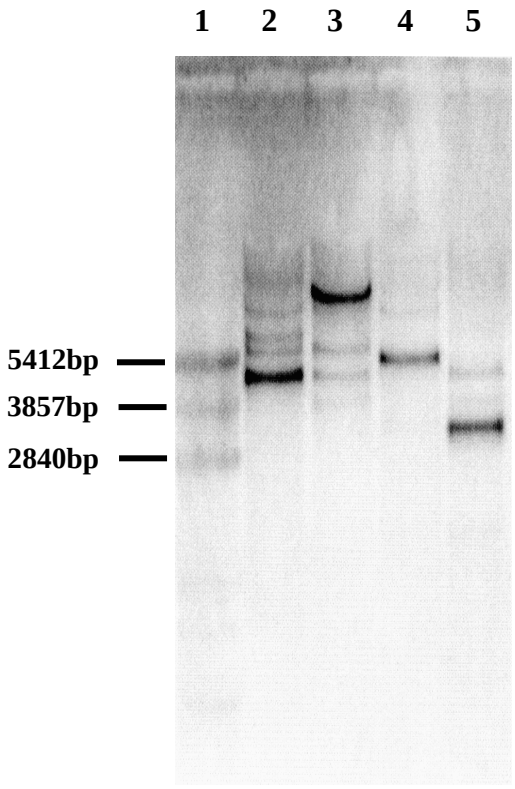


Figure 4.1 Southern blot analysis of grapevine genomic DNA using the V-PPase cDNA fragment M10A as radioactive probe. Lanes 2,3,4 and 5 represent genomic DNA digested with *Eco*R I, *Eco*R V, *Hin*d III and *Xba* I respectively. Lane 1 represents a DNA molecular weight marker.

Using LR-iPCR, an extension of the M10A clone was amplified and designated as vpp2 (AJ557256). The predicted amino acid sequence of this fragment showed 93 % identity to the same region of the VVPP1 cDNA (AF257777) isolated by Terrier *et al*. (1998). A BLAST search (Altschul *et al*. 1997) indicated that vpp2 (AJ557256) has strong similarity to the vacuolar pyrophosphatase genes of: *Pyris communis* PVP3 (96 %), *Nicotiana tabacum* TVP31 (97 %), *Hordeum vulgare* HVP1 (96 %), *Arabidopsis thaliana* AVP3 (95 %) and *Oryza sativa* OVP1 (93 %) and OVP2 (92 %). Amino acid alignment analysis deduced from 111 bp (partial region of exon1 identified during isolation of the VPP promoter) of the first exons from VVPP1 and VPP (AJ544719) revealed 48,65 % identity (Figure 4.2).

```
VVPP1       1 MGVMGDAFTQLLIPVAALVGIGFALLQWLLVSKVKVS     37
                |    |  |  | |  | ||| |   || ||| || |
VPP         1 MAILSDLGTEILVPACAIVGIVFSVVQWILVSRVKLS     37
```

Identity = 48,65 %

Figure 4.2 Amino acid alignment analysis deduced from 111 bp coding sequence of the first exons of VPP (AJ544719) and VVPP1 (AF257777) respectively.

The clear amino acid variation within the first exon of VPP and VVPP1 confirms the existence of at least two V-PPase isoforms in *Vitis vinifera* L and would therefore allow selective discrimination on RNA gel blots.

V-PPase expression on cellular level

In situ hybridisation (Figure 4.4.1) confirmed the expression pattern of the V-PPase as previously observed by RNA gel blot analysis (Venter *et al*. 2001). Berries collected two, 4, 6 and 8 weeks after anthesis are typical of grape berries prior to the onset of ripening with similar levels of reducing sugars and sucrose, and high acid levels (Figure 4.3). Stage 5 (10 weeks after anthesis) is characterised by a very rapid increase in fructose and glucose and a decrease in malate levels. Stages 7 and 8 represent fully ripened fruits where total sugars reached a maximum plateau and the acid levels are low.

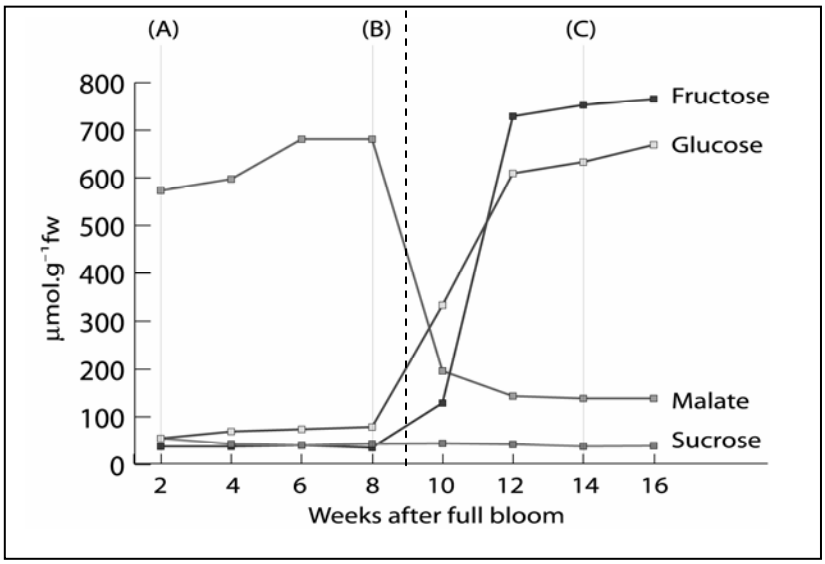


Figure 4.3 Changes in soluble sugar and malate concentrations of grape berries during development. (A), (B) and (C) represent the three stages of development utilised in the *in situ* hybridisation experiments, where (A) = stage 1, (B) = stage 4 and (C) = stage 7. A dashed line indicates *Véraison*.

In situ hybridisation experiments were conducted on berries two, 8 and 14 weeks post flowering. Distribution of the M10A V-PPase transcript in the berry is depicted in figure 4.4.1 and figure 4.4.2 represents a diagram of the cross section of the mature grape berry to assist in the analysis. Even though the M10A gene was expressed in all the developmental stages of the grape berry investigated (Figure 4.4.1 and Venter *et al*. 2001), the distribution of expression changed during development.

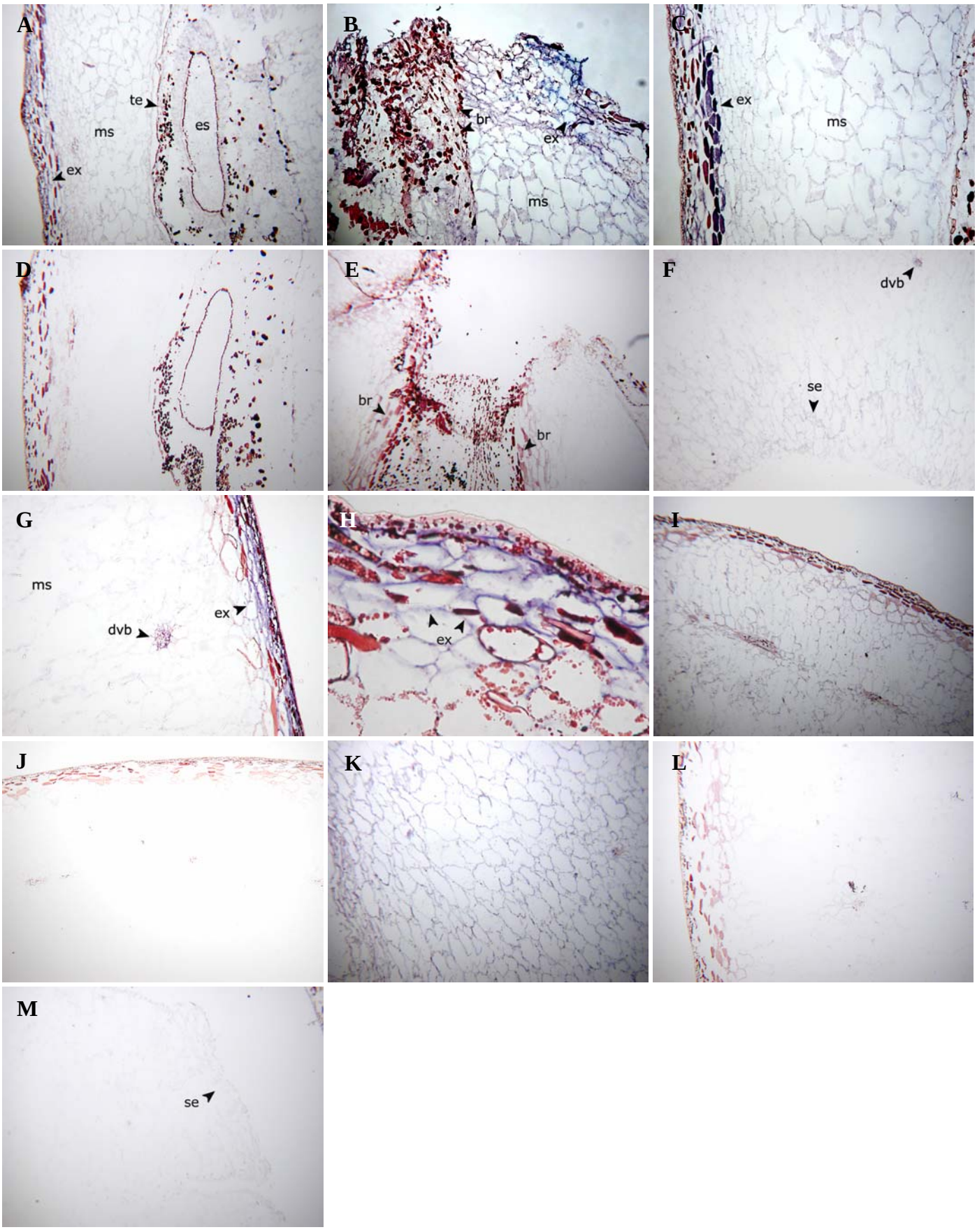

A
te
ms
es
ex
B
br
ex
ms
C
ex
ms
D
E
br
br
F
dvb
se
G
ms
ex
dvb
H
ex
I
J
K
L
M
se

Figure 4.4.1 *In situ* hybridisation expression of V-PPase mRNA in grape berry tissue performed with stages 1, 4 and 7. Results of the experiment on the stage 1 tissue are provided in (A) –(E), (D) and (E) are the negative controls for this stage. (F), (G), (H) and (J) are the results of the stage 4 experiment, with (J) being the negative control of this tissue. (I), (K), (L) and (M) are the results of the experiment performed with the stage 7 tissue, with (L) and (M) representing the negative controls. In all instances the blue colouring represents the binding of the probe to the target RNA. Ex = exocarp, se = septum, te = testa, ms = mesocarp, es = endosperm, br = brush, dvb = dorsal vascular bundle.

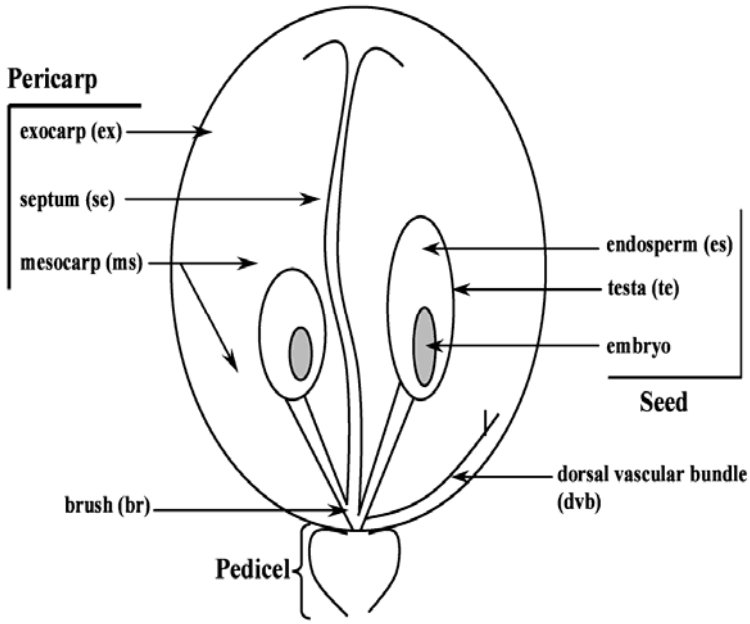


Figure 4.4.2 A simplified, anatomical representation of the grape berry.

In stage 1, which is 2 weeks post flowering (Figure 4.4.1A, B & C), the gene was expressed at a higher level in the vascular bundle tissue of the brush (Figure 4.4.1B), where the fruit is connected to the stem and the levels of solutes such as sucrose, tartrate, phenols, inorganic ions and potassium, are high (Coombe 1987). Photosynthetic activity (Pandey and Farmahan 1977) as well as auxin and/or auxin-like substances (Farmahan and Pandey 1976) are at the highest during this developmental phase coinciding with rapid cell division, cell enlargement (Coombe and McCarthy 2000) and accumulation of malic and tartric acids in the vacuole with a pH of 2.5 (Terrier *et al*. 1998). Expression of the M10A V-PPase transcript in this developmental stage

was strongly visible in the exocarp and mesocarp (Figure 4.4.1C). These results are consistent with transcriptional activation of the V-PPase in developing tissue of mung bean (Nakanishi and Maeshima 1998). A low level of expression could also be observed in the developing seeds of this stage (Figure 4.4.1A).

Alternatively, expression of the M10A transcript in the stage 4 tissue (8 weeks post flowering) (Figure 4.4.1F, G & H) was different, with a lower level of expression in the mesocarp (Figure 4.4.1F) and higher expression levels in the exocarp (Figure 4.4.1G & H) and in the vascular bundles (Figure 4.4.1G). Expression could however still be observed clearly in the parenchyma cells of the inner part of the mesocarp and the septum (Figure 4.4.1F). Stage 4, represents the so-called 'lag'-phase characterized by slow growth, primary accumulation of malic acid (Coombe and McCarthy 2000) and a decline in photosynthetic activity (Pandey and Farmahan 1977). It was suggested that the metabolism of malic acid is rapid in the vascular bundles before and during ripening leading to malate movement, by gradients, towards the vascular bundle tissue (Coombe 1987). Additionally, high levels of V-PPase expression in the vascular bundles could play a role in maintaining a balance between acid storage and passive diffusion of vacuolar content in order to prevent lethal over-acidification of the cytoplasm (Terrier *et al*. 2001).

The last developmental stage investigated (14 weeks post flowering) (Figure 4.4.1I & K) also revealed a changed distribution of expression of the M10A V-PPase cDNA. This ripening stage is associated with high levels of abscisic acid (ABA; Coombe and Hale 1973) and gibberellin (GA)-like substances (Farmahan and Pandey 1976), a decrease in acidity reaching pH of 3.5 (Terrier *et al*. 1998), a rapid accumulation of sugars and amino acids and expansion of flesh cells (Coombe 1992). The expression pattern became uniform, and no difference could be observed in the level of expression amongst the different tissues within the berry. Activity of the V-PPase proton pump in this stage could lead to an increase in tonoplast permeability which in turn would explain the increase of pH (Terrier *et al*. 2001). Photosynthetic and respiratory

activity decline during berry development. However, associated with the noticeable expansion of flesh cells, respiratory activity reveals a slight increase during this 'post-lag' phase to accommodate the biochemical changes and major influx of different compounds (Pandey and Farmahan 1977).

Effect of stress treatments on V-PPase transcript activity

Berries collected for stress treatments confirmed to be in the herbaceous stage prior to *véraison*, i.e. they displayed similar levels of fructose, glucose and sucrose, and high malic acid content (Table 4.1). Figure 4.5 shows that V-PPase transcript levels were the highest in berries subjected to cold shock and osmotic stress (0.4 M and 0.8 M Mannitol, respectively). Expression levels of the V-PPase gene during heat shock and hypoxia were vaguely detectable.

Table 4.1 Sucrose, glucose, fructose and malate levels isolated from grape berries. Each value is the average ± SD of three extractions.

Sucrose	Glucose	Fructose	Malate
	mol.g^{-1} fresh mass		
18.6 ± 0.74	20.5 ± 0.42	20.0 ± 0.11	276.0 ± 9.82

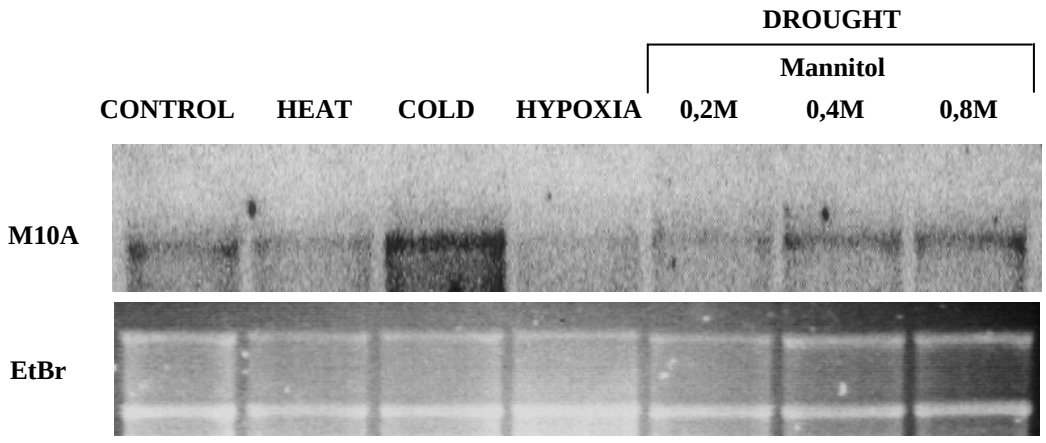


Figure 4.5 Transcript levels of **M10A** in response to different stress treatments. Panel **EtBr** shows approximately equal amounts of intact total RNA (5 μg/lane).

The low transcriptional activity of the M10A V-PPase during heat shock at 37 °C (Figure 4.5) is consistent with the results obtained by Maeshima (1991) that showed progressive inactivation of the V-PPase in mung bean at a temperature above 30 °C. These results contradict the progressive increase of V-PPase activity in grape when temperature rose to 65 °C followed by a rapid decline and inactivity after 70 °C (Terrier *et al*. 1998). This data could support our hypothesis of the existence of multiple V-PPase isoforms in grapevine exhibiting different thermo-tolerant characteristics. Drought stress as well as heat shock was conducted at a constant 37 °C, thus, drought stress induction could represent a combination of both heat and drought stress. Transcriptional activity of M10A was lower during heat shock than during a combination of drought and heat shock (Figure 4.5). This phenomenon of different transcript expression patterns involved with a combination of stress treatments rather than with a specific stress induction was demonstrated on tobacco plants where genes, specifically associated with pathogen defence, photosynthesis and sugar metabolism, reveal to have different activity in response to heat, drought and a combination of both respectively (Rizhsky *et al*. 2002).

Results obtained during hypoxia treatment might indicate a down-regulation of this V-PPase transcript (Figure 4.5). With exceptions, where the V-PPase was induced by anoxia in *Oryza sativa* (Carystinos *et al*. 1995) and alcohol dehydrogenase by hypoxia in *Arabidopsis thaliana* (Peng *et al*. 2001), a vast majority of genes reveal a rapid down-regulated activity during hypoxia and/or anoxia (Zeng *et al*. 1998). The authors chose hypoxia as a stress condition instead of total oxygen deprivation (anoxia), as it was thought that the chances of grapes on the vine to be totally deprived from oxygen was highly unlikely. It has previously been demonstrated with sugar-sensitive sucrose synthase that a commonality of up-regulated gene expression exists between low oxygen stress (hypoxia and/or anoxia) and sugar availability (Zeng *et al*. 1998). From those results together with the M10A V-PPase transcript maintaining a constant level of expression in herbacous as well as ripe berries (Venter *et al*. 2001) and therefore not coinciding with the notably sugar increase at *véraison*, we suggest that the down-

regulation of M10A during hypoxia in young berries (Figure 4.5) could be independent from sugar availability.

Chill-induced enzyme activity of the V-PPase has been reported in mung bean (*Vigna radiata*) hypocotyls (Darley *et al*. 1995) and rice (*Oryza sativa*) seedlings (Carystinos *et al*. 1995). Coinciding with the transcriptional increase in V-PPase activity during chilling in rice (Carystinos *et al*. 1995), we have shown that the M10A V-PPase transcriptional activity was up-regulated during cold shock in grapes (Figure 4.5). In concert with other environmental constraints, chilling can alter cytoplasmic and vacuolar pH leading to the reduction of ATP levels consumed by glycolytic reactions or by the inactivation of the V-ATPase (Davies 1997, Rea and Poole 1993). During this stress situation it is believed that the V-PPase act as an alternative energy 'back-up' system by utilizing inorganic pyrophosphate (PPi) when ATP levels are diminished and subsequently preventing cellular damage by recovering pH homeostasis and stabilizing the vacuolar pH gradient to maintain the vital activities of the vacuole (Taiz and Zeiger 1998, Davies 1997, Rea and Poole 1993).

The V-PPase 5'-flanking region: isolation, determination of promoter activity and computational sequence analysis with putative correlation to gene expression patterns

A 5 kb genomic sequence was isolated using LR-iPCR. Additionally, a genomic PCR fragment (approx.5.4 kb) was amplified from a positive phage isolated after a second round library screening. A 510 bp promoter sequence was identified from the consensus sequence derived from the genomic products. This fragment was used as a template for inverse PCR using primers 10Apro1frw (5'-GACGTGGCCTCTTTTGATTAC-3') and 10Apro1rev (5'-GCCAAAGGCAACTCCATTATTC-3' towards the start codon). An extended promoter sequence, 1567 bp upstream from the start codon, was amplified and isolated from grapevine genomic DNA using primers 10Aprom-frw (5'-CCCCATCTAGAGGTCTCTAAACAAACTTACC-3') and 10Aprom-rev (5'-CAAAGGATCCCATGGACGGACGGAC-3') that were designed for direct cloning of the

promoter into the pGFP expression vector. Particle bombardment analyses using grape and strawberry tissues were unsuccessful, however, young and mature green peppers (*Capsicum* spp.) proved to be an efficient alternative specie for transient evaluation of promoter activity in non-climacteric tissues. GFP-reporter gene expression revealed the activity of the isolated promoter (Figure 4.6A), designated as VPP-promoter (AJ544719), to be visually similar to the CaMV 35S promoter (Figure 4.6B), thus confirming promoter activity on a transient level when compared to positive control (Figure 4.6B) as well as to the promoterless negative control (Figure 4.6C).

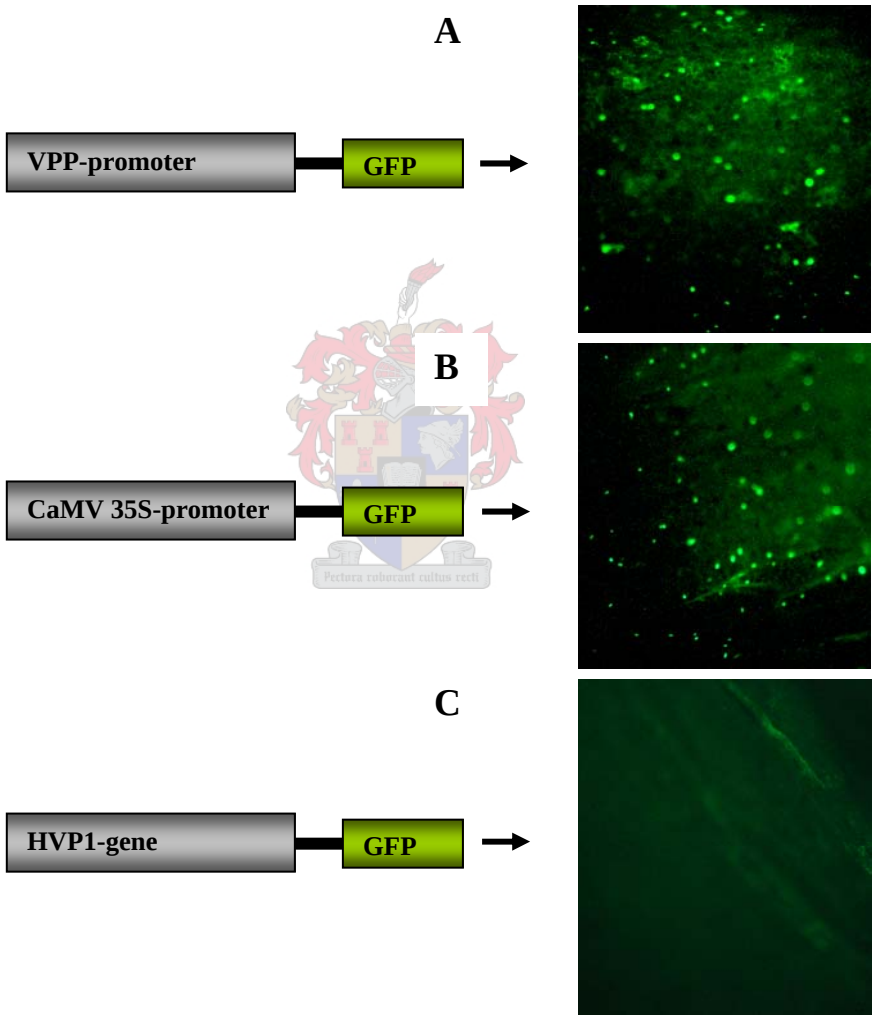


Figure 4.6 Green fluorescent protein (GFP)-reporter gene expression to evaluate promoter activity in green pepper (*Capsicum* spp.) after particle bombardment. (**A**) Indicates activity of the isolated VPP promoter (AJ544719). (**B**) And (**C**) represents reporter gene constructs using the CaMV 35S promoter (E05206) as positive control and the *Hordeum vulgare* HVP1 gene (AB032839) as a promoterless negative control respectively.

The promoter sequence was analysed in the (+)-strand for putative *cis*-motifs responsive to tissue-specificity, hormonal induction and environmental stress conditions (shown in figure 4.7A, 4.7B and table 4.2). Coinciding with RNA gel blot results (Figure 4.5), exception was made for analysis in the (-)-strand where two distinct motifs (indicated in figure 4.7A as LTRE), for low temperature and drought responsiveness, were identified by both PLACE and PlantCARE. These motifs, C repeat DRE, contain the core sequence CCGAC (Baker *et al*. 1994, Jiang *et al*. 1996, Kim *et al*. 2002) and are usually associated with the combined response to cold, drought, abscisic acid (ABA) and/or salt stress (Baker *et al*. 1994, Shinozaki and Yamaguchi-Shinozaki 2000). Recently it has been shown that light signalling mediates C-repeat DRE to activate cold and drought responsiveness in *Arabidopsis thaliana* (Kim *et al*. 2002). In contrast to low level or no responsiveness to heat shock treatment (Figure 4.5), a cluster of eight heat shock elements (HSEs), residing within a few hundred base pairs, area –768 bp to –1440 bp, as exhibited by most eukaryotic heat responsive genes (Schöffl *et al*. 1998) was identified. The presence of HSE sequences in the VPP promoter indicate the putative involvement of heat shock transcription factors (HSFs), however, a plausible reason consistent with heat shock results (Figure 4.5), could be the inactivation of HSFs by phosphorylation during heat stress (Schöffl *et al*. 1998). The mechanism of signalling pathways involved with the heat shock response is poorly understood, but it is suggested that heat shock proteins are essential to cells during different stages of development (Schöffl *et al*. 1998).

GC-motifs, known for anoxia responsiveness (Manjunath and Sachs 1997), were identified in the VPP promoter of grapevine. Results from oxygen-deficient but not -deprived induction, hypoxia revealed no presence of the V-PPase M10A transcript (Figure 4.5). Groups of GT-elements in close proximity (–132 bp to –144 bp, -421 bp to –488 bp and –1270 bp to –1421 bp) were identified. GT-elements with a core sequence G(A/G)(A/T)AA(A/T), closely related to the binding factors of GATA-boxes, of which two were identified in VPP, are known to be responsive to light and associated with cell-type-specific transcription (Argüello-Astorga and

Herrera-Estrella 1998, Zhou 1999, Teakle *et al*. 2002). It was previously demonstrated that a combination of GT- and GATA-elements is necessary for light responsiveness and that the GT-elements were necessary but not sufficient for light induction (Zhou 1999). The most abundant putative *cis*-elements identified in the (+)-strand of the VPP promoter, have previously been shown to be responsible for tissue-and/or cell-type-specific regulation of expression. These elements include the core motif ATATT, for high level expression in roots (Elmayan and Tepfer 1995) and the plant-specific Dof motif with the core sequence AAAG, associated with endosperm specific expression (Yanagisawa and Schmidt 1999, Yanagisawa 2002). The presence of these two motifs is consistent with expression patterns of the M10A transcript in roots (Venter *et al*. 2001) and endosperm (Figure 4.4.1A). Putative roles of Dof proteins to be associated with light, auxin, defence and gibberellin responsiveness have previously been described (Yanagisawa 2002) and could therefore play a cooperative role with other factors during grape ripening.

Two putative abscisic acid (ABA) responsive elements (indicated as ABREs, Figure 4.7A) were identified. Necessary for a diverse range of regulatory cues, functional responsiveness of the ABRE motif varies depending on the sequences flanking the conserved sequence ACGT (Busk and Pagès 1998). Previous investigations have revealed the promoters of chill-and drought-responsive genes, induced by exogenous ABA, to contain ABRE-motifs. However, it is suggested that the ABA-signalling pathway is primarily involved with dehydration stress rather than cold shock (Shinozaki and Yamaguchi-Shinozaki 2000). The presence of these ABRE elements could be associated with drought responsiveness of the V-PPase transcript (Figure 4.5) and therefore be mediated by an ABA-dependant pathway (Shinozaki and Yamaguchi-Shinozaki 2000). As described by Jiang *et al*. (1996), the presence of LTRE-motifs (containing a CCGAC-core) in the VPP promoter (Figure 4.7A) could mainly be responsible for activation of the M10A transcript expression during cold shock (Figure 4.5) and that the low temperature responsiveness is independent of an ABA-biosynthetic pathway (Jiang *et al*. 1996).

Multiple gibberellin (GA)-responsive *cis*-motifs have previously been characterized (Gómez-Cadenas *et al*. 2001). We have identified one such motif, namely GARE (GA-responsive element) with the core sequence TAACA(A/G)A residing between common promoter elements, the CAAT-box and the TATA-box, 154 bp from the putative transcriptional start site. The GARE motif together with another GA-responsive element, the pyrimidine box, identified in the (-)-strand (data not shown), revealed to be partially involved in sugar repression in the rice α-amylase gene (Morita *et al*. 1998). Identification of a putative GARE motif might indicate that a combination of gibberellin and differential sugar signals could trigger the activity of the V-PPase transcript (Morita *et al*. 1998). A putative auxin responsive element, known as AuxRE (Ulmasov *et al*. 1995), has been identified 11 bp upstream from the ATG. Acting as a molecular 'trigger' (Guilfoyle *et al*. 1998), previous research has implicated auxin and the effect of auxin-like compounds to play a role in the control of grape berry maturation (Coombe and Hale 1973, Davies *et al*. 1997). Hormone responsive elements AuxRE, ABRE and GARE in the VPP promoter suggest that fluctuating levels of endogenous auxin, ABA and GA during grape development (as described earlier) could trigger the VPP promoter. Therefore regulating the expression of this V-PPase transcript (Figure 4.4.1) to the hormonal status in specific stages of berry ripening.

The possibility of other motifs to be present in a larger part of the VPP promoter was not ruled out. Analysis using PLACE and PLantCARE revealed no sucrose responsive elements/boxes (Grierson *et al*. 1994, Tsukuya *et al*. 1991), previously identified in the promoters of the hexose transporter gene (Fillion *et al*. 1999) and the dihydroflavonol reductase gene (Gollop *et al*. 2002) in grapevine, to be present in the VPP promoter. However, 'steady-state' expression pattern of the M10A transcript in all developmental stages of berry ripening (Venter *et al*. 2001) as well as the absence of putative sucrose responsive elements in the available sequence of the VPP promoter suggests that the V-PPase transcript used in this study does not have a ripening related role that coincide with the rapid increase of sugars at the onset of ripening. Ethylene responsive elements (ERE) have been identified in the VPP promoter (results not shown),

however, there is no significant evidence that ethylene plays a role during berry development (Coombe 1992) therefore, the putative influence of ethylene on V-PPase activity has not been discussed. Previous investigations have implicated elements such as potassium, calcium and magnesium in relation to specific plant organs and cell compartmentation, to play a significant role on the V-PPase activity (Rea and Pool 1993, Maeshima 2000). The putative correlation of regulatory data and V-PPase transcriptional activity to these elements as well as to other stress, environmental and ripening-related factors, not discussed during this study, would help to further elucidate the complex nature of the V-PPase proton pump during fruit development.

CNS analysis of V-PPase promoters

The VISTA program (Mayor *et al*. 2000) was used to align and compare V-PPase promoters of arbitrary length between *Vitis vinifera* (VPPprom), *Arabidopsis thaliana* (AVP3prom) and *Oryza sativa* (OVP2prom) for analysis of conserved noncoding sequences (CNSs). Promoter sequences (VPPprom as the 'base' sequence) plus 111 bp of the first exon of each sequence was compared (VPPprom/AVP3prom, VPPprom/OVP2prom and AVP3prom/OVP2prom) using constant conservation criteria of 70 % identity over a 10 bp, 15 bp, 20 bp, 25 bp, 30 bp, 50 bp and 100 bp range respectively. Each CNS identified from two compared promoters was submitted to the databases PLACE (Higo *et al*. 1999) and PlantCARE (Rombauts *et al*. 1999). The highest CNS frequency of putative known *cis*-motif identity was obtained with the VISTA criteria of ≥70 % identity in a 20 bp window. Figure 4.8 demonstrates the VISTA output of aforementioned criteria where each peak represents a conserved segment. VISTA analyses of the VPPprom/AVP3prom comparison indicate that the V-PPase promoters of two dicot plants are more conserved than VPPprom/OVP2prom and AVP3prom/OVP2prom comparisons (Figure 4.8). The functional implication of conserved regions within promoter sequences of the same gene in different plant species is unclear (Guo and Moose 2003). These results suggest that the promoters acting as molecular 'switches' regulate V-PPase transcriptional activity (using PPi as an alternative energy source to maintain crucial cell functions) in a distinct mode and are 'tailor-made' according to plant-specie and requirements.

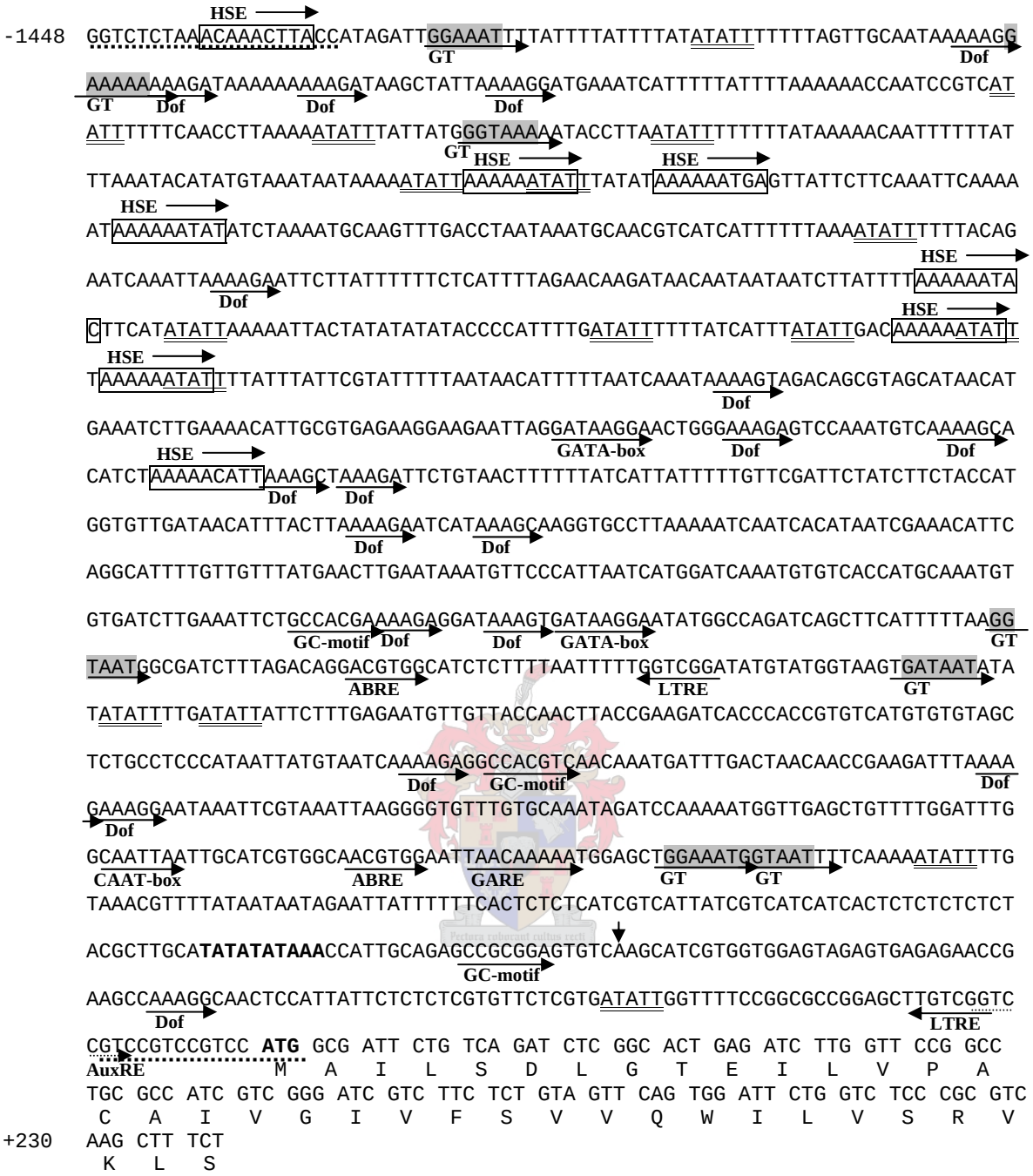


Figure 4.7A Nucleotide sequence of the VPP promoter (AJ544719) and partial sequence of the first exon. Putative TATA-box and transcription start site (sequence at +1) is indicated in bold and downward arrowhead respectively. A putative CAAT-box is located 150 bp upstream and ATG 142 bp downstream from the TATA-box respectively. Arrows indicate direction of each putative transcription factor binding site/*cis*-motif analysed within the promoter region. Promoter analysis and key functions of each *cis*-element are described in the text and table 4.2. Dashed underlined nucleotides represent primer sequences used for promoter amplification from genomic DNA. Amino acid translation of the first exon is indicated with single letter amino acid code beneath each codon.

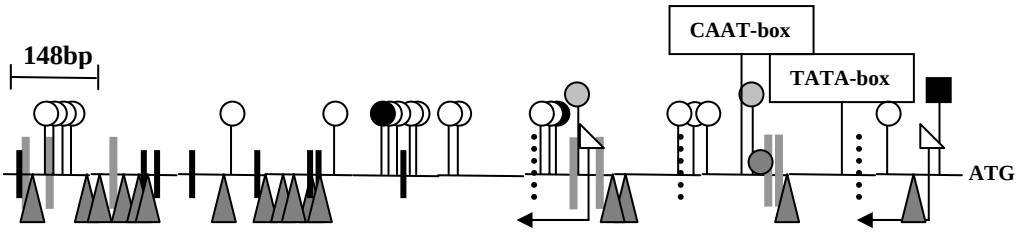


Figure 4.7B *In silico* map of the VPP promoter shows putative transcription factor binding sites analysed in the (+)-strand of the promoter, except for the motif 'LTRE', identified in the (-)-strand indicated with arrows. Key symbols of each motif appear in table 4.2.

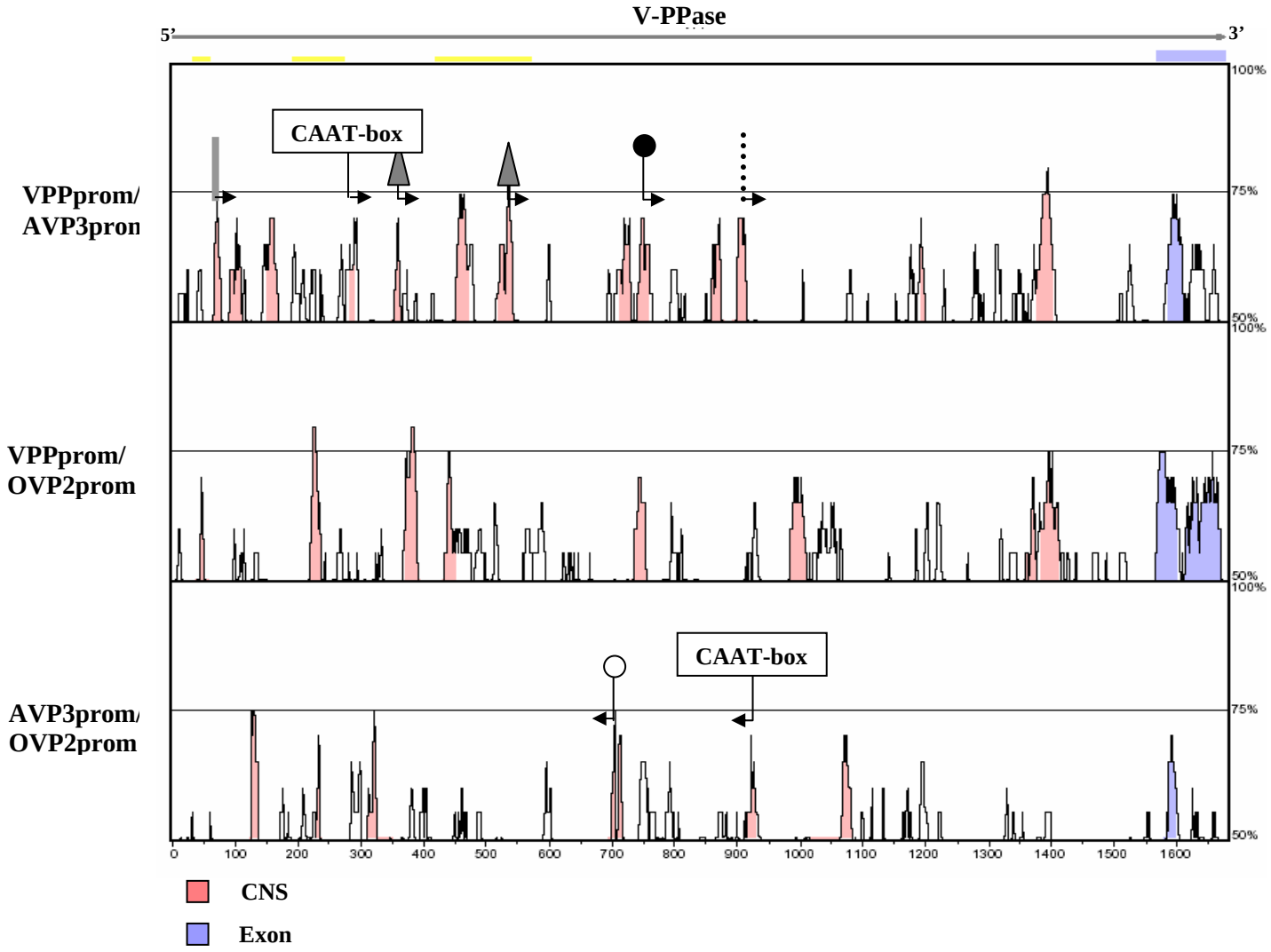
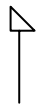


Figure 4.8 VISTA graphical output showing peaks of similar conserved sequences within compared V-PPase promoters. Key symbols of putative known *cis*-motifs are plotted (see table 4.2).

Table 4.2 List of key symbols and description of putative transcription factor binding sites identified during computational promoter analysis. *Cis*-elements and their functions were identified using plant transcription factor databases, PLACE and PlantCARE.

Cis-Motif	Key	Description	Reference/s
LTRE		Low temperature responsive element Core sequence ACCGAC a.k.a. C-repeat DRE element Involved in drought responsiveness	Baker et al. 1994 Kim et al. 2002
ABRE		Abscisic acid responsive element	Busk and Pagès, 1998 (and references therein)
GATA-box		Cis-element for light-and tissue specific responsiveness	Teakle et al. 2002
GARE		Gibberellin-responsive element Partially involved with sugar repression	Gómez-Cadenas et al, 2001 Morita et al. 1998
GC-motif		Anoxic inducibility	Manjunath and Sachs, 1997
HSE		Heat shock element	Schöffl et al. 1998
Dof		Dof-core element, unique to plant Associated with endosperm specific expression	Yanagisawa and Schmidt, 1999 Yanagisawa 2002
GT-element		Light responsiveness – G(A/G)(A/T)AA(A/T)-core Cell-type specificity Stabilize TFIIA-TBP-TATA complex	Zhou 1999 Le Gourrierec and Zhou, 1999
ATATT		Core motif, associated with high level expression in roots	Elmayan and Tepfer, 1995
AuxRE		Auxin responsive element	Guilfoyle et al. 1998 Ulmasov et al, 1995

4.5 Conclusion

The V-PPase gene (AJ430532) that we have isolated was expressed in all stages of grape berry development. Despite this apparent constitutive expression pattern *in situ* hybridisation clearly showed that expression was largely restricted to certain cell types in the grape berry. This spatial expression pattern changes during ripening. This could imply that the transcriptional activity of this particular V-PPase clone is controlled by the energy status of developing cells during different stages of ripening in the berry. However, it is also evident that the particular gene is also induced by cold and water stress.

Integration of the promoter data with gene expression events allowed correlation of transcription factor binding sites to known ripening induced changes, cell-and/or tissue-specificity as well as environmental stress in the grape berry. The use of this integrative approach and the *in silico* mapping of regulatory regions could allow for a more accurate strategy to conduct sequential deletion analysis for genetic improvement of grapevine.

Acknowledgments

Financial support from Winetech, the National Research Foundation and the department of Trade and Industry and assistance from Fletcher Hiten, Hennie Groenewald, Anita Burger, Charmaine Vermaak and Marna van der Merwe are acknowledged.

References

Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. 1990. Basic local alignment search tool. Nucleic Acids Res 215: 403-410.

Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.

Argüello-Astorga,G.and Herrera-Estrella,L.1998. Evolution of light-regulated plant promoters. Annu.Rev.Plant Physiol.Plant Mol.Biol 49: 525-555.

Baker, S.S., Wilhelm, K.S. and Thomashow, M.F.1994. The 5'-region of Arabidopsis thaliana cor15a has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. Plant Mol Biol 24:701-713.

Ballesteros,E., Donaire,J.P. and Belver,A. 1996. Effects of salt stress on H(+)-ATPase and H(+)-PPase activities of tonoplast-enriched vesicles isolated from sunflower roots. Physiol Planta 97: 259-268.

Boss,P.K., Davies,C. and Robinson,S.P. 1996. Expression of anthocyanin biosynthesis pathways genes in red and white grapes. Plant Mol Biol 32: 565-569.

Bower,R., Elliott,A.R., Potier,B.A.M. and Birch,R.G. 1996. High-efficiency, microprojectile mediated co-transformation of sugarcane, using visible or selected markers. Mol Breed 2: 239-249.

Busk,P.K. and Pagès,M. 1998. Regulation of abscisic acid-induced transcription. Plant Mol Biol 37:425-435.

Carystinos,G.D., MacDonald,H.R., Monroy,A.F., Dhindsa,R.S. and Poole,R.J. 1995. Vacuolar H(+)-translocating pyrophosphatase is induced by anoxia or chilling in seedlings of rice . Plant Physiol 108: 641-649.

Coombe,B.G. 1987. Distribution of solutes within the developing grape berry in relation to its morphology. Am J Enol Vitic 38(2): 120-126.

Coombe,B.G. and Hale,C.R. 1973. The hormone content of ripening grape berries and the effects of growth substance treatments. Plant Physiology 51: 629-634.

Coombe,B.G. Research on development and ripening of the grape berry. Am.J.Enol.Vitic. 43[1], 101-110. 1992.

Coombe,B.G. and McCarthy,M.G. 2000. Dynamics of grape berry growth and physiology of ripening. Austr J Grape and Wine Res 6: 131-135.

Darley,C.P., Davies,J.M., and Sanders,D. 1995. Chill-Induced Changes in the Activity and Abundance of the Vacuolar Proton-Pumping Pyrophosphatase from Mung Bean Hypocotyls. Plant Physiol 109: 659-665.

Davies,C. and Robinson,S.P. 1996. Sugar Accumulation in Grape Berries: Cloning of two putative vacuolar invertase cDNAs and their expression in grapevine tissues. Plant Physiology 111: 275-283.

Davies,C., Boss,P.K. and Robinson,S.P. 1997. Treatment of grape berries, a nonclimacteric fruit with a synthetic auxin, retards and alters the expression of developmentally regulated genes. Plant Physiology 115: 1155-1161.

Davies,C., Wolf,T. and Robinson,S.P. 1999. Three putative sucrose transporters are differentially expressed in grapevine tissues. Plant Science 147: 93-100.

Davies,C. and Boss,P.K. 2000. The use of molecular techniques to study and manipulate the grapevine: Why and how? Austr J Grape and Wine Res 6: 159-167.

Davies,J.M. 1997. Vacuolar energization: pumps, shunts and stress. J Exp Bot 48(308): 633-641.

Drozdowicz,Y.M. and Rea,P.A. 2001. Vacuolar H(+) pyrophosphatases: from the evolutionary backwaters into the mainstream. Trends Plant Sci. 6: 206-211.

Elliott,A.R., Campbell,J.A., Dugdale,B., Brettell,R.I.S. and Grof,C.P.L. 1999. Green-fluorecent protein facilitates rapid in vivo detection of genetically transformed plant cells. Plant Cell Rep 18: 707-714.

Elmayan,T. and Tepfer,M. 1995. Evaluation in tobacco of the organ specificity and strength of the rol D promoter, domain A of the 35S promoter and the 35S^2 promoter. Transgenic Res 4:388-396.

Famiani,F., Walker,R.P., Tecsi,L., Chen,Z.H., Proietti,P. and Leegood,R.C. 2000. An immunohistochemical study of the compartmentation of metabolism during the development of grape (Vitis vinifera L.) berries. J Exp Bot 51: 675-683.

Farmahan,H.L. and Pandey,R.M. 1976. Hormonal regulation of the lag phase in seeded and seedless grapes (Vitis vinifera L.). Vitis 15: 227-235.

Fillion,L., Ageorges,A., Picaud,S., Coutos-Thévenot,P., Lemoine,R., Romieu,C. and Delrot,S. 1999. Cloning and expression of a Hexose transporter gene expressed during the ripening of grape berry. Plant Physiol 120: 1083-1093.

Gollop,R., Even,S., Colova-Tsolova,V. and Perl,A. 2002. Expression of the grape dihydroflavonol reductase gene and analysis of its promoter region. J Exp Bot 53(373) 1397-1409.

Gómez-Cadenas,A., Zentella,R., Sutliff,T.D. and Ho,T.-H.D. 2001. Involvement of multiple *cis*-elements in the regulation of GA responsive promoters: Definition of a new *cis*-element in the Amy32b gene promoter of barley (*Hordeum vulgare*). Physiol Planta 112: 211-216.

Grierson,C., Du,J.S., Zabala, M.T., Beggs,K., Smith,C., Holdsworth,M. and Bevan,M. 1994. Separate cis sequences and trans factors direct metabolic and developmental regulation of a potato tuber storage protein gene. Plant J 5:815-826.

Guilfoyle,T., Hagen,G., Ulmasov,T. and Murfett,J. 1998. How does auxin turn on genes? Plant Physiol 118: 341-347.

Guo,H. and Moose,S.P. 2003. Conserved noncoding sequences among cultivated cereal genomes identify candidate regulatory sequence elements and patterns of promoter evolution. Plant Cell 15: 1143-1158.

Higgins,D., Thompson,J., Gibson,T., Thompson,J.D., Higgins,D.G. and Gibson,T.J. 1994. CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through

sequence weighting,position-specific gap penalties and weight matrix choice. Nucleic Acids Res22: 4673-4680.

Higo,K., Ugawa,Y., Iwamoto,M. and Higo,H. 1998. PLACE: a database of plant cis-acting regulatory DNA elements. Nucleic Acids Res. 26: 358-359.

Iocco,P., Franks,T. and Thomas, M.R. 2001. Genetic transformation of major wine grape cultivars of *Vitis vinifera* L. Transgenic Res 10: 105-112.

Jiang,C., Iu,B. and Singh,J. 1996. Requirement of a CCGAC cis-acting element for cold induction of the BN115 gene from winter Brassica napus. Plant Mol Biol 30: 679-684.

Kim,H.J., Kim,Y.K., Park,J.Y. and Kim,J. 2002. Light signalling mediated by phytochrome plays an important role in cold-induced gene expression through the C-repeat/dehydration responsive element (C/DRE) in Arabidopsis thaliana. Plant J 29: 693-704.

Le Gourrierec,J., Li,Y.-F. and Zhou,D.-X. 1999. Transcriptional activation by *Arabidopsis* GT-1 may be through interaction with TFIIA-TBP-TATA complex. Plant J 18: 663-668.

Lerchl,J., König,S., Zrenner,R. and Sonnewald,U. 1995. Molecular cloning, characterization and expression analysis of isoforms encoding tonoplast-bound proton-translocating inorganic pyrophosphatase in tobacco. Plant Mol Biol 29: 833-840.

Maeshima,M. 1991. H+-translocating inorganic pyrophosphatase of plant vacuoles. FEBS Lett 196: 11-17.

Maeshima,M. 2000. Vacuolar H(+)-pyrophosphatase. Biochim Biophys Acta 1465: 37-51.

Manjunath,S. and Sachs, M.M. 1997. Molecular characterization and promoter analysis of the maize cytosolic glyceraldehyde 3-phosphate dehydrogenase gene family and its expression during anoxia. Plant Mol Biol 33(1): 97-112.

Mayor,C., Brudno,M. Schwartz,J.R., Poliakov,A., Rubin,E.M., Frazer,K.A., Pachter,L.S. and Dubchak,I. 2000. VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics 16(11): 1046-1047.

Milner,I.D., Ho,L.C. and Hall,J.L. 1995. Properties of proton and sugar transport at the tonoplast of tomato (*Lycopersicon esculentum*) fruit. Physiol Planta 94: 399-410.

Morita,A., Umemura,T., Kuroyanagi,M., Futsuhara,Y., Perata,P. and Yamaguchi,J. 1998. Functional dissection of a sugar-repressed α-amylase gene (Ramy1A) promoter in rice embryos. FEBS Lett 423:81-85.

Murashige,T. and Skoog,F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Planta 15: 473-497.

Nakanishi,Y. and Maeshima,M. 1998. Molecular cloning of vacuolar H(+)-pyrophosphatase and its developmental expression in growing hypocotyl of mung bean. Plant Physiol 116: 589-597.

Pandey,R.M. and Farmahan,H.L. 1977. Changes in the rate of photosynthesis and respiration in leaves and berries of Vitis vinifera grapevines at various stages of berry development. Vitis 16: 106-111.

Peng,H.-P., Chan,C.-S., Shih,M.-C. and Yang,S.F. 2001. Signaling events in the hypoxic induction of alcohol dehydrogenase gene in *Arabidopsis*. Plant Physiol 126: 742-749.

Ratajczak,R. 2000. Structure, function and regulation of the plant vacuolar H+-translocating ATPase. Biochim Biophys Acta 1465: 17-36.

Rea,P.A. and Poole,R.J. 1993. Vavuolar H^+-translocating pyrophosphatase. Annu Rev Plant Physiol Plant Mol Biol 44: 157-180.

Reese,M.G. and Eeckman,F.H. 1995. Novel Neural Network Algorithms for Improved Eukaryotic Promoter Site Recognition. Accepted talk for The seventh international Genome sequencing and analysis conference, Hyatt Regency, Hilton Head Island, South Carolina September 16-20.

Rezaian,M.A. and Krake,L.R. 1987. Nucleic acid extraction and virus detection in grapevine . J Vir Methods 17: 277-285.

Rizhsky,L., Liang,H. and Mittler,R. 2002. The combined effect of drought stress and heat shock on gene expression in tobacco. Plant Physiol 130: 1143-1151.

Robinson,S.P. and Davies,C. 2000. Molecular biology of grape berry ripening. Austr J Grape and Wine Res 6: 175-188.

Rombauts,S., Déhais,P., Van Montagu,M. and Rouzé,P. 1999. PlantCARE, a plant *cis*-acting regulatory element database. Nucleic Acids Res 27: 295-296.

Roubelakis-Angelakis,K.A. 2001. Molecular biology and biotechnology of the grapevine. 1st ed. Kluwer academic publishers, Dordrecht.

Sakakibara,Y., Kobayashi,H. and Kasamo,K. 1996. Isolation and characterization of cDNAs encoding vacuolar H^+-pyrophosphatase isoforms from rice (*Oryza sativa* L.). Plant Mol Biol 31: 1029-1038.

Sambrook,J., Fritsch,E.F. and Maniatis,T. 1989. Molecular Cloning: A laboratory manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Schöffl,F., Prändl,R. and Reindl,A. 1998. Regulation of the Heat-Shock Response. Plant Physiol 117: 1135–1141.

Shinozaki,K. and Yamaguchi-Shinozaki,K. 2000. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. Curr Opin Plant Biol 3: 217-223.

Steenkamp,J., Wiid,I., Lourens,A. and Van Helden,P. 1994. Improved method for DNA extraction from *Vitis vinifera*. Am J Enol Vitic 45: 102-106.

Tanaka,Y., Chiba,K., Maeda,M., Maeshima,M., 1993. Molecular cloning of cDNA for vacuolar membrane proton-translocating inorganic pyrophosphatase in Hordeum vulgare. Biochem Biophys Res Commun 190 (3): 1110-1114.

Taiz,L. and Zeiger,E. 1998. Solute Transport. *In* Plant Physiology. Second ed. Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts. pp. 125-152.

Teakle,G.R., Manfield,I.W., Graham,J.F. and Gilmartin,P.M. 2002. Arabidopsis thaliana GATA factors: organisation, expression and DNA-binding characteristics. Plant Mol Biol 50 :43-57.

Terrier,N., Deguilloux,C., Sauvage,F.X., Martnoia,E. and Romieu,C. 1998. Proton pumps and anion transport in *Vitis vinifera*: The inorganic pyrophosphatase plays a predominant role in the energization of the tonoplast. Plant Physiol Biochem 36(5): 367-377.

Terrier,N., Sauvage,F.X., Ageorges,A. and Romieu,C. 2001. Changes in the acidity and in proton transport at the tonoplast of grape berries during development. Planta 213: 20-28.

Terrier,N. and Romieu,C. 2001. Grape berry acidity. *In* Molecular biology and biotechnology of the grapevine. 1st ed. Kluwer Academic Publishers, Dordrecht. pp. 35-57.

Tsukaya,H., Ohshima,T., Naito,S., Chino,M. and Komeda,Y. 1991. Sugar-dependant expression of the CHS-A gene for chalcone synthase from *Petunia* in transgenic *Arabidopsis*. Plant Physiol 97: 1414-1421.

Ulmasov,T. Liu,Z.-B., Hagen,G. and Guilfoyle,T.J. 1995. Composite structure of auxin responsive elements. Plant Cell 7: 1611-1623.

Venter,M., Burger,A.L. and Botha,F.C. Identification of genes differentially expressed during berry ripening in *Vitis vinifera* cv. Chardonnay. Vitis 40(4): 191-196.

Vivier,M.A. and Pretorius,I.S. 2000. Genetic improvement of grapevine: Tailoring grape varietiesfor the third millennium - A review. S Afr J Enol Vitic 21: 5-26.

Yanagisawa,S. and Schmidt,R.J. 1999. Diversity and similarity among recognition sequences of Dof transcription factors. Plant J 17: 209-214.

Yanagisawa,S. 2002. The Dof family of plant transcription factors. Trends Plant Sci. 7: 555-560.

Zeng,Y., Wu,Y., Avigne,W.T and Koch,K.E. 1998. Differential regulation of sugar-sensitive sucrose synthases by hypoxia and anoxia indicate complementary transcriptional and posttranscriptional responses. Plant Physiol 116: 1573-1583.

Zhou,D.X. 1999. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. Trends in Plant Sci 4:210-214.

Chapter 5

Concluding remarks & future prospects

Grapevine is considered as a valuable crop. However, many biological processes in grapevine, especially fruit ripening, remain poorly understood. Different physiological, metabolic and genetic studies have been conducted to further elucidate the ripening process. Results obtained from these investigations have shed light on specific aspects of development regarding specific changes on certain levels e.g. focus on a certain biosynthetic pathway that has an effect on size and/or colour of the grape berry. At least part of these developmental changes are controlled by the precise regulation of differential gene expression. Thus, to gain further insight into these changes, 1) it was necessary to identify differentially expressed sequences at different stages of fruit development that in turn would allow for the 2) isolation of developmental stage- and/or tissue-specific promoters. The use of these promoters could have a major impact on programs aimed at the genetic manipulation of grapevine. In addition, the study and characterisation of promoter elements could help elucidate the molecular mechanisms that are directly associated with ripening-stage and fruit-specific gene expression.

Predictions to understand gene regulatory events must ultimately anticipate experimental confirmation. However, as mentioned in chapter 1, for this study it was not realistic to evaluate fruit-bearing, transgenic grapevine for fruit-specific promoter activity. Therefore, we focussed on the combination of gene expression data (and/or knowledge of gene function) with computational promoter data (when compared to plant *cis*-element databases) that could provide a putative but accurate framework for the deciphering of coordinated gene activity during grape development. From this approach we envisage a more refined sequential deletion and induction analyses for future promoter evaluation in stable transformed plants. In chapter 2 the principles, advantages and applications of an integrative strategy that relies on previous experimental

designs and lessons learned from relatively 'simple' model organisms such as *E. coli* and *S. cerevisiae* are discussed. This chapter emphasizes the importance of combining regulatory data (*cis*-motifs and transcription factors), bioinformatics and gene expression profiles under a specific condition to form a better understanding of complex, multi-cellular eukaryotes, focussed especially on plants. From the analysis of previous research discussed in chapter 2, it was apparent that the integrative approach will not only advance the understanding of complex biological mechanisms but could furthermore facilitate a more accurate prediction of gene identity and/or association (but not necessarily function).

The first phase of this study, described in chapter 3, was the application of a high throughput gene expression technology, cDNA-AFLP, in *Vitis vinifera* L. Identification of a vast amount of expressed genes revealed cDNA-AFLPs to be invaluable for plant species where the prerequisite of genome sequencing have not been met. This mRNA fingerprinting system allows for identification of genes in different tissues and organs as well as during different stages of development to be analysed simultaneously. The isolation of pure and intact total RNA played an important part of this experimental strategy, where in grapes with high polyphenol content, the RNA isolation process proved to be challenging especially for older grapes that account for later stages of development. A major advantage of cDNA-AFLPs over other methods e.g. chip-based methods or cDNA library construction, is that transcript analysis can be conducted on gene sequences with no prior knowledge of any transcriptional activity and the cloning of cDNA fragments is not a prerequisite for initial transcript screening. However, as concluded in chapter 3, candidate transcripts identified from the PAGE gel represent a mixed template, therefore it is essential to clone the fragment for subsequent expression analysis. In this study cDNA-AFLPs was successfully used to identify genes expressed during a specific condition i.e. during development in grapes. However, from this analysis it is envisaged that cDNA-AFLPs could be useful to identify large sets of grapevine genes in response to various conditions such as hormone treatment, different sugars levels, pH fluctuations and/or various stress conditions. With a limited amount of template material needed (using the pre-

amplification step, chapter 3), the powerful and sensitive nature of this transcript profile technology is revealed. The cDNA-AFLP system could play an important role for future gene identification and manipulation programs in *Vitis vinifera* L.

From the cDNA clones isolated in chapter 3, we selected a cDNA transcript encoding vacuolar pyrophosphatase (V-PPase), as the candidate gene to be used for transcriptional and promoter analysis in this study. Other clones revealed more 'attractive' tissue- and stage-specific gene expression patterns during RNA gel blot analysis for candidate gene selection. However, grape berry acidity is of vital importance for stability in wine and needs to receive significant attention (Terrier and Romieu 2001). For that reason, we envisaged that analysis of the V-PPase promoter and transcriptional activity would be valuable for further elucidation of the molecular characteristics regulating pH variations during berry ripening. The differential transcriptional profiles and stress-responsive features of the V-PPase, previously investigated and exhibited during this study, revealed this gene to be an exciting target for study and genetic enhancement in grapevine. Therefore, the next phase of this study entailed the 1) isolation and computational analysis of the V-PPase promoter region and 2) putative correlation of regulatory data with gene expression profiles and grape ripening events.

Promoter isolation in grapevine, where the genomic sequence is limited, proved to be challenging during this study. From the different promoter isolation strategies used in plants, described in chapter 2, we used inverse PCR and genomic DNA library screening in parallel, which proved to be effective for obtaining a genomic sequence upstream from the gene in grapevine DNA (chapter 4). Long-range inverse PCR (LR-iPCR) revealed to be a rapid method for ‘jumping’ several hundred to thousands of base pairs from an original cDNA sequence. In this study, described in chapter 4, we have successfully isolated a 1567 base pair promoter sequence of the V-PPase gene designated as VPP. Additionally we identified the proximal promoter region (identified from sequence analysis) of a cDNA transcript (clone M8.A, chapter 3) that was highly similar to a ‘grape ripening induced protein, Grip 31’ previously identified by

Davies and Robinson (2000) using differential screening. Several attempts to obtain a larger part of the 'Grip 31' promoter region, using established as well as new combinations of methods, were unsuccessful. For this study LR-iPCR proved to be a rapid method with less labour intensiveness compared to genomic DNA library screening. However, because of the sensitivity and specificity of primer design, long range inverse PCR revealed to be a high-risk method yielding uncertain or no results when PCR conditions were not optimal.

In chapter 4 putative correlations and conclusions on the basis of integrating data regarding gene expression, computational promoter analysis and comparative analysis with previously described functions and cues of the V-PPase gene and grape ripening respectively were made. From the results and analysis discussed in chapter 4, it became apparent that this integrative strategy allowed for a more accurate explanation of regulatory events when expression patterns, specifically stress responsiveness of the V-PPase, coincided with motifs, putatively characterized when compared to plant *cis*-motif databases, within the VPP promoter region. Although not a replacement for experimental verification, we presented a more accurate understanding of regulatory control by correlating grape ripening events with known gene function (in this study V-PPase using PPi as an alternative energy source to maintain pH homeostasis) and transcriptional localization on tissue and cellular level. Additionally, transcriptional response to cold shock and drought stress revealed this V-PPase, gene (and promoter, chapter 4) as an essential target to be used for chill and drought tolerance in a grapevine-engineering program.

This study was based on an established strategy implementing a system for the large-scale identification of candidate genes expressed under a certain condition with the purpose to isolate their promoters that control the site and level of gene activity. This approach is effective but time-consuming, especially with a seasonal crop like grapevine, where the evaluation of specifically promoters depends on plants in the field bearing fruit. Other non-climacteric fruit systems such as the strawberry, with a shorter fruit-bearing cycle, are available for evaluation of

molecular ‘tools’ from grapevine and currently we are in the process of preparing the VPP promoter-reporter gene construct for transformation in strawberry plants. Alternative ‘reverse genetic’ strategies such as promoter trapping and chromatin immunoprecipitation assays (ChiPs) have been described for plants such as *Arabidopsis* and strawberry (chapter 2). Although not currently optimised for a wide variety of plant species, they allow for rapid identification of large sets of promoter sequences and could play an important part in future genetic engineering programs for *Vitis vinifera*, particularly for the isolation of fruit- and developmental stage-specific promoters.

Nevertheless, grapevine genetic enhancement is still at an early stage and commercial release of genetically engineered grapevine depends on field-trial evaluation that could range between 5 and 10 years. Therefore, the ability to integrate genetic data from genes and promoters to be used for transformation and evaluation could be considered more valuable at this early stage for the understanding of grapevine genetics (Kikkert *et al*. 2001). The biological activity of the VPP promoter was confirmed on a transient level during this study. Additionally we hypothesized that certain transcriptional events such as heat and cold shock have a conserved nature in plant and mammalian systems. Therefore, using the VPP promoter with putatively identified heat and cold shock elements, we attempted to evaluate and induce promoter activity (analysing reporter gene GFP) in *Xenopus laevis* oocytes with the use of microinjection but experiments were unsuccessful. Although analysis of promoters in mature fruit was not realistic for this project, stress responsive promoter induction can be evaluated in transformed grapevine calli. We are currently in the process of establishing a grapevine transformation system in our laboratory that would enable us to analyse stress induced promoter activity using reporter genes GFP or GUS in particle bombarded calli. Focused on the information-gaining strategy, we applied data from computational promoter analysis to induce fruit with stress treatments according to selected responsive regions putatively identified within the VPP promoter. From these investigations, although part of a more time-consuming strategy, we suggest that a more accurate conclusion can be drawn from utilizing regulatory data to understand transcriptional

events exhibited from cDNA-AFLPs, RNA gel blots and/or higher resolution systems such as *in situ* hybridisation (chapter 3 and 4).

The unprecedented advance of genetics, genomics and lately proteomics has yielded a myriad of new biological data and technologies. The genome sequence drafts of *Arabidopsis* and rice have provided researchers with a global map of genetic information in dicots and monocots for large-scale comparison, elucidation and integration. In this plant genetic revolution, it may seem that grapevine research have a 'less important' role to play when compared to 'model' organisms. However, intricacies regarding development, cultivar variation (Davies and Boss 2000) and of course all elements responsible for good wine, make grapevine an exciting model for non-climacteric as well as climacteric fruit, to study and manipulate. This project entailed the use of established as well as relatively new techniques and ideas to serve as a small 'stepping stone' in the burgeoning and exciting field of grapevine molecular biology.

LITERATURE CITED

Aharoni A, Vorst O. 2002. DNA microarrays for functional plant genomics. Plant Mol Biol 48: 99-118.

Aiyar A. 2000. The use of CLUSTAL W and CLUSTAL X for multiple sequence alignment. Methods Mol Biol 132: 221-241.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Nucleic Acids Res 215: 403-410.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.

Archer E. 1981. Rypwording en oesmetodes. In: Burger JD, Deist J, eds. Wingerdbou in Suid-Afrika. Cape Town: Maskew Miller. p 463-476. (Translation: Archer, E.; 1981: Ripening and Harvesting methods. Viticulture in South Africa.)

Arguello-Astorga G, Herrera-Estrella L. 1998. Evolution of light-regulated plant promoters. Annu Rev Plant Physiol Plant Mol Biol 49: 525-555.

Bachem CWB, Van der Hoeven RS, De Bruijn MS, Vreugdenhil D, Zabeau M, Visser RGF. 1996. Visualization of differential gene expression using a novel method of RNA fingerprinting based on AFLP: Analysis of gene expression during potato tuber development. Plant J 9: 745-753.

Bachem CWB, Oomen RJFJ, Visser RGF. 1998. Transcript imaging with cDNA-AFLP: A

step-by-step protocol. Plant Mol Biol Rep 16: 157-173.

Bailey TL, Elkan C. 1995. Unsupervised learning of multiple motifs in biopolymers using Expectation Maximization. Machine Learning 21: 51-80.

Bajic VB, Seah SH, Chong A, Zhang G, Koh JL, Brusic V. 2002. Dragon Promoter Finder: recognition of vertebrate RNA polymerase II promoters. Bioinformatics 18: 198-199.

Baker SS, Wilhelm KS, Thomashow MF. 1994. The 5'-region of Arabidopsis thaliana cor15a has *cis*-acting elements that confer cold-, drought- and ABA-regulated gene expression. Plant Mol Biol 24: 701-713.

Ballesteros E, Donaire JP, Belver A. 1996. Effects of salt stress on H(+)-ATPase and H(+)-PPase activities of tonoplast-enriched vesicles isolated from sunflower roots. Physiol Planta 97: 259-268.

Benfrey PN, Chua N-H. 1990. The Cauliflower Mosaic Virus 35S promoter: Combinatorial regulation of transcription in plants. Science 25: 959-966.

Boss PK, Davies C, Robinson SP. 1996. Expression of anthocyanin biosynthesis pathways genes in red and white grapes. Plant Mol Biol 32: 565-569.

Bower R, Elliott AR, Potier BAM, Birch RG. 1996. High-efficiency, microprojectile mediated co-transformation of sugarcane, using visible or selected markers. Mol Breed 2: 239-249.

Breyne P, Zabeau M. 2001. Genome-wide expression analysis of plant cell cycle modulated genes. Curr Opin Plant Biol 4: 136-142.

Burke TW, Kadonaga JT. 1996. Drosophila TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. Genes Dev 10: 711-724.

Busk PK, Pagès M. 1998. Regulation of abscisic acid-induced transcription. Plant Mol Biol 37: 425-435.

Carystinos GD, MacDonald HR, Monroy AF, Dhindsa RS, Poole RJ. 1995. Vacuolar H(+)-translocating pyrophosphatase is induced by anoxia or chilling in seedlings of rice. Plant Physiol 108: 641-649.

Chabouté ME, Clement B, Sekine M, Philipps G, Chaubet-Gigot N. 2000. Cell cycle regulation of the tobacco ribonucleotide reductase small subunit gene is mediated by E2F-like elements. Plant Cell 12: 1987-2000.

Chen W, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan S, Zou G, Whitham SA, Budworth PR, Tao Y, Xie Z, Chen X, Lam S, Kreps JA, Harper JF, Si-Ammour A, Mauch-Mani B, Heinlein M, Kobayashi K, Hohn T, Dangl JL, Wang X, Zhu T. 2002. Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell 14: 559-574.

Cho HT, Cosgrove DJ. 2002. Regulation of root hair initiation and expansin gene expression in *Arabidopsis*. Plant Cell 14: 3237-3253.

Christensen AH, Sharrock RA, Quail PH. 1992. Maize polyubiquitin genes: structure, thermal

perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mol Biol 18: 675-689.

Colova-Tsolova V, Perl A, Krastanova S, Tsvetkov I, Atanassov A. 2001. Genetically engineered grape for disease and stress tolerance. In: Roubelakis-Angelakis KA, ed. Molecular biology and biotechnology of the grapevine. Dordrecht: Kluwer Academic Publishers. p 411-432.

Coombe BG, Hale CR. 1973. The hormone content of ripening grape berries and the effects of growth substance treatments. Plant Physiol 51: 629-634.

Coombe BG. 1987. Distribution of solutes within the developing grape berry. Am J Enol Vitic 38: 120-126.

Coombe BG. 1992. Research on development and ripening of the grape berry. Am J Enol Vitic 43(1): 101-110.

Coombe BG, McCarthy MG. 2000. Dynamics of grape berry growth and physiology of ripening. Australian Journal of Grape and Wine Research 6: 131-135.

Darley CP, Davies JM, Sanders D. 1995. Chill-Induced Changes in the Activity and Abundance of the Vacuolar Proton-Pumping Pyrophosphatase from Mung Bean Hypocotyls. Plant Physiol 109: 659-665.

Davies C, Boss PK, Robinson SP. 1997. Treatment of grape berries, a nonclimacteric fruit with a synthetic auxin, retards and alters the expression of developmentally regulated genes. Plant Physiol 115: 1155-1161.

Davies C, Robinson SP. 2000. Differential screening indicates a dramatic change in mRNA profiles during grape berry ripening. Cloning and characterization of cDNAs encoding putative cell wall and stress response proteins. Plant Physiol 122: 803-812.

Davies C, Boss PK. 2000. The use of molecular techniques to study and manipulate the grapevine: Why and how? Australian Journal of Grape and Wine Research 6:159-167.

Davies JM. 1997. Vacuolar energization: pumps, shunts and stress. J Exp Bot 48(308): 633-641.

Donson J, Fang Y, Espiritu-Santo G, Xing W, Salazar A, Miyamoto S, Armendarez V, Volkmuth W. 2002. Comprehensive gene expression analysis by transcript profiling. Plant Mol Biol 48: 75-97.

Drozdowicz YM, Rea PA. 2001. Vacuolar H(+) pyrophosphatases: from the evolutionary backwaters into the mainstream. Trends Plant Sci 6: 206-211.

Edwards JW, Coruzzi GM. 1990. Cell specific gene expression in plants. Annu Rev Genet 24: 275-303.

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

Elmayan T, Tepfer M. 1995. Evaluation in tobacco of the organ specificity and strength of the rolD promoter, domain A of the 35S promoter and the 35S2 promoter. Transgenic Res 4: 388-396.

Eulgem T, Rushton PJ, Robatzek S, Somssich IE. 2000. The WRKY superfamily of plant transcription factors. Trends Plant Sci 5: 199-206.

Famiani F, Walker RP, Tecsi L, Chen ZH, Proietti P, Leegood RC. 2000. An immunohistochemical study of the compartmentation of metabolism during the development of grape (Vitis vinifera L.) berries. J Exp Bot 51: 675-683.

Farmahan HL, Pandey RM. 1976. Hormonal regulation of the lag phase in seeded and seedless grapes (*Vitis vinifera* L.). Vitis 15: 227-235.

Fessele S, Maier H, Zischek C, Nelson PJ, Werner T. 2002. Regulatory context is a crucial part of gene function. Trends Genet 18: 60-63.

Fillion L, Ageorges A, Picaud S, Coutos-Thévenot P, Lemoine R, Romieu C, Delrot S. 1999. Cloning and expression of a Hexose transporter gene expressed during the ripening of grape berry. Plant Physiol 120: 1083-1093.

Fowler S, Thomashow MF. 2002. Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14: 1675-1690.

Gantet P, Memelink J. 2002. Transcription factors: tools to engineer the production of pharmacologically active plant metabolites. Trends Pharmacol Sci 23: 563-569.

Giovannoni J. 2001. Molecular biology of fruit maturation and ripening. Annu Rev Plant Physiol Plant Mol Biol 52: 725-749.

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller

P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D , Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296: 92-100.

Gollop R, Even S, Colova-Tsolova V, Perl A. 2002. Expression of the grape dihydroflavonol reductase gene and analysis of its promoter region. J Exp Bot 53: 1397-1409.

Gomez-Cadenas A, Zentella R, Sutliff TD, Ho TH. 2001. Involvement of multiple *cis*-elements in the regulation of GA responsive promoters: Definition of a new *cis*-element in the Amy32b gene promoter of barley (*Hordeum vulgare*). Physiol Planta 112: 211-216.

Grierson C, Du JS, de Torres ZM, Beggs K, Smith C, Holdsworth M, Bevan M. 1994. Separate cis sequences and trans factors direct metabolic and developmental regulation of a potato tuber storage protein gene. Plant J 5: 815-826.

Guilfoyle T, Hagen G, Ulmasov T, Murfett J. 1998. How does auxin turn on genes? Plant Physiol 118: 341-347.

Guilfoyle TJ. 1997. The structure of plant gene promoters. In: Setlow, JK, ed. Genetic Engineering. New York, Plenum Press. (19), p 15-47.

Guo H, Moose SP. 2003. Conserved noncoding sequences among cultivated cereal

genomes identify candidate regulatory sequence elements and patterns of promoter evolution. Plant Cell 15: 1143-1158.

Habu Y, Fukada-Tanaka S, Hisatomi Y, Iida S. 1997. Amplified restriction fragment length polymorphism-based mRNA fingerprinting using a single restriction enzyme that recognizes a 4-bp sequence. Biochem Biophys Res Comm 234: 516-521.

Hehl R, Wingender E. 2001. Database-assisted promoter analysis. Trends Plant Sci 6: 251-255.

Higgins D, Thompson J, Gibson T, Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673-4680.

Higo K, Ugawa Y, Iwamoto M, Higo H. 1998. PLACE: a database of plant *cis*-acting regulatory DNA elements. Nucleic Acids Res 26: 358-359.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res 27:297-300.

Holtorf S, Apel K, Bohlmann H. 1995. Comparison of different constitutive and inducible promoters for the overexpression of transgenes in Arabidopsis thaliana. Plant Mol Biol 29: 637-646.

Huerta AM, Glasner JD, Gutierrez-Rios RM, Blattner FR, Collado-Vides J. 2002. GETools: gene expression tool for analysis of transcriptome experiments in *E.coli*. Trends Genet 18: 217-218.

Hughes JD, Estep PW, Tavazoie S, Church GM. 2000. Computational identification of *cis*-regulatory elements associated with groups of functionally related genes in *Saccharomyces cerevisiae*. J Mol Biol 296: 1205-1214.

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

Jefferson RA, Kavanagh TA, Caboche M. 1987. GUS fusions: ß-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J 6: 3901.

Jiang C, Iu B, Singh J. 1996. Requirement of a CCGAC *cis*-acting element for cold induction of the BN115 gene from winter *Brassica napus*. Plant Mol Biol 30: 679-684.

Kay R, Chan A, Daly M, McPherson J. 1987. Duplication of the CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236: 1299-1302.

Kikkert JR, Thomas MR, Reisch BI. 2001. Grapevine genetic engineering. In: Roubelakis-Angelakis KA, ed. Molecular biology and biotechnology of the grapevine. Dordrecht: Kluwer Academic Publishers. p 393-410.

Kim HJ, Kim YK, Park JY, Kim J. 2002. Light signalling mediated by phytochrome plays an important role in cold-induced gene expression through the C-repeat/dehydration responsive element (C/DRE) in *Arabidopsis thaliana*. Plant J 29: 693-704.

Kim EJ, Zhen RG, Rea PA. 1994. Heterologous expression of plant vacuolar pyrophosphatase in yeast demonstrates sufficiency of the substrate-binding subunit for proton transport. Proc Natl Acad Sci U S A 91:6128-6132.

Klok EJ, Wilson IW, Wilson D, Chapman SC, Ewing RM, Somerville SC, Peacock WJ, Dolferus R, Dennis ES. 2002. Expression profile analysis of the low-oxygen response in *Arabidopsis* root cultures. Plant Cell 14: 2481-2494.

Knudsen S. 1999. Promoter2.0: for the recognition of Pol II promoter. Bioinformatics 15: 356-361.

Koch MA, Weisshaar B, Kroymann J, Haubold B, Mithell-Olds T. 2001. Comparative genomics and regulaory evolution: Conservation and function of the *Chs* and *Apetala3* promoters. Mol Biol Evolution 18(10): 1882-1891.

Kuhn E. 2001. From library screening to microarray technology: Strategies to determine gene expression profiles and to identify differentially regulated genes in plants. Annals of Botany 87: 139-155.

Lawrence CE, Altschul SF, Boguski MS, Liu JS, Neuwald AF, Wootton JC. 1993. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. Science 262: 208-214.

Le Gourrierec J, Li Y-F, Zhou D-X. 1999. Transcriptional activation by *Arabidopsis* GT-1 may be through interaction with TFIIA-TBP-TATA complex. Plant J 18: 663-668.

Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, Hannett NM, Harbison CT, Thompson CM, Simon I, Zeitlinger J, Jennings EG, Murray HL, Gordon DB, Ren B, Wyrick JJ, Tagne JB, Volkert TL, Fraenkel E, Gifford DK, Young RA. 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. Science 298: 799-804.

Lerchl J, König S, Zrenner R, Sonnewald U. 1995. Molecular cloning, characterization and expression analysis of isoforms encoding tonoplast-bound proton-translocating inorganic pyrophosphatase in tobacco. Plant Mol Biol 29: 833-840.

Li Z, Jayasankar S, Gray DJ. 2001. Expression of a bifunctional green fluorescent protein (GFP) fusion marker under the control of three constitutive promoters and enhanced derivatives in transgenic grape (*Vitis vinifera*). Plant Sci 160: 877-887.

Liang P, Pardee AB. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-970.

Maeshima M. 1991. H+-translocating inorganic pyrophophatase of plant vacuoles. FEBS Lett 196: 11-17.

Maeshima M. 2000. Vacuolar H(+)-pyrophosphatase. Biochim Biophys Acta 1465: 37-51.

Manjunath S, Sachs MM. 1997. Molecular characterization and promoter analysis of the maize cytosolic glyceraldehyde 3-phosphate dehydrogenase gene family and its expression during anoxia. Plant Mol Biol 33: 97-112.

Martinez E. 2002. Multi-protein complexes in eukaryotic gene transcription. Plant Mol Biol 50: 925-947.

Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, Dubchak I. 2000. VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics 16(11): 1046-1047.

Meissner R, Chague V, Zhu Q, Emmanuel E, Elkind Y, Levy AA. 2000. Technical advance: a high throughput system for transposon tagging and promoter trapping in tomato. Plant J 22: 265-274.

Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM. 2002. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics 18: 333-334.

Miller AJ, Short SR, Hiratsuka K, Chua N-H, Kay SA. 1992. Firefly luciferase as a reporter of regulated gene expression in higher plants. Plant Mol Biol Rep 10: 345-361.

Milner ID, Ho LC, Hall JL. 1995. Properties of proton and sugar transport at the tonoplast of tomato (*Lycopersicon esculentum*) fruit. Physiol Planta 94: 399-410.

Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. 2002. Network motifs: simple building blocks of complex networks. Science 298: 824-827.

Morita A, Umemura T, Kuroyanagi M, Futsuhara Y, Perata P, Yamaguchi J. 1998. Functional dissection of a sugar-repressed alpha-amylase gene (RAmy1 A) promoter in rice embryos. FEBS Lett 423: 81-85.

Moseyko N, Zhu T, Chang H-S, Wang X, Feldman LJ. 2002. Transcription profiling of the early gravitropic response in *Arabidopsis* using high-density oligonucleotide probe microarrays. Plant Physiol 130: 720-728.

Müller M, Irkens-Kiesecker U, Rubinstein B, Taiz L. 1996. On the mechanism of hyperacidification in lemon. Comparison of the vacuolar H(+)-ATPase activities of fruits and epicotyls. J Biol Chem 271:1916-1924.

Murashige T, Skoog F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Planta 15: 473-497.

Nakamura M, Tsunoda T, Obokata J. 2002. Photosynthesis nuclear genes generally lack TATA-boxes: a tobacco photosystem I gene responds to light through an initiator. Plant J 29: 1-10.

Nakanishi Y, Maeshima M. 1998. Molecular cloning of vacuolar H(+)-pyrophosphatase and its developmental expression in growing hypocotyl of mung bean. Plant Physiol 116: 589-597.

Novina CD, Roy AL. 1996. Core promoters and transcriptional control. Trends Genet 12(9): 351-355.

Nunan KJ, Davies C, Robinson SP, Fincher GB. 2001. Expression patterns of cell wall-modifying enzymes during grape berry development. Planta 214: 257-264.

Ochman H, Gerber AS, Hartl DL. 1988. Genetic applications of an inverse polymerase chain reaction. Genetics 120: 621-623.

Ohler U, Niemann H. 2001. Identification and analysis of eukaryotic promoters: recent computational approaches. Trends Genet 17:56-60.

Pan S, Czarnecka-Verner E, Gurley WB. 2000. Role of the TATA binding protein-transcription factor IIB interaction in supporting basal and activated transcription in plant cells. Plant Cell 12: 125-136.

Pandey RM, Farmahan HL. 1977. Changes in the rate of photosynthesis and respiration in leaves and berries of *Vitis vinifera* grapevines at various stages of berry development. Vitis 16: 106-111.

Park PJ, Butte AJ, Kohane IS. 2002. Comparing expression profiles of genes with similar promoter regions. Bioinformatics 18: 1576-1584.

Peng HP, Chan CS, Shih MC, Yang SF. 2001. Signaling events in the hypoxic induction of alcohol dehydrogenase gene in *Arabidopsis*. Plant Physiol 126 :742-749.

Périer RC, Praz V, Junier T, Bonnard C, Bucher P. 2000. The eukaryotic promoter database (EPD). Nucleic Acids Res 28: 302-303.

Pilpel Y, Sudarsanam P, Church GM. 2001. Identifying regulatory networks by combinatorial analysis of promoter elements. Nat Genet 29: 153-159.

Pryer KM, Schneider H, Zimmer EA, Ann BJ. 2002. Deciding among green plants for whole genome studies. Trends Plant Sci 7: 550-554.

Puente P, Wei N, Deng XW. 1996. Combinatorial interplay of promoter elements constitutes the minimal determinants for light and developmental control of gene expression in *Arabidopsis*. EMBO J 15: 3732-3743.

Qin L, Prins P, Jones JT, Popeijus H, Smant G, Bakker J, Helder J. 2001. GenEST, a powerful bidirectional link between cDNA sequence data and gene expression profiles generated by cDNA-AFLP. Nucleic Acids Res 29: 1616-1622.

Ratajczak R. 2000. Structure, function and regulation of the plant vacuolar H(+)-

translocating ATPase. Biochim Biophys Acta 1465: 17-36.

Rea PA, Poole RJ. 1993. Vavuolar H^+-translocating pyrophosphatase . Annu.Rev.Plant Physiol Plant Mol.Biol. 44: 157-180.

Reese,M.G. and Eeckman,F.H. 1995. Novel Neural Network Algorithms for Improved Eukaryotic Promoter Site Recognition. Accepted talk for The seventh international Genome sequencing and analysis conference, Hyatt Regency, Hilton Head Island, South Carolina September 16-20.

Ren B, Cam H, Takahashi Y, Volkert T, Terragni J, Young RA, Dynlacht BD . 2002. E2F integrates cell cycle progression with DNA repair, replication, and.G2/M-checkpoints. Genes Dev 16: 245-256.

Rezaian MA, Krake LR. 1987. Nucleic acid extraction and virus detection in grapevine. J Vir Methods 17: 277-285.

Riechmann JL, Ratcliffe OJ. 2000. A genomic perspective on plant transcription factors. Curr Opin Plant Biol 3: 423-434.

Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ, Ghandehari D, Sherman BK, Yu G. 2000. *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. Science 290: 2105-2110.

Rizhsky L, Liang H, Mittler R. 2002. The combined effect of drought stress and heat shock on gene expression in tobacco. Plant Physiol 130: 1143-1151.

Robinson SP, Davies C. 2000. Molecular biology of grape berry ripening. Australian Journal of Grape and Wine Research 6: 175-188.

Roeder RG. 1996. The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem Sci 21: 327-335.

Rombauts S, Déhais P, Van Montagu M, Rouzé P. 1999. PlantCARE, a plant *cis*-acting regulatory element database. Nucleic Acids Res 27: 295-296.

Roubelakis-Angelakis KA. 2001. Molecular biology and biotechnology of the grapevine. 1st ed. Dordrecht: Kluwer academic publishers. 474 p.

Rushton PJ, Reinstädler A, Lipka V, Lippok B, Somssich IE. 2002. Synthetic plant promoters containing defined regulatory elements provide novel insights into the pathogen- and wound-induced signaling. Plant Cell 14: 749-762.

Sagerström CG, Sun BI, Sive HL. 1997. Subtractive cloning: past, present and future. Annu Rev Biochem 66: 751-783.

Sakakibara Y, Kobayashi H, Kasamo K. 1996. Isolation and characterization of cDNAs encoding vacuolar H^+-pyrophosphatase isoforms from rice (*Oryza sativa* L.). Plant Mol Biol 31: 1029-1038.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning: A laboratory manual. 2nd ed. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Schenk PM, Remans T, Sagi L, Elliott AR, Dietzgen RG, Swennen R, Ebert PR, Grof CP,

Manners JM. 2001. Promoters for pregenomic RNA of banana streak badnavirus are active for transgene expression in monocot and dicot plants. Plant Mol Biol 47: 399-412.

Scherf M, Klingenhoff A, Werner T. 2000. Highly specific localization of promoter regions in large genomic sequences by PromoterInspector: a novel context analysis approach. J Mol Biol 297: 599-606.

Schoffl F, Prandl R, Reindl A. 1998. Regulation of the heat-shock response. Plant Physiol 117: 1135-1141.

Schwechheimer C, Zourelidou M, Bevan MW. 1998. Plant transcription factor studies. Annu.Rev.Plant Physiol Plant Mol.Biol. 49: 127-150.

Seki M, Narusaka M, Kamiya A, Ishida J, Satou M, Sakurai T, Nakajima M, Enju A, Akiyama K, Oono Y, Muramatsu M, Hayashizaki Y, Kawai J, Carninci P, Itoh M, Ishii Y, Arakawa T, Shibata K, Shinagawa A, Shinozaki K. 2002. Functional annotation of a full-length *Arabidopsis* cDNA collection. Science 296: 141-145.

Sessions A, Yanofsky MF, Weigel D. 2000. Cell-cell signaling and movement by the floral transcription factors LEAFY and APETALA1. Science 289: 779-782.

Shen-Orr SS, Milo R, Mangan S, Alon U. 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. Nat Genet 31: 64-68.

Shinozaki K, Yamaguchi-Shinozaki K. 2000. Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. Curr Opin Plant Biol 3: 217-223.

Simon I, Barnett JHN, Harbison CT, Rinaldi NJ, Volkert TL, Wyrick JJ, Zeitlinger J, Gifford DK, JaakkolaT.S., Young RA. 2001. Serial regulation of transcriptional regulators in the yeast cell cycle. Cell 106: 697-708.

Singh KB. 1998. Transcriptional regulation in plants: The importance of combinatorial control. Plant Physiol 118: 1111-1120.

Singh R, Cheah S-C. 2000. Differential gene expression during flowering in the oil palm (*Elaeis guieensis*). Plant Cell Rep 19: 804-809.

Sparvoli F, Martin C, Scienza A, Gavazzi G, Tonelli C. 1994. Cloning and molecular analysis of structural genes involved in flavonoid and stilbene biosynthesis in grape (*Vitis vinifera* L.). Plant Mol Biol 24: 743-755.

Steenkamp J, Wiid I, Lourens A, Van Helden P. 1994. Improved method for DNA extraction from *Vitis vinifera*. Am J Enol Vitic 45: 102-106.

Suzuki Y, Maeshima M, Yamaki S. 1999. Molecular cloning of vacuolar H(+)-pyrophosphatase and its expression during the development of pear fruit. Plant Cell Physiol 40: 900-904.

Taiz L, Zeiger E. 1998. Solute Transport. In: Taiz L, Zeiger E, eds. Plant Physiology. Sunderland, Massachusetts: Sinauer Associates, Inc., Publishers. p 125-152.

Tanaka Y, Chiba K, Maeda M, Maeshima M. 1993. Molecular cloning of cDNA for vacuolar membrane proton-translocating inorganic pyrophosphatase in *Hordeum vulgare*. Biochem Biophys Res Commun 190: 1110-1114.

Teakle GR, Manfield IW, Graham JF, Gilmartin PM. 2002. *Arabidopsis thaliana* GATA factors: organisation, expression and DNA- binding characteristics. Plant Mol Biol 50: 43-57.

Terrier N, Deguilloux C, Sauvage FX, Martnoia E, Romieu C. 1998. Proton pumps and anion transport in *Vitis vinifera*: The inorganic pyrophosphatase plays a predominant role in the energization of the tonoplast. Plant Physiol Biochem 36(5): 367-377.

Terrier N, Romieu C. 2001. Grape berry acidity. In: Roubelakis-Angelakis KA, ed. Molecular biology and biotechnology of the grapevine. Dordrecht: Kluwer Academic Publishers. p 35-57.

Terrier N, Ageorges A, Abbal P, Romieu C. 2001. Generation of ESTs from grape berry at various developmental stages. J Plant Physiol 158: 1575-1583.

Terrier N, Sauvage FX, Ageorges A, Romieu C. 2001. Changes in the acidity and in proton transport at the tonoplast of grape berries during development. Planta 213: 20-28.

The Arabidopsis genome initiative A. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815.

Thijs G, Lescot M, Marchal K, Rombauts S, De Moor B, Rouze P, Moreau Y. 2001. A higher-order background model improves the detection of promoter regulatory elements by Gibbs sampling. Bioinformatics 17: 1113-1122.

Trindade LM, Horvath B, Bachem C, Jacobsen E, Visser RGF. 2003. Isolation and functional

characterization of a stolon specific promoter from potato (*Solanum tuberosum* L.). Gene 303: 77-87.

Tsukaya H, Ohshima T, Naito S, Chino M, Komeda Y. 1991. Sugar-dependant expression of the CHS-A gene for chalcone synthase from *Petunia* in transgenic *Arabidopsis*. Plant Physiol 97: 1414-1421.

Ulmasov T, Liu ZB, Hagen G, Guilfoyle TJ. 1995. Composite structure of auxin response elements. Plant Cell 7: 1611-1623.

Venter M, Burger AL, Botha FC. 2001. Identification of genes differentially expressed during berry ripening in *Vitis vinifera* cv. Chardonnay. Vitis 40(4): 191-196.

Vivier MA, Pretorius IS. 2000. Genetic improvement of grapevine: Tailoring grape varietiesfor the third millennium - A review. S Afr J Enol Vitic 21: 5-26.

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res 23: 4407-4414.

Wang H, Tang W, Zhu C, Perry SE. 2002. A chromatin immunoprecipitation (ChIP) approach to isolate genes regulated by AGL15, a MADS domain protein that preferentially accumulates in embryos. Plant J 32: 831-843.

Weinmann AS, Yan PS, Oberley MJ, Huang TH-M, Farnham PJ. 2002. Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. Genes Dev 16: 235-244.

Whittaker A, Botha FC. 1997. Carbon partitioning during sucrose accumulation in sugarcane internodal tissue. Plant Physiol 115: 1651-1659.

Wingender E, Chen X, Hehl R, Karas H, Liebich I, Matys V, Meinhardt T, Pruss M, Reuter I, Schacherer F. 2000. TRANSFAC: an integrated system for gene expression regulation. Nucleic Acids Res 28: 316-319.

Wyrick JJ, Young RA. 2002. Deciphering gene expression regulatory networks. Curr Opin Genet Dev 12: 130-136.

Yanagisawa S, Schmidt RJ. 1999. Diversity and similarity among recognition sequences of Dof transcription factors. Plant J 17: 209-214.

Yanagisawa S. 2002. The Dof family of plant transcription factors. Trends Plant Sci 7: 555-560.

Yoshida K, Shinmyo A. 2000. Transgene expression systems in plant, a natural bioreactor. J Biosci Bioeng 90(4): 353-362.

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W,

Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296: 79-92.

Zeng Y, Wu Y, Avigne WT, Koch KE. 1998. Differential regulation of sugar-sensitive sucrose synthases by hypoxia and anoxia indicate complementary transcriptional and posttranscriptional responses. Plant Physiol 116: 1573-1583.

Zhang H, Zhang R, Liang P. 1996. Differential screening of gene expression difference enriched by differential display. Nucleic Acids Res 24: 2454-2455.

Zhou DX. 1999. Regulatory mechanism of plant gene transcription by GT-elements and GT-factors. Trends Plant Sci 4: 210-214.

Zhu Z, Pilpel Y, Church GM. 2002. Computational identification of transcription factor binding sites via a transcription-factor-centric clustering (TFCC) algorithm. J Mol Biol 318: 71-81.