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

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

University of Cape Town

CLONING AND CHARACTERISATION OF LEA1-EM GENES IN THE RESURRECTION PLANT *XEROPHYTA HUMILIS*

by

Nqobile A. C. Ngubane

University of Cape Town

A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Molecular and Cell Biology, University of Cape Town

December 2008

TABLE OF CONTENTS i

List of Tables iv

List of Figures v

Acknowledgements vii

Declaration viii

Abstract ix

Abbreviations x

1. Introduction and Background 1

1.1. Chapter Overview 1

1.2. Desiccation Tolerance 1

1.3. Resurrection Plants 2

1.4. Mechanisms of Desiccation Tolerance 3

1.4.1. Antioxidants that counter damage by reactive oxygen species 3

1.4.2. Sugars that replace water and form glasses 4

1.4.3. Proteins that stabilize macromolecules and membranes, LEA proteins. 5

1.5. Summary and Aims 12

2 Isolation and cloning of LEA1-Em genes from *X. humilis* 13

2.1 Introduction 13

2.2 Materials and Method 15

2.2.1 Plant Material and Growth Conditions 15

2.2.2 RNA Extraction for LEA1-Em isolation 16

2.2.3 cDNA synthesis 16

2.2.4 Degenerate Primer Design 16

2.2.5 Degenerate Primer PCR 17

2.2.6 Cloning 17

2.2.7 Screening and sequencing 18

2.2.8 5'and 3' Rapid amplification of cDNA ends (RACE) gene specific primer (GSP) design 18

2.2.9 5'and 3' Rapid Amplification of cDNA Ends (RACE) PCR 19

2.2.9 Primer Design for cloning of full length Open Reading Frame (ORF) sequences ... 21

University of Cape Town

2.3 Results 22
2.3.1 Degenerate Primer Design 22
2.3.2 Optimization of degenerative PCR and cloning 24
2.3.3 5' and 3' RACE Cloning 27
2.3.4 XHLEA1-1 Sequence 33
2.3.4 XHLEA1-2 Sequence 35
2.3.5 XHLEA1-4 Sequence 37
2.3.6 XHC00797a Sequence 39
2.3.7 XHC00797b Sequence 40
2.4 Discussion 42
2.4.1 Sequence Analysis 42
2.4.2 Alternative splicing of mRNA transcript XHC00797a and XHC00797b 44
2.5 Conclusion 45

3 Changes in mRNA transcript abundance of *X. humilis* LEA1-Em during dehydration and rehydration 46

3.1 Introduction 46
3.1.1 Quality and measurement of the RNA and starting cDNA 46
3.1.2 Inhibitory components 47
3.1.3 The reverse transcriptase reaction 48
3.1.4 Optimising the PCR reaction 48
3.1.5 Normalisation of gene expression 48
3.2 Materials and Methods 49
3.2.1 Plant material and growth conditions 49
3.2.2 Abiotic treatment of plants 50
3.2.3 RNA Extraction, RNA quantification and reverse transcription 50
3.2.4 Primer design, primer concentration optimisation 51
3.2.5 Real-time (RT)-PCR 52
3.2.6 Optimisation of RT-PCR 52
3.2.7 Data analysis 52
3.3 Results 53
3.3.1 RNA integrity 53

University of Cape Town

3.3.2 PCR Primer Design 55
3.3.3 PCR Primer Optimisation 57
3.3.4 RT-PCR for the reference gene R1 59
3.3.5 Expression of *X.humilis* XHLEA1-1 in seed, leaf and root during the abiotic stress treatment. 61
3.3.6 Expression of *X.humilis* XHLEA1-4 in seed and in leaf and root during the abiotic stress treatment. 63
3.3.7 Expression of *X.humilis* XHC00797a RNA from seed, leaf and root during the dehydration and rehydration cycle. 66
3.4 Discussion 70
3.5 Conclusion 72
4 General Conclusions 73
5 References 75

University of Cape Town

List of Tables

Table 1.1: Classification, properties and proposed roles of LEA proteins, reported in various sources. This table is taken from Berkjak *et al.*, 2007 8

Table 2.1: List of LEA1-EM genes 13

Table 2.2: List of primers used for 5' and 3' RACE PCR 19

Table 2.3: The controls used for 5' and 3' RACE PCR 20

Table 2.4: List of primers used for full length amplification of the Open Reading Frame ... 21

Table 2.5: Degenerate primers used for isolation of LEA1-Em from *X. humilis*. 22

Table 2.6: Summary of RACE PCR sequencing results. 33

Table 3.1: Primer sequences for real-time PCR 51

Table 3.2: RNA samples with their RNA Intergrity numbers (RIN) 53

Table 3.3: Primer specificity check between *X. humilis* LEA1-Em genes. 56

University of Cape Town

List of Figures

Figure 1.1: Evolutionary hypothesis of desiccation tolerance in vegetative tissue. 2

Figure 1.2: Summary of the predicted functions of LEAs. 7

Figure 2.1: Multiple sequence alignments of selected LEA1-Em's, listed in Table 2.1. 23

Figure 2.2: Amplification of LEA1-Em from *X. humilis* cDNA using degenerate primers. .. 25

Figure 2.3: Amplification of LEA1-Em from *X. humilis* genomic DNA using degenerate primers. 26

Figure 2.4: Multiple alignment of the sequences isolated using degenerate primers............. 27

Figure 2.5: Agarose gel analysis of (A) 3'RACE PCR with root cDNA as template and (B) 5'RACE PCR with root cDNA as template.. 28

Figure 2.6: Agarose gel analysis of RACE-PCR products from leaf and seed cDNA template. 29

Figure 2.7: Partial sequence of XHC00797 gene identified from microarray study on *X.humils* 30

Figure 2.8: The 5' RACE amplification of XHC00797 gene from leaf, root and seed cDNA. 30

Figure 2.9 : The 3' RACE amplification of XHC00797 gene from leaf, root and seed cDNA 31

Figure 2.10: Alignment of XHC00797 partial sequence with 5'RACE PCR products 32

Figure 2.11: Cloning of XHLEA1-1 full length open reading frame (ORF) sequence using 3' RACE ready cDNA. 34

Figure 2.12: Primary structure of the full length XHLEA1-1 gene. 34

Figure 2.13: Cloning of XHLEA1-2 full length sequences using gene specific primers FLXHM6gsp1 and FLXHM6gsp1 (Table 2.4). 35

Figure 2.14: Primary structure of the full length XHLEA1-2 gene. 36

Figure 2.15: Cloning of XHLEA1-4 full length ORF sequence using 3' RACE ready cDNA and the FLXHB1gsp1 primer. 37

Figure 2.16: Primary structure of the full length XHLEA1-4 gene. 38

Figure 2.17: (A) Primary structure of the full length XHC00979a, and (B) amino acid sequence alignment of the two motifs. 39

Figure 2.18: Primary sequence of the full length XHC00797a gene. 40

University of Cape Town

Figure 2.19: Primary structure of the full length XHC00797b gene. 41

Figure 2.20: Sequence homology tree of LEA1-Em sequences from *X. humilis*. 42

Figure 2.21: Alignment of the five LEA1-Em amino acid sequences isolated from *X. humils*. 43

Figure 3.1: Electropherogram from RNA of root and seed samples 54

Figure 3.2: Electropherogram from RNA of leaf samples 55

Figure 3.3: RT-PCR Primer specific for real time quantification of XHC00797a 57

Figure 3.4: Specificity of RT-PCR primers 58

Figure 3.5 (A) Melting curve of XHLEA1-4 at 200nM primer concentration (B) Melting curve of XHLEA1-4 at 400nM primer concentration 59

Figure 3.6: (A) Melting curve of housekeeping gene (R1) using undiluted samples (B) Melting curve of housekeeping gene (R1) using diluted samples 60

Figure 3.7: (A) Standard curve of housekeeping gene (R1) using undiluted samples (B) Standard curve of housekeeping gene (R1) using diluted samples 61

Figure 3.8: Amplification of XHLEA1-1 gene from pooled cDNA samples. 62

Figure 3.9: Amplification of XHLEA1-1 from tissue of different RWC cDNA samples. 63

Figure 3.10: Melting curve for XHLEA1-4 mRNA quantification 64

Figure 3.11: Standard curve of XHLEA1-4 using diluted samples 64

Figure 3.12 (A) Expression of XHLEA1-4 normalized to starting RNA (B) Expression of XHLEA1-4 normalized to housekeeping gene (R1) 65

Figure 3.13 (A) Expression of XHLEA1-4 in root tissue normalized to starting RNA (B) Expression of XHLEA1-4 in root tissue normalized to housekeeping gene (R1) 66

Figure 3.14 (A) Melting curve for XHC00797a using undiluted samples (B) Melting curve for XHC00797a using diluted samples 67

Figure 3.15: (A) Standard curve for XHC00797a using undiluted samples (B) Standard curve for XHC00797a using diluted samples 68

Figure 3.16 (A) Expression for XHC00797a normalized to starting RNA (B) Expression for XHC00797a normalized to housekeeping gene (R1) 69

University of Cape Town

Acknowledgements

I would like to thank my supervisors; Prof. Jill Farrant and Assoc. Prof. Nicola Illing, for their guidance, encouragement and support throughout my thesis work. I also would like to thank respective members of their groups, past and present, for scientific inputs and making my stay a fulfilling one.

I am indebted to Prof Jill Farrant, in her capacity as head of department, for giving me the opportunity to study in MCB, at the University of Cape Town. My special gratitude goes to Inonge Mulako for her cheerful assistance when I joined the lab, Pei-yin Ma and Di James for their DNA primer and sequencing services, respectively.

On a personal note, I wish to thank God for His presence in my life. I express my utmost gratitude to my family: Dad, mum, Baso, Mlondi, Nhlanhla, Kiki and Siya. I thank you for all your prayers, support and guidance. I especially thank my immediate friends Nancy, Thoko, Tshego, Nelani for their unfailing belief in me.

For financial support I would like to thank the Thabo Mbeki Leadership in Africa Scholarship and the National Research Fund Scarce Skills. I wish to acknowledge project funding from the National Research Foundation (SA).

University of Cape Town

Declaration

I declare that the dissertation being submitted is my own unaided work, apart from the normal guidance of my supervisors. It has not been submitted in the past for any degree or examination at this or any other University

I am presenting the thesis for examination for the Degree of Master of Science.

Signed by ___________________________

___________day of____________________2008

.

University of Cape Town

Abstract

The presence and expression patterns of orthologues of LEA group 1 genes has been characterised in the resurrection plant, *Xerophyta humilis*. The group I LEAs (Em1 and Em6) were first identified as proteins that were abundantly and specifically expressed during the desiccation and germination phase of angiosperm seed development. The group I LEA genes are characterised by the presence of one or more tandemly repeated 20-amino acid motifs that are particularly rich in Gly residues. In Arabidopsis, the group I LEA genes (AtEM1 and AtEM6) have been used as a model to study the regulation of gene expression by ABA (abscisic acid) during seed development. Phenotypic analysis of AtEM6 T-DNA insertion mutants has shown that AtEM6 plays a role in buffering the rate of dehydration during the later stages of seed maturation, but does not otherwise affect plant development. Since the LEA1 genes are specifically associated with seed maturation and not abiotic stress responses in desiccation sensitive plants, we have targeted the characterisation of *X. humilis* LEA1 genes, to test the hypothesis that evolution of desiccation tolerance in *X. humilis* is a consequence of activation of seed specific genes in vegetative tissue. Degenerate PCR primers designed to conserved regions of LEA1 genes were used to amplify five LEA1 orthologues from cDNA prepared from *X. humilis* desiccated seed, root and leaves. The full-length cDNAs of these orthologues was cloned by 5′ and 3′ RACE PCR. The three *X. humilis* LEA1 orthologues XhLEA1-1, XhLEA1-2 and XhLEA1-4 respectively have one, two and four of the 20 amino acid motif repeats. A fourth LEA1 orthologue, XHC00797a was identified in a microarray screen for mRNA transcripts that are up-regulated during desiccation in *X. humilis* leaves. Its isoform XHC00797b was later isolated from leaf tissue. Two LEAs, XHLEA1-4 and XHC00797a were used for mRNA expression analysis using quantitative RT-PCR. Both XHLEA1-4 and XHC00797a mRNA transcripts were significantly induced by desiccation stress. The findings from this study established a link between the LEA1-EM, previously characterized as seed specific, and desiccation tolerance in vegetative tissue of resurrection plant *X. humilis*. This opens exciting new perspectives for identification of genes previously associated with acquisition of desiccation tolerance in seeds, as candidates for acquisition of desiccation tolerance of vegetative tissue.

University of Cape Town

Abbreviations

ABA	abscisic acid
bp	base pair(s)
cDNA	copy DNA
DEPC	diethylpyrocarbonate
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
EDTA	ethylenediaminetetra-acetic acid
g	grams
h	hour
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase(s)
kDa	kilodalton(s)
LEA	late embryogenis abundant
l	litre(s)
LB	Luria-Bertani
M	molar concentration
μg	microgram(s)
mg	milligram(s)
μl	microlitre(s)
ml	millilitre(s)
mM	millimolar
mRNA	messenger RNA
ORF	open reading frame
PCR	polymerase chain reaction
RNA	ribonucleic acid
ROS	reactive oxygen species
RWC	relative water content
U	unit(s) of enzymatic activity
UTR	untranslated region

University of Cape Town

1. Introduction and Background

1.1. Chapter Overview

This chapter will introduce the current understanding of desiccation tolerance in relation to resurrection plants. First, a general introduction and a hypothesis on the evolution of desiccation tolerance (DT) is presented. This is followed by an introduction to resurrection plants and mechanisms of desiccation tolerance. In particular, late embryogenesis abundant (LEA) proteins are discussed in terms of their classification, structure and proposed functions. Experimental findings are reviewed with a view to give perspective to the data presented in subsequent chapters of this dissertation. LEA proteins are putative targets in the understanding of desiccation tolerance, and the one of their groups, namely LEA1-Em form the basis of this thesis. Finally, the aims of the thesis are mentioned as a guide for subsequent chapters.

1.2. Desiccation Tolerance

Although the vegetative tissue of most organisms cannot survive severe water deficit stress, the seeds of most plants (termed orthodox) and vegetative tissues of a small group of plants (termed resurrection plants) have developed mechanisms to survive it. Such mechanisms enable cells to resume normal cellular function and activities upon rehydration. This ability to withstand intense and prolonged desiccation (equivalent to the loss of 95% of cellular water content) is called desiccation tolerance (DT) (Bewley, 1979). Desiccation tolerance in vegetative tissue is found mainly in less complex organisms like algae, lichens and bryophytes, although some 300 species of angiosperms also have vegetative drought tolerance (Gaff, 1977; Oliver *et al.*, 2000). Gymnosperms are the only major class of vascular plants that does not have a species that has desiccation tolerant vegetative tissue (Oliver *et al.*, 2000).

Oliver *et al.* (2000) hypothesizes that vegetative desiccation tolerance was primitively present in bryophytes, the basal most living clade of land plants. This was lost during the evolution of tracheophytes in preference for the advantages afforded to plants by increased growth rates, structural and morphological complexity, and the mechanisms that conserve water within the plant while maintaining efficient carbon fixation. The genes that had evolved for cellular protection and repair in bryophytes were recruited for different but related processes such as

University of Cape Town

response to water stress, and the desiccation tolerance of reproductive propagules. Oliver *et al.* (2000) goes on to postulate that the mechanism of desiccation tolerance in seeds may have evolved secondarily from more primitive forms of vegetative desiccation tolerance and once these mechanisms were established in seeds, the developmentally induced cellular protection system became available for induction in the vegetative tissues by environmental cues that are related to drying (Figure 1.1).

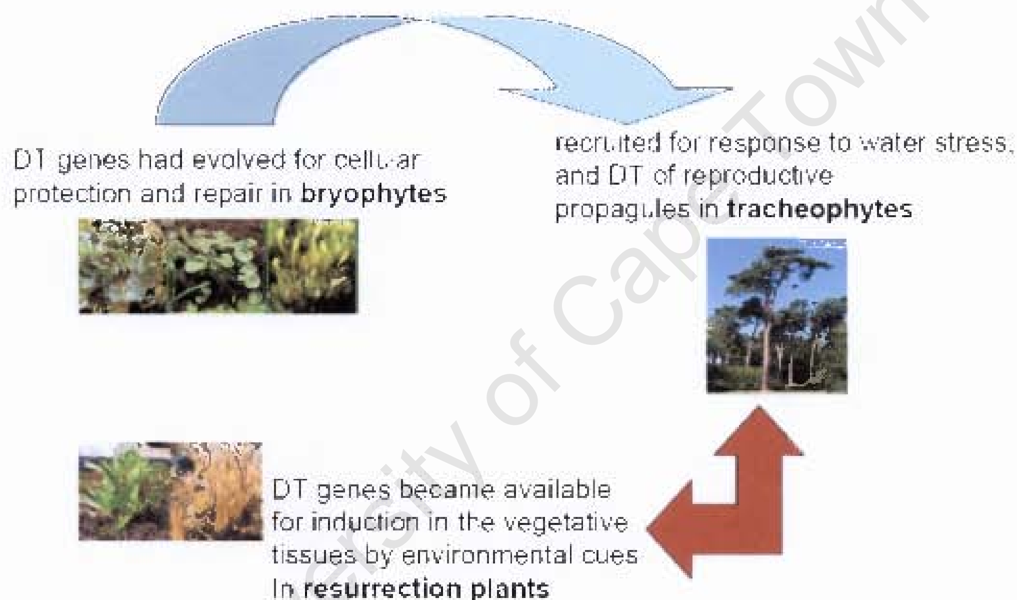


Figure 1.1: Evolutionary hypothesis of desiccation tolerance in vegetative tissue.
The flow diagram summarising the hypothesis on evolution of desiccation tolerance in vegetative tissue of resurrection plants Oliver *et al.* (2000). The tracheophyte and bryophyte pictures were adapted from http://www.answers.com/topic/vascular-plant?cat=technology and http://www.science.siu.edu/landplants/LPSeminar/bryophytes.GIF respectively.

University of Cape Town

1.3. Resurrection Plants

There is a special group of plants known as resurrection plants which possess unique and effective mechanism for coping with drought stress by being desiccation tolerant (Scott, 2000). Fully matured leaves of these plants can lose up to 95 % of their water content and then upon rehydration, the leaves are rehydrated and are photosynthetically active within 24 - 72 h

(Bernacchia *et al*., 1996; Sherwin and Farrant, 1996; Farrant, 2007). These plants can remain quiescent for considerable periods of time, and upon rehydration they can resurrect (Scott, 2000), for example the aquatic plant *Chamegegis intrepidus* is known to survive up to 11 months in the dehydrated state (Hartung *et al*., 1998). Resurrection plants are naturally found in arid climates such as southern Africa, southern America, and western Australia (Gaff, 1987) and are frequently found in shallow, sandy soils (Sherwin and Farrant, 1995) as a result of their ability to survive desiccation. Resurrection plants that are angiosperms have been subdivided into two groups: homoiochlorophyllous, that retain their chlorophyll during drying (such as *Craterostigma wilmsii* and *Myrothamnus fabelifolia*), and poikilochlorophyllous, those that lose chlorophyll on drying, (such as *Xerophyta viscosa*) (Tuba *et al.*, 1993; 1994; Sherwin and Farrant, 1995). Some of the mechanisms that are used by resurrection plants to tolerate desiccation are discussed below.

University of Cape Town

1.4. Mechanisms of Desiccation Tolerance

There are various mechanisms employed by the plant to tolerate desiccation and prevent lethal damage to the different cellular components including membranes, proteins and cytoplasm. These include antioxidants that counter damage by reactive oxygen species, sugars that replace water and form glasses and proteins that stabilize macromolecules and membranes (Alpert, 2006).

1.4.1. Antioxidants that counter damage by reactive oxygen species

Reactive oxygen species (ROS) are natural by-products of metabolism. They are to a greater degree present in chloroplasts and mitochondria. ROS are atoms or molecules with an unpaired electron that are readily donated and thus highly reactive, and are cited as being the most damaging consequence of desiccation stress (Hendry 1992, Smirnoff 1993, Kranner and Grill 1996, Kranner and Birtic 2005, Kranner *et al.,* 2006). During desiccation the metabolic processes shut down. During this time, by-products of continued respiration and light harvesting lead to the accumulation of high energy intermediates that leak out of the mitochondria and plastids. These intermediates form reactive oxygen species and free radicals (Puntarulo *et al*.,

1991; Hendry 1993; Leprince *et al.,* 1993, 1994, 1995). As a result plants have antioxidant molecules and scavenging systems (Larson 1988; Hendry, 1993). Included in the enzymatic free-radical scavenging system is the superoxide dismutase (SOD), which catalyses the dismutation of superoxide (O_2^-) into H_2O_2 and O_2, and those that are involved in the detoxification of H_2O_2 (ie catalase, glutathione reductase and peroxidases) (Berjek *et al.*, 2007). The removal of cytotoxic products resulting from oxidative events is considered to be of prime importance in vegetative tissue for the survival of abiotic stress because genes that encode for enzymatic antioxidants are upregulated during drying (Ingram and Bartels, 1996). Sherwin and Farrant (1996) have shown that in *C. wilmsii* and *X. viscosa*, the activity of ascorbate peroxidase, SOD and glutathione reductase increases during dehydration. In addition to that, *C. wilsmii* was also found to accumulate large amounts of anthocyanins, which have antioxidant capabilities.

1.4.2. Sugars that replace water and form glasses

Soluble sugars have been shown to accumulate in different tissues, in seeds (LePrince *et al.* 1993), pollen (Hoekstra *et al.,* 1992) and in desiccation tolerant vegetative tissue (Bewley and Krochko, 1982; Ingram and Bartels, 1996; Oliver and Bewley, 1997), response to dehydration. It is mostly sucrose, along with raffinose family oligosaccharides (RFO's) such as stachyose, raffinose and cyclitols (Horbowicz and Obendorf, 1994) that accumulate in desiccation tolerant tissue.

There are two roles that sugars are thought to be involved in desiccation. That is in water replacement hypothesis and vitrification. In the water replacement role, it has been postulated that during desiccation soluble sugars interact with the polar head group and replace water molecules, allowing phospholipid molecules to largely retain the original spacing between one another (Buitink *et al.*, 2002). Crowe *et al.* (1992) proposed that the hydroxyl group of trehalose may substitute for water to maintain hydrophilic interaction in membranes and proteins during dehydration by using liposomes and isolated proteins. However, this sugar is not present in most higher plants (Berjak *et al.,* 2007; Paul 2007) and it has been shown that sucrose is present, and accumulates in DT plants (Bianchi *et al.*, 1991, 1992, 1993; Drennan *et al.,* 1993; Kaiser *et al.,* 1985; Suan *et al.*, 1991). There is no evidence that it does so by interaction of OH groups with membranes. Instead, it is now believed that sugar plays a buffering and "osmotic spacer role" as

University of Cape Town

water becomes limiting (reviewed in Berjak *et al.,* 2007). At $\psi \leq$ -60 MPa cell membranes move closer together such that repulsive forces counteracting the close approach of opposing hydrophilic surfaces become increasingly operative. However concomitantly, if membrane surfaces come to be positioned sufficiently close to one another, lateral compression can result, with the possible consequence of transition of some phospholipids to the gel phase, and even demixing of membrane components and exclusion of integral proteins. Sugars within the aqueous phase between opposing membrane surfaces, will help to counteract the close approach of the membranes, acting physically as volumetric and osmotic spacers (Koster and Bryant, 2005).

The second mechanism involves sugars as major contributors to vitrification, the formation of a biological glass, of the cytoplasm of dry cells (Leopold *et al.*, 1994). It has been previously proposed that high concentrations of sugars lead to vitrification of the cytoplasm during desiccation which prevents crystallization (Burke, 1986). The glass formation in the cytoplasm has been postulated to maintain the structural and functional integrity of macromolecules (Crowe *et al.,* 1986). Glass formation alone is not sufficient to confer desiccation tolerance since desiccation sensitive tissues are capable of forming cytoplasmic glasses (Sun *et al.,* 1994; Buitink *et al.,* 1996). However, there is no doubt that cytoplasmic vitrification plays an important role in the survival of anhydrobiotic organisms during desiccation and subsequent dry storage (Sun et al., 1994).

University of Cape Town

1.4.3. Proteins that stabilize macromolecules and membranes, LEA proteins.

The LEA proteins are known to be present in abundance during the late stage of embryogenesis in seed development. This stage is associated with the acquisition of desiccation tolerance is seeds (Cumming 1999; Kermode and Finch-Savage, 2002). They have also been shown to be present in vegetative tissue of the DT grass *Sporobolus stapfianus* (Kuang *et al.,* 1995), the dicot *C. plantagineum* (Piatkowski *et al.,* 1990) and monocot *X. humilis* (Collet *et al.,* 2004). LEA proteins represent a broad class of proteins of highly conserved genes that are expressed in a wide range of plants. The classification of LEA proteins has previously been independently

done by different authors who have divided them into five groups. These groups are namely, group 1, 2, 3, 4 and 5 (Dure, 1993; Ingram and Bartels, 1996, Zhang and Zhao, 2003). This classification was mainly based on their amino acid composition and mRNA homology. Recently though, Wise (2003) and Wise and Tunnacliffe (2004) have grouped the LEAs according to a computational method called Protein and Oligonucleotide Probability Profile (POPP) which allows proteins to be compared based on similarities in their peptide composition as opposed to similarities in their sequences, this has led to the definition of super families (SF) of LEA proteins. This classification was used for the purposes of this study. Table 1.1 taken from Berjek *et al.,* (2007) summarizes the various classifications that have been used by different authors while referencing the author, while the last eight rows of the table include the classification by Wise and Tunnacliffe (2004).

LEAs have been shown to respond to dehydration, low temperature, salinity and exogenous abscisic acid (ABA), indicating their responsiveness to cellular dehydration. Additional to their responsiveness to cellular dehydration, LEA proteins have also been shown to be absent in some dehydration sensitive seed (Farrant *et al.,* 1992), validating their association with desiccation tolerance.

LEA proteins are usually of low molecular weight ~ 10 - 30 kDa and are presumably involved in protecting higher plants from environmental stress damage particularly dehydration. LEA proteins are mostly hydrophilic proteins and are generally unstructured in solution. These proteins unfolded state in solution has made it difficult to assign structure and determine function for which they can be classified and named (Farrant, 2007). Although their precise functions have not been extensively investigated, some functions of LEA proteins have been postulated mainly in reference to their amino acid composition. These include protection of the subcellular milieu by acting *inter alia* as ion scavengers, acting as molecular chaperones, forming fibre-like structures, which may act as a cytoskeleton to avoid excessive shrinking of the cells during desiccation and facilitating glass formation (Bray, 1993; Cuming *et al.*, 1999; Berjak, 2006; Mtwisha *et al.*, 2006; Berjak *et al.*, 2007). Figure 1.2 shows a summary presentation of the functions mentioned above.

University of Cape Town

A database search of LEA protein sequences using the FoldIndex unfolded protein prediction tool (Prilusky *et al.*, 2005) showed that LEA proteins from group 1, 2, 3 are at least 50 % unfolded, and many of the smaller proteins are predicted to be totally unfolded (Tunnacliffe and Wise, 2007). Although these proteins usually occur as unstructured in solution, environmental conditions, and/or the binding to a partner molecule or cation can induce folding. Several LEA proteins have been shown to gain structure when dried. This has been demonstrated with LEAM from pea mitochondria (Tolleter *et al.*, 2007). They are typically found or localized in the cytosol and nuclear regions. The hydrophilic and α-helix nature confers LEA's with thermal stability, this structure is thought to be involved in ion enrichment during dehydration (Zhang and Zhao, 2003).

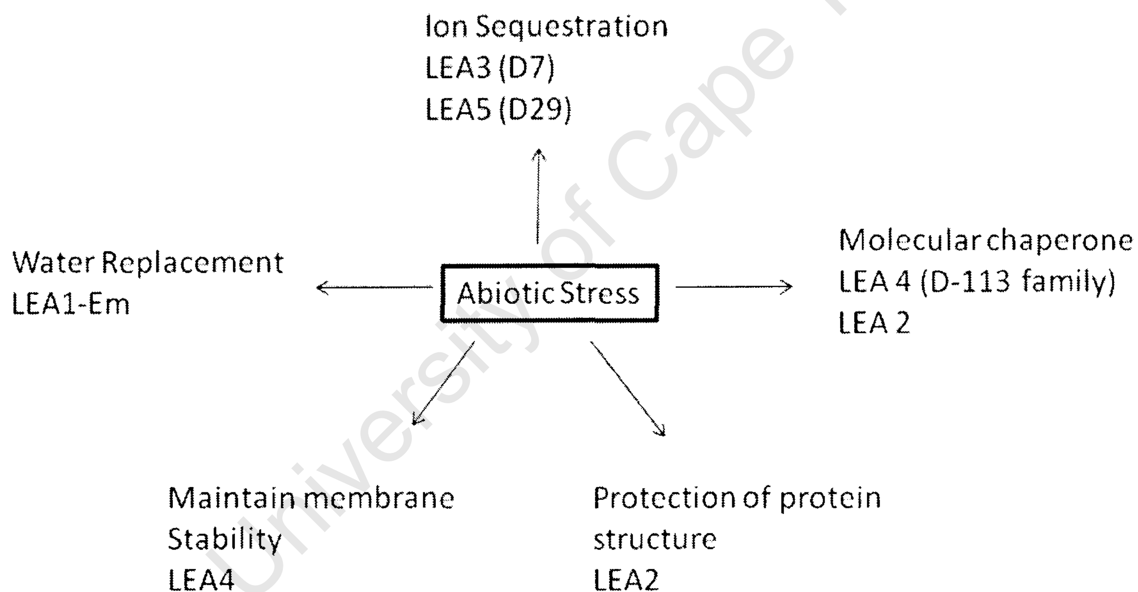


Figure 1.2: Summary of the predicted functions of LEAs.

The roles of LEAs during the dehydration and rehydration cycle are shown in the above schematic. The diagram above shows different LEAs grouped according to their associated functions during dehydration. These include their involvement in ion sequestration, acting as molecular chaperones, protection of protein structure, maintenance of membrane stability and water replacement.

University of Cape Town

Table 1.1: Classification, properties and proposed roles of LEA proteins, reported in various sources. This table is taken from Berkjak *et al.*, 2007 .

Group	Super-family (Wise, 2003)	Family	Pfam domain (Illing et al, 2005)	Motif / Consensus	Postulated roles & distributions (where applicable)	Source
1 (I)[+]		D-19	PF00477	GGQTRREQLGEEGYSQMGRK	Water binding	Bray (1993); Cuming (1999)
2 (II)[+]		D-11 dehydrins	PF00257	K segment at carboxy terminus: EKKGIMDKIKEKLPG or minor variants, but all characterised by KIKEKLPG S (serine-rich) segment in some dehydrins; Y segment in most - (V/T)DEYGNP - near amino terminus	All seedling tissue [water-stressed or not] but quantitative differences related to tissue type Localised [also] near membranes of protein bodies in maize scutellum; endomembrane localisation in onion epidermal cells. Postulate interactions of dehydrins with types of surfaces, rather than specific macromolecular classes K-segment: amphiphilic α-helix; hydrophobic interaction with partially denatured proteins, membranes; synergistic action with compatible solutes in stabilising macromolecules & protoplasm; solubilising agents [detergent properties]; dehydrin-sugar interaction* Distribution: nucleus and cytoplasm; nuclear location depends on tissue and dehydrin concerned	Close *et al.* (1993); Close (1996; 1997) [+]Mtwisha *et al.* (2006) * Blackman *et al.* (1992)
3 (III)[+]		D-7	PF004238	11-mer repeating TAQAAKEKAGE	Repeating elements postulated to exist as amphiphilic α-helices; speculated function of both D-7 & D-29 (Gp 5) is ion sequestration	Bray (1993); Dure (1993)
4 (IV)[+]		D-113	PF03760		May preserve membrane structure	Baker *et*

University of Cape Town

						al. (1988); Bray (1993); Dure (1993)
5 (V)'		D-29		11-mer repeats; some homology with D-7	Sequesters ions	
6 (VI)'					α-helical conformation; along with some Group 3 members, are the only other group to have some structure in water	Mtwisha *et al.* (2006)
1a	4			GGQTRREQLGEEGYSQMGRK		Wise and Tunnacliffe (2004)
1b	6					
2a	1			DEYGNP (Y domain)	Produced late in embryogenesis; not associated with cold-tolerance	
2a	10			EEKK (K domain)	As for 2a	
2b	3			S_n (S segt)	Associated with cold-tolerance; a few not produced during embryogenesis	
3a	2			TAQAAKEKAGE		
3b	5					
6	7			?		
Members of 'conventional' groups 4 & 5 incorporated into LEA groups 2 & 3, according to Wise and Tunnacliffe (2004)						

The group 1 LEA proteins (D19), which are the main focus of this study, will be discussed in detail in chapter 2. As an introduction they are a highly conserved group of LEAs which are characterized by a conserved sequence made up of 20 amino acid residues that can be repeated in tandem up to four times and these can absorb a large amount of water (Close, 1996).

The group 2 LEAs (D11), are also called dehydrins (dehydration induced proteins; Close *et al.*, 1989). They have a conserved stretch of 15 amino acids, namely, EKKGIMDKIKEKLPG in their carboxyl terminal region. These proteins are thought to play a role as molecular chaperones during metabolism and protecting protein structure during drought (Rizhsky *et al.*, 2002; Wisniewski *et al.*, 2001). The POPP classification (Wise and Tunnacliffe, 2004) of group 2 characterises this group as having repeats of either or both DEYGNP and EEKK. This classification also splits this group into two subgroups, group 2a and 2b (Table 1.1). Group 2a, includes all group 2 proteins that are not associated with cold tolerance and all that are present late in embryogenesis; while group 2b includes those that are generally not associated with cold stress and with a few that are not produced during late embryogenesis (Wise and Tunnacliffe, 2004). Group 2a overall charge is neutral or basic, with an over-representation of glycine. Group 2b has similar levels of basic residues to group 2a, however, this group also has increased levels of acidic residues (Tunnacliffe and Wise, 2007). Jensen *et al.* (1998) demonstrated that the movement of group 2b proteins from maize, DHN1/Rab 17, between the nucleus and the cytoplasm is controlled by phosphorylation of its serine stutter. Consequently, the removal of this sequence results in the lack of phosphorylation and retention of the proteins in the cytoplasm. This group of LEAs like most other LEAs are predominantly unfolded as shown by examples from maize, the resurrection plant *Craterostigma plantagineum*, cowpea, *Citrus* and *Arabidopsis* (Ceccardi *et al.*, 1994; Lisse *et al.*, 1996; Ismail *et al.*, 1999; Hara *et al.*, 2001; Mouillon *et al.*, 2006).

The group 3 LEA's (D7 from cotton) (Baker *et al.*, 1988) are made up of 11 amino acid namely, TAQAAKEKAGE , in 13 repeats and have an amphypathic α-helix structure, they are involved in ion enrichnment during dehydration in higher plants (Zhang and Zhao, 2003). The POPP classification (Wise and Tunnacliffe, 2004) splits group 3 LEAs into two subgroups, group 3a and 3b (Table 1.1). Two closely related group 3 LEA proteins that contained a putative

University of Cape Town

endoplasmic reticulum (ER) signal peptide at the N-terminus and a putative ER-retention signal at the C-terminus were shown to be located in the ER (Hsing *et al.*, 1995). These groups of proteins are predominantly unfolded in the hydrated state, for example a group 3 LEA protein from bulrush pollen (Wolkers *et al.*, 2001) and one from an anhydrobiotic nematode *Aphelecus avenae* (Goyal *et al.*, 2003). It has however been demonstrated that a small (8kDa) group 3 LEA protein from *Typha latifolia* became largely α-helical when dried rapidly; slow drying resulted in both α-helical and intermolecular β-sheet structures (Wolkers *et al.*, 2001). The same response was observed with group 3 LEA proteins AavLEA1 from nematode *Aphelenchus avenae* (Goyal *et al.*, 2003) and LEAM from pea mitochondria (Tolleter *et al.*, 2007), which also fold upon drying.

The group 4 LEA proteins (e.g. Dll3, cotton LEA14) have no repeating motif in their sequence and contain a conserved region at the N-terminal to form an amphiphathic α-helical structure (Hong *et al.*, 2005). These proteins are thought to maintain membrane stability because they are able to form adaptative conformational changes during drying and dehydration (Chaves *et al.*, 2003).

The Group 5 LEA proteins (D29) have low amino acid residue specificity and this is predicted to be related to their roles in seed maturation, dehydration and combining concentrated ions (Han *et al.*, 1996; Chen *et al.*, 2002; Yu *et al.*, 2002). Group 5 as well as group 4 disappear under the POPP classification. The former members of the group are redistributed into LEA groups 2 and 3 (Wise and Tunnacliffe, 2004).

There is not much reported on group 6 of LEAs. Bray and Lay (1994) classify D34 from cotton (Baker *et al.*, 1988) and rab28 from maize (Pla *et al.*, 1991) as members of this group. The rab28 and D34 proteins have 55 % similarity and are both rich in alanine and valine whereas cysteine is absent in both cases (Pla *et al.*, 1991). They are however postulated to have a α-helical conformation. Together with group 3 LEAs they are thought to have some structure in water (Mtwisha *et al.*, 2006).

University of Cape Town

1.5. Summary and Aims

Global transcriptome analysis has offered a basic framework to test different hypothesis on the evolution of desiccation tolerance and expression of LEA proteins. Such a study analyzing the *Arabidopsis* transcriptome (Illing *et al.*, 2005) has demonstrated that the LEA protein gene set divides roughly into those with seed specific expression and those expressed in vegetative tissues and some a combination of seed and vegetative tissue expression. Illing *et al.* (2005) also identified LEA genes belonging to superfamilies -1, -6 and -9 to be only significantly expressed during seed development and suggested that these LEA groups might be uniquely associated with defence against severe water loss such as would occur in desiccation-tolerant angiosperm, resurrection plants or orthodox seeds. Group 1 LEAs have been demonstrated to be essential in slowing down the rate of drying in Arabidopsis seed by characterizing the T-DNA knockout ATEM6 knockout mutant (Manfre *et al.*, 2006). This group of LEAs have however not been characterized in the vegetative tissue of resurrection plants. This study, characterizing the LEA1-Em genes in *X. humilis*, aims test the hypothesis that desiccation tolerance in vegetative tissue of resurrection plants is a consequence of the appropriation of seed specific genes. This will be done by:

1. Isolating LE1-Em orthologues from *X. humilis* vegetative tissue and seed.

2. Isolating and cloning the 5'and 3' UTR sequences of all the orthologues isolated.

3. Investigating the expression patterns for LEA1-Em orthologues in seed, and during desiccation and rehydration in leaves and roots of *X. humilis*.

University of Cape Town

2 Isolation and cloning of LEA1-Em genes from *X. humilis*

2.1 Introduction

The group 1 LEAs are named LEA1-Em. They obtained their name , early methionine labelled (Em) as they were the first strongly labeled polypeptide observed in wheat embryos when incubated in the presence of [^{35}S] Met (Cumming and Lane, 1979). Several LEA1-Em genes have been isolated from different plants. The national centre for biotechnology information (ncbi) database contains several entries of LEA1-Em orthologues isolated from at least nine different plant species. Some of the LEA1-Em orthologues are listed in Table 2.1 with the nomenclature used in this study for LEA1-Em. The nomenclature adopted in this study was derived by taking the first letter of the genus and species followed by the LEA group which is LEA1 and the number of times that the 20 amino acid motif, used to classify the LEA1 group, is repeated in that sequence.

Table 2.1: List of LEA1-EM genes

Source- plant name	Accession number (NCBI)	Number of repeats	Nomenclature used in this study:
Triticum aestivum (bread wheat)	CAA36323	1	TALEA1-1
Glycine max (soybean)	AAB71224	1	GMLEA1-1
Brassica napus (rape)	CAA05711	1	BNLEA1-1
Robinia pseudoacacia	AAB39474	1	RPLEA1-1
Robinia pseudoacacia	AAB39473	2	RPLEA1-2
Vigna radiate	AAB07225	1	VRLEA1-1
Vigna radiate	AAB07224	2	VRLEA1-2
Arabidopsis thaliana ATEM1	NP_190749	4	ATLEA1-1
Arabidopsis thaliana ATEM6	NP_181546	1	ATLEA1-4
Oryza sativa (japonica cultivar-group)	CAA44836	1	OSLEA1-1
Secale cereal (rye)	CAB88086	1	SCLEA1-1
Gossypium hirsutum (upland cotton)	CAA38374	2	GHLEA1-2

As mentioned before in the introduction, this group of LEA proteins has a conserved sequence made up of 20 amino acid, that can be repeated in tandem up to four times (Close, 1996). LEA-1s are highly hydrophilic and are predicted to have a high percentage of charged amino acids and

University of Cape Town

glycine (Tunnacliffe and Wise, 2007). This has led to the prediction that they have high capacity for water binding and thus thermal stability.

The precise structure and function of group 1 LEAs remain unknown. On the basis of their amino acid sequence, these proteins are predicted to largely exist as random coil structures in aqueous environments (Gilles *et al.*, 2007). LEA-1 proteins are found in plants where their function is thought to be the prevention of protein aggregation under water stress conditions (Goyal *et al.*, 2005). Russouw *et al.* (1995) showed using far UV circular dichroism (CD) that LEA1 from pea (*Pisum sativum*) was almost entirely unstructured, having only an estimated 2 % of protein folding as α-helix in water. Eom *et al.* (1996) also showed that the Em homologue from carrot (*Daucus carota*) has no secondary or tertiary structure when examined by proton nuclear magnetic resonance.

The exact function of Em proteins is not known. However, they have been associated with DT of seeds and vegetative tissue. Manfre *et al.*, (2006), has recently shown using a *Arabidopsis* mutant with a T-DNA insertion in the ATEM6 locus, that in the absence of ATEM6 there is premature drying and maturation of seeds at the distal ends of the silique. In transgenic rice, the expression of LEA1 protein (PMA959) in vegetative tissue has been shown to enhance cellular integrity when plants are subjected to salt stress (Cheng *et al.*, 2002). They are assumed to protect cellular or molecular structures during desiccation (Dure *et al.*, 1989, 1993, Kermode, 1997). This has also been demonstrated in yeast where expression of Em protein attenuates the growth inhibition that is normally observed in media of high osmolarity and mitigates the detrimental effects of low water potential in a non-specific manner (Swire-Clark and Marcotte, 1999). The Em proteins are also presumed to act as a "molecular shield" to prevent formation of damaging protein aggregates during water stress. In addition it was demonstrated that in the presence of trehalose, Em proteins exhibit a protective and synergistic effect and do not behave as classical molecular chaperones (Goyal *et al.*, 2005).

Studies published to date, which investigates the expression of LEA1-Em genes are limited to seeds (Illing *et al.*, 2005; Manfre *et al.*, 2006), the expression of these genes in desiccation tolerant vegetative tissue is yet to be published. In the model plant *Arabidopsis* the two LEA1-

University of Cape Town

Em genes, encode for similar protein, the main differences being the number of repeats of the conserved 20-amino acid motif (ATEM1 having four copies and ATEM6 one copy). They are however, temporally and spatially expressed differently during embryo maturation. ATEM1 mRNA is mainly expressed in the provascular tissue with the strongest expression in the root tip whereas ATEM6 expression is essentially in all regions of the embryo with the strongest expression in the apical meristem and provascular tissue (Vicient *et al.,* 2000). ATEM1 mRNA is also expressed 2 days prior to the expression of ATEM6 (Bies *et al.,* 1998). Morris *et al.* (1990) has shown that this group of LEAs is induced by osmotic stress in the developing wheat embryos in culture, while their accumulation is triggered by abscisic acid (Williamson *et al.*, 1985; Marcotte, 1988).

Messenger RNA and reporter gene expression studies of the two *Arabidopsis* LEA1-Em genes have shown that these genes are only expressed during the later stages of seed maturation (Hughes *et al.,* 1989, 1991; Raynal *et al.,* 1989). Recently it has been shown using microarray data from Arabidopsis that both ATEM1 and ATM6 genes are seed specific and have low expression in vegetative tissue in response to abiotic stress (Illing *et al.,* 2005).

The seed specificity of this group of LEAs makes them an ideal candidate to investigate the hypothesis, proposed in the introduction, that desiccation tolerance in *X. humilis* arose through the activation of seed-specific genes in vegetative tissues in response to water loss. The first step in testing this hypothesis was cloning the full-length cDNAs for LEA1-Em orthologues from *X. humilis*. This was done using 5' and 3' rapid amplification of cDNA ends (RACE). RACE is a molecular biology technique used to obtain the full length sequence of an RNA transcript.

2.2 Materials and Method

2.2.1 Plant Material and Growth Conditions

Mature *X. humilis* plants were collected from Pilanesberg Nature Reserve (North West, South Africa). They were grown in a mixture of their natural soil and potting soil and were maintained in the Botany greenhouse with no supplementary lighting. Whole plants of *X. humilis* were dried

University of Cape Town

by withholding water and allowing the plant to dry out naturally. Hydrated tissue samples were taken at full turgor and dried tissue samples were taken at the relative water content (RWC) less or equal to 5 %.

2.2.2 RNA Extraction for LEA1-Em isolation

Approximately 100 mg of *X. humilis* leaf samples, were ground in liquid nitrogen in a motar and pestle and transferred while frozen to a 2 ml eppendorf tube containing 10 times the volume (1 ml) of TRI REAGENT® (ie: 1 ml per 100 mg of leaf material). Total RNA was isolated from the tissue using TRI REAGENT® (Molecular Research Centre, USA) according to the manufacturer's instructions, with the following amendments. Plant cells were lysed by agitation (vortex) in 1 ml TRI REAGENT. A high salt precipitation [0.8 M tri-sodium citrate, 1.2 M NaCl, 50 % (v/v) isopropanol] at -20 °C for an hour was performed in replacement to the standard 100% isopropanol precipitation, in order to remove polysaccharides. RNA was resuspended in 20 μl nuclease-free water and stored at -80 °C until analysis and further use. The concentration of the RNA was determined using the ND-1000 spectrophotometer (NanoDrop Technologies, USA). RNA samples were run on agarose gels in the presence of formaldehyde (Merk, USA). Gel pictures were used to confirm the integrity of the RNA. RNA was taken to be intact if the 28S RNA band was approximately twice the intensity of the 18S RNA band.

University of Cape Town

2.2.3 cDNA synthesis

The total RNA (5 μg) was heated to 65 °C for 10 min in the presence of 1 μl of 50 μM oligo(dT) primer, 1 μl of 10 mM dNTP mix, in a final volume of 13 μl with nuclease free water. The reaction was placed on ice for 2 min, prior to the addition of 4 μl 5X First Strand (FS) buffer (Invitrogen, USA), 1 μl 0.1 M DTT, 1 μl RNase Inhibitor (Invitrogen, USA) and 1 μl Superscript RT III (Invitrogen, USA) to give a final volume of 20 μl. Reactions were incubated at 50 °C for 90 min and inactivated at 70 °C for 15 min.

2.2.4 Degenerate Primer Design

The BLAST algorithm (http://ncbi.nlm.nih.gov/http://ncbi.nlm.nih.govt/) was used to search for LEA1-Em and Em-like gene sequences in the GenBank database. Identified sequences were

downloaded to the BioEdit programme and amino acid sequences were aligned by ClustalW. Twelve LEA1-Em sequences isolated from nine different plant species were used for designing primers. Conserved domains were identified following the ClustalW alignment, and were used to design degenerate primers. DNAMAN (version 5.2.10) was used to predict the melting temperatures (Tm) of the degenerate primers and their specificity and self complementarity. The primers were synthesised in the Department of Molecular and Cell Biology Oligonucleotide Synthesis Unit.

2.2.5 Degenerate Primer PCR

All PCR reactions were made up in deionised water in 50 µl volumes containing 1 µl cDNA/genomic DNA template, 1 µl primers (10 mM) and 6 µl $MgCl_2$ (25 mM), 2 µl dNTPs (10 mM), 10 µl PCR buffer (10X) and 0.5 µl Supertherm polymerase enzyme (Southern Cross Biotechnology, South Africa).

The annealing temperatures tested for degenerative primer optimization were 50.6 °C, 52.1 °C, 54.5 °C, 56 °C. The PCR cycles were: 94 °C (3 min), [94 °C 20 sec); 50 °C - 56 °C (40 sec); 72 °C (40 sec)] x 30, 72 °C (1 min). A negative control was included for all reactions. 20 µl of the PCR product was electrophoresed on agarose gels.

2.2.6 Cloning

All PCR products were cloned into pGEM-T (Promega, USA) prior to sequencing. 20 µl of the PCR reaction was electrophoresed on a 2 % low melting point agarose gel alongside 100 bp ladder (Fermentas, USA). Specific bands were selected, cut from the gels under long wavelength UV light with a sterile blade. The agarose was melted at 70 °C for 10 min and PCR product purified using the Gel purification kit (Qiagen, USA) as described by the manufacturer. 3 µl of the sample was used for cloning of the fragment into pGEM-T Easy vector (Promega, USA) as described by the manufacturer. The ligation reactions were used to transform competent *E. coli* DH5α cells according to the protocol for transformation for pGEM-T ligation reactions (Promega, USA). Transformants were plated onto Luria broth (LB) agar plates

University of Cape Town

containing 100 μg/ml ampicillin, 10 μg/ml IPTG and 5 μg/ml X-gal and incubated at 37 °C overnight.

2.2.7 Screening and sequencing

Colonies were picked using a micropipette tip and placed directly into the 20 μl PCR reaction, containing 1X *Taq* polymerase Buffer, 1.5 mM $MgCl_2$, 40 μM dNTP mixture (Roche, Germany), 0.2 μM of each primer forward primer T7 (5'TAA-TAC-GAC-TCA-CTA-TAG-GG-3') and reverse primer SP6 (5'-ATT-TAG-GTG-ACA-CTA-TAG-AA-3') and 0.025 U/μl Supertherm *Taq* polymerase. PCR conditions were 94 °C (5 min), [94 °C 40 sec); 52 °C (40 sec); 72 °C (40 sec)] x 30, 72 °C (1 min). No template controls were employed. Aliquots (20 μl) of PCR reactions were loaded onto a 2 % agarose gel and electrophoresed at 100 V for one hour. Products were visualized and photographed for future reference. Insert sizes, were estimated from comparison with the 100 bp DNA marker (Fermentas, USA).

University of Cape Town

The positive colonies were inoculated into 5 ml Luria broth (LB) with ampicillin and incubated at 37 °C overnight. The plasmids were isolated from the *E. coli* DH5α overnight culture using the High Pure Plasmid Isolation Kit (Roche, Germany) according to the manufactures instruction. Each pGEMT-Easy plasmid construct was sequenced using the forward primer T7 and reverse primer SP6. The nucleotide sequence of each gene was determined on both forward and reverse direction where sequencing was successful. Sequence alignments were performed using DNAMAN software (version 5.2.10).

2.2.8 Design of 5'and 3' Rapid amplification of cDNA ends (RACE) gene specific primer (GSP)

The sequences from the PCR products were used to design primers for 5' and 3' RACE to amplify the full-length sequence of LEA-1 cDNAs identified by PCR using degenerate primers. The melting temperatures (Tm) of the GSP were kept ≥65°C and determined using the DNAMAN programme. The GSPs were designed to produce overlapping RACE products. Primers were also aligned to check for self-complementarity and complementarity to the RACE kit universal primers. All 5' and 3' RACE primers were synthesised in the Department of

Molecular and Cell Biology Oligonucleotide Synthesis Unit. The primers used for RACE amplification are listed below (Table 2.2).

Table 2.2: List of primers used for 5' and 3' RACE PCR

Target Sequence	Primer	Primer Name	Sequence (5'-3')	Primer Length	Primer Tm*
Sequence 1 (XHLEA1-1)	3' RACE PCR	XHT2gsp1	ggataccaaagaaatgggaagccgcgg	27bp	74.9°C
	5' RACE PCR	XHT2gsp2	ggggttccaagctgttcacgtctagtctga	30bp	73.0°C
Sequence 2 (XHLEA1-2)	3' RACE PCR	XHM6gsp1	ggtgaagctcaagagcaccttgctgaa	27bp	71.4°C
	5' RACE PCR	XHM6gsp2	cgcccttgtggcccatttctttg	23bp	72.1°C
Sequence 3 (XHLEA1-4)	3' RACE PCR	XHB1gsp1	tagcaagggagggcagacgaggaa	25bp	71.0°C
	5' RACE PCR	XHB1gsp2	gctcctgataccctctgtccccaa	25bp	71.1°C
XHC00797	3' RACE PCR	XHCgsp1	gggccccttcttttccctcgtc	22bp	70°C
	5' RACE PCR	XHCgsp2	gcgatctcgtcggagataggtggcgt	26bp	71.7°C

*Tm= thermodynamic melting temperature (°C).

University of Cape Town

2.2.9 5'and 3' Rapid Amplification of cDNA Ends (RACE) PCR

Total RNA (1 µg) was mixed in a final volume of 5 µl with 1 µl of 5'CDS (5'–AAGCAGTGGTATCAACGCAGAGTACGCGGG–3') and 1 µl of SMART II oligo for 5'CDNA and 3' CDS (5'–AAGCAGTGGTATCAACGCAGAGTAC(T)30V N–3') for 3'cDNA. The contents were mixed and incubated at 70 °C for 2 minutes. The tubes were then cooled for 2 minutes and the following was added to the total of 10 µl. 2 µl 5X First strand buffer, 1 µl DTT (20 mM), 1 µl dNTP (10 mM) and 1 µl Powerscript Reverse transcriptase (Clontech, USA). The mixture was incubated at 42 °C for 1 h, and reaction was made up to 100 µl with Tricine-EDTA buffer and stopped by heating at 72 °C for 7 minutes.

All PCR reactions were done using the Clontech SMART™ RACE cDNA amplification kit (Clontech, USA). The positive and negative controls, summarised in Table 2.3, were included

for all RACE PCRs. The positive control used the human placental total RNA as a template, and the 5'-RACE TFR Primer (10 μM) and 3'-RACE TFR Primer (10 μM) were used for RACE amplification (Clontech, USA).

Table 2.3: The controls used for 5' and 3' RACE PCR .

Tube number	5' RACE sample	5'TFR (PCR + control)	GSP1+2 (internal +control)	UPM only (+ control)	GSP1 only (+ control)
5'RACE Ready cDNA	+	+ (human placental cDNA)	+	+	+
UPM	+	+	-	+	-
GSP1	+	-	+	-	+
GSP2	-	-	+	-	-
Control 5' RACE TFR primer	-	+	-	-	-
Master Mix	+	+	+	+	+

University of Cape Town

Target cDNA derived from the RT reaction was amplified by PCR using the 5'/3'RACE primer (Table 2.2) and the Universal primer mix (UPM) (Long (0.4 μM): 5'–CTAATACGACTCACTAT AGGGCAAGCAGTGGTATCAACGCAGAGT–3' and Short (2 μM): 5'–CTAATACGACTCA CTATAGGGC–3'). The primers used for each sequence are listed in Table 2.2. Target cDNA was combined with 0.2 μmol of each PCR primer, 1X universal primer mix, 200 μM dNTPs, 1 μl of Taq DNA polymerase in PCR buffer and made up to the final volume of 50 μl with PCR-Grade water. PCR was 30 cycles of 94 °C for 5 sec, 68 °C for 10 sec, and 72 °C for 2 min, followed by an extension step at 72 °C for 7 min. 1 μg of cDNA was used for each PCR. The PCR products were excised from a 2 % agarose gel, purified with a PCR purification kit (Qiagen, USA). All fragments were cloned into pGEM-T Easy vector and the insert fragments were sequenced.

2.2.9 Primer Design for cloning of full length Open Reading Frame (ORF) sequences

Primers were designed to amplify and the clone the full length open reading frame of XHLEA1-1, XHLEA1-2 and XHLEA1-4. The primers were based on the 3' and the 5' RACE sequences of the cloned cDNA. The primers were designed by manual inspection. Primers were then checked on DNAMAN for self- complementarity, annealing temperature and primer dimer formation. The primers were synthesised in the Department of Molecular and Cell Biology Oligonucleotide Synthesis Unit are listed in Table 2.4. For XHLEA1-1 and XHLEA1-4, the gene specific primers failed to amplify, the forward primer was used together with the Universal primer mix (UPM) (Long(0.4 µM):5'–CTAATACGACTCACTAT AGGGCAAGCAGTGGTATCAACGCAGAGT–3' and Short (2 µM): 5'–CTAATACGACTCA CTATAGGGC–3') and the 3' RACE ready cDNA to amplify the full length open reading frame.

Table 2.4: List of primers used for full length amplification of the Open Reading Frame

Target Sequence	Primer	Primer Name	Sequence (5'-3')	Primer Length	Primer Tm*
Sequence 1 (XHLEA1-1)	Full length PCR	FLXHT2gsp1	GCAGCATCCACGCCTGGAG AAC	22bp	70.6°C
Sequence 2 (XHLEA1-2)	Full length PCR	FLXHM6gsp1	GAGAAGCAGCAGCTGAGG GTTAGTAGT	26bp	65°C
	Full length PCR	FLXHM6gsp2	TTTATTAACCTAACATAAC CCTAGCTGATTG	31bp	62.1°C
Sequence 3 (XHLEA1-4)	Full length PCR	FLXHB1gsp1	CTCATTGTCATTTTCGTTGT GAGCAGG	27	68.6°C

University of Cape Town

2.3 Results

2.3.1 Degenerate Primer Design

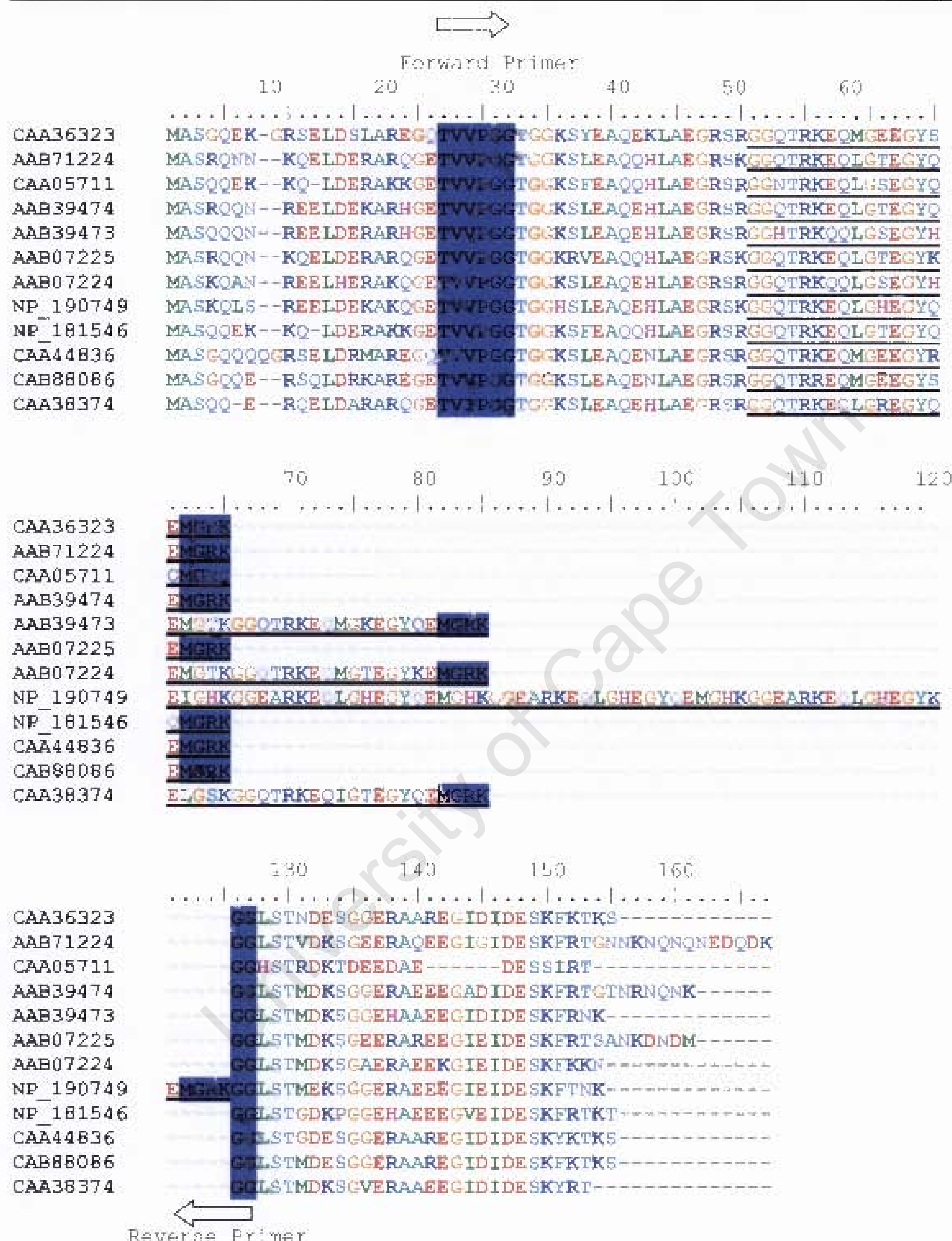
LEA1-Em and Em-like sequences were obtained from the GenBank database and aligned in order to identify conserved domains. The alignments (Figure 2.1) show the conserved sequences highlighted in blue that occur outside the 20 amino acid motif repeats. The sequences, forward (TVVPGG) and reverse (MGRKGGL) primers are ideal because they are highly conserved, and not repeated in a single gene. Such primers eliminate the possibility of obtaining multiple PCR products from a single gene amplification. The details of the primer designed are summarised below (Table 2.5).

Table 2.5: Degenerate primers used for isolation of LEA1-Em from *X. humilis.*

Target Gene	Primer	Sequence (5'-3')	Length	Tm*	Expected Product Size
LEA1-Em	Forward	GACCGTCGT(CG)CC(CA)GGIGGIAC	21	59.9°C	100-300bp
	Reverse	GGAGXCCXCC(CT)TTX(CT)GXCCCAT	22	53.6°C	

*Tm= thermodynamic melting temperature (°C).

University of Cape Town



University of Cape Town

Figure 2.1: Multiple sequence alignments of selected LEA1-Em's, listed in Table 2.1.

The amino acid sequences were downloaded onto the BioEdit programme and aligned using ClustalW. The underlined sequenced present the conserved 20 amino acid domain. The blue highlighted sequences represent the forward and reverse primer positions. The direction of the primer is shown by an arrow below the primer region.

2.3.2 Optimization of degenerative PCR and cloning

A range of annealing temperatures was tested for the optimisation of the PCR. The optimal annealing temperature was found to be 52 °C (data not shown). The expected product size from the *Arabidopsis* ATEM6 cDNA sequence was 138 bp and ATEM1 being 315 bp. The PCR product size of approximately 200 bp was obtained from the PCR using the desiccated leaf tissue cDNA (Figure 2.2). This size was considered to be within the expected size range. Lane 2 is the no template control which resulted in no PCR product indicating absence of contaminants in the PCR reaction. However, positive PCR products, with the product size of about 330 bp, in negative RT control (Lane 3, Figure 2.2) suggested that there were additional LEA-1 orthologues in *X. humilis* that were amplified by a carry-over of genomic DNA in the reaction. Similar sized PCR products were also present in PCR on cDNA template, and were interpreted as a consequence of genomic contamination. Lane 6 was a positive control using Arabidopsis genomic DNA as a template, this was used to confirm the success of the PCR. The possible reason for the presence of two higher molecular weight bands instead of the expected sizes of, 138 bp and 315 bp, is likely a result of the presence of an intron in the template as genomic DNA was used as a template in the PCR. The 200 bp band from the PCR on cDNA template from desiccated leaf tissue (lane 4 and 5) was cloned into pGEM-T Easy vector. Five colonies were confirmed by colony PCR to contain the insert of the expected size of approximately 200 bp and one clone was sequenced in both the forward and reverse direction.

University of Cape Town

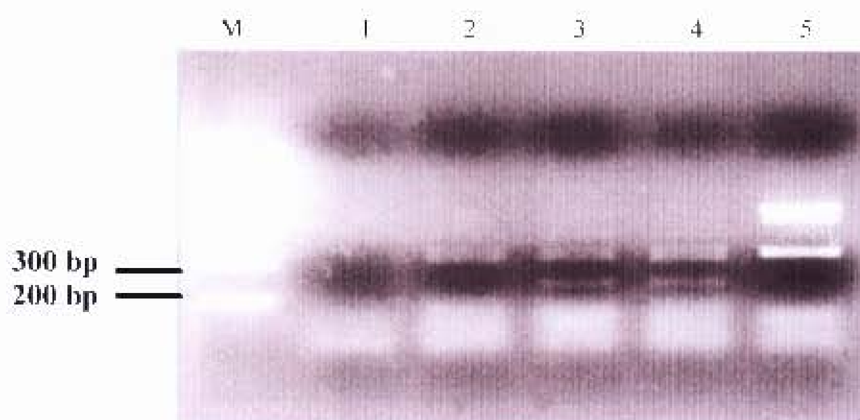


Figure 2.2: Amplification of LEA1-Em from *X. humilis* cDNA using degenerate primers.
Lane M: Low molecular weight maker, Lane 1: contains a negative control (no template), Lane 2: contains reverse transcription negative control, Lane 3and 4: shows a PCR product obtained using *X. humilis* desiccated leaf cDNA as template, while Lane 5: contains positive control using Arabidopsis genomic DNA as a template.

The results of the sequence isolated from the desiccated leaf tissue cDNA had an 83 % maximum identity to the *Arabidopsis* ATEM1 with the E value of 7e-22 and a 75 % maximum identity to ATEM6 with the E value 3e-14. These results confirmed that the isolated sequence was indeed a LEA1-Em orthologue. To check for paralogues of LEA1-Em in the genome of *X. humilis* the same PCR was done at the same annealing temperature of 52 °C (Figure 2.3) using genomic DNA as a template. Lane 3 and 4 PCR had *X. humilis* genomic DNA as a template and resulted in two strong bands of approximately 250 bp and 330 bp, as well as a much fainter band of 280 bp between the two strong bands. The bands obtained from this PCR were also cloned into pGEM-T Easy vector, the inserts confirmed by colony PCR and sequenced. The size difference in these PCR products could be due to the different number of the 20 amino acid repeats associated with this group of LEAs. Lane 2 is the negative control and had no amplification indicating the absence of non specific amplification while lane 5 is a positive control using the *Arabidopsis* genomic DNA to confirm the success of the PCR. The presence of amplified product confirms that the PCR was successful.

University of Cape Town

```
Xerophyta genomic sequence 1     1   CGGCGGAAAGAGCCTCGAAGCTCAACAACACCTCGCCGAA  40
Xerophyta cDNA sequence          1   GGGCGGGAAGAGCCTCGAAGCTCAGCAACACCTCGCTGAA  40
Xerophyta genomic sequence 2     1    GGCGGGAAAAGCCTTGAAGCTCAAGAGCACCTTGCTGAA  39
Xerophyta genomic sequence 3     1   TGGCGGGAAAAGCCTTGAAGCTCAAGAGCACCTTGCTGAA  40

Xerophyta genomic sequence 1    41   GGTCGTAGCAAGGGAGGGCAGACGAGGAAGGAACAGTTGG  80
Xerophyta cDNA sequence         41   GGTCGTAGCAAGGGAGGGCAGACGAGGAAGGAACAGCTGG  80
Xerophyta genomic sequence 2    40   GGACGGAGCCGTGGAGGTCAGACTAGACGTGAACAGCTTG  79
Xerophyta genomic sequence 3    41   GGACGGAGCCGCGGAGGTCAGACTAGACGTGAACAGCTTG  80

Xerophyta genomic sequence 1    81   GGACAGAGGGGTATCAGGAGATGGG---CCACA-AGGGCG 116
Xerophyta cDNA sequence         81   GGACAGAGGGGTATCAGGAGATGGG---CCACA-AGGGCG 116
Xerophyta genomic sequence 2    80   GAACTGAAGGATACAAAGAAATGGG---CCACA-AGGGCG 115
Xerophyta genomic sequence 3    81   GAACTGAAGGATACAAAGAAATGGAAGCCGCGGAGGTCA 120

Xerophyta genomic sequence 1   117   GCCTCCA                                  123
Xerophyta cDNA sequence        117   GCCTCCA                                  123
Xerophyta genomic sequence 2   116   GCCTCCA                                  122
Xerophyta genomic sequence 3   121   GACTAGACGTGAACAGCTTGGAACTGAAGGATACAAGGAA 160

Xerophyta genomic sequence 1   124                            123
Xerophyta cDNA sequence        124                            123
Xerophyta genomic sequence 2   123                            122
Xerophyta genomic sequence 3   161   ATGGGCCGCAAAGGCGGCCTCCA  183
```

Figure 2.4: Multiple alignment of the sequences isolated using degenerate primers.

The highlighted sequences are the forward and reverse GSP for RACE PCR. The primer positions for RACE PCR are highlighted in pink for sequence 1, red for sequence 2 and blue for sequence 3.

2.3.3 5' and 3' RACE Cloning

Rapid amplification of cDNA ends (RACE) was used to isolate and clone full length sequences of all the isolated partial sequences. The RACE PCR was done using cDNA isolated from leaf, root and seed tissue including all the appropriate controls (Table 2.3). Although 3' RACE products were successfully amplified for XHLEA1-2 and XHLEA1-4 from root cDNA template (Figure 2.5A), a 5' RACE product was only identified for XHLEA1-4 (Figure 2.5B). RACE products were not amplified for XHLEA1-1 from root cDNA template.

University of Cape Town

The mRNA transcripts for XHLEA1-1, XHLEA1-2 and XHLEA1-4 were all identified in seed cDNA by 5' RACE, although the PCR products for 3' RACE were only successfully identified for XHLEA1-2 and XHLEA1.4 in seed. In leaf cDNA template 3' RACE product was only identified for XHLEA1-4 (Figure 2.6).

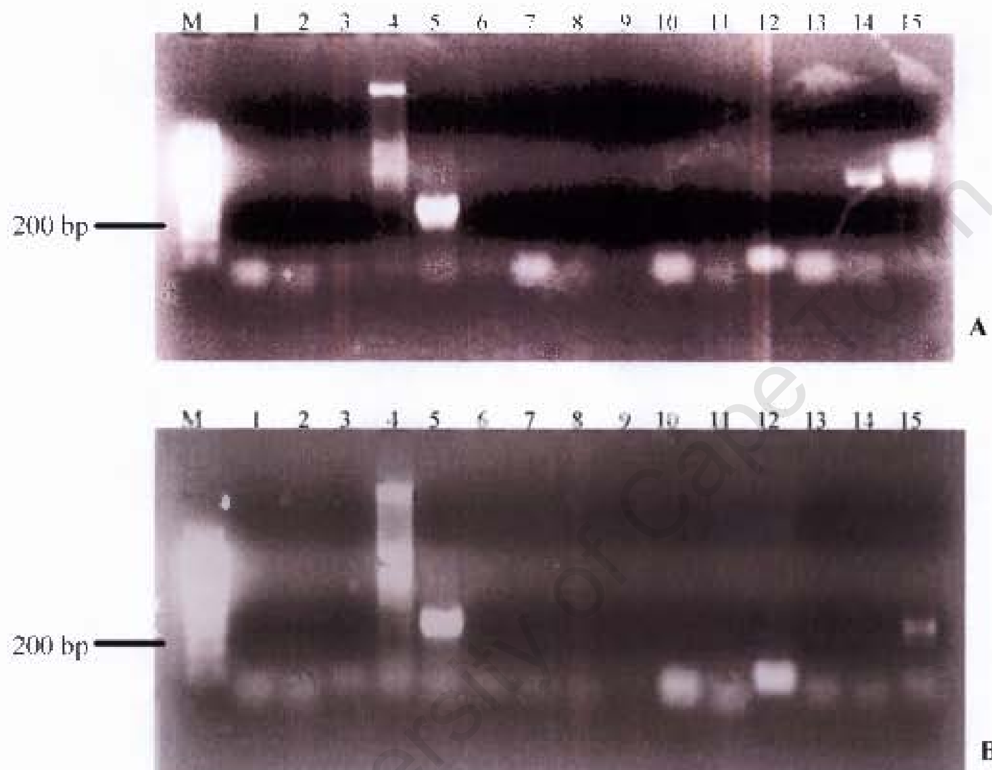


Figure 2.5: Agarose gel analysis of (A) 3'RACE PCR with root cDNA as template and (B) 5'RACE PCR with root cDNA as template

Gel A and B were loaded is a similar pattern, the only difference was Lane 13, 14 and 15 that used 5' RACE GSP (XHT2gsp2, XHM6gsp2, XHB1gsp2) respectively. Lane M: O'GeneRuler 100 bp DNA ladder. Lane 1 to Lane 3: contains a negative control (no template) for each primer set (Table 2.2). Lane 4 and 5: contains RACE PCR positive control using the human placental cDNA with TFR primers (Clontech). Lane 6: contains the UPM primer negative control (no GSP). Lane 7 to 9: shows GSP negative control (no UPM). Lane 10-12: contains the GSP positive control using both 3'and 5'RACE GSP primers. Lane 13 to 15 are the positive 3' RACE PCR products i.e. Lane 13 using the primers (XHT2gsp1 + UPM), Lane 14 (XHM6gsp1+UPM) and Lane 15 (XHB1gsp1+UPM).

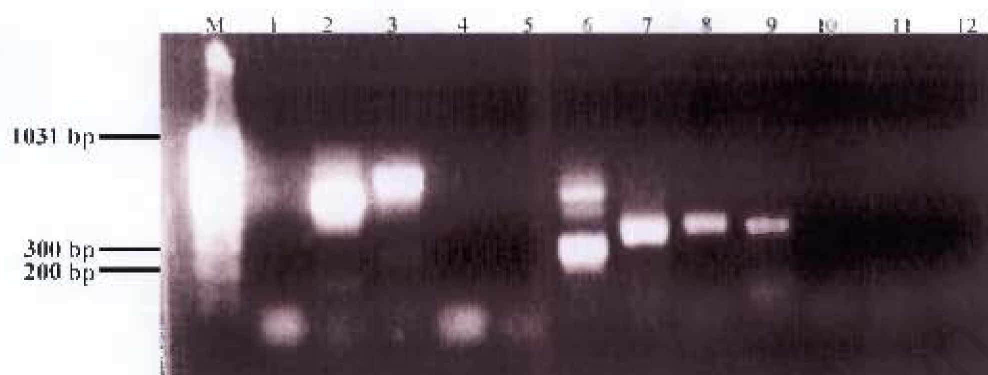


Figure 2.6: Agarose gel analysis of RACE-PCR products from leaf and seed cDNA template.

Lane M: O'GeneRuler 100 bp DNA ladder. Lane 1 to Lane 3: shows the 3' RACE PCR product from seed cDNA using XHT2gsp1, XHM6gsp1 and XHB1gsp1 respectively. Lane 4 to Lane 6: contains the 3' RACE PCR product from leaf cDNA using XHT2gsp1, XHM6gsp1 and XHB1gsp1 respectively. Lane 7 to Lane 9: shows the 5' RACE PCR product from seed cDNA using XHT2gsp2, XHM6gsp2 and XHB1gsp2 respectively. Lane 10 to Lane 12: shows the 5' RACE PCR product from leaf cDNA using XHT2gsp2, XHM6gsp2 and XHB1gsp2 respectively.

A partial sequence of a LEA1-Em orthologue (XHC00797) was identified in a cDNA library used in a concurrent microarray study on *X. humilis* (Illing and Shen, personal communication). This clone was derived from a cDNA library, and thus contained the full-length 3' sequence (Figure 2.7). The partial sequence was used to design a primer (Figure 2.7) for 5' RACE to isolate and clone full length sequence of XHC00797 using cDNA isolated from leaf, root and seed tissue.

The mRNA transcript were identified in all three tissue, seed, root and leaf (Figure 2.8), however sequencing results were only successful for the root and seed sequence at this point.

```
GCACGAGGGGGCGAGGGAGGGCGAGACCGTCGTCCCCGGCGGCACAGGCGGAAAG
AGCCTCGAAGCTCAAGAACACCTCGCTGACGGACGGAGCCGTGGAGGGCAGACGC
GCAGGGATCAGCTGGGATCAGAAGGGTACAGTGAGCTTGGCCGCATGGGCGGACA
GAGCGGCGGCTCGTATTCCGATGAGATGGCCACAGGCGGGGTGTTGGGGCCGGAT
CTCGGAGCGGGGGAGACCGTCGTACCTGGAGGCGCCGGCGGCAAGAGTGTTGAAG
CCCAAGAGAATCTCGCCAGAGGTCTGCTAGATTTGCTCTTATCTGTTACCTACTA
ATCTTTTCATGCACTCAATAACTGTTCATTAGCACAATTGCAATGGATGAAGGAC
GGAGACGCGGCGGGGAGACGCGCAGGGAGCAGCTCGGAACCGAAGGTTACAGTGC
ACTCGGCCGCACAGGGGGAAGCAGCTGAACGCCACCTATCTCCGACGAGATCGCC
GGAGAGTTGAGACACGTATTTATACGGATCTACCGCTGTAACGATGAATATATAGT
ATATAAGTATGTAAACGTAGTGTGTTTTTAGGCGTTCTGTCTTGTGTGGCTCGTT
TCTACGTACGATGTTTTGTGCTTTGCTTTCTGTGTGTGGAATAAATTAAAAGTA
```

Figure 2.7: Partial sequence of XHC00797 gene identified from microarray study on *X.humils*

The 5'RACE primer sequence is shown in a red font. The solid arrow below the primer shows its direction.

University of Cape Town

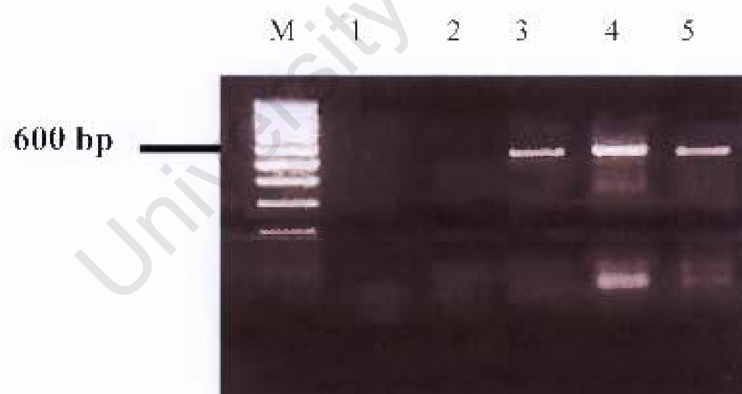


Figure 2.8: The 5' RACE amplification of XHC00797 gene from leaf, root and seed cDNA.

Lane M: O'GeneRuler 100 bp DNA ladder. Lane 1 and Lane 2: contains a negative control (no template), Lane 3 to 5: shows a PCR product obtained from leaf, root and seed cDNA respectively. The sequence for GSP used for this reaction is XHCgsp2.

The sequencing results from the 5' RACE amplification indicated that the sequence amplified during the full length amplification was different from the original XHC00797 partial sequence, even though there was high similarity. The 3'RACE amplification was done to compare the 3'UTR regions of all the sequences. The primer for the 3'RACE (highlighted in blue in Figure 2.10) was designed to amplify an amplicon including the open reading frame (ORF) and the 3'UTR (Figure 2.9). The resulting sequences from the 3'UTR showed that the 3'UTR region was conserved in all the sequences. However, the sequence information revealed that XHC000797 isolated from leaf had a different repeat pattern to XHC00797 isolated from root and seed, hence the renaming of the sequencing, XHC00797a for the seed and root sequence and XHC00797b for the leaf sequence.

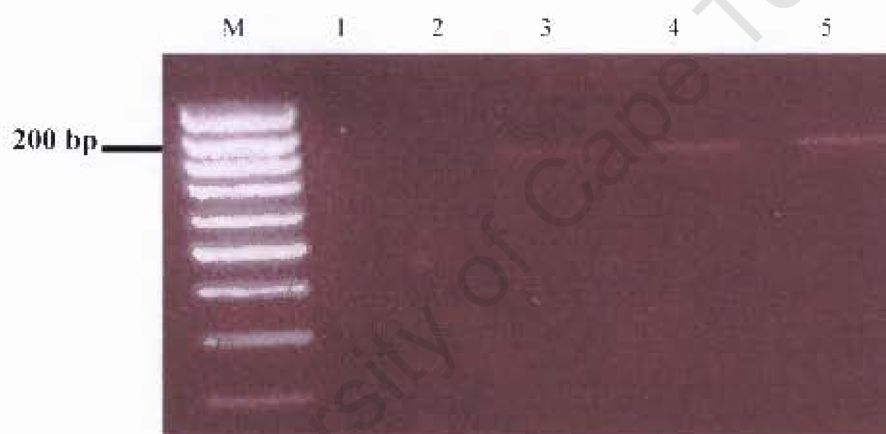
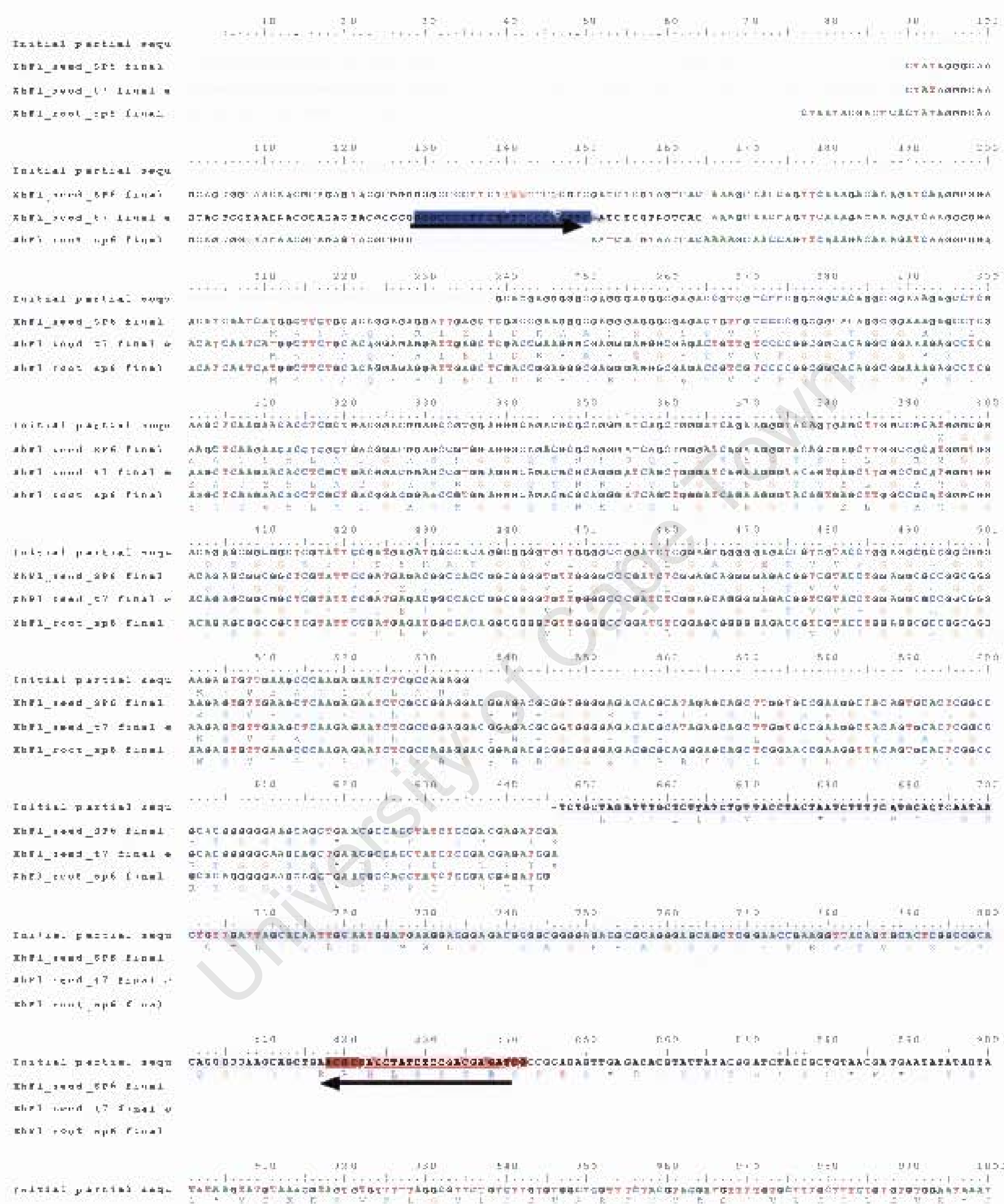


Figure 2.9 : The 3' RACE amplification of XHC00797 gene from leaf, root and seed cDNA

Lane M: O'GeneRuler 100 bp DNA ladder. Lane 1: contains a negative control (no template), Lane 2 to 4: shows a PCR product obtained from leaf, root and seed cDNA respectively. The sequence for GSP used for this reaction is XHCgsp1.

All fragments were cloned into pGEM-T Easy vector and the insert fragments were sequenced. The sequence information was obtained by sequencing in both the forward and reverse direction. The resulting sequence is summarized in Table 2.6 with reference to the tissue from which it was isolated.

University of Cape Town



University of Cape Town

Figure 2.10: Alignment of XHC00797 partial sequence with 5'RACE PCR products

The position of the 5' RACE and 3' RACE primer are highlighted in red and blue respectively. The primer direction is indicated by arrows below the primer sequences.

Table 2.6: Summary of RACE PCR sequencing results.

	Tissue Type	5' RACE sequence	3'RACE sequence
XHLEA1-1	Seed	✓	☒
	Root	☒	☒
	Leaf	☒	☒
XHLEA1-2	Seed	✓	✓
	Root	☒	✓
	Leaf	☒	☒
XHLEA1-4	Seed	✓	✓
	Root	✓	✓
	Leaf	☒	✓
XHC00797	Seed	✓	✓
	Root	☒	✓
	Leaf	✓	✓

* the tick indicate successful cloning and sequencing from the specific tissue and the cross indicatesunsuccessful cloning

2.3.4 XHLEA1-1 Sequence

The initial RACE PCR for cloning XHLEA1-1 resulted in a single sequence from the 5'RACE using seed cDNA. The sequence from the 5'RACE was used to design primers to amplify the full length sequence and the 3'UTR region. The mRNA transcripts for the full length 3'RACE PCR was only identified from seed cDNA (Figure 2.11). The sequence had one 20 amino acid motif repeat used to classify the group 1 LEAs (Figure 2.12). The coding region encodes a polypeptide of 92 amino acids with a predicted molecular of 9.91 kDa and a predicted pI of 5.67 (DNAMAN).

University of Cape Town

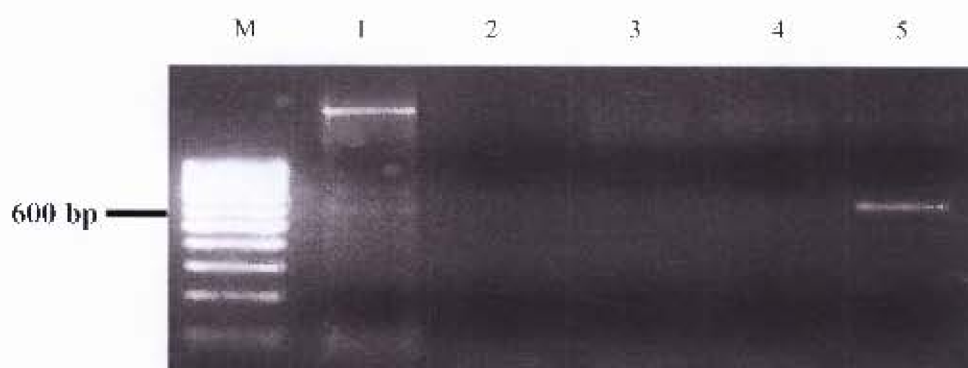


Figure 2.11: Cloning of XHLEA1-1 full length open reading frame (ORF) sequence using 3' RACE ready cDNA.

Lane M: O'GeneRuler 100 bp DNA ladder. Lane 1: shows a positive control using the human placental cDNA while Lane 2: contains a negative control (no template). Lane 3 to 5: shows a XHLEA1-1 full length ORF PCR product obtained from leaf, root and seed cDNA respectively.

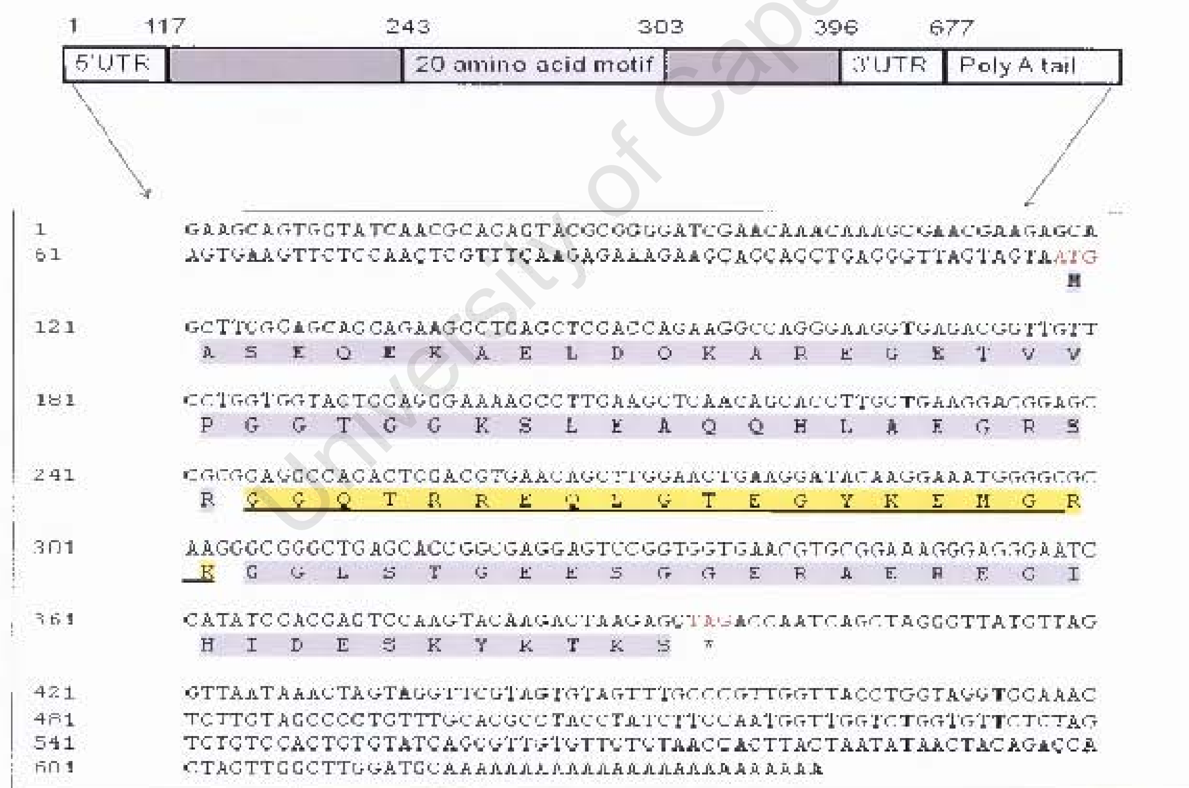


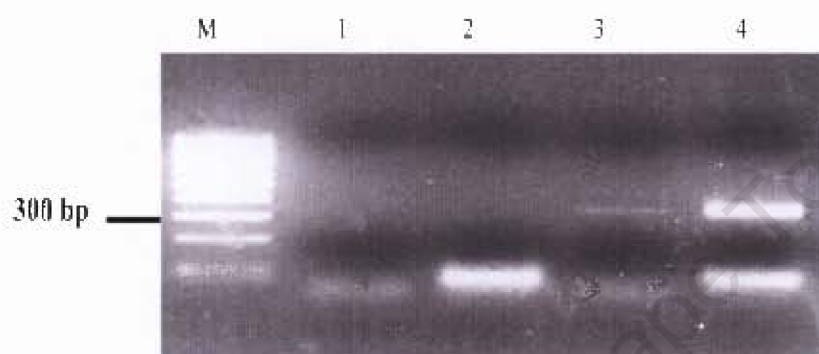
Figure 2.12: Primary structure of the full length XHLEA1-1 gene.

The key amino acid motif present in XHLEA1-1 gene together with its sequence are shown in the schematic above. The ORF is highlighted in grey with the putative start and stop codon.

University of Cape Town

2.3.4 XHLEA1-2 Sequence

XHLEA1-2 ORF full length sequence was isolated, using two GSP, from both seed and root tissue (Figure 2.13). The sequence has two of the 20 amino acid motif repeats used to classify the group 1 LEAs (Figure 2.14). The coding region encodes a polypeptide of 112 amino acids with a predicted molecular of 12.01 kDa and a predicted pI of 5.03 (DNAMAN).



University of Cape Town

Figure 2.13: Cloning of XHLEA1-2 full length sequences using gene specific primers FLXHM6gsp1 and FLXHM6gsp1 (Table 2.4).

Lane M: O'GeneRuler 100 bp DNA ladder. Lane 1: contains a negative control (no template). Lane 3 to 5: shows a XHLEA1-2 full length ORF PCR product obtained from leaf, root and seed cDNA respectively.

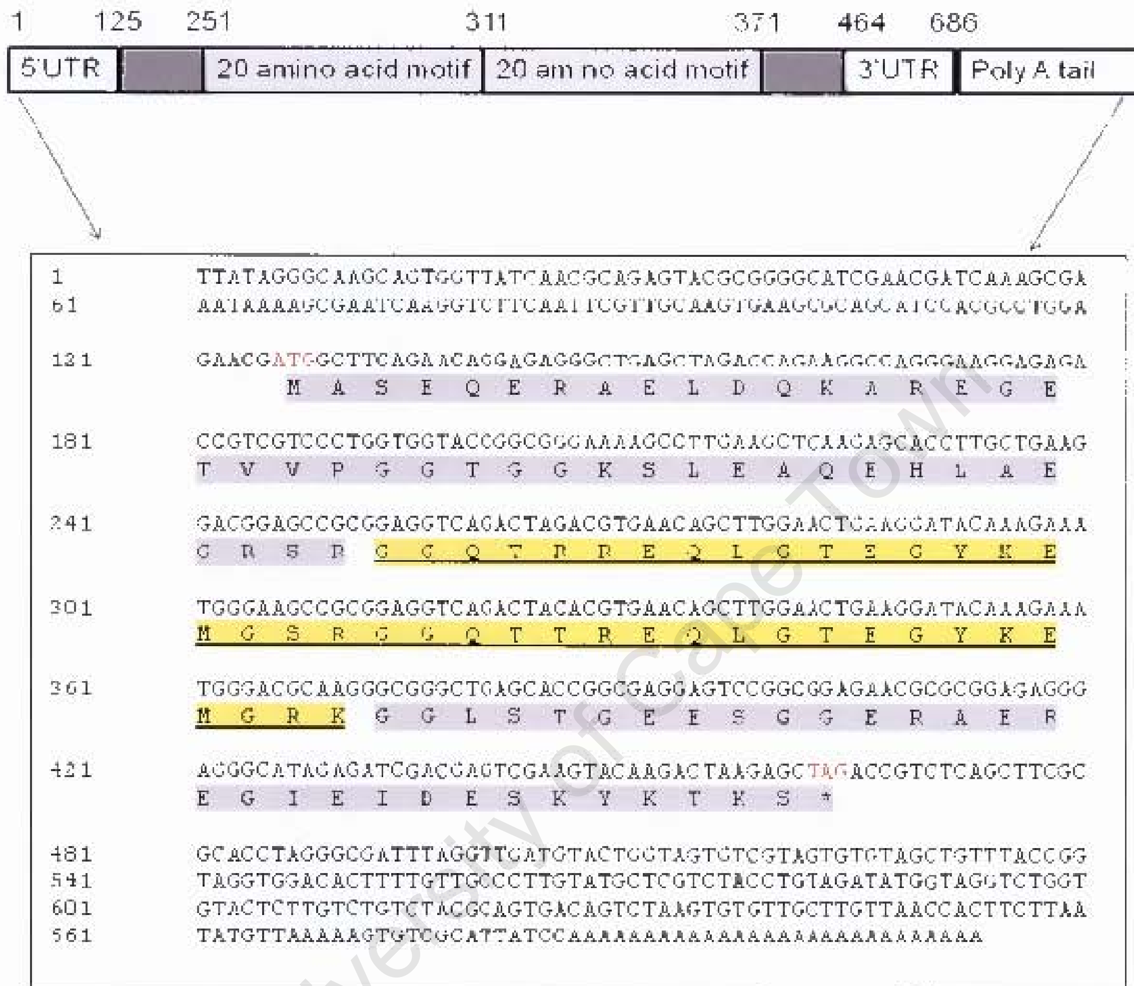


Figure 2.14: Primary structure of the full length XHLEA1-2 gene.

The key amino acid motif present in XHLEA1-2 gene together with its sequence are shown in the schematic above. The ORF is highlighted in grey with the putative start and stop codon presented in red and the 20 amino acid motif is repeated four times which is highlighted in yellow and underlined.

2.3.5 XHLEA1-4 Sequence

For XHLEA1-4, 3' RACE sequences were isolated from all three plant tissues; leaf, seed and root. The 5' RACE sequences were identified from seed and root tissue. However, the mRNA for the isolation of the full length ORF was identified from root and seed tissue (Figure 2.15). The combined sequence gives a full length cDNA sequence of 841 bp (Figure 2.16). The sequence has four of the 20 amino acid motif repeated in tandem (Figure 2.16). The coding region encodes a polypeptide of 152 amino acids with a predicted molecular of 16.58 kDa and a predicted pI of 5.75 (DNAMAN).

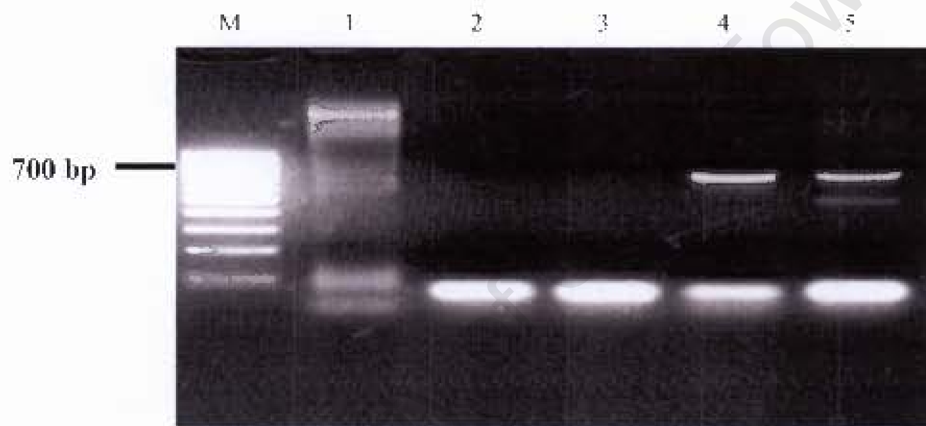


Figure 2.15: Cloning of XHLEA1-4 full length ORF sequence using 3' RACE ready cDNA and the FLXHB1gsp1 primer.

Lane M: O'GeneRuler 100 bp DNA ladder. Lane 1: shows a positive control using the human placental cDNA while Lane 2: ,contains a negative control (no template). Lane 3 to 5: shows a XHLEA1-4 full length ORF PCR product obtained from leaf, root and seed cDNA respectively

University of Cape Town

```
1        133     259                                        499     592      840
| 5'UTR |        | 20 aa motif | 20 aa motif | 20 aa motif | 20 aa motif |        | 3'UTR | Poly A tail |
```

```
1     AATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGAAGAAGCTG
121   GTTGTGAGCAGGCATGGCTTCCCATCAAGAAAGGGCTGAGATGGACCGTAGGGCCAGGGA
                     M  A  S  H  Q  E  R  A  E  M  D  R  R  A  R  E

181   AGGTGAGACTGTCGTCCCCGGTGGTACCGGCGGAAAGAGCCTCGAAGCTCAACAACACCT
61     G  E  T  V  V  P  G  G  T  G  G  K  S  L  E  A  Q  Q  H  L

241   CGCCGAAGGTCGTAGCAAGGGAGGGCAGACGAGGAAGCAACAGTTGGGGACAGAGGGGTA
81     A  E  G  R  S  K  G  G  Q  T  R  K  R  Q  L  G  T  E  G  Y

301   TCAGGAGATGGGCCACAAAGGCGGTGAGACGAGGAGGGAACAGCTGGGTTCAGAAGGGTA
101    Q  E  M  G  H  K  G  G  E  T  R  R  E  Q  L  G  S  E  G  Y

361   CAGTGAAATCGGGCACAAAGGAGGGGAGGCGAGGAGAGAACAACTGGGTTCCGAAGGGTA
121    S  E  I  G  H  K  G  G  E  A  R  R  E  Q  L  G  S  E  G  Y

421   CAGTGAAATCGGGCACAAAGGTGGTGAGACGAGGAGGGATCAGATTGGGTCGGAGGAATA
141    S  E  I  G  H  K  G  G  E  T  R  R  D  Q  I  G  S  E  E  Y

481   CCGTGAGATGGGTCGCAAAGGTGGTCTGTCGACCAAGGACGAGTCGGGTGGAGAAACGTGC
161    R  E  M  G  R  K  G  G  L  S  T  K  D  E  S  G  G  E  R  A

541   TGCCCGGGAGGGCATAGAGATCGACGAGTCCAAGTACAAGACCAATGTTTAAGGTCTTGC
181    A  R  R  G  I  E  I  D  E  S  K  Y  K  T  N  V  *

601   GACGTTCTTAGTTCGTGGTCTTTAGTAGTTGTGTGTCGTGTCGTGTATGGTATGTGTGTT
661   AAGATGGTGCTGTTACGATTTTGTGAAGGCCAGCTTGGTCTCGACTCTTCTGGGAGTCGT
721   ATGATAGTTGAGTAACCAAGCACCAGGTAGGTTTGTGTCTTGAAGTTGTTTATGTGCTTT
781   GTTTTAGATCATCTGGTACGTTAATGGAAGTAAATGGCTCTAGCTATTTTAATATGGTGA
841   AAAAAAAAAAAAAAAAAAAAAAAAA
```

Figure 2.16: Primary structure of the full length XHLEA1-4 gene.

The key amino acid motif present in XHLEA1-4 gene together with its sequence are shown in the schematic above. The ORF is highlighted in grey with the putative start and stop codon presented in red and the 20 amino acid motif is repeated four times which is highlighted in yellow and underlined.

University of Cape Town

2.3.6 XHC00797a Sequence

XHC00797a full length sequence, was isolated from both seed and root tissue. The 20 amino acid motif (indicated in grey in Figure 2.17) used to classify the group I LEAs is repeated twice in this sequence. There is an atypical repeat pattern in this sequence, as another amino acid motif is present. The second amino acid repeat of 26 amino acids (indicated in red in Figure 2.17) is also repeated twice in the sequence. The coding region encodes a polypeptide of 136 amino acids with a predicted molecular weight of 13.71 kDa and a predicted pI of 5.11 (DNAMAN). The combined sequence gives a full-length cDNA sequence of 681 bp (Figure 2.18).

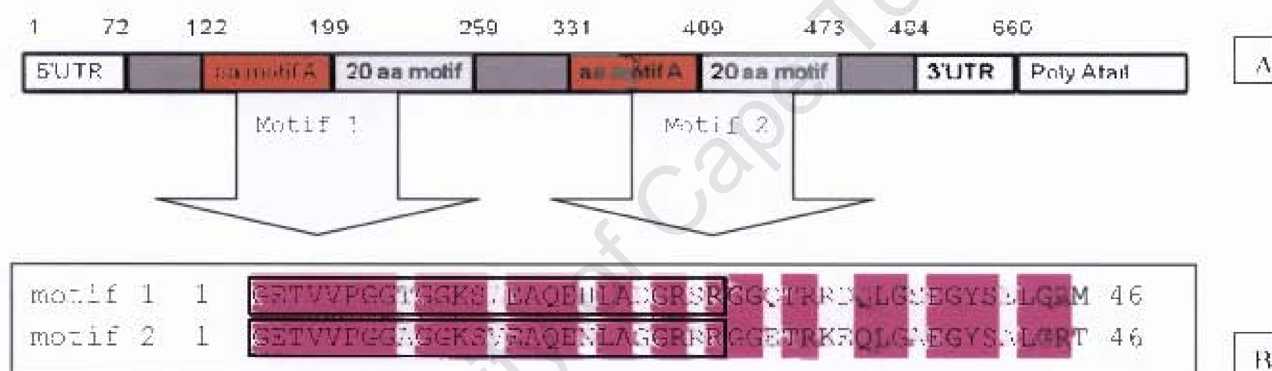


Figure 2.17: (A) Primary structure of the full length XHC00979a, and (B) amino acid sequence alignment of the two motifs.

The typical 20 amino acid motif of LEA1-EM group is highlighted in grey. The amino acid motif highlighted in red indicates a unique motif for this group of LEAs. The sequence of the unique motif is boxed in panel B. The conserved amino acid among the two motifs is highlighted in pink.

University of Cape Town

```
1        TTCTTTTGCCTGGTGGATCTGGTAGTCACAAAGCAACCAGTTCAAAGACAAAGATCAAGG
61       GGCAACATCAATCATGGCTTCTGCACAGGAGAGGATTGAGCTCGACCGAAGGGCGAGGGA
                       M  A  S  A  Q  E  R  I  E  L  D  R  R  A  R  E

121      GGGCGAGACTGTTGTCCCCGGCGGCACAGGCGGAAAGAGCCTCGAAGCTCAAGAACACCT
           G  E  T  V  V  P  G  G  T  G  G  K  S  L  E  A  Q  E  H  L

181      CGCTGACGGACGGAGCCGTGGAGGGCAGACGCGCAGGGATCAGCTGGGATCAGAAGGGTA
           A  D  G  R  S  R  G  G  Q  T  R  R  D  Q  L  G  S  E  G  Y

241      CAGTGAGCTTGGCCGCATGGGTGGACAGAGCGGCGGCTCGTATTCCGATGAGACGGCCAC
           S  E  L  G  R  M  G  G  Q  S  G  G  S  Y  S  D  E  T  A  T

301      CGGCGGGGTGTTGGGGCCCGATCTCGGAGCAGGGGAGACGGTCGTACCTGGAGGCGCCGG
           G  G  V  L  G  P  D  L  G  A  G  E  T  V  V  P  G  G  A  G

361      CGGGAAGAGTGTTGAAGCTCAAGAGAATCTCGCCGGAGGACGCAGACGCGGTGGGGAGAC
           G  K  S  V  E  A  Q  E  N  L  A  G  G  R  R  R  G  G  E  T

421      GCGCAAAGAGCAGCTTGGTGCCGAAGGCTACAGTGCACTCGGCCGCACGGGGGGAAGCAG
           R  K  E  Q  L  G  A  E  G  Y  S  A  L  G  R  T  G  G  S  S

481      CTGAACGCCACATATCTCCGACGAGATCGCCGGAAAGTTGAGACACGTATTATGCGGATC
           *

541      TACAGCTGTAACGATGAATATATATTTATATAAGTATGTAAACGTAGTGTGTTTTTAGGCG
501      TTCTGTCTTGTGTGGCTCGTTTCTACGTACGACGTTTTGTGCTTTACTTTCTCTGTGTGG
561      AATAAATTAAAAGCATATTACAAAAAAAAAAAAAAAAAAAAAAAAAA
```

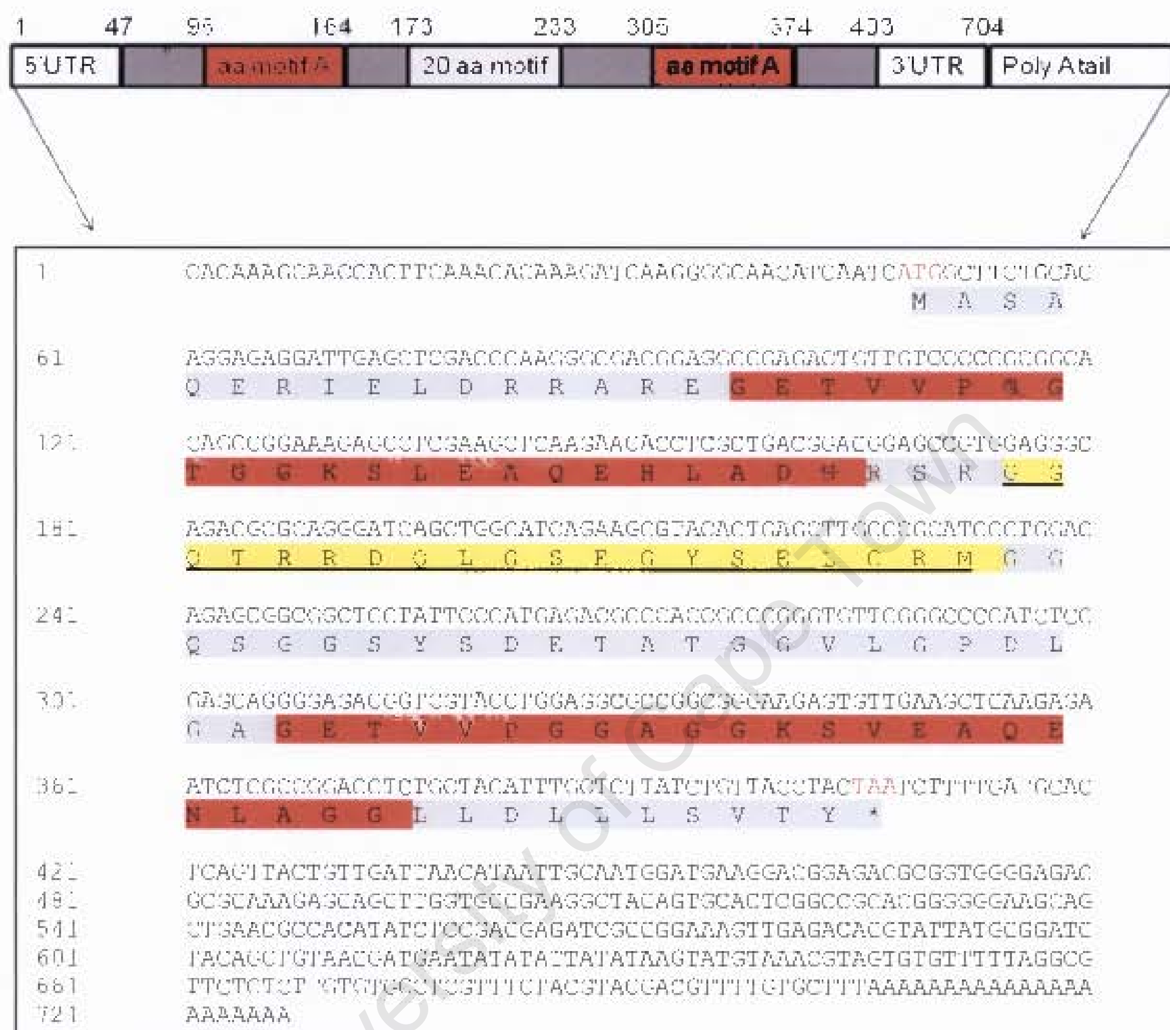
Figure 2.18: Primary sequence of the full length XHC00797a gene.

The key amino acid motif present in XHC00797a gene together with its sequence are shown above. The ORF is highlighted in grey. The two motifs (Figure 2.15) are highlighted in red and yellow. The typical 20 amino acid motif of the LEA1-Em group is highlighted in yellow and underlined.

University of Cape Town

2.3.7 XHC00797b Sequence

The XHC00797b full length sequence was isolated from leaf tissue. The 20 amino acid motif repeat used to classify the group I LEAs is only present once in the sequence. However, the 23 amino acids motif (GETVVPGGTGGKSLEAQEHLADG) is repeated twice in the sequence. The combined sequence gives a full-length cDNA sequence of 705bp (Figure 2.19). The coding region encodes a polypeptide of 119 amino acids with a predicted molecular of 12.02 kDa and a predicted pI of 4.43 (DNAMAN).



University of Cape Town

Figure 2.19: Primary structure of the full length XHC00797b gene.

The key amino acid motifs present in XHC00797b gene together with its sequence are shown in the schematic above. The ORF is highlighted in grey with the putative start and stop codon presented in red. The 20 amino acid motif is repeated four times and is highlighted in yellow. The newly identified motif is highlighted in red.

2.4 Discussion

The five cDNA clones of LEA1-Em were obtained using degenerate and RACE PCR. These clones were between 625 and 865 base pairs, with a 67 – 97 % identity with each other (Figure 2.20). The clones represented four paralogues of LEA1-Em genes present in *X. humilis*. All isolated sequences possess the conserved 20 amino acid motif that is characteristic of group 1 LEAs.

2.4.1 Sequence Analysis

BLAST results showed a percentage similarity of 56 and 77 % between LEA1 genes isolated from *X.humilis* and those from *Arabidopsis*. However, multiple sequence alignment (Figure 2.21) revealed the difference in the pattern of the repeated amino acid motif. It was concluded that XHLEA1-1,2 and 4 can be grouped together as they only have the 20 amino acid motif (highlighted in grey and underlined in Figure 2.21) typically associated with the LEA1-Em group. However, XHC00797a and XHC00797b, have an additional amino acid repeat sequence (highlighted in red in Figure 2.21), which is unique for this group. The repeat is 26 amino acids long in XHC00797a and 23 amino acids long in XHC00797b. The overall sequence similarity between the two transcripts is very high. (Figures 2.20 and 2.21).

University of Cape Town

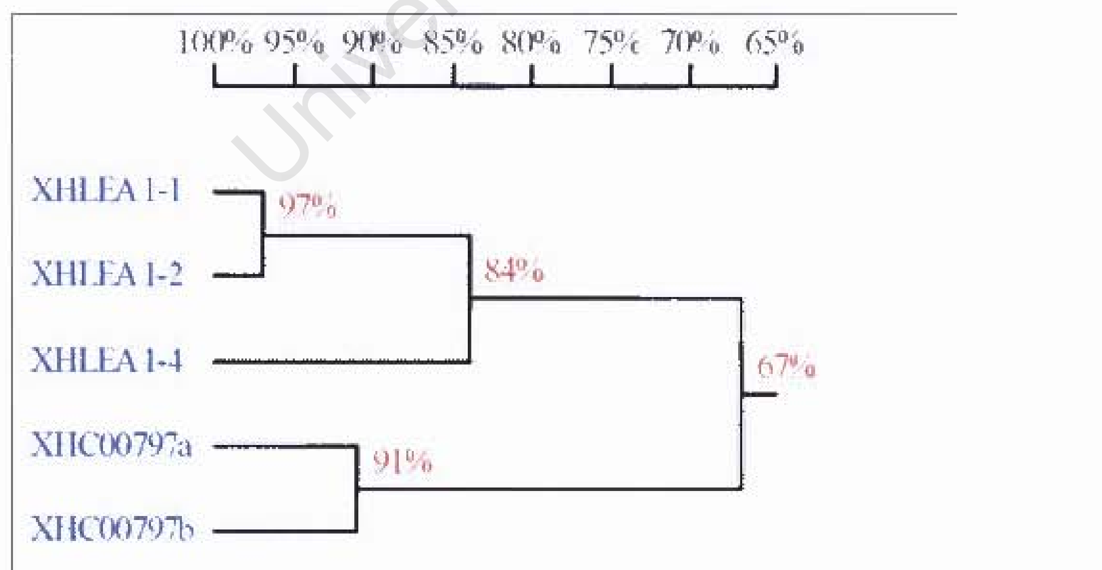
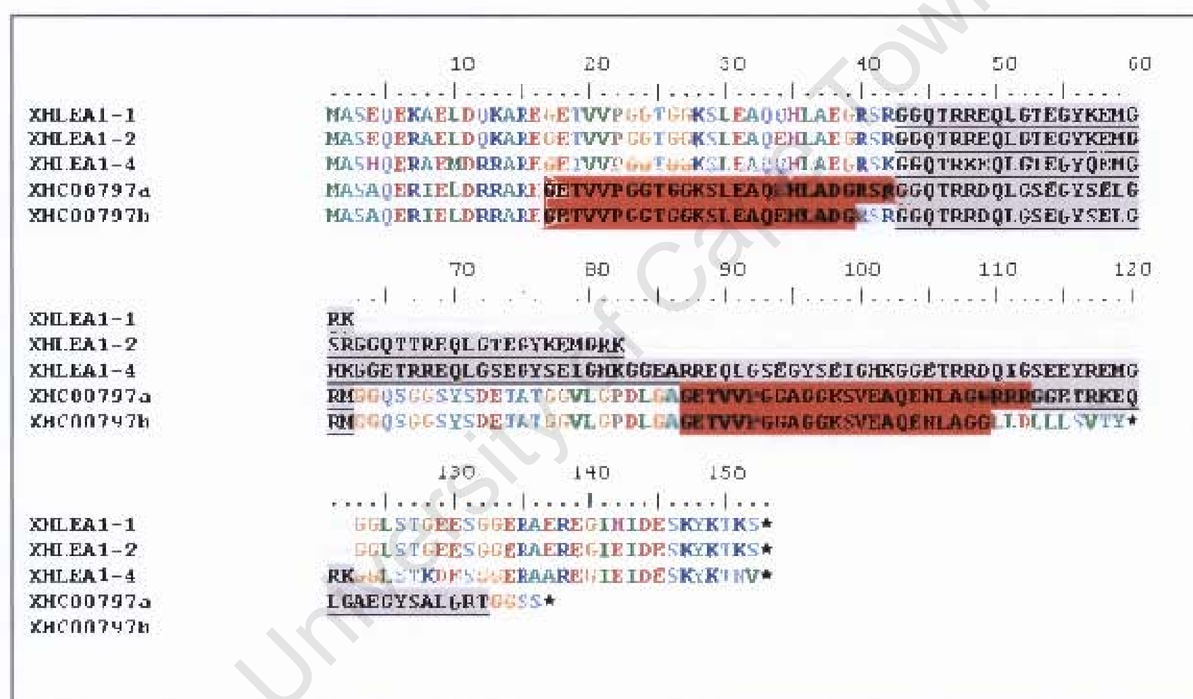


Figure 2.20: Sequence homology tree of LEA1-Em sequences from *X. humilis*.

XHLEA1-1, 2 and 4 are very similar to the *Arabidopsis* LEA1-Em orthologues. Their sequences are highly conserved, similar to the *Arabidopsis* LEA1s containing one (ATEM6) and four (ATEM1) repeats of the 20 amino acid motif repeat. A blast search of sequences in the GenBank database identified other dicots with similar sequences. For example the dicot *Robinia pseudoacacia*, commonly known as the black locust, two LEA1-Em containing one and two of the 20 amino acid motif repeats respectively. All the other LEA's on GenBank including the ones in listed in Table 2.1 are highly conserved.



University of Cape Town

Figure 2.21: Alignment of the five LEA1-Em amino acid sequences isolated from *X. humilis*.

The amino acids highlighted in grey are the typical 20 amino acid motif associated with LEA1-Em group while the red highlighted sequence is the newly identified motif.

XHC00797a and XHC00797b clones are different from any other LEA1 gene in GenBank. This sequence is classified together with the LEA1 as it possesses the 20 amino acid motif characteristic of the LEA1-Em group (highlighted in grey in Figure 2.21) but it also contains an

additional repeated sequence which has not been identified before. No similar repeat structure has been identified from the monocot rice, as well as the dicot *Arabidopsis* genome. Both these clone are most likely a result of alternative splicing of the gene.

2.4.2 Alternative splicing of mRNA transcript XHC00797a and XHC00797b

Alternative splicing is less documented in plants as opposed to mammals. The result of alternative splicing is that diverse forms of mRNA and translation products can be a product of one gene. It can be constitutive or under different developmental stages, cell differentiation and physiological conditions. Recently, bioinformatics has also suggested that approximately 4 % of barley genes (Zhang *et al*., 2004), 11 % of *Arabidopsis* (Iida *et al.*, 2004) and 10 % of rice genes (Kikuchi *et al*., 2003) undergo alternative splicing. It is then possible that XHC00797a and XHC00797b are the product of alternative splicing of the same primary gene transcript.

The alternative splicing of XHC00797 seems at this point to be tissue specific as XHC00797a was isolated from seed and root while XHC00797b was isolated from leaf. Factors that have been noted to affect alternative splicing in plants include cell cycle, various tissue developmental stages, cultivars and pathways of carbon assimilation. However, it must be noted that some alternative splicing are constitutive with similar ratios of variant mRNA in different cells, whereas others are alternatively spliced dependent on tissue and developmental stages (Gue *et al.*, 2007). It is also possible that the cell type and age can affect the splice site selection and the relative concentration of the mRNA transcript in different tissues and developmental stages (Zhang *et al.,* 2003).

Alternative splicing also contributes to genome complexity and protein diversity since one gene can produce several proteins which are structurally and/or functionally different or identical, or inter-inhibiting. This could be beneficial to the vegetative tissue of resurrection as the plants is continuously adapting to various degrees of water availability.

Alternative splicing has also been demonstrated to regulate the sublocalisation of a protein within an organelle. This has been demonstrated by de la Fuente van Bentem *et al.,* 2003 on

University of Cape Town

tomato's LePP5 gene which after splicing produces a 55 kDA LEPP5 which is localized in both the nucleus and the cytoplasm, while the 62 kDa isoform is targeted to the endoplasmic reticulum, including the nuclear envelope. Localisation of the LEA1-Em proteins in desiccated tissue will lead to some indication whether this is the case with the XHC00797a and XHC00797b transcripts.

Connections have been found between alternative splicing and some human diseases (Caceres and Kornblihtt, 2002), but whether alternate splicing of mRNA contributes to some abnormal phenotype or susceptibility of plants is still unknown (Gue *et al.*, 2007).

2.5 Conclusion

Following the completion of the isolation and cloning of the LEA1-Em from *X. humilis*, research focus was shifted to the expression of LEA1-Em in vegetative tissue in response to dehydration. Messenger RNA expression analysis of the five isolated LEA1-Em genes, was done to further characterize the LEA1-Em in vegetative tissue. From the cloning study, the mRNA transcripts were detected in seed as well as vegetative tissue either root or leaf for all clones. To elucidate whether the LEA1-Em expression is responding to dehydration, RT-PCR experiments were done and are discussed in the subsequent chapter.

University of Cape Town

3 Changes in mRNA transcript abundance of *X. humilis* LEA1-Em during dehydration and rehydration

3.1 Introduction

The real-time reverse transcription polymerase reaction (RT-qPCR) is a very sensitive technique that is used to quantify mRNA transcript abundance from very small amounts of RNA even at the level of the content of a single cell (Freeman *et al.,* 1999). Nolan *et al.* (2006) summarises RT-qPCR as a combination of three steps: (i) the reverse transcriptase (RT) - dependent conversion of RNA into DNA, (ii) the amplification of the cDNA using PCR and (iii) the detection and quantification of amplified products in real time. This technique combines the amplification and detection steps of the PCR reaction in one tube by using fluorescent reporter dyes. For each PCR cycle the assay measures the increase in fluorescent signal which is proportional to the amount of DNA produced in that cycle. The reactions are characterised by the threshold cycle (Ct), which is the PCR cycle at which fluorescence first rises above a defined or threshold background fluorescence. It is the correlation between the fluorescence and the amount of amplified product that allows for accurate quantification of target molecules, while retaining the sensitivity and specificity of conventional PCR assays (Nolan *et al.,* 2006).

There are however some considerations to be made in order to get quantitative data that is reliable and reproducible. The main things to consider are (i) the quality and accurate measurement of the starting RNA material and starting cDNA concentrations (ii) the presence of inhibitory components in the samples (iii) the reverse transcriptase reaction and the priming of the cDNA and finally (iv) the proper optimisation of the PCR.

3.1.1 Quality and measurement of the RNA and starting cDNA

RNA is rapidly digested by ubiquitous RNase enzymes, resulting in shorter fragments of RNA which can potentially compromise RT-PCR results. RNA integrity has been previously evaluated by using agarose gel electrophoresis stained with ethidium bromide, which produces a characteristic banding pattern (Sambrook *et al.,* 1989).

University of Cape Town

The banding pattern usually comprises of a 28S, 18S. RNA is interpreted to be of good quality when the 28S: 18S band ratio is 2:1. The limitation of this approach is its subjectivity as it relies on human interpretation of relative intensities from photographs on RNA gels.

An alternative approach is now available which uses the Agilent 2100 bioanalyser for the separation of RNA samples. The instrument is a bioanalytical device using microfluidics technology that provides electrophoretic separations in an automated and reproducible manner (Mueller *et al.*, 2000). Using a system like the bioanalyser allows RNA samples of different quality to be distinguishable by examining electrophoretic traces and assigning integrity values or integrity categories (Imbeaud *et al.,* 2005). The bioanalyser fulfils an important need for a reliable, reproducible, and standardised approach to classify the quality of RNA samples, because of the critical influence RNA integrity has on RT-PCR. The discrepancies that arise during cDNA synthesis also need to be considered. The ideal situation is when the cDNA synthesis is done in a single batch to avoid any potential variability caused by differences in the efficiency of target and reference gene transcription (Whelan *et al.*, 2003), as was the case in this study.

3.1.2 Inhibitory components

Inhibitory components can negatively affect the sensitivity and kinetics of the RT-PCR (Radstrom *et al.*, 2004; Guy *et al.,* 2003). The inhibitors can possibly be traces of reagents that are used during nucleic acid extraction or co-purified components from the biological sample like polyphenolics in plant tissues or bile and urea in animal tissues. Inhibitors can lead to inaccurate results of RT-PCR or create a false negative (Nolan *et al.*, 2006).

If the PCR is to be used for quantitative purposes the presence of inhibitors in the test sample may lead to the underestimation of mRNA in the test sample, especially in cases where the standard curve is not generated with the same sample (Nolan *et al.,* 2006). The assessment of inhibitors in a sample can be assessed by a serial dilution of the sample and then comparing the PCR efficiency (Stahlberg *et al.,* 2003). Alternatively, one can use mathematical algorithms that can provide PCR efficiency (Tichopad *et al.*, 2003; Liu *et al.,* 2002). Yet, another way to detect inhibitors and also indicate template loss during the processing is to use internal amplification controls (IAC) that co-purify and co-amplify with the target nucleic acid (Pasloske *et al.,* 1998).

University of Cape Town

3.1.3 The reverse transcriptase reaction

The efficiency of the reverse transcription directly affects the accuracy of RT-PCR. The efficiency depends on the target and the choice of reverse transcriptase. For comparative purposes it is important to use the same enzyme and priming strategy (Stahlberg *et al.,* 2004). Both oligo-dT and random priming allow a representative pool of cDNA to be produced during a single reaction (Nolan *et al.*, 2006). However it has been shown that in priming using random hexamer primers, not all targets in the sample are efficiently transcribed and there is no linear correlation between input target and cDNA yield when specific targets are measured (Buston and Nolan, 2004; Lacey *et al.,* 2005). Alternatively, oligo-dT priming should be used on intact RNA, with polyA sequences, to enable a more representative transcription of cDNA.

3.1.4 Optimising the PCR reaction

The different primer sequences and concentration will affect the binding of the primer to the target DNA. It is therefore very important to use primers at a concentration that results in optimal hybridisation and priming. Accurate identification of amplified products and distinguishing them from primer dimers can be accurately done by first checking results of PCR reaction on an agarose gel, and subsequently by checking the melting curve of the real-time PCR reaction. The temperature where half of the DNA helical structure is lost is called the melting temperature (Tm). It depends on the size and nucleotide composition. The use of fluorescent dye allows for the calculation (-dF/dT) of the melt curve. These peaks are equivalent to the bands on the electrophoresis gel (Nolan *et al.*, 2006). Longer target amplicon products will melt at a higher temperature than shorter primer dimers.

3.1.5 Normalisation of gene expression

To correct for sample-to-sample variation, measurements of mRNA transcript abundance must be normalised. Starting material will usually differ in tissue mass or cell number, RNA integrity or quantity or experimental treatment. As a result, real time PCR results are usually normalised against a control gene. The mRNA transcript abundance of the ideal control gene should be unchanged throughout the different experimental conditions.

University of Cape Town

Housekeeping genes are traditionally used as controls in gene expression, as they are thought to have stable expression. It is however necessary to validate the stability of a housekeeping gene in the experimental conditions. Ribosomal RNA (rRNA) can sometimes be used as housekeeping genes, the main limitation being that rRNAs are transcribed with a different polymerase than mRNA, so this might introduce variability as changes in polymerase activity may not affect both types of RNA expression equally (Spanakis, 1993). Total RNA concentration can be use to normalise gene expression. There are however problems associated with this approach as it does not correct for differences in RNA quality and reverse transcriptase efficiency. Normalisation is also dependent on the accuracy of the RNA quantification.

Real-time PCR was used to quantify changes in mRNA transcript abundance of *X. humilis* LEA1 genes. The above conditions were considered when designing the RT-PCR experiment. To ensure the quality and quantification of the starting RNA the Bionalyser was used in addition to the denaturing gel analysis. The RNA from all the biological samples were pooled for the reference sample, and this sample was used to generate the standard curve to minimise variation that might be due to the presence of any inhibitory components from the extraction method. The reverse transcription reaction was standardised for all reactions. This was accomplished by using the same reverse transcriptase enzyme and priming with oligo dT primers to allow for transcription of a representative pool of cDNA. Furthermore, sample-to-sample variation of mRNA transcript abundance was normalised using a housekeeping gene, R1, that has been previously shown to be relatively unchanged throughout the dehydration and rehydration cycle of *X. humilis* (Walford, unpublished data).

University of Cape Town

3.2 Materials and Methods

3.2.1 Plant material and growth conditions

Mature *X. humilis* plants were collected and grown as described in Chapter 2. Prior to abiotic stress treatment and sampling, the plants were equilibrated to the growth room conditions room [16 hr light (350 μmol.m^2.s^{-1}); 8 hr dark) for two weeks.

3.2.2 Abiotic treatment of plants

The experiment was conducted on 3 trays (60 cm X 30 cm) of plants, each containing 15 - 25 plants, sampled randomly. Twenty leaves were collected at the same time everyday for two weeks until plants were fully dehydrated (~5 % RWC). For rehydration, the same number of leaves was sampled every six hours until the plant was fully hydrated (~100 % RWC). Leaf samples were prepared by splitting individual leaves in half down the midrib, the one half being used to determine the RWC while the other half for RNA extraction was wrapped in aluminium foil, and quickly frozen in liquid nitrogen. RWC was determined using the formula: RWC= [Absolute water content/ absolute water content at full turgor] x 100 %. The absolute water content was calculated using the formula: Absolute water content = [(FW-DW)/DW] where FW is the fresh weight and DW is the dry weight.

The leaf sample in liquid nitrogen was stored at -80 °C until RNA extractions were performed. This method allows for accurate measurement of RWC for each leaf sample. Leaf samples were collected at full turgor and were pooled into the following ranges thereafter 80 %, 60 %, 40 - 45 %, 10 - 15 %, 5 %, according to the measurements of individual RWCs. Rehydration samples were pooled together at the following RWC 25 – 35 %, 40 % and 100 %. Root samples were sampled at full turgor and when fully desiccated, 5 % RWC.

3.2.3 RNA Extraction, RNA quantification and reverse transcription

Total RNA was extracted from three half leaves (approximately 100 mg) of plant material using TRI-REAGENT as described in Chapter 2, section 2.2.3. The same mass of seed and root tissue was used for RNA extraction. The concentration and integrity of RNA was analysed using the Agilent 2100 bioanalyser (Agilent Technologies, USA). To remove contaminating DNA, the RNA was treated with DNase I (Ambion, USA). 2.5 µg of total RNA was incubated at 37 °C for one hour in the presence of 4 µl of DNaseI buffer, 4 U of DNaseI and 0,5 µl of RNase inhibitor in a total volume of 20 µl. The DNaseI enzyme was inactivated by incubating the sample at 75°C for 10 minutes.

University of Cape Town

For first strand cDNA synthesis, the 20 µl reaction from the DNaseI treatment was incubated overnight at 50 °C in the presence of 1 µl of 50 µM oligo(dT) primer, 2 µl of 10 mM dNTP mix, 8 µl 5XFirst Strand (FS) Buffer, 2 µl of 0.1 M DTT, 1 µl of RNAse inhibitor (Invitrogen, USA), and 1 µl Superscript RT III (Invitrogen, USA). The cDNA synthesis reaction was inactivated by incubation at 70 °C for 15 min.

3.2.4 Primer design, primer concentration optimisation

All the primers used for SYBR real-time RT-PCR are listed in Table 3.1. The LEA1 primers were based on the 3' UTR sequence of the cloned cDNA of LEA1-Em genes to minimize cross-complementarity between the LEA1 genes. The primers were designed using manual inspection. Primers were also aligned using DNAMAN to check for self - complementarity and primer dimer formation and sequence specificity. The primers were synthesised in the Department of Molecular and Cell Biology Oligonucleotide Synthesis Unit.

The R1 gene encoding for a *X. humilis* secretory carrier membrane protein, was used as a constitutively expressed control gene to normalize starting cDNA template concentrations, Primers for the R1 gene were provided by Sally-Ann Walford (personal communication) and are listed in Table 3.1.

University of Cape Town

Table 3.1: Primer sequences for real-time PCR

Gene	Primer Sequences (5'- 3')	
R1 (housekeeping gene)	5'CCATGTACGCGATGCTTCTA 3'	F
	5'AGCGTGTGTAAAGTCATCCTG 3'	R
XHLEA1-1	5'GGAATCAGCTAGGGTTATGTTAGGT 3'	F
	5'TAGAGAACACCAGACCAACCATT 3'	R
XHLEA1-2	5'GCGATTTAGGTTGATGTACTGGTAGT 3'	F
	5'GGGTTTAACATATTAAGAAGTGGTTAACAA 3'	R
XHLEA1-4	5'GCTTGCGACGTTCTTAGTTCGT 3'	R
	5'GATAAACAACTTCAAGACACAAACCT 3'	R
XHC00797	5'CGGATCTACCGCTGTAACGAT3'	F
	5'GGGTACTTTTAATTTATTCCACACAG3'	R

3.2.5 Real-time (RT)-PCR

Quantification of mRNA transcript abundance was measured by real time PCR, using a Corbett Rotor Gene 6000 real time rotary analyser (Corbett, Australia). Detection of real-time RT-PCR products was done by using the SYBR®Green SensiMix kit (Celtic Molecular Diagnositcs, South Africa). The total volume of the PCR reaction was 25 µl, containing 5 µl cDNA template, 1 µl primers (10 µM), 12.5 µl 2X SensiMix, 0.5 µl 50X syber green and 5 µl deionised water. PCR cycling conditions comprised an initial cycle at 95 °C for 10 min, followed by 40 cycles at 95 °C for 10 s, 60 °C for 15 sec and 72 °C for 20 sec. Three technical repeats were performed for each biological sample. At the end of each PCR run, a melting curve was generated to identify non specific PCR products and primer-dimers. Specific PCR products can be discriminated from these potential artefacts by using melting curve analysis, in which the Tm of the product is determined and results in a single fluorescent peak.

3.2.6 Optimisation of RT-PCR

The PCR conditions and specificity was initially tested with the conventional PCR assay. The reference cDNA was obtained by pooling all biological samples. Both the reference cDNA and cDNA from seed were used to first check the PCR reactions. The reference cDNA was later used to construct the standard curve. Primer concentration of target and reference gene was determined using serial dilutions of pooled cDNA as template. The optimal concentration used was 400 nM. The $MgCl_2$ final concentration was 3 mM.

University of Cape Town

3.2.7 Data analysis

To calculate the relative abundance of cDNA of the LEA1-Em genes the standard curve was constructed with the reference cDNA pool. Four dilutions (concentrated, 1:5, 1:20, 1:50) , each represented by three technical replicates, were used to construct the standard curve. The data was analysed with the Rotor Gene version 6.0 (Corbet Research) software. The threshold cycle (Ct) values of the triplicate PCRs were averaged and used for quantification of transcripts. The transcript levels of the target genes were normalised against a housekeeping gene, secretory carrier membrane protein gene (R1), transcript levels.

3.3 Results

3.3.1 RNA integrity

The quality of RNA that was used for cDNA synthesis for quantitative RT-PCR was checked using both denaturing agarose gels (data not shown) and the Agilent 2100 Bioanalyser. Electrophorograms showed good RNA integrity with clear 28S and 18S rRNA peaks (Figure 3.1 and Figure 3.2) and the Bioanalyser RIN values were between 7.1 and 8.4 with the exception of the seed RNA of 6.1 confirming good RNA quality (Table 3.2).

Biocompare product literature (http://www.biocompare.com/techart.asp?id=1522) on the interpretation of RIN values suggests that RNA samples with RIN between 7 and 10 are the highest quality and can be used for all types of experiments whereas samples with the RIN values between 5 and 7 can still be used for many other types of experiments, such as PCR. Several RNA samples extracted from seed were analysed, and the best RIN value of 6.1 was used as this was deemed still suitable to use for gene expression profiling. The same author from Biocompare, James F. Eliason, further suggests that the RIN values below 5 can still be used for some gene expression profiling techniques for example, PCR using short amplicons of 100 bp of which the amplicon size for this study was in the same range.

University of Cape Town

Table 3.2: RNA samples with their RNA Intergrity numbers (RIN)

Tissue type	RWC	RNA Integrity number (RIN)
Root	95 - 100 %	8.4
Root	5 %	8.4
Seed	Mature seed	6.1
Leaf	100 %	8.1
Leaf	80 %	7.4
Leaf	60 %	7.7
Leaf	40 – 45 %	7.9
Leaf	10 – 15 %	7.2
Leaf	5 %	7.3
Leaf	After rehydration 25 – 35 %	7.3
Leaf	After rehydration 46 – 55 %	7.5
Leaf	Fully rehydrated 100 %	7.1

* The RIN values were determined by the Agilent 2100 Bionalyser (Agilent Technologies, USA).

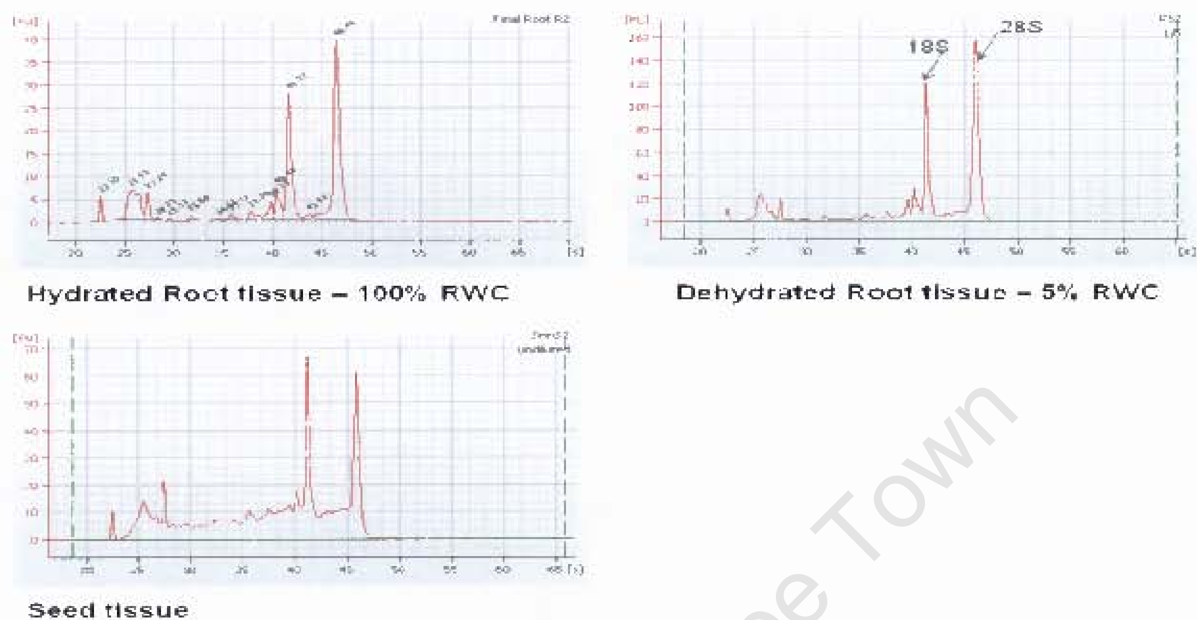
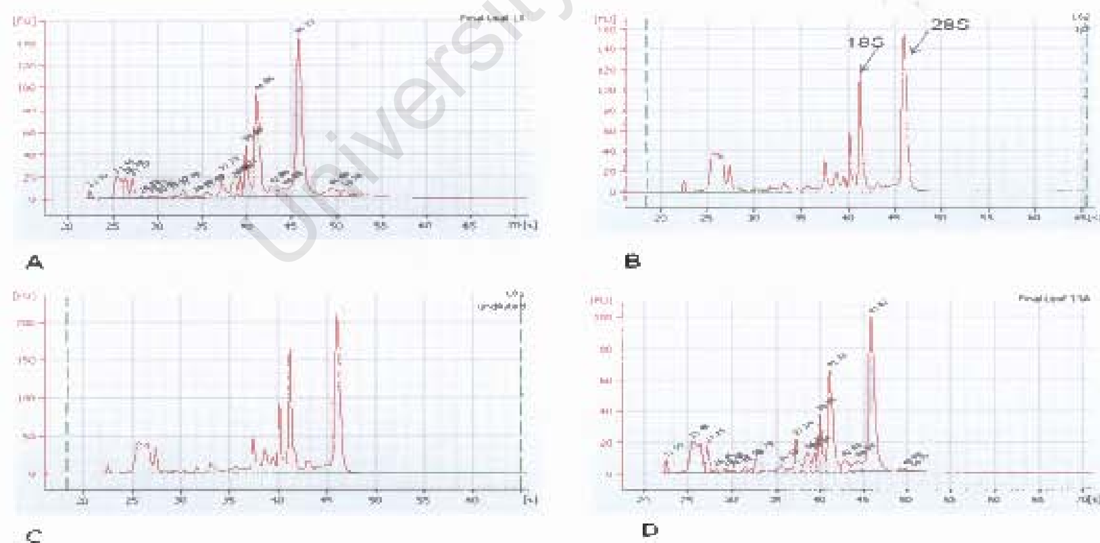


Figure 3.1: Electropherogram from RNA of root and seed samples

The Agilent 2100 Bioanalyser (Agilent Technologies, USA) electrograms showing RNA quality are shown above. The 18S and 28 S subunits peak positions are indicated with an errow in the electropherogram for dehydrated (5 % RWC) root RNA sample.

University of Cape Town



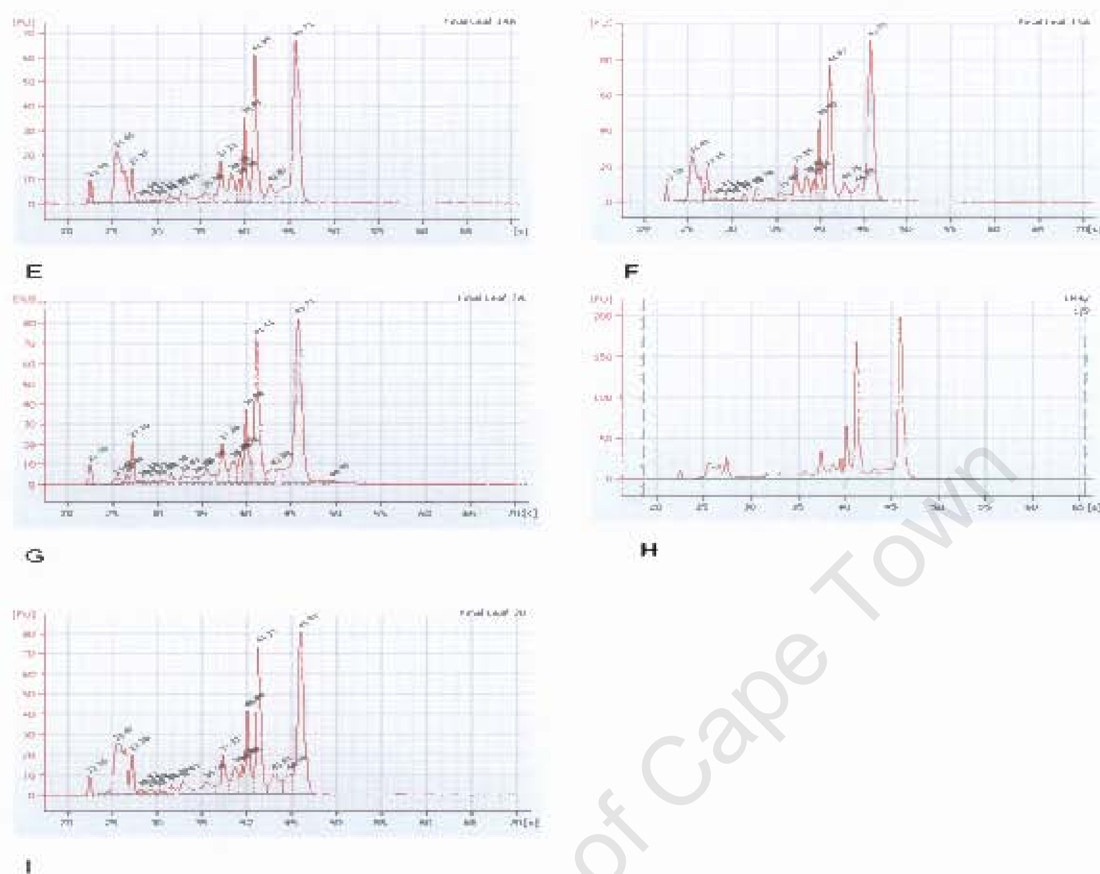


Figure 3.2: Electropherogram from RNA of leaf samples

The electrograms are from RNA samples extracted from leaf tissue of different RWC during the dehydration and rehydration cycle. Electrogram A to F; shows electrograms from RNA of leaf tissue at different RWC during dehydration at RWC of A- 100 %, B- 80 %, C-60 %, D- 40 - 45 %, E- 10 - 15 %, F- 5 %. While electrogram G to I are the rehydration samples at RWC of G- 25 – 35 %, H- 46 - 55 % and I- 100 %.

University of Cape Town

3.3.2 PCR Primer Design

Real-time PCR primers were designed for all four LEA1-Em (XHLEA1-1, XHLEA1-2, XHLEA1-4 and XHC00797) genes isolated from *X. humilis*. The LEA1-Em genes are highly conserved (Figure 2.20 and 2.21), as a result it was very important to check for primer specificity for RT-PCR. Specificity of each primer set for its targeted LEA1 orthologue was checked by aligning the primer (BioEdit Sequence alignment editor) to all the orthologues isolated (Table 3.3).

Primer alignment to target DNA indicated that all primers had a100 % specificity to their target LEA1-Em gene, indicated by the green arrows (Table 3.3). The red dotted arrows indicate a possibility of an alternate product, this will however be differentiated by size and has a lower complimentarity to the target DNA. Where there is red solid arrows, no PCR product can be formed.

Table 3.3: Primer specificity check between *X. humilis* LEA1-Em genes.

Gene Name	XHLEA1-1	XHLEA1-2	XHLEA1-4	XHC00797
XHLEA1-1 (95 bp amplicon)		Low complimentarity F = 57 % R= 64,71 % With 53 bp possible amplicon		
XHLEA1-2 (128 bp amplicon)				Low complementarity F = 60 % R= 56 % 123 bp possible amplicon
XHLEA1-4 (133 bp amplicon)				
XHC00797a (128 bp amplicon)				

* The black lines represent the DNA sequence, the green arrows represent 100 % complementarity to DNA sequence, red arrows represent no product formation and the dotted red line indicate possible amplicon with percentage of primer complementarity to DNA sequence and possible amplicon size.

The primers for XHC00797a were designed for the 3'UTR region. As discussed in chapter 2 the 3' UTR for XHC00797a and XHC00797b are very similar. However, a sequence at the end of the 3'UTR was present in XHC007907a sequence and absent in XHC00797b (highlighted in green - Figure 3.3). This allowed for primer specificity for XHC00797a. The possibility that the primers could amplify both sequences cannot be ruled out due to their position on the sequence. According to the sequencing information, the primers designed only amplified XHC0797a as indicated in Figure 3.2. Both the forward primer (highlighted in grey in Figure 3.3) and the

University of Cape Town

reverse primer (highlighted in green in Figure 3.3) had complementarity to XHC00797a giving a PCR product size of 128bp. When the same primer sequence is aligned against XHC00797b together with the forward primer no PCR product can be formed.

University of Cape Town

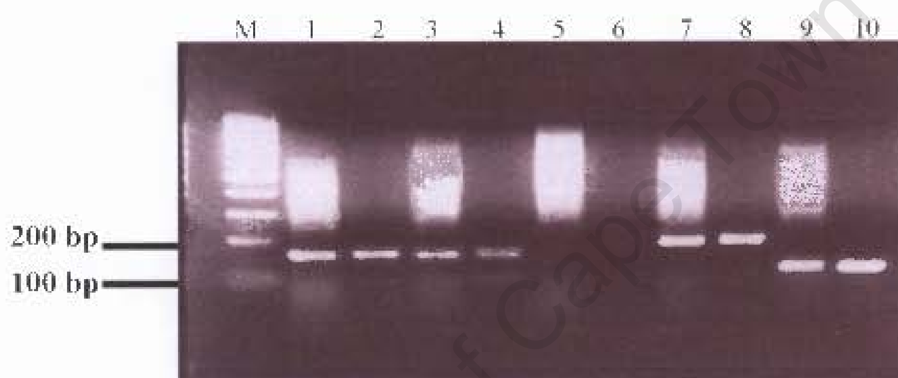
Figure 3.3: RT-PCR Primer specific for real time quantification of XHC00797a

Sequence alignment for the 3'UTR of XHC0097a and XHC0097b, showing primer specificity for single gene product amplification for XHC00797a. The forward primer is highlighted in grey and the reverse primer in green, resulting in an amplicon size of 128bp.

3.3.3 PCR Primer Optimisation

After the confirmation of the primers on BioEdit sequence alignment, the primers were firstly used for PCR reaction with the same conditions that would later be used for real-time PCR and PCR products were checked on an agarose gel (Figure 3.4). The amplified products were in the expected size range (Table 3.3). Primers for XHLEA1-2 (lane 5 and 6) failed to amplify the product, and because of the sequence similarity no other specific primer could be designed.

Smearing of DNA on agarose gel might be a result of various things. The most likely possibility in this case might be that there was a high amount of salts. Several people claim that high Mg^{++} can lead to smearing (Rapley, 2000). The reason it is visible in the pooled sample and not in the seed sample may be because the pooled sample has a higher concentration of these salts as it combines all 12 samples used for RT-PCR analyses as opposed to a single sample from seed. Contaminating RNA or protein from the RNA extraction and reverse transcription is also a possible explanation for the smearing.



University of Cape Town

Figure 3.4: Specificity of RT-PCR primers.

Lane M: O'GeneRuler 100 bp DNA ladder. Primers specificity of the RT-PCR primers was determined using pooled reference cDNA and seed cDNA as template. The odd numbered lanes contain the pooled cDNA while the even numbered lanes used seed cDNA. The RT-PCR GSP tested were for genes XHC00797 (Lane 1 and 2), XHLEA1-1 (Lane 3 and 4), XHLEA1-1 (Lane 5 and 6), XHLEA1-1 (Lane7 and 8), and R1 (Lane 9 and 10) respectively.

3.3.3.1 PCR primer concentration optimization

Two primer concentrations, 200 µM and 400 µM, were tested. The optimum primer concentration was 400 µM. 200 µM gave a fainter band on the agarose gel (data not shown) indicating a less efficient amplification of the target gene. This was also observed on the RT-PCR reaction an example of XHLEA1-4 is shown in Figure 3.5 with the lower primer concentration (200 µM) giving a maximum fluorescent signal at 0.4 versus the 3.5 from the higher primer concentration of 400 nM

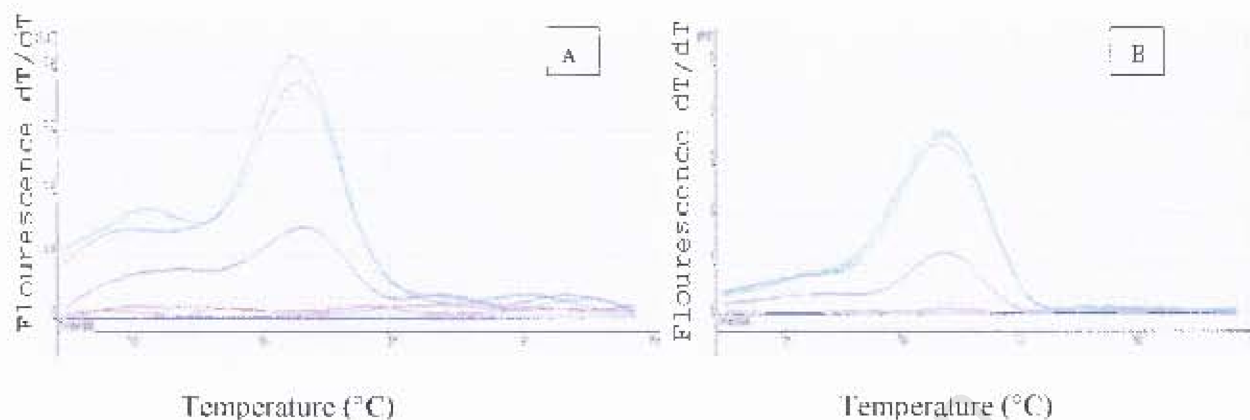


Figure 3.5 (A) Melting curve of XHLEA1-4 at 200nM primer concentration (B) Melting curve of XHLEA1-4 at 400nM primer concentration

Panel A shows the highest fluorescent signal of 0.4 at the final primer concentration of 200nM for the amplification of XHLEA1-4. While Panel B highest fluorescent signal is 3.5.

University of Cape Town

3.3.4 RT-PCR for the reference gene R1

The reference gene (R1) amplification was initially confirmed on agarose gel (Figure 3.4) and followed by RT-PCR. The RT-PCR melting curve of R1 (Figure 3.6) confirms the agarose gel results by showing a single fluorescent peak, which is indicative of single PCR product. No primer dimer artefacts were present in the PCR.

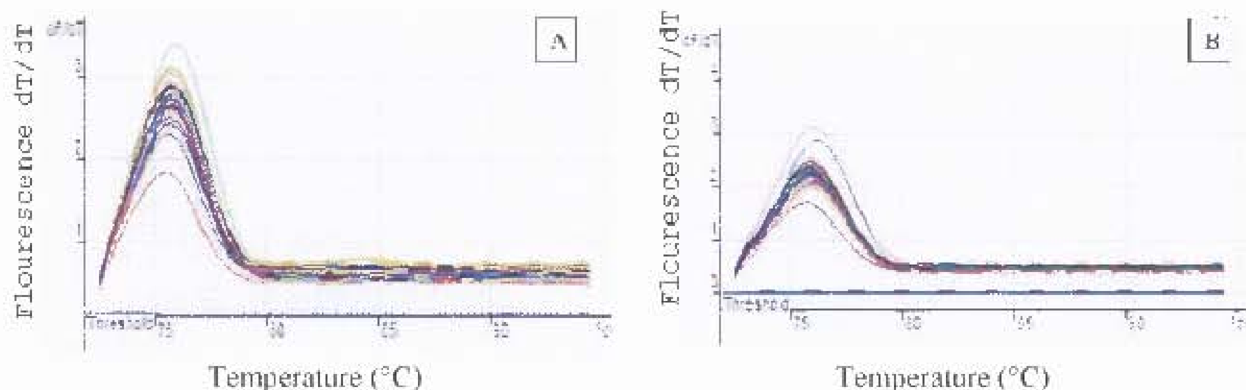


Figure 3.6: (A) Melting curve of housekeeping gene (R1) using undiluted samples (B) Melting curve of housekeeping gene (R1) using diluted samples

Panel A shows a melting curve for the housekeeping gene (R1) on the all the RNA samples tested. The single fluorescent peak indicates a single PCR product with no primer dimer artefacts. The samples that were too concentrated and fell outside the standard curve were diluted and Panel B shows the melting curve of those sample with a similar trend to Panel A.

University of Cape Town

The standard curve for R1 (Figure 3.7) was generated to calculate the concentration of mRNA transcripts in different samples. Some of the more concentrated samples were outside the standard curve, those samples were diluted and the RT-PCR run repeated for accuracy. The concentrations calculated from this standard curve were used to normalize gene expression of the LEA1-Em genes in the same samples of different RWC.

Given that an equal amount of starting RNA for each biological sample was used and additionally, the housekeeping gene has been previously shown to be stable during the dehydration cycle by Sally-Ann Walford (personal communication), the reason for the levels of R1 mRNA transcript abundance fluctuating across the biological samples can be attributed to the presence of different inhibitory components that might be present in different biological samples. This is the main reason of using a housekeeping gene to account for the variability in amplification between samples.

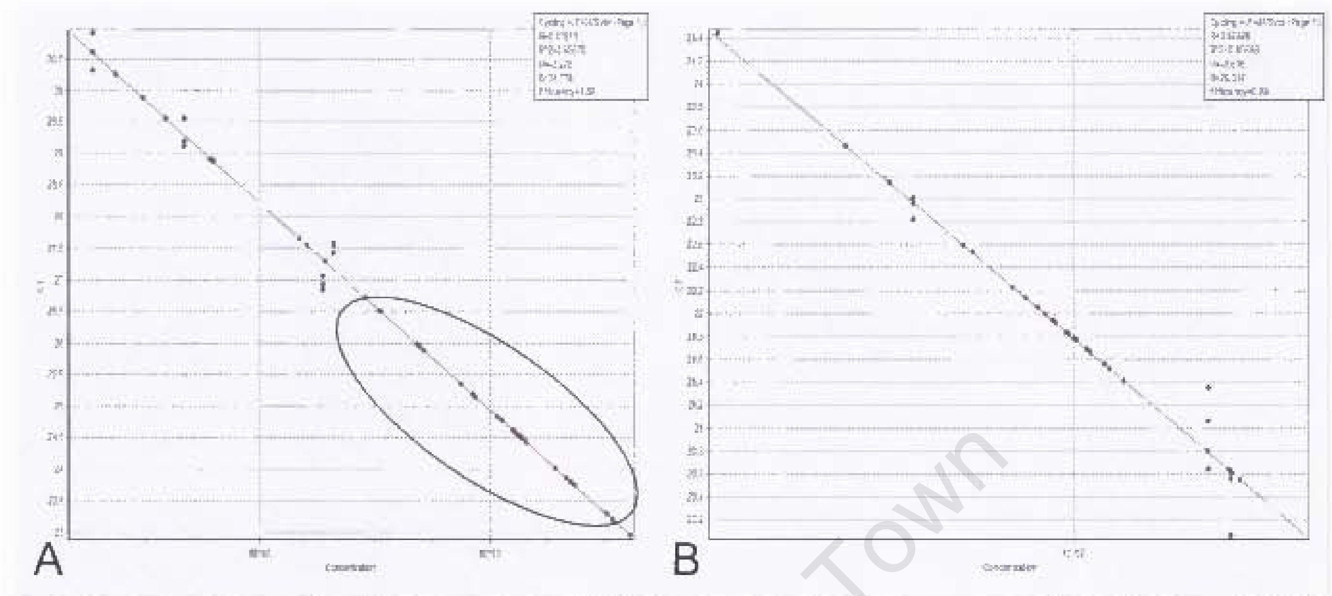


Figure 3.7: (A) Standard curve of housekeeping gene (R1) using undiluted samples (B) Standard curve of housekeeping gene (R1) using diluted samples

Panel A shows a standard curve for the housekeeping gene (R1) on the all the RNA samples tested. The circled samples in Panel A were too concentrated and fell outside the standard curve. The samples were diluted and quantification repeated using the standard curve in Panel B.

3.3.5 Expression of *X.humilis* XHLEA1-1 in seed, leaf and root during the abiotic stress treatment.

While the pooled samples for construction of the standard curve in the agarose gel analysis of the PCR, showed single product amplification for XHLEA1-1 (Figure 3.8), the test samples at most of the RWC (60 %, 40 - 45 %, 10 – 15 %, 5 %, and after rehydration samples RWC of 25 – 35 %, 40 %, 100 % and dehydrated root 5 %) failed to give a PCR product (Figure 3.9). The quantitative RT-PCR confirmed the agarose gel analysis as it resulted in no amplification except for the sample at 80 % RWC which was the only positive sample on the agarose gel (Figure 3.9). There are a few explanations to why the PCR failed for some samples while it worked for others. This might have occurred due to the following reasons, the presence of PCR inhibitors carried over from sample preparation, it is however unlikely that this is the only factor as the same sample was used successfully for the quantification of all the LEA genes in the study. What is more likely to be a contributing factor is the cDNA efficiency in transcribing this transcript. The

University of Cape Town

efficiency of the cDNA synthesis is known to be affected by the structure and concentration of the RNA template.

The secondary structure of the RNA and the presence of protein complexes on the target RNA can interfere with the reaction by causing enzyme pausing, dissociation, or skipping over looped regions (Liss, 2002). There is differing abilities amongst the different reverse transcriptase to read through the secondary structure (Brooks *et al.*, 1995) which might have reduced the RT reaction efficiency. Poor reverse transcription of this orthologue might well be the cause of the failure of the PCR reaction for this gene. The success in amplifying the product with the pooled samples in the standard curve might simple be due to the templates that were present in the seed and the 80 % RWC leaf samples. Primer dimers formed in the negative control (Figure 3.8 and 3.9, lane1). The lower the concentration of template, meaning the higher the dilution of the sample more primer dimers were observed similar to the negative control. Due to the inconsistencies in the amplification of PCR products with the XHLEA1-1 primers, no expression analysis using RT-PCR were further done.

Figure 3.8: Amplification of XHLEA1-1 gene from pooled cDNA samples.

Lane M. O'GeneRuler 100 bp DNA ladder. Lane 1: shows negative control (no DNA template). Lane 2: is an empty lane. Lane 3: shows PCR product obtained using concentrated pooled sample. Lane 4, 5, 6, 7, 8: shows the product obtained using various dilution of pooled concentrated pooled DNA sample. The dilutions were 1:5, 1:20, 1:50 and 1:100 for Lanes 4, 5, 6, 7 and 8 respectively.

University of Cape Town

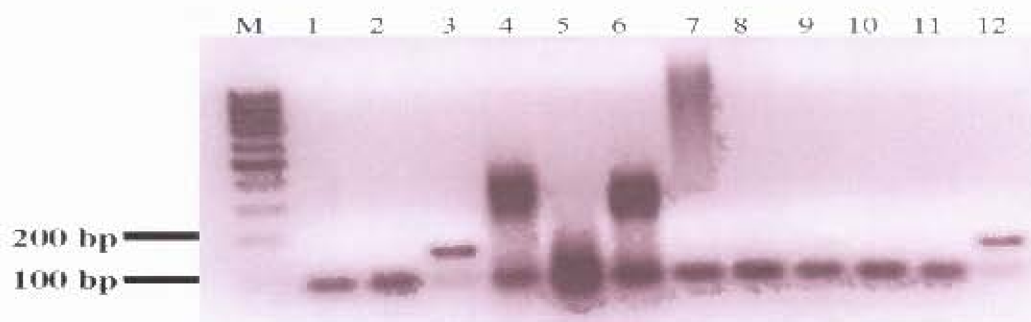


Figure 3.9: Amplification of XHLEA1-1 from tissue of different RWC cDNA samples.
Lane M: O'GeneRuler 100 bp DNA ladder . Lane 1 shows negative control (no DNA template). Lane 2 to 10: shows PCR product obtained using cDNA from leaf tissue at different RWC during the dehydration and rehydration cycle. Lane 2 to 7 are the dehydration samples at RWC of 100 %, 80 %, 60 %, 40-45 % , 10 - 15 % , 5 %. While Lane 8 to 10 are the rehydration samples at RWC of 25 - 35 %, 40 % and 100 % respectively. A dehydrated root sample at the RWC of 5 % was loaded in Lane 11. The cDNA from seed was included as a positive control in Lane 12.

3.3.6 Expression of *X.humilis* XHLEA1-4 in seed and in leaf and root during the abiotic stress treatment.

Melting curve analysis (Figure 3.10 and 3.11) of XHLEA1-4 revealed that the Tm characteristic of the seed and root XHLEA1-4 was consistently lower (+/- 81.47 °C) than that of the similar gene from leaf tissue (+/- 83.11 °C). The Tm values are known to be affected by both amplicon size and GC content (Ririe *et al.*, 1997), the primers for this gene were designed to give the same amplicon size of 133 bp, indicating that there might a mutation in the sequence for the gene in the root and seed which can either be a change in nucleotide or an insertion or deletion.

University of Cape Town

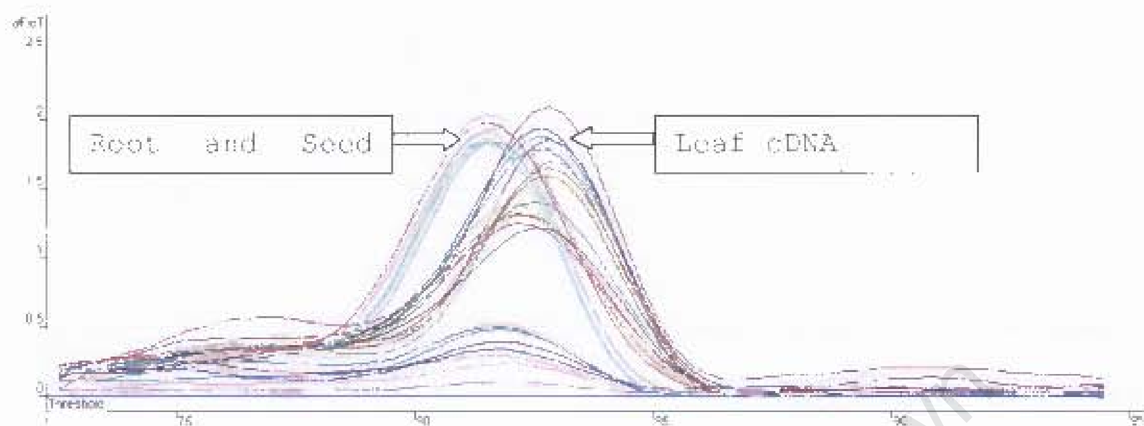


Figure 3.10: Melting curve for XHLEA1-4 mRNA quantification

The melting curve shows two fluorescent peaks. The first peak is a result of amplification of RNA from seed and root sample. While the second peak is amplification of RNA from leaf samples. The different peaks indicate different melting temperatures (Tm).

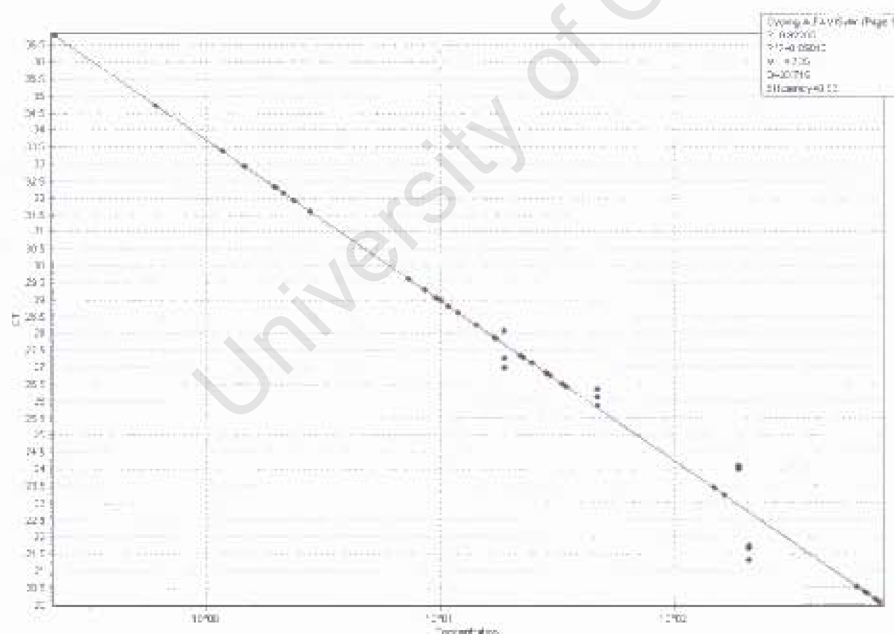


Figure 3.11: Standard curve of XHLEA1-4 using diluted samples

Standard curve (4 dilution points) for XHLEA1-4 on the all the RNA samples tested. There were four dilution, done in triplicate used to construct the standard curve. The dilution points are represented by the blue spots. The samples that were too concentrated and fell outside the standard curve were diluted and quantification repeated. The samples that were too dilute to and were outside the standard curve were not quantified and reported as not detectable.

University of Cape Town

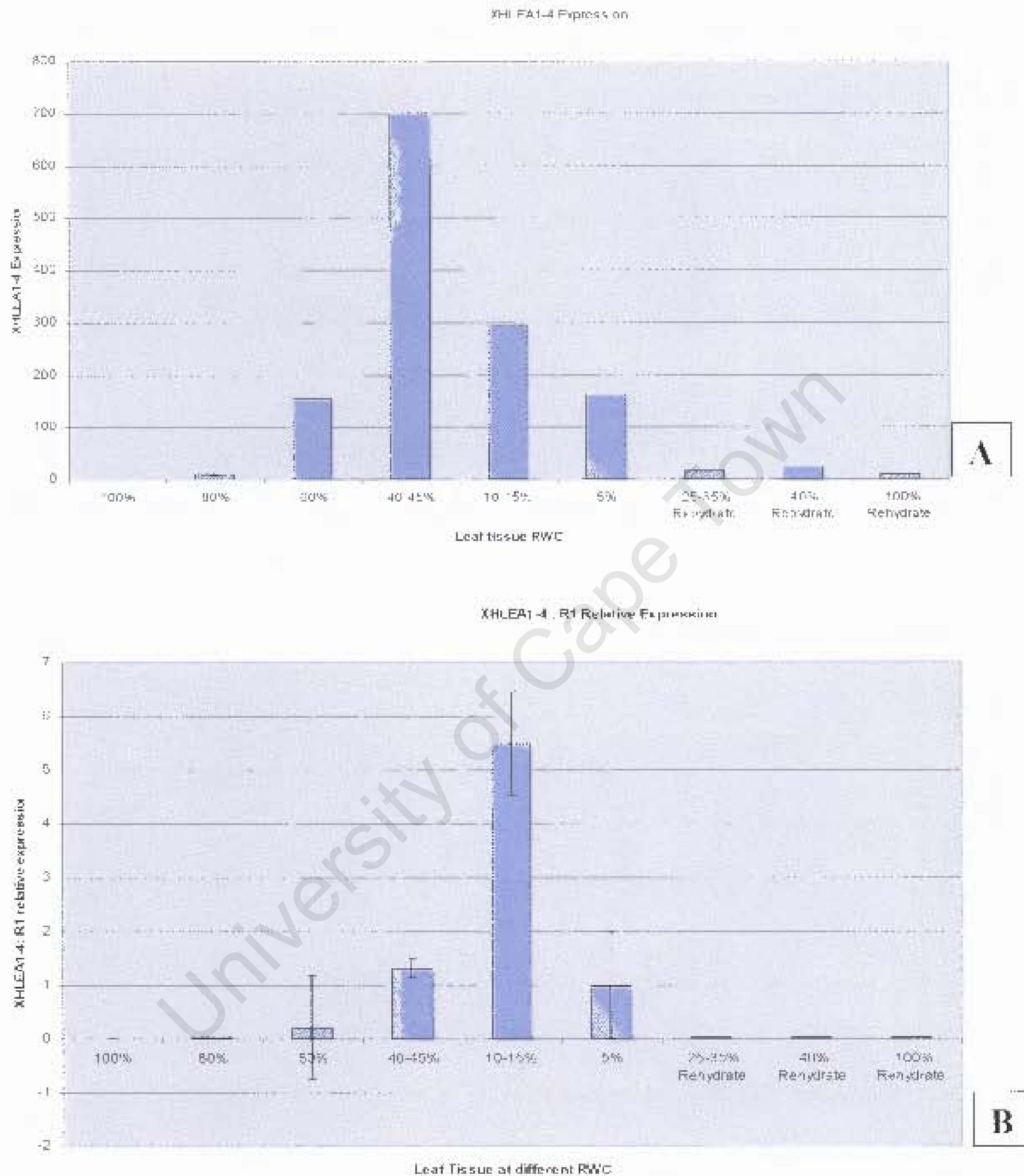


Figure 3.12 (A) Expression of XHLEA1-4 normalized to starting RNA (B) Expression of XHLEA1-4 normalized to housekeeping gene (R1)

The x-axis shows the RWC of the samples during the dehydration and rehydration cycle. Starting with a fully rehydrated sample of 100 % RWC to a dehydrated sample at 5 % RWC. This is followed by rehydration samples at different RWC until fully rehydrated at 100 % RWC. The student t-test was applied between samples and the differences were significant ($p \leq 0.05$, n=3 per RWC). The error bars represent standard deviation (SD).

University of Cape Town

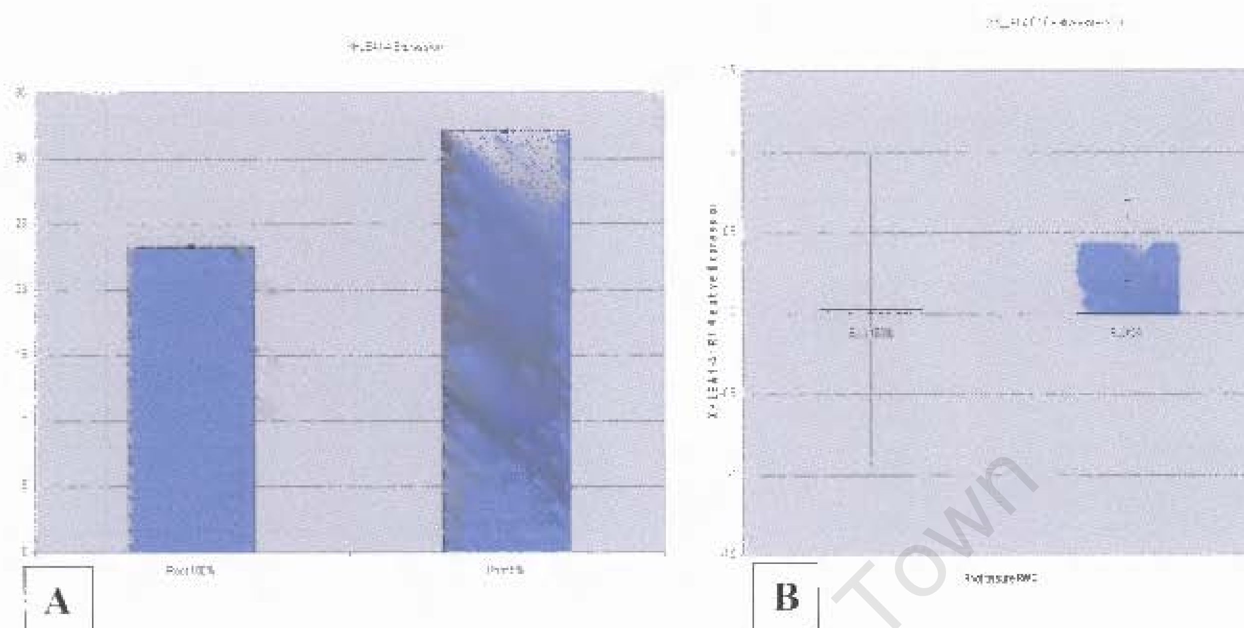


Figure 3.13 (A) Expression of XHLEA1-4 in root tissue normalized to starting RNA (B) Expression of XHLEA1-4 in root tissue normalized to housekeeping gene (R1)

The graph shows a comparison of XHLEA1-4 RNA levels from hydrated (100 % RWC) and dehydrated (5 % RWC) root samples. Panel A RNA levels are normalised to starting RNA while panel B is normalised to the housekeeping gene (R1). The student t-test was applied between the two groups of samples and the difference was significant ($p < 0.05$, n=3 per RWC). The error bars represent standard deviation (SD).

Panel A of Figure 3.12 shows the expression of XHLEA1-4 normalised to the amount of starting RNA. Panel B shows the relative expression of these genes after normalisation to the housekeeping gene R1. Both results show that the gene is highly induced by dehydration. XHLEA1-4 is also induced by dehydration in root (Figure 3.13).

3.3.7 Expression of XHC00797a RNA from seed, leaf and root tissue from *X.humilis* during the dehydration and rehydration cycle.

The melting curve analysis (Figure 3.14) for XHC00797 showed a single Tm for all amplification products. The single fluorescence peak of all samples further supports primer specificity and the absence genomic contamination observed in the agarose gel analysis (Figure 3.4) .

University of Cape Town

The two standard curves (Figure 3.15) were used for the quantification of XHC00797a, they represent the first standard curve with the original cDNA samples (Panel A), while panel B represents the samples that were outside the standard curve after dilution.

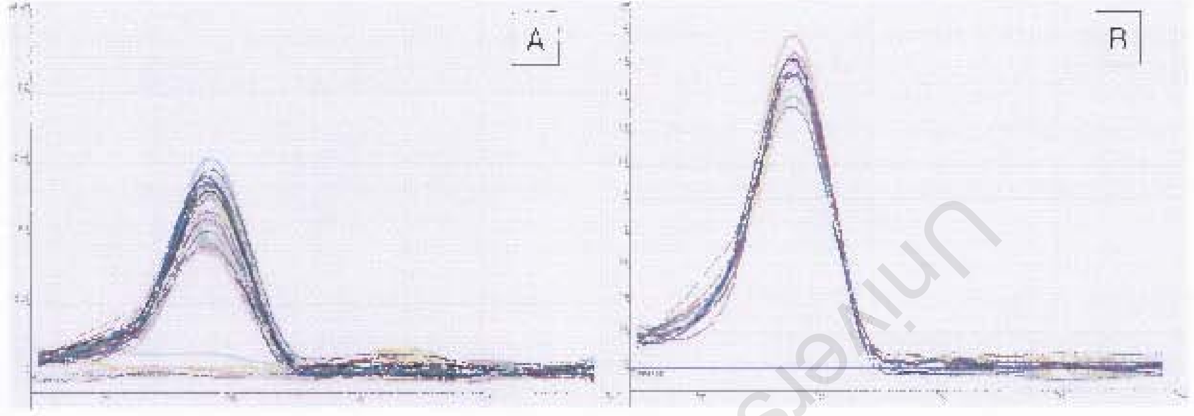


Figure 3.14 (A) Melting curve for XHC00797a using undiluted samples (B) Melting curve for XHC00797a using diluted samples

Panel A shows a melting curve for the XHC00797a on the all the RNA samples tested. The single fluorescent peak indicates a single PCR product with no primer dimer artefacts. The samples that were too concentrated and fell outside the standard curve were diluted and Panel B shows the melting curve of those sample with a similar trend to Panel A.

University of Cape Town

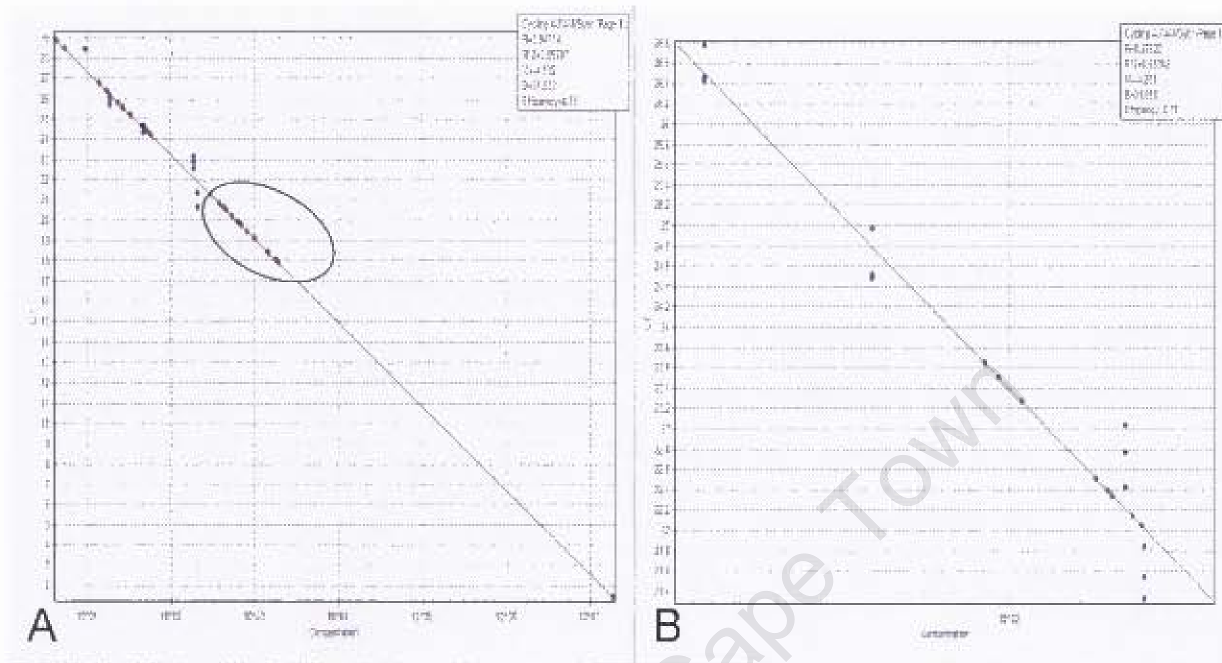


Figure 3.15: (A) Standard curve for XHC00797a using undiluted samples (B) Standard curve for XHC00797a using diluted samples

Panel A shows a standard curve for the XHC00797a on the all the RNA samples tested. The circled samples in Panel A were too concentrated and fell outside the standard curve. The samples were diluted and quantification repeated using the standard curve in Panel B

University of Cape Town

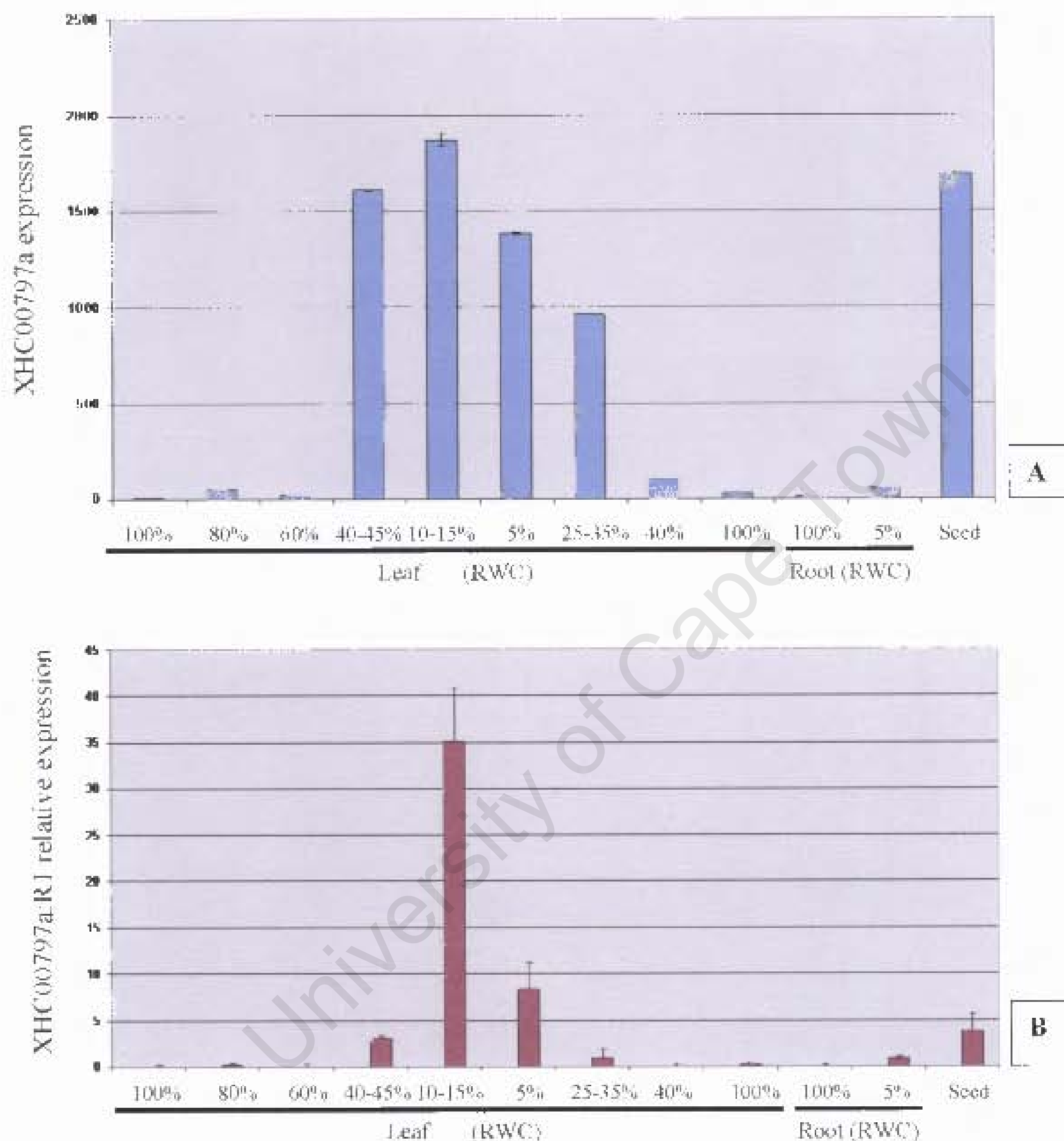


Figure 3.16 (A) Expression for XHC00797a normalized to starting RNA (B) Expression for XHC00797a normalized to housekeeping gene (R1)

The x-axis shows the RWC of the samples during the dehydration and rehydration cycle. Starting with a fully rehydrated expression from leaf tissue at 100 % RWC to a dehydrated leaf sample at 5 % RWC. This is followed by rehydration leaf samples at different RWC until fully rehydrated at 100 % RWC. The hydrated root (100 % RWC), dehydrated root (100 % RWC) and seed samples are included in Panel B. The student t-test was applied between

samples and the differences were significant ($p \leq 0.05$, n=3 per RWC). The error bars represent standard deviation (SD).

Figure 3.16 shows for each point of the RWC change in the level of XHC00797a transcript in the leaf, root and seed tissue. The XHC00797a gene is highly induced by dehydration. The initial peak in transcript abundance is observed at the RWC of between 40 - 45 %. There is a noticeable increase in transcript levels, ≥ 35 fold at the RWC of 10 – 15 % and a noticeable drop in the level of mRNA transcripts when the leaf is full dehydrated at 5 % RWC. However, the mRNA transcripts levels at the RWC of 5 % still exceed its mRNA transcript level in seed. At the RWC above 40% after rehydration, the mRNA transcripts of XHC00797a could not be accurately detected, and were recorded as undetectable. This might be due to lower concentration, below the detectable limit of the assay.

The gene transcripts were undetected in the hydrated root tissue and were induced by dehydration on dry root tissue (5 % RWC). The R1 cDNA was amplified from the same sample as the rest of the target genes allowing accurate gene expression normalization.

3.4 Discussion

In order to understanding how resurrection plants adapt to desiccation, we have studied the LEA1-Em gene response to desiccation. Desiccation stress results in changes in the cellular program, which involves significant transcriptional alterations aimed at increasing the chances of survival. Increased levels of protection and repair mechanisms would be expected. One way seeds have been described as responding to desiccation is by the biosynthesis of LEA proteins. In *Arabidopsis* seeds the T-DNA knockout of LEA1-Em resulted in acceleration of drying during seed development, which implicated LEA1-Em in slowing the rate of drying during seed development (Manfre *et al.*, 2006). However the role of LEA1-Em in desiccation tolerance has not been fully elucidated. The characterisation of the T-DNA mutant (Manfre *et al.*, 2006), their presence in desiccation seeds versus their absence in desiccation sensitive vegetative tissue suggests an important role of these proteins in the acquisition of desiccation tolerance.

Research has begun to elucidate the role of LEA1-Em proteins, for example using bioinformatics to predict the functions of the two LEA1-Em proteins from mung bean (Subramanian and

University of Cape Town

Ayyanar, 2006). Gilles *et al.* (2007) characterised a recombinant LEA1-Em and demonstrated that the N-terminal helical domain is important for the protection of enzyme activity from the effects of drying. In this study the LEA1-Em orthologues isolated from *X. humilis* were used to test the hypothesis that seed specific genes are upregulated in vegetative tissue of desiccation tolerant plants, resulting in the acquisition of desiccation tolerance.

The RT-PCR experiments in this study demonstrated that orthologues of the LEA1-Em genes, shown to be seed specific in *Arabidopsis* (Illing *et al.*, 2005) are upregulated in response to desiccation in the vegetative tissue of resurrection plant *X. humilis*.

Induction of the LEA1-Em cDNA in our experiments in *X. humilis* was dramatically desiccation dependent. Leaf tissue appear to constitutively express very low amounts of the LEA1-Em mRNA transcripts at relative water contents above 80% and are capable of increasing the mRNA levels >30-fold following continued water stress.

However, the possibility that the high fold increase seen with XHC00797a is cause by the amplification of both XHC007967a and XHC00797b cannot be excluded with certainty. This is because the primer design for RT-qPCR was towards the end of the transcript in the 3' UTR region. However, according to figure 3.3, this is unlikely the case. We suggest that that *X. humilis* employs alternative splicing for its LEA1-Em transcript, it will also be interesting to determine whether induction in different tissue type is specific to a single transcript or whether both XHC00797a as well as XHC00797b are co-expressed. Interestingly, the tissue specific two fluorescent peaks in the XHLEA1-4 melting curve (Figure 3.10) might also be an indication that this gene also has an isoform that is tissue specific.

This study shows a correlation of the accumulation of LEA1-Em mRNA transcript that is initiated following exposure of *X. humilis* to desiccation stress. Changes in the levels of known LEA1-Em mRNA transcripts were documented, which are significantly induced in response to stress due to desiccation.

University of Cape Town

In a recent study (Hundertmark and Hincha, 2008) have shown that AtEm6, a LEA-1 from *Arabidopsis*, is one of the genes highly expressed in both seed and non-seed tissue. This study also demonstrated that this group of genes is upregulated in response to salinity, cold, and abscisic acid (ABA). The findings of this study make a good foundation for future studies, comparing AtEM1 mRNA levels in non-seed tissue of a desiccation sensitive plant, *Arabidopsis* versus a desiccation tolerant plant. The function of this group of genes might be affected or enhanced by the amount of transcript present in the tissue during desiccation stress.

3.5 Conclusion

In this chapter, the evidence presented showed that LEA1-Em genes, known to be specifically upregulated in seed in response to desiccation are also transcribed in vegetative tissue of the desiccation tolerant plant, *X. humilis*. Further studies on protein expression resulting from these changes is required, as well as the characterization of the LEA1-Em gene knockout phenotypes in Arabidopsis mutants. Recently developed methods, such as small interfering RNA (siRNA) can be applied to generate LEA1-Em knockout mutants in Arabidopsis to be characterized.

University of Cape Town

4 General Conclusions

In conclusion, this study reports the first isolation of LEA1-Em genes from a resurrection plant *X. humilis*. The isolation and cloning results also suggest that some of the LEA1-Em mRNA transcripts could be generated through alternative splicing. Alternative splicing is a powerful tool that a plant can use to improve protein diversity. Further investigations can be done to investigate environmental, physiological and other factors that regulate the alternative splicing of LEA1-Em genes in *X. humilis*. The different roles of individual LEA1-Em genes in desiccation tolerance acquisition can also be investigated. To further test whether alternative splicing is characteristic of this group of LEAs in vegetative tissue, isolation of LEA1-Em orthologues from other desiccation tolerant vegetative tissue and their characterisation would prove to be useful in understanding this group of LEAs.

Since the LEA1-Em genes had been previously characterised as being seed specific in a desiccation sensitive plant, *Arabidopsis*, the aim of the study was to use this group of LEAs to test the hypothesis that the acquisition of desiccation tolerance in vegetative tissue of resurrection plants is a consequence of the appropriation of seed specific genes. Recently, Fisher (2008) using phylogenetic studies, addressed the question of the origin of DT in vegetative tissue. The study involved the use of a complete genome sequences from the moss *Physcomitrella patens*, the angiosperm *Arabidopsis thaliana* and EST data from the DT moss *Tortula ruralis* and the DT angiosperms *Craterostigma plantagineum* and *Xerophyta humilis.* The study found that, all the ancestral nodes expression pattern was not exclusively stress related; they included seed and/or pollen expression either exclusively or in combination. Thus, her findings contradict the hypothesis that modified vegetative DT arose via a modified stress tolerance pathway and support the hypothesis that modified vegetative DT in angiosperms arose through the adoption of mechanisms related DT in reproductive propagules. Fisher (2008) supports the hypothesis tested in this study.

Following the successful full length cloning of LEA1-Em genes and using RT-PCR the aim of the thesis was fulfilled. The RT-PCR experiments demonstrated the correlation between the desiccation cycle and the upregulation of LEA1-Em genes in both leaf and root tissue. The findings from this study provide leads to the working hypothesis of desiccation tolerance,

University of Cape Town

implicating seed specific genes to have an important role in acquisition of desiccation tolerance of vegetative tissue. Targeting seed specific genes linked to desiccation in seed may lead to identification of critical gene and protein information important for acquisition of desiccation tolerance in vegetative tissue of plants.

University of Cape Town

5 References

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

Baker, J., Van dennSteele, C., and Dure III, L. (1988) Sequence and characterization of 6Lea proteins and their genes from cotton. *Biomedical and Life Sciences* **11**: 277-291.

Berjak, P. (2006) Unifying perspectives of some mechanisms basic to desiccation tolerance across life forms. *Seed Science Research* **16**: 1-15.

Berjak, P., Farrant, J.M. and Pammenter, N.W. (2007) Seed desiccation-tolerance mechanisms. Pp 151-192. IN: Plant Desiccation Tolerance Eds: Matthew A. Jenks and Andrew J. Wood, Blackwell Publishing, Oxford, UK. ISBN-13: 978-0-8138-1263-2/2007.

Bernacchia, G., Salamini, F., and Bartels, D. (1996) Molecular Characterization of the Rehydration Process in the Resurrection Plant Craterostigma plantagineum. *Plant Physiology* **111**: 1043-1050.

Bewley, J.D. (1979) Physiological Aspects of Desiccation Tolerance. *Annual Review of Plant Physiology* **30**: 195-238.

Bewley J.D. and Krochko J.E. (1982) Desiccation-tolerance. In: O.L. Lange, C.B. Osmond and H. Ziegler, Editors, Physiological Plant Ecology II. Water Relations and Carbon Assimilation, Springer, New York (1982), pp. 362–368.

Bianchi G, Gamba A, Murelli C, Salamini F and Bartels D (1991) Novel carbohydrate metabolism in the resurrection plant Craterostigma plantagineum. *Plant Journal* **1**: 355–359

Bianchi G, Gamba A, Murelli C, Salamini F and Bartels D (1992) Low molecular weight solutes in desiccated and ABA treated calli and leaves of Craterostigma plantagineum. *Phytochemistry* **31**: 1917–1922

Bianchi G, Gamba A, Limiroli R, Pozzi N, Elster R, Salamini F and Bartels D (1993) The unusual sugar composition in leaves of the resurrection plant Myrothamnus flabellifolia. *Physiologia Plantarum* **87**: 223–226

Bies, N., Aspart, L., Carles, C., Gallois, P., and Delseny, M. (1998) Accumulation and degradation of Em proteins in Arabidopsis thaliana; evidence for post-transcriptional controls. *Journal of Experimental Botany* **49**: 1925-1933.

University of Cape Town

Blackman, S.A., Obendorf, R.L., and Leopold, A.C. (1992) Maturation Proteins and Sugars in Desiccation Tolerance of Developing Soybean Seeds. *Plant Physiologia* **100**: 225-230.

Bray, D., and Lay, S. (1994) Computer simulated evolution of a network of cell-signaling molecules. *Biophysical Journal* **66**: 972-977.

Bray, E.A. (1993) Molecular Responses to Water Deficit. *Plant Physiologia* **103**: 1035-1040.

Brooks, E.M., Sheflin, L.G., and Spaulding, S.W. (1995) Secondary structure in the 3' UTR of EGF and the choice of reverse transcriptases affect the detection of message diversity by RT-PCR. *Biotechniques* **19**: 806-812, 814-805.

Buitink, J., Walters-Vertucci, C., Hoekstra, F.A., and Leprince, O. (1996) Calorimetric Properties of Dehydrating Pollen (Analysis of a Desiccation-Tolerant and an Intolerant Species). *Plant Physiologia* **111**: 235-242.

Buitink, J., Hoekstra, F.A, Leprince, O. (2002) Biochemistry and Biophysics of tolerance systems, in: M. Black, H.W. Pritchard (Eds.), Desiccation and Survival in Plants: Drying without Dying, CABI Publishing, Oxon, UK, 2002, pp. 111–146. [45]

Burke, M. J. (1986) The glassy state and survival of anhydrous biological systems. Pp. 358–363. In: Leopold, A. C. (ed), Membranes, metabolism and dry organisms. Cornell University Press, Ithaca, New York

Bustin, S.A., and Nolan, T. (2004) Pitfalls of Quantitative Real-Time Reverse-Transcription Polymerase Chain Reaction. *J Biomol Tech* **15**: 155-166.

Caceres, J.F., and Kornblihtt, A.R. (2002) Alternative splicing: multiple control mechanisms and involvement in human disease. *Trendsin Genetics* 18: 186-193.

Ceccardi, T.L., Meyer, M.C., and Close, T.J. (1994) Purification of a maize dehydrin. *Protein Express Purification* **5**: 266-269.

Chaves, M.C., Maroco, J.P., and pereira, J.S. (2003) Understanding plant responses to drought from genes to the whole plant. *Functional Plant Biology* **30**: 239-264.

Chen, Z.Q., Liu, Q., Zhu, Y.S., and Li, Y.X. (2002) Assessment on predicting methods for membrane protein transmembrane regions. *Acta Biochem Biophys Sin* **34**: 285-290.

Cheng, Z., Targolli, J., Huang, X., and Wu, R. (2002) Wheat LEA genes, PMA80 and PMA1959, enhance dehydration tolerance of transgenic rice (Oryza sativa L.). *Molecular Breeding* **10**: 71-82.

University of Cape Town

Close, T.J., Kortt, A.A., and Chandler, P.M. (1989) A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. *Plant Molecular Biology* **13**: 95-108.

Close, T. J., Fenton, R. D., Yang, A., Asghar, R., DeMason, D. A., Crone, D. E., Meyer, N. C. & Moonan, F.(1993). Dehydrin: The protein. Pp. 104–118. In: Close, T. J. & Bray, E. A. (eds), Plant responses to cellular dehydration during environmental stress. Current Topics in Plant Physiology. American Society of Plant Physiologists. Series, Vol. 10, Rockville, Maryland.

Close, T.J. (1996) Dehydrins: emergence of a biochemical role of a family of plant dehydrin proteins. *Physiologia Plantarum* **97**: 795-803.

Close, T. (1997) A commonality in the response of plants to dehydration and low temperature. *Physiologia Plantarum* **100**: 291-296.

Collett, H., Shen, A., Gardner, M., Farrant, J.M., Denby, K.J., and Illing, N. (2004) Towards transcript profiling of desiccation tolerance in *Xerophyta humilis*: Construction of a normalized 11 k *X. humilis* cDNA set and microarray expression analysis of 424 cDNAs in response to dehydration. *Physiologia Plantarum* **122**: 39-53.

Crowe, J.H., and Crowe, L.M. (1986) Water and carbohydrate interactions with membranes: studies with infrared spectroscopy and differential scanning calorimetry methods. *Methods in Enzymology* **127**: 696-703

Crowe, J.H., Hoekstra, F.A., and Crowe, L.M. (1992) Anhydrobiosis. *Annuual Review of Physiol* **54**: 579-599.

Cuming, A.C., and Lane, B.G. (1979) Protein Synthesis in Imbibing Wheat Embryos. *European Journal of Biochemistry* **99**: 217-224.

Cuming, **A. C.** (1999) LEA proteins. *In* P. R. Shewry and R. Casey (eds.), *Seed proteins*, pp. 753–780. Kluwer Academic, Dordrecht, NL

de la Fuente van Bentem, S., Vossen, J.H., Vermeer, J.E., de Vroomen, M.J., Gadella, T.W., Jr., Haring, M.A., and Cornelissen, B.J. (2003) The subcellular localization of plant protein phosphatase 5 isoforms is determined by alternative splicing. *Plant Physiology* **133**: 702-712.

Drennan PM, Smith MT, Goldsworthy D and van Staden J (1993) The occurrence of trehalose in the leaves of the desiccation-tolerant Angiosperm Myrothamunus flabellifolius Welw. *J Plant Physiology* **142**: 493–496

University of Cape Town

Dure, L., 3rd, Crouch, M., Harada, J., Ho, T.D., Mundy, J., Quatrano, R., Thomas, T., and Sung, Z.R. (1989) Common amino acid sequence domains among the LEA preoteins of higher plants. *Plant Molecular Biology* **12**: 475-486.

Dure, L., 3rd (1993) A repeating 11-mer amino acid motif and plant desiccation. *Plant Journal* **3**: 363-369.

Eom, J., Baker, W., Kintanar, A., and Wurtele, A. (1996) The embryo-specific EMB-1 protein of Daucus carota is flexible and unstructured in solution. *Plant Science* **115**: 17-24.

Farrant, J.M., Pammenter, N.W., and Berjak, P. (1992) Development of the recalcitrant (homoiohydrous) seeds of *Avicennia marina* (Forsk.) Vierrh.: anatomical, ultrastructural and biochemical events associated with development from histodifferentiation to maturation. *Annals of Botany* **7**: 75-86.

Farrant, J.M. (2007) Mechanisms of desiccation tolerance in Angiosperm resurrection plants. pp 51-90. IN: Plant Desiccation Tolerance Eds: Matthew A. Jenks and Andrew J. Wood, Blackwell Publishing, Oxford, UK. ISBN-13: 978-0-8138-1263-2/2007

Fisher, K.M. (2008) Bayesian reconstruction of ancestral expression of the LEA gene families reveals propagule-derived desiccation tolerance in resurrection plants. *American Journal of Botany* **95**: 506-515.

Freeman, W.M., Walker, S.J., and Vrana, K.E. (1999) Quantitative RT-PCR: Pitfalla and Potential. *Biotechniques* **26**: 112-125.

Gaff, D.F. (1977) Desiccation tolerant vascular plants of southern Africa. *Oecologia* **31**: 95-109.

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

Gilles, G.J., Hines, K.M., Manfre, A.J., and Marcotte, W.R., Jr. (2007) A predicted N-terminal helical domain of a Group 1 LEA protein is required for protection of enzyme activity from drying. *Plant Physiology and Biochemistry* **45**: 389-399.

Goyal, K., Tisi, L., Basran, A., Browne, J., Burnell, A., Zurdo, J., and Tunnacliffe, A. (2003) Transition from Natively Unfolded to Folded State Induced by Desiccation in an Anhydrobiotic Nematode Protein. *Journal of Biological Chemistry* **278**: 12977-12984.

Goyal, K., Walton, L.J., and Tunnacliffe, A. (2005) LEA proteins prevent protein aggregation due to water stress. *Biochemical Journal* **388**: 151-157.

University of Cape Town

Guo Xiao-Qin, Zhang Hong-Zhi and Li De-Bao (2007) Alternative splicing of pre-mRNA in plants. *Chinese Journal of Agricultural Biotechnology,* **4**: 1-8

Guy, R.A., Payment, P., Krull, U.J., and Horgen, P.A. (2003) Real-time PCR for quantification of Giardia and Cryptosporidium in environmental water samples and sewage. *Applied and Environtal Microbiology* **69**: 5178-5185.

Han, B., and Kermode, A.R. (1996) Dehydrin-like proteins in castor bean seeds and seedlings are differentially produced in response to ABA and water-deficit-related stresses. *Journal of Experimental Bot*any **47**: 933-939.

Hara, M., Terashima, S., and Kuboi, T. (2001) Characterization and cryoprotective activity of cold-responsive dehydrin from Citrus unshiu. *Journal of Plant Physiology* **158**: 1333.

Hartung, W., Schiller, P., and Karl-Josef, D. (1998) Physiology of poikilohydric plants. *Progress in Botany* **59**: 299-327.

Hendry, G.A.F., Finch-Savage, W.E., Thorpe, P.C., Atherton, N.M., Buckland, S.M., Nilsson, K.A., and Seel, W.E. (1992) Free radical processes and loss of seed viability during desiccation in the recalcitrant species Quercus robur L. *New Phytologist* **122**: 273-279.

Hendry, G.A.F. (1993) Oxygen, free radical processes and seed longevity. *Seed Science Research* **3**: 141-153.

Hoekstra, F.A., Crowe, J.H., and Crowe, L.M. (1992) Germination and ion leakage are linked with phase transitions: of membrane lipids during imbibition of Typha latifolia pollen. *Physiologia Plantarum* **84**: 29-34.

Hong-Bo, S., Zong-Suo, L., and Ming-An, S. (2005) LEA proteins in higher plants: structure, function, gene expression and regulation. *Colloids Surf B Biointerfaces* **45**: 131-135.

Horbowics, M., and Obendorf, R.L. (1994) Seed desiccation tolerance and storability: dependence of flatulence-producing oligosaccharides and cyclitols: review and survey. *Seed Science Research* **4**: 385-405.

Hsing, Y., Chen, Z., Shih, M., Hsieh, J., and Chow, T. (1995) Unusual sequences of group 3 LEA mRNA inducible by maturation or drying in soybean seeds. *Plant Molecular biology* **29**: 863-868.

Hughes, D.W., and Galau, G.A. (1989) Temporally modular gene expression during cotyledon development. *Genes and Development.* **3**: 358-369.

University of Cape Town

Hughes, D.W., and Galau, G.A. (1991) Developmental and Environmental Induction of Lea and LeaA mRNAs and the Postabscission Program during Embryo Culture. *Plant Cell* **3**: 605-618.

Hundertmark, M., and Hincha, D.K. (2008) LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis thaliana. *BMC Genomics* **9**: 118.

Illing, N., Denby, K.J., Collett, H., Shen, A., and Farrant, J.M. (2005) The Signature of Seeds in Resurrection Plants: A Molecular and Physiological Comparison of Desiccation Tolerance in Seeds and Vegetative Tissues. *Integrative and Comparative. Biology* **45**: 771-787.

Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A., and Auffray, C. (2005) Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic AcidsResearch* **33**: e56-.

Ingram, J., and Bartels, D. (1996) The Molecular Basis Of Dehydration Tolerance In Plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 377-403.

Iida K, Seki M, Sakurai T. (2004) Genome-wide analysis of alternative pre-mRNA splicing in Arabidopsis thaliana based on full-length cDNA sequences. *Nucleic Acids Research* **32**: 5096–5103.

Ismail, A.M., Hall, A.E., and Close, T.J. (1999) Purification and partial characterization of a dehydrin involved in chilling tolerance during seedling emergence of cowpea. *Plant Physiology* **120**: 237-244.

Jensen, A.B., Goday, A., Figueras, M., Jessop, A.C., and Pages, M. (1998) Phosphorylation mediates the nuclear targeting of the maize Rab17 protein. *Plant Journal***13**: 691-697.

Kaiser K, Gaff DF and OutlawWHJr (1985) Sugar contents of leaves of desiccation-sensitive and desiccation-tolerant plants. *Naturwissenschaft* **72**: 608–609

Kermonde, A.R. (1997) Approaches to elucidate the basis of desiccation-tolerance in seeds. *Seed Science Research* **7**: 75-96.

Kermode, A.R. and Finch-Savage, W.E. (2002) "Desiccation sensitivity in orthodox and recalcitrant seeds in relation to development." In Desiccation and Survival in Plants: Drying without Dying. edited by M. Black and H.W. Pritchard, pp. 149-184. Wallingford UK: CABI publishing.

University of Cape Town

Kikuchi S, Satoh K, Nagata T. (2003) Collection, mapping and annotation of over 28,000 cDNA clones from japonica rice. *Science* **301**: 376–379.

Koster, K.L. and Bryant, G. (2005) "Dehydration in model membranes and protoplasts: contrasting effects at low, intermediate and high hydrations." In Cold Hardiness in Plants: Molecular Genetics, Cell Biology and Physiology, edited by T.H.H. Chen, M. Uemura and S. Fujikawa, pp. 219-234. Wallingford, UK: CAB International.

Kranner, I., and Grill, D. (1996) Significance of thiol-disulfide exchange in resting stages of plant development. *Botanica Acta* **109**: 8-14.

Kranner, I., Cram, W.J., Zorn, M., Wornik, S., Yoshimura, I., Stabentheiner, E., and Pfeifhofer, H.W. (2005) Antioxidants and photoprotection in a lichen as compared with its isolated symbiotic partners. *Proceedings of the National Academy of Science U S A* **102**: 3141-3146.

Kranner, I., Birtic, S., Anderson, K.M., and Pritchard, H.W. (2006) Glutathione half-cell reduction potential: a universal stress marker and modulator of programmed cell death? *Free Radical Biology and Medicine* **40**: 2155-2165.

Kuang, J., Gaff, D.F., Gianello, R.D., Blomstedt, C.K., Neale, A.D., and Hamill, J.D. (1995) Changes in in vivo Protein Complements in Drying Leaves of the Desiccation-Tolerant Grass Sporobolus stapfianus and the Desiccation-Sensitive Grass Sporobolus pyramidalis. *Functional Plant Biology* **22**: 1027-1034.

Lacey, H.A., Nolan, T., Greenwood, S.L., Glazier, J.D., and Sibley, C.P. (2005) Gestational profile of Na+/H+ exchanger and Cl-/HCO3- anion exchanger mRNA expression in placenta using real-time QPCR. *Placenta* **26**: 93-98.

Larson, R.A. (1988) The antioxidants of higher plants. *Phytochemistry* **27**: 969-978.

Leopold, A.C.; Sun, W.Q.; Bernal-Lugo, I. (1994) The glassy state in seeds: analysis and function *Seed Science Research* **4(3)**: 267-274

Leprince O., (1993) Desiccation tolerance. *In* GAF Hendry, JP Grime, eds, Methods in Comparative Plant Ecology. Chapman and Hall, London, pp 48-53

Leprince, O., Atherton, N.M., Deltour, R., and Hendry, G. (1994) The Involvement of Respiration in Free Radical Processes during Loss of Desiccation Tolerance in Germinating Zea mays L. (An Electron Paramagnetic Resonance Study). *Plant Physiology* **104**: 1333-1339.

University of Cape Town

Leprince, O., Vertucci, C.W., Hendry, G.A.F., and Atherton, N.M. (1995) The expression of desiccation-induced damage in orthodox seeds is a function of oxygen and temperature. *Physiologia Plantarum* **94**: 233-240.

Liss, B. (2002) Improved quantitative real-time RT-PCR for expression profiling of individual cells. *Nucleic Acids Research* **30**: e89.

Lisse, T., Bartels, D., Kalbitzer, H.R., and Jaenicke, R. (1996) The recombinant dehydrin-like desiccation stress protein from the resurrection plant Craterostigma plantagineum displays no defined three-dimensional structure in its native state. *Journal of Biological Chemistry* **377**: 555-561.

Liu, W., and Saint, D.A. (2002) Validation of a quantitative method for real time PCR kinetics. *Biochemical and Biophysical Research Communications* **294**: 347-353.

Manfre, A.J., Lanni, L.M., and Marcotte, W.R., Jr. (2006) The Arabidopsis group 1 LATE EMBRYOGENESIS ABUNDANT protein ATEM6 is required for normal seed development. *Plant Physiology* **140**: 140-149.

Marcotte, W.R., Bayley, C.C., and Quatrano, R.S. (1988) Regulation of a wheat promoter by abscisic acid in rice protoplasts. *Nature* **335**: 454.

Morris, P.C., Kumar, A., Bowles, D.J., and Cuming, A.C. (1990) Osmotic stress and abscisic acid induce expression of the wheat Em genes. *European Journal of Biochemistry* **190**: 625-630.

Mouillon, J.-M., Gustafsson, P., and Harryson, P. (2006) Structural Investigation of Disordered Stress Proteins. Comparison of Full-Length Dehydrins with Isolated Peptides of Their Conserved Segments. *Plant Physiology.* **141**: 638-650.

Mtwisha L, Farrant JM, Brandt W and Lindsey GG (**2006**). Water stress proteins. pp 531-542. In: *Drought Adaptaion in Cereals.* Ed J-M Ribaut. Food Products Press; The Haworth Press, New York.

Mueller, O., Hahnenberger, K., Dittmann, M., Yee, H., Dubrow, R., Nagle, R., and Ilsley, D. (2000) A microfluidic system for high-speed reproducible DNA sizing and quantitation. *Electrophoresis* **21**: 128-134.

Nolan, T., Hands, R.E., and Bustin, S.A. (2006) Quantification of mRNA using real-time RT-PCR. *Nature Protocols* **1**: 1559-1582.

University of Cape Town

Oliver, M.J., and Bewley, J.D. (1997) Desiccation tolerance of plant tissues: A mechanistic overview. *Horticulture Review* **18**: 171-214.

Oliver, M.J., Tuba, Z., and Mishler, B.D. (2000) The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology* **151**: 85-100.

Pasloske, B.L., Walkerpeach, C.R., Obermoeller, R.D., Winkler, M., and DuBois, D.B. (1998) Armored RNA technology for production of ribonuclease-resistant viral RNA controls and standards. *Journal of Clinical Microbiology* **36**: 3590-3594.

Paul, M. (2007) Trehalose 6-phosphate. *Current Opinion in Plant Biology* **10**: 303-309.

Piatkowski, D., Schneider, K., Salamini, F., and Bartels, D. (1990) Characterization of Five Abscisic Acid-Responsive cDNA Clones Isolated from the Desiccation-Tolerant Plant Craterostigma plantagineum and Their Relationship to Other Water-Stress Genes. *Plant Physiology* **94**: 1682-1688.

Pla, M., Gomez, J., Goday, A., and Pages, M. (1991) Regulation of the abscisic acid-responsive gene rab28 in maize viviparous mutants. *Journal of Molecular Biology* **230**: 394-400.

Prilusky, J., Felder, C.E., Zeev-Ben-Mordehai, T., Rydberg, E.H., Man, O., Beckmann, J.S., Silman, I., and Sussman, J.L. (2005) FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **21**: 3435-3438.

Puntarulo, S., Galleano, M., Sanchez, R.A., and Boveris, A. (1991) Superoxide anion and hydrogen peroxide metabolism in soybean embryonic axes during germination. *Biochimica et Biophysica Acta* **1074**: 277-283.

Radstrom, P., Knutsson, R., Wolffs, P., Lovenklev, M., and Lofstrom, C. (2004) Pre-PCR processing: strategies to generate PCR-compatible samples. *Molecular Biotechnology* **26**: 133-146.

Rapley R.(Ed) (2000) The Nucleic Acid Protocols Handbook. Published by Humana Press, ISBN 9780896034594 : p1050

Raynal, M., Depigny, D., Cooke, R., and Delseny, M. (1989) Characterization of a Radish Nuclear Gene Expressed during Late Seed Maturation. *Plant Physiology.* **91**: 829-836.

Ririe, K.M., Rasmussen, R.P., and Wittwer, C.T. (1997) Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Analitical Biochemistry* **245**: 154-160.

University of Cape Town

Rizhsky, L., Liang, H., and Mittler, R. (2002) The Combined Effect of Drought Stress and Heat Shock on Gene Expression in Tobacco. *Plant Physiology.* **130**: 1143-1151.

Russouw, P.S., Farrant, J.M., Brandt, W., and Lindsey, G.C. (1995) Isolation and characterisation of a heat soluble protein from pea (*Pisum sativum*) embryos. *Seed Science Research* **5**: 137-144.

Sambrook, J.; Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: a laboratory manual.* 2nd ed. N.Y., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, 1659 p. ISBN 0-87969-309-6

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., and Ragg, T. (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology* **7**: 3.

Scott, P. (2000) Resurrection Plants and the Secrets of Eternal Leaf. *Annals of Botany* **85**: 159-166.

Sherwin, H.W., and Farrant, J.M. (1995) Surviving desiccation. *Veld and Flora* **81**: 119-121.

Sherwin, H.W., and Farrant, J.M. (1996) Differences in Rehydration of Three Desiccation-tolerant Angiosperm Species. *Annals of Botany* **78**: 703-710.

Smirnoff, N. (1993) The role of active oxygen in the response of plants to water deficit and desiccation. *New Phytology* **125**: 27-58.

Spanakis, E. (1993) Problems related to the interpretation of autoradiographic data on gene expression using common constitutive transcripts as controls. *Nucleic Acids Research.* **21**: 3809-3819.

Stahlberg, A., Aman, P., Ridell, B., Mostad, P., and Kubista, M. (2003) Quantitative real-time PCR method for detection of B-lymphocyte monoclonality by comparison of kappa and lambda immunoglobulin light chain expression. *Clinical Chemistry* **49**: 51-59.

Stahlberg, A., Hakansson, J., Xian, X., Semb, H., and Kubista, M. (2004) Properties of the reverse transcription reaction in mRNA quantification. *Clinical Chemistry* **50**: 509-515.

Subramanian, R., and Ayyanar, M. (2006) Prediction of functions for two LEA proteins from mung bean. *Bioinformation* **1**: 133-138.

Suan R, Cuevas A, Valpuesta V and Reid MS (1991) Arbutin and sucrose in leaves of the resurrection plant Myrothamnus flabellifolia. *Phytochemistry* **30**: 2555–2556

University of Cape Town

Sun, W.Q., Irving, T.C., and Leopold, A.C. (1994) The role of sugar, vitrification and membrane phase transition in seed desiccation tolerance. *Physiologia Plantarum* **90**: 621-628.

Swire-Clark, G.A., and Marcotte, W.R., Jr. (1999) The wheat LEA protein Em functions as an osmoprotective molecule in Saccharomyces cerevisiae. *Plant Molecular Biology* **39**: 117-128.

Tichopad, A., Dilger, M., Schwarz, G., and Pfaffl, M.W. (2003) Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucleic Acids Research* **31**: e122.

Tolleter, D., Jaquinod, M., Mangavel, C., Passirani, C., Saulnier, P., Manon, S., Teyssier, E., Payet, N., Avelange-Macherel, M.H., and Macherel, D. (2007) Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell* **19**: 1580-1589.

Tuba, Z., Lichtenthaler, H.K.K., Csintalan, Z., and Pocs, T. (1993) Regreening of desiccated leaves of the poikilochlorophyllous Xerophyta scabrida rehydration. *Journal of plant physiology* **142**: 103-108.

Tuba, Z., Lichtenthaler, H.K., Csintalan, Z., Nagy, Z., and Szente, K. (1994) Reconstitution of chlorophylls and photosynthetic CO2 assimilation upon rehydration of desiccated poikilochlorophyllous plant Xerophyta scabrida (Pax) Th. Dur. et. *Planta* **192**: 414-420.

Tunnacliffe, A., and Wise, M.J. (2007) The continuing conundrum of the LEA proteins. *Naturwissenschaften* **94**: 791-812.

Yu, J.N., Zhang, J.F., and Gao, J.F. (2002) Effect of water stress on induced protein expression of wheat seedlings and suspension in cultures. *Acta Bot Boreal-Occident Sin* **52**: 865-870.

Vicient, C.M., Hull, G., Guilleminot, J., Devic, M., and Delseny, M. (2000) Differential expression of the Arabidopsis genes coding for Em-like proteins. *Journal of Experimental Botany* **51**: 1211-1220.

Whelan, J.A., Russell, N.B., and Whelan, M.A. (2003) A method for the absolute quantification of cDNA using real-time PCR. *Journal of Immunological Methods* **278**: 261-269.

Williamson, J.D., Quatrano, R.S., and Cuming, A.C. (1985) Em polypeptide and its messenger RNA levels are modulated by abscisic acid during embryogenesis in wheat. *European Journal of Biochemistry* **152**: 501-507.

Wisniewski, K., and Zagdanska, B. (2001) Genotype-dependent proteolytic response of spring wheat to water deficiency. *Journal of Experimental. Botany* **52**: 1455-1463.

University of Cape Town

Wise, M.J. (2003) LEAping to conclusions: a computational reanalysis of late embryogenesis abundant proteins and their possible roles. *BMC Bioinformatics* **4**: 52.

Wise, M.J., and Tunnacliffe, A. (2004) POPP the question: what do LEA proteins do? *Trends Plant Science* **9**: 13-17.

Wolkers, W.F., McCready, S., Brandt, W.F., Lindsey, G.G., and Hoekstra, F.A. (2001) Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses in vitro. *Biochimica et Biophysica Acta* **1544**: 196-206.

Yang, Q.H., Ye, W.H., Song, S.Q., and Yin, S.H. (2002) Seed desiccation tolerance and its relationship to seed types and developmental stages. *Acta Bot Boreal-Occident Sin* **22**: 1518-1525.

Zhang, Z., Honda, C., Kita, M., Hu, C., Nakayama, M., and Moriguchi, T. (2003) Structure and expression of spermidine synthase genes in apple: two cDNAs are spatially and developmentally regulated through alternative splicing. *Molecular Genetics and Genomics* **268**: 799-807.

Zhang, L.S., and Zhao, W.M. (2003) LEA protein functions to tolerance drought of the plant. *Plant Physiology Communications* **39**: 61-66.

Zhang, H., Sreenivasulu, N., Weschke, W., Stein, N., Rudd, S., Radchuk, V., Potokina, E., Scholz, U., Schweizer, P., Zierold, U., Langridge, P., Varshney, R.K., Wobus, U., and Graner, A. (2004) Large-scale analysis of the barley transcriptome based on expressed sequence tags. *Plant Journal* **40**: 276-290.

University of Cape Town