Identification and molecular characterisation of unknown genes involved in desiccation tolerance mechanisms in the resurrection plant *Craterostigma plantagineum*

Valentino Giarola

Identification and molecular characterisation of unknown genes involved in desiccation tolerance mechanisms in the resurrection plant *Craterostigma plantagineum*

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Valentino Giarola

aus

San Pietro di Morubio, Italien

Bonn, 2014

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen
Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Dorothea Bartels
2. Gutachter: Prof. Dr. Peter Dörmann

Tag der Promotion: 2. Oktober 2014
Erscheinungsjahr: 2014

A Linda

donna unica e straordinaria,

nonché futura madre di mia figlia,

che grazie alla sua presenza,

ai suoi innumerevoli sacrifici

e al suo costante supporto

ha reso tutto questo possibile

CONTENTS

CONTENTS

SUMMARY

In higher plants desiccation tolerance mainly occurs in seeds and in vegetative tissues of a small group of species termed resurrection plants. Among resurrection plants, *Craterostigma plantagineum* has been extensively studied to understand the molecular basis of desiccation tolerance. The mechanisms that enable desiccation tolerance in this species are not fully understood, however many of the genes that are induced during desiccation have been reported. Recently, transcriptome data from this plant was obtained thus providing the gene sequence information for molecular studies. Here *C. plantagineum* is used as an experimental system to further characterise desiccation tolerance mechanisms in this plant. RT-qPCR is a very sensitive and accurate technique to study gene expression but cannot be performed without any prior validation of genes used to normalise data. To allow RT-qPCR expression studies, putative reference genes were identified in the transcriptome and their stability in different plant tissues was investigated. Next, transcriptome data from *C. plantagineum* was used as a source to discover unknown genes involved in desiccation tolerance mechanisms in this plant.

In CHAPTER 2, the stability of eight commonly used reference genes (*ACT*, *CAC*, *EF1α*, *EIF5A*, *GAPDH*, *TKT3*, *PP2AA3*, *YLS8*) was analysed in leaves and roots during dehydration and rehydration and in ABA-treated, dried and ABA-treated dried callus in order to identify the most suitable genes for expression analysis in these tissues. *YLS8*, *EIF5A* and *TKT3* were identified as the optimal reference genes for normalizing gene expression data in leaves whereas *EF1α*, *EIF5A* and *GAPDH* for roots and *YLS8*, *CAC* and *GAPDH* for callus. This study represents the first evaluation of gene stability in a resurrection plant. Validated genes allow the use of RT-qPCR for quantitative gene expression studies in *C. plantagineum*.

In resurrection plants desiccation tolerance mainly relies on the induction of cellular protection mechanisms during dehydration. It is believed that the accumulation of sucrose and protective proteins (e.g. LEA proteins) contribute to protection mechanisms in *C. plantagineum*. Important but less characterised aspects, i.e., folding of cell walls and the accumulation of protective proteins in the chloroplasts during dehydration as well as protection and repair mechanisms during rehydration are thought to be essential to the desiccation tolerance phenomenon in *C. plantagineum*.

Although, the rehydration process is part of the survival strategy of resurrection plants, it has been scarcely explored and underlying mechanisms are unknown. Here unknown genes involved in leaf folding, chloroplast protection and rehydration-specific responses were identified and molecularly characterised.

In CHAPTER 3, the involvement of an apoplastic glycine-rich protein (CpGRP1) in the cell wall shrinking and folding mechanisms was investigated. It was shown that the *CpGRP1* transcript and protein accumulate in dehydrating leaves thus suggesting a role in the leaf dehydration process. Promoter studies demonstrated that the *CpGRP1* gene expression is mainly regulated on the transcriptional level, is independent of ABA and involves a drought responsive *cis*-element (DRE). The CpGRP1 protein shows sequence similarities with the *Arabidopsis thaliana* pathogenesis-related protein 3 (AtGRP-3) which is known to interact with cell wall associated protein kinases (AtWAKs). Two AtWAK orthologs were identified in *C. plantagineum* and an interaction was demonstrated between one of them, i.e., CpWAK1 and CpGRP1. Cell expansion in *A. thaliana* requires WAKs. We showed that *CpWAKs* genes are downregulated during dehydration. Our data suggest that CpWAKs might be working similarly to AtWAKs in the control of the cell volume and that cell folding during dehydration and expansion during rehydration may require an interaction of CpWAKs and CpGRP1. Cell wall pectins and dehydration-induced pectin modifications are predicted to be involved in the activity of the CpGRP1-CpWAK1 complex. However, further experiments are required to demonstrate the biological relevance of the CpGRP1/CpWAK1 interaction. The identification of additional CpGRP1 interaction partners would also help to understand the role of CpGRP1 and its relation to pectins.

In CHAPTER 4, the characterisation of a cysteine-rich rehydration-responsive 1 protein (CpCRP1) and of an early dehydration responsive 1 protein (CpEDR1) is reported. It was shown that CpCRP1 is secreted in the cell apoplasm and is likely cross-linked to other cell wall components. On the other hand, CpEDR1 is imported into the chloroplasts where it is predicted to interact with other proteins or with membranes. Orthologs of *CpCRP1* and *CpEDR1* were exclusively identified in *L. brevidens* and *L. subracemosa* which belong to the same family as *C. plantagineum* thus suggesting a recent evolution of the genes in this family. Genes occurring only in a species or in closely related species are defined as taxonomically restricted genes (TRGs). TRGs may contribute to the adaptation of organisms to different

environmental conditions. *CpCRP1* represents the first report about a TRG associated with the rehydration process. This suggests that rehydration-specific mechanisms may also contribute to desiccation tolerance in resurrection plants. *C. plantagineum* maintains the photosynthetic machinery intact during dehydration and thus it is more exposed to oxidative stress. Our data suggests that *CpEDR1* is a TRG likely involved in protection mechanisms in the chloroplasts. The results of this study suggest that recently evolved genes are involved in desiccation tolerance mechanisms in *C. plantagineum*. Comparative analyses using *C. plantagineum* and other members of the Linderniaceae family may provide additional evidence of the contribution of unique genes to desiccation tolerance in *C. plantagineum*.

As indicated by the results of this thesis, desiccation tolerance is a complex trait which is far to be completely understood. Currently, the *C. plantagineum* transcriptome represents a promising source for the further identification of genes involved in desiccation tolerance mechanisms.

CHAPTER 1

General Introduction

1. Desiccation tolerance in plants: definition, distribution and evolution

Water is essential for life. In plants, water usually constitutes more than the 90 % of the fresh weight, acts as solvent, participates as reactant in biochemical reactions and is required to maintain the cell turgor (Kramer & Boyer, 1995). Plants are constantly loosing water to the surrounding environment through leaves due to their disequilibrium with the atmosphere. Different levels of adaptation have been evolved in plants to survive in a water-limited environment, i.e., desiccation avoidance, desiccation resistance and desiccation tolerance (Le & McQueen-Mason, 2006). The majority of land plants survives dry periods by producing specialised structures (e.g. pollen, seeds and spores) able to cope with the substantial water loss and thus desiccation avoidance is the most frequent adaptation observed in plants. Desiccation resistant plants have developed particular morphological structures to retain the cellular water mainly by reducing water flux, e.g., through reduction of leaf growth, waxes and hairs, or increase water uptake with specialized roots.

Desiccation tolerance is defined as the ability of an organism to dry to equilibrium with dry air and to resume normal metabolic function after rehydration (Bewley, 1979; Alpert, 2005; Wood, 2005; Alpert, 2006; Wood & Jenks, 2007).

In the plant kingdom the desiccation tolerance trait is widely distributed but uncommon (Bewley & Krochko, 1982; Oliver & Bewley, 1997). Desiccation tolerance is observed mostly in ancient plant clades such as algae, lichens and bryophytes. Among vascular land plants this trait predominantly occurs in ferns and fern allies but is rare in angiosperms. About 350 desiccation tolerant species, the so-called "resurrection plants", have been identified so far within angiosperms (Tuba & Lichtenthaler, 2011).

Desiccation tolerance is thought to be a very old trait in plant evolution. Using phylogenetic evidence Oliver *et al.* (2000) postulated that the colonization of the land by ancient plants was accomplished through the evolution of vegetative desiccation tolerance. The increase in complexity and the evolution of a water transport system in

vascular plants eventually led to the loss of desiccation tolerance in vegetative organs (Oliver *et al.*, 2000). However, independent evolution (or re-evolution) events of the desiccation tolerant trait occurred in *Selaginella,* ferns and within angiosperms (Oliver *et al.*, 2000).

2. Strategies used by land plants to survive desiccation

Vegetative desiccation tolerant plants use different types of protection mechanisms which are connected to their variable sensitivity to the dehydration rate. Primitive organisms such as algae, bryophytes and lichens survive very fast drying. Studies mainly conducted using the moss *Tortula ruralis* suggest that a combination of constitutive protection measures during drying and the activation of pre-existing repair mechanisms upon rehydration allow desiccation tolerant bryophytes to survive desiccation (Oliver *et al.*, 2000; Oliver, 2008). In *T. ruralis* high constitutive levels of sucrose (Oliver, 2008) and class II LEA proteins (Bewley *et al.*, 1993) were observed and are supposed to confer constitutive protection to desiccation. Repair mechanisms seem mainly regulated through translational control in *T. ruralis*. Transcripts involved in these mechanisms are accumulated but prevented from translation until the first two hours of rehydration through the formation of mRNP particles (Wood & Oliver, 1999). Recently, the analysis of the transcriptome of *T. ruralis* revealed that LEA proteins are the most abundant proteins expressed during rehydration (Oliver *et al.*, 2004).

In resurrection plants desiccation tolerance relies predominantly on the induction of cellular protection systems during dehydration (Gaff, 1989; Oliver *et al.*, 1998; Cushman & Oliver, 2011). For this reason resurrection plants can revive from desiccation only when water loss occurs at a sufficient low rate (Bartels & Salamini, 2001). In contrast to less complex organisms, these plants possess morphological and physiological characteristics to slow-down water loss and to reduce dehydration rates (Farrant *et al.*, 1999).

3. Desiccation tolerance in resurrection plants

Resurrection plants re-evolved specific cellular protection mechanisms that collectively confer plant desiccation tolerance (Farrant, 2000; Farrant *et al.*, 2007; Moore *et al.*, 2013). These mechanisms involve a combination of genetic, metabolic and antioxidant systems together with macromolecular and structural stabilizing processes (Moore *et al.*, 2009).

The accumulation of non-reducing sugars and the synthesis of protective proteins, mainly LEA proteins are observed in both desiccation tolerant seeds and vegetative tissues of resurrection plants and are believed to participate in protection mechanisms during desiccation (Ingram & Bartels, 1996). Besides regulating the water-holding capacity of cells, sugars are supposed to protect cells through a process termed "glass phase formation" or "vitrification". A glass is a supersaturated liquid with the mechanical properties of a solid that is able to slow down chemical reactions or even to stop them altogether. In this amorphous and metastable glassy state, conformational changes of proteins and membrane fusions are prevented (Crowe *et al.*, 1998). It is assumed that with the progression of the water stress severity, water is replaced by sugar molecules which form hydrogen bounds with polar groups of macromolecules thus stabilising their native conformation (Webb, 1965; Clegg *et al.*, 1982). At a particular point during dehydration, molecular mobility enormously decreases and the cytoplasm vitrifies. In seeds and resurrection plants various soluble carbohydrates may be present in fully hydrated tissues, but sucrose usually accumulates in the dried state. For example, desiccation in the leaves of *C. plantagineum* is accompanied by conversion of the C8-sugar 2-octulose (90% of the total sugars in hydrated leaves) into sucrose, which eventually accounts for about the 40% of the dry weight (Bianchi *et al.*, 1991). In addition LEA proteins also possess properties similar to non-reducing sugars and likewise contribute to glass formation and together with other mechanisms to desiccation tolerance (Hoekstra *et al.*, 2001; Wolkers *et al.*, 2001).

Apart from LEA proteins previous studies in *C. plantagineum* revealed the presence of unique genes linked to desiccation tolerance, i.e., *CDT-1* (Furini *et al.*, 1997) and *CpEdi-9* (Rodrigo *et al.*, 2004). *CDT-1* is a small regulatory non-coding RNA which was shown to confer constitutive desiccation tolerance to *C. plantagineum* callus whereas the dehydration-responsive gene *CpEdi-9* is supposed to protect cells of

seeds and leaf phloem tissues from desiccation mediated damage (Rodrigo *et al.*, 2004). Genes which do not share any similarity to genes in other species are found in every eukaryotic genome and are defined as orphan or taxonomically restricted genes (TRGs) (Wilson *et al.*, 2005). Increasing evidence suggests that TRGs together with the variation of regulatory networks permit organisms to evolve and adapt to different environmental conditions (Khalturin *et al.*, 2009).

3.1. LEA proteins

LEA proteins were initially identified as the major class of proteins accumulating during the maturation of cotton (*Gossypium hirsutum*) seeds (Dure *et al.*, 1981). Subsequent studies reported the presence of similar proteins in seeds of many other species but also in vegetative tissues of desiccation tolerant and sensitive plants under cold, drought, high salinity stresses or exogenous ABA treatment (Ramanjulu & Bartels, 2002). Lately, LEA proteins were also found in other non-plant organisms like bacteria (Stacy & Aalen, 1998; Battista *et al.*, 2001; Dure, 2001), cyanobacteria (Close & Lammers, 1993), slime molds (Eichinger *et al.*, 2005), fungi (Sales *et al.*, 2000; Katinka *et al.*, 2001; Abba *et al.*, 2006). In addition LEA and LEA-like proteins were reported in animals, e.g., nematodes (Browne *et al.*, 2002; Browne *et al.*, 2004; Gal *et al.*, 2004), rotifers (Tunnacliffe *et al.*, 2005; Denekamp *et al.*, 2009; Denekamp *et al.*, 2010), tardigrades (Forster *et al.*, 2009) and arthropods such as the larva of the African midge *Polypedilum vandeplanki* (Kikawada *et al.*, 2006), the Arctic springtail *Onychiurus arcticus* (Clark *et al.*, 2007), the embryos of the brine shrimp *Artemia franciscana* (Sharon *et al.*, 2009) and the Antarctic midge *Belgica Antarctica* (Teets *et al.*, 2012).

Structure

LEA proteins are characterised by high hydrophilicity and by a biased amino acid composition due to the overrepresentation of certain amino acid residues such as Gly, Ala, Glu, Lys/Arg, and Thr and the lack or a low proportion of Cys and Trp residues. Most of the LEA proteins belong to the group of the so-called "hydrophilins" (Garay-Arroyo *et al.*, 2000). Hydrophilins are characterised by a glycine content higher than 6 % and a hydrophilicity index higher than 1. The first LEA protein which was structurally characterised was the wheat Em protein (McCubbin *et al.*, 1985). Using biophysical

techniques, McCubbin *et al.* (1985) showed that this protein lacked compactness, had an asymmetrical or flexible conformation and just a little secondary structure with as much as 70% of the protein behaving as a random coil. The extensive lack of secondary structures are largely typical for most hydrophilic LEA proteins and led to classify these proteins as "natively unfolded", "intrinsically disordered" or "intrinsically unstructured" (Uversky *et al.*, 2000; Dunker *et al.*, 2001; Tompa, 2002).

Classification

Two systems are mainly used to classify LEA proteins. The first system was introduced by Dure *et al.* (1989) and groups LEA proteins according to the conservation of short amino acid sequence motifs which occur one or several times within the protein sequence. Recently, a computational-based method which classifies LEAs based on peptide composition was developed and led to the definition of superfamilies of LEA proteins (Wise, 2002; Wise & Tunnacliffe, 2004). Currently, there is no general consensus in the scientific community concerning which method should be preferentially used to classify these proteins. Battaglia *et al.* (2008) argue that the system based on conserved motifs would better underline structural, functional and evolutionary relationships among LEA proteins. The classification of Dure initially based on three groups was extended with the discovery of new LEAs over time and presently divides these proteins into seven groups (Battaglia *et al.*, 2008). According to the terminology first introduced by Cuming (1999) and followed by Battaglia (2008) each group is represented by the proteins which were initially used to described these groups. Groups 1 to 4 correspond to the first LEA proteins described from cotton: group 1 (D-19), group 2 (D-11), group 3 (D-7/D-29), group 4 (D-113). All the uncommon LEAs belong to group 5 (D-34, D-73, D-95). The LEA18 (Colmenero-Flores *et al.*, 1999) and ASR1 (Silhavy *et al.*, 1995; Rossi *et al.*, 1996) are used to define group 6 and 7, respectively. Group 1 LEA proteins are characterised by a high proportion of glycine (approximately 18 %), glutamate and glutamine and a hydrophilic 20-amino-acid motif. Group 2 LEA proteins, also known as dehydrins, are the best-studied. Mainly three different motifs, named K-, the Y- and S-segment are observed in this group (Close *et al.*, 1989; Close *et al.*, 1993; Campbell & Close, 1997). While the K-segment is characteristic of this group, the Y- and S-segments are not necessarily present in dehydrin sequences. The K-segment is a conserved Lys-rich 15-residue motif,

EKKGIMDKIKEKLPG, which is predicted to form an amphipathic α-helix and can occur up to 11 times in a single sequence. The Y-segment motif whose conserved sequence is [V/T]D[E/Q]YGNP is usually located at the N-terminus of group 2 proteins and may occur in one to 35 tandem copies. Finally, the S-segment consists of a stretch of serine residues which can undergo phosphorylation. Group 3 LEA proteins contain multiple copies of an 11-amino-acid motif. Dure *et al.* (1993) predicted that this motif may form an amphipathic α-helix with possibilities for intra- and inter-molecular interactions. Group 4 LEA proteins are conserved in their N-terminal region which is about 70 to 80 residues long and is predicted to form α-helices. The less conserved C-terminal part is variable in size and is supposed to assume a random coil structure (Dure, 1993). In contrast to other groups, LEA proteins in group 5 contain more hydrophobic residues and are not soluble after boiling suggesting that they adopt globular structures (Cuming, 1999). LEA proteins in group 6 are small in size (7-14 kD) and are characterised by the presence of four motifs, two of which (1 and 2) are highly conserved (Battaglia *et al.*, 2008). Characteristics of LEAs in the group 7 are the reduced size, the heat stability, the lack of structures and the presence of three highly conserved regions one of which contains a nuclear localisation signal (Silhavy *et al.*, 1995). Over time, many new proteins have been defined as LEAs on the basis of either biochemical properties or sequence similarities alone. In the case of no obvious sequence relatedness to previously recognised LEA proteins, such proteins have led to a designation of new groups or at least to a recognition that they fall outside the existing categories. These proteins are often referred to as "LEA-like" proteins (Tunnacliffe & Wise, 2007).

Function

There is a strong association in different organisms between expression of LEAs and tolerance to water stress thus suggesting that these proteins represent a common adaptation to water deficit (Garay-Arroyo *et al.*, 2000). Ectopic expression studies have shown that genes encoding some plant LEA proteins can improve the water stress tolerance of transgenic plants (Xu *et al.*, 1996; Sivamani *et al.*, 2000) and yeast (Imai *et al.*, 1996; Swire-Clark & Marcotte, 1999; Zhang *et al.*, 2000). Chilling tolerance has also been linked to the accumulation of LEA proteins (Danyluk *et al.*, 1998; Ismail *et al.*, 1999b; Ismail *et al.*, 1999a; Puhakainen *et al.*, 2004; Nakayama *et al.*, 2007)

whereas the deletion of a hydrophilin and of a dehydrin gene caused an osmosensitive phenotype in *E. coli* (Garay-Arroyo *et al.*, 2000) and in the moss *Physcomitrella patens* (Saavedra *et al.*, 2006), respectively. On the other hand, some studies have not shown a correlation between the overexpression of LEA genes and increasing chilling or water stress tolerance (Iturriaga *et al.*, 1992; Kaye *et al.*, 1998). This may indicate either that not all LEA proteins play a key role in the plant stress tolerance, or that, as it has been shown for transgenic strawberries, they need a particular background to be functional (Houde *et al.*, 2004).

Since the discovery of their accumulation during the acquisition of desiccation tolerance in seeds, LEAs have been linked to desiccation tolerance (Cuming, 1999). Even though many LEA proteins from many different species were reported only a restricted group has been functionally and structurally characterised. These proteins are found in different plant tissues (Rorat *et al.*, 2004), several subcellular compartments like cytoplasm, nucleus, vacuole, mitochondria, chloroplasts (Schneider *et al.*, 1993; Rinne *et al.*, 1999; Borovskii *et al.*, 2002; Heyen *et al.*, 2002; Tunnacliffe & Wise, 2007; Mehta *et al.*, 2009) or in connection with plasma membranes (Danyluk *et al.*, 1998). Molecular and biochemical features together with their broad localisation strongly suggest a protective role for these proteins (Bartels & Salamini, 2001) but underlying mechanisms are far from being completely understood. LEA proteins are predicted to have several protective functions (Wise & Tunnacliffe, 2004; Reyes *et al.*, 2005; Hincha & Thalhammer, 2012). Two well-studied functions of LEA proteins concern the stabilisation of enzymes and membranes during freezing and desiccation. The ability of several dehydrins to protect freeze-labile enzymes such as lactate dehydrogenase or malate dehydrogenase has been shown by *in vitro* cryoprotection assays (Lin & Thomashow, 1992; Honjoh *et al.*, 2000; Hara *et al.*, 2001; Bravo *et al.*, 2003; Sanchez-Ballesta *et al.*, 2004; Reyes *et al.*, 2005; Dang *et al.*, 2014). It has been proposed that these proteins play a role in preventing the desiccation- or freezing-induced aggregation and inactivation of enzymes (Goyal *et al.*, 2005; Chakrabortee *et al.*, 2007; Kovacs *et al.*, 2008; Boucher *et al.*, 2010; Chakrabortee *et al.*, 2012).
Several plant and animal LEAs undergo folding into mainly α-helical structures when they bind a cation, when dried or when in contact with certain phospholipid membranes (Tunnacliffe *et al.*, 2010; Hand *et al.*, 2011). Regions of LEA proteins were predicted to form amphipathic α-helical structures *in silico*. These structures are known for their ability to interact with membranes (Segrest *et al.*, 1992) and a membrane binding

activity was demonstrated by Fourier-transform IR spectroscopy for both rotifers and plant proteins (Pouchkina-Stantcheva *et al.*, 2007; Rahman *et al.*, 2010; Thalhammer *et al.*, 2010; Tolleter *et al.*, 2010; Popova *et al.*, 2011). A clear evidence of a membrane-stabilizing function *in vivo* has been shown for the cold-induced chloroplast LEA protein COR15A from *Arabidopsis* (Artus *et al.*, 1996).

Other LEA activities, i.e., protection of DNA, phosphorylation dependent ion binding, radical scavenging and stabilisation of the glassy state in the dry state have also been described (Tunnacliffe & Wise, 2007; Tunnacliffe *et al.*, 2010). The questions of how different LEA protein structures and functions are related to each other and how these structures select for target molecules remain to be answered (Hincha & Thalhammer, 2012).

3.2. Morphological changes and modifications in cell wall properties

Folding of leaves during dehydration is a distinctive characteristic of resurrection plants (Gaff, 1989; Farrant, 2000; Farrant *et al.*, 2003; Willigen *et al.*, 2003; Farrant *et al.*, 2007). Leaves of *C. plantagineum* and *C. wilmsii* gradually curl inward with the progression of dehydration and eventually become extensively folded exposing only the abaxial surfaces of the most external leaves of the whorl to the light (Sherwin & Farrant, 1998). Leaf folding was proposed as a mechanism used to shade chlorophyll and thus avoid ROS formation in water-limiting conditions (Farrant, 2000). It was demonstrated that this process is strictly correlated to the ability of *C. wilmsii* to revive from desiccation (Farrant *et al.*, 2003). At the microscopic level, cell walls of resurrection plants appear extensively contracted and folded upon desiccation (Gaff, 1989; Farrant, 2000; Farrant *et al.*, 2003; Willigen *et al.*, 2003; Farrant *et al.*, 2007). During dehydration and rehydration considerable fluctuations of cell volume takes place thereby causing severe mechanical stress which needs to be managed by resurrection plants to avoid irreversible damage (Farrant, 2000). The shrinking and folding of cell walls is supposed to counteract the dehydration-induced mechanical stress allowing cell walls of desiccated tissues to be mechanically stabilised (Iljin, 1957; Farrant, 2000) but the underlying mechanisms are largely unknown. In *Myrothamnus flabellifolia* arabinan-rich pectins and arabinogalactan proteins are proposed to maintain cell wall flexibility during desiccation (Moore *et al.*, 2013). It has been proposed for *Craterostigma* ssp. that dehydration-induced modifications of existing cell

wall components might play a major role to increase cell wall flexibility during dehydration (Moore *et al.*, 2013). Expansin activity, xyloglucan remodelling and calcium ion accumulation have been proposed to take part in the cell wall folding process (Vicré *et al.*, 1999; Jones & McQueen-Mason, 2004; Vicré *et al.*, 2004). Cell walls of *Craterostigma* are similar in pectin content and composition to other dicot plants and no significant differences were observed between untreated and dehydrated leaf cell walls (Vicré *et al.*, 2004; Moore *et al.*, 2013). It is hypothesised that in *C. plantagineum* small but still undefined differences might be involved in preparing plant cell walls to desiccation (Moore *et al.*, 2013).

3.3. Chloroplastic proteins with protective roles

Resurrection plants down-regulate photosynthesis-related genes and shut down photosynthesis upon desiccation (Bockel *et al.*, 1998). Some resurrection plants, termed as homoiochlorophyllous plants retain chlorophyll and maintain structurally intact thylakoid membranes during desiccation whereas plants defined as poikilochlorophyllous degrade chlorophyll and dismantle photosystem complexes (Tuba *et al.*, 1998). However, homoiochlorophyllous plants do undergo changes in the morphology and membrane lipid composition of chloroplasts (Schneider *et al.*, 1993; Navari-Izzo *et al.*, 1995; Quartacci *et al.*, 1997; Thomson & Platt, 1997; Farrant, 2000; Gasulla *et al.*, 2013). Homoiochlorophyllous desiccation tolerant plants recover more rapidly from desiccation than poikilochlorophyllous plants but need additional mechanisms to protect them from reactive oxygen species which are generated due to light. It has been proposed that homoiochlorophyllous plants undergo leaf folding and anthocyanin accumulation to minimize the interaction of chlorophyll with light and thus prevent photo-oxidative damage (Sherwin & Farrant 1998; Farrant 2000). Moreover, different classes of proteins are induced, and accumulate, in chloroplasts upon desiccation and seem to be involved in protection mechanisms (Schneider *et al.*, 1993; Alamillo & Bartels, 1996; Neale *et al.*, 2000). Genes encoding proteins with high similarity with early light-inducible proteins (ELIPs) were reported in *C. plantagineum* and *Sporobolus stapfianus*. ELIPs are known to bind to thylakoid membranes and are involved in the protection of the photosynthetic apparatus from photo-oxidative damage (Hutin *et al.*, 2003). Other dehydration- and ABA-induced proteins with hypothetical protective roles were reported in *C. plantagineum*, i.e., the plastid-targeted

proteins (CpPTPs) (Phillips *et al.*, 2002) and the LEA-like dehydration stress protein 21 (dsp 21) and 34 (dsp 34) (Schneider *et al.*, 1993). CpPTPs are a small nuclear gene family encoding proteins with DNA-binding activity supposed to take part in plastid DNA protection and gene regulation during desiccation (Phillips *et al.*, 2002). Moreover, the prediction of coiled coil structures in CpPTPs suggests that these proteins may interact with other plastid proteins to perform their function (Phillips *et al.*, 2002). Detoxifying enzymes such as aldehyde dehydrogenase were also reported to accumulate in dehydrating chloroplasts (Kirch *et al.*, 2001).

3.4. ABA-dependent and -independent regulation of protective mechanisms

The phytohormone abscisic acid (ABA) plays a major role in seed development and abiotic stress responses including dehydration (Leung & Giraudat, 1998). ABA is involved in many of the metabolic changes leading to desiccation tolerance (Giraudat *et al.*, 1994). During dehydration the level of ABA increases and leads to the activation of several dehydration-regulated genes such as LEA proteins (Ingram & Bartels, 1996). Essential regulatory elements were identified through the study of promoters from genes which were responsive to both stress and ABA.

The ABRE (ABA responsive element), containing the core sequence ACGT, was for the first time identified in the promoter region of the LEA *Em* gene of wheat (Marcotte *et al.*, 1989). The presence of the core sequence alone is not sufficient to confer ABA responsiveness but nucleotides flanking the ACGT were shown to be required for the specificity of the ABA responses. It was demonstrated that this element plays a key role in ABA signalling during seed development and under abiotic stress (Bartels & Sunkar, 2005; Yamaguchi-Shinozaki & Shinozaki, 2005). In *C. plantagineum*, e.g., up to 7-fold increase of ABA levels is observed in dried leaves and a large number of dehydration- and ABA-induced genes were reported (Bartels *et al.*, 1990). ABREs were identified in the regulatory regions of several of these genes, i.e., *CDeT6-19* (Michel *et al.*, 1994), *CDeT27-45* (Michel *et al.*, 1993), *CpEdi-9* (Rodrigo *et al.*, 2004), *CpC2* (Ditzer & Bartels, 2006), *CDeT11-24* (Velasco *et al.*, 1998; van den Dries *et al.*, 2011) and their requirement for the promoter ABA responsiveness was clearly proved by site-directed mutagenesis for *CDeT11-24* gene (van den Dries *et al.*, 2011). The presence of non-functional ABREs has also been reported for the *CDeT6-19* promoter (Michel *et al.*, 1994).

14

Besides ABA other factors are responsible for the induction of genes regulated upon dehydration. Frank *et al.* (1998) for example showed that the transcription factor *CPHB-1* from *C. plantagineum* is induced by drying but is insensitive to ABA. A second *cis*-element termed DRE/CRT/LTRE (drought responsive/C-repeat/low temperature response) is found in the promoter of different drought-, high salinity- and cold-responsive genes (Lata & Prasad, 2011) and is the target of DREBs (dehydration responsive element binding) transcription factors. The first DRE element was identified in the promoter of the *Arabidopsis* desiccation- and cold-responsive *rd29A* gene (Yamaguchi-Shinozaki & Shinozaki, 1994). A DRE element was shown to be important for the activity of *CDeT11-24* promoter in *C. plantagineum*. There is increasing evidence that ABA has similar signalling roles in resurrection plants and in *Arabidopsis* during dehydration. In *Arabidopsis* the promoter activity and expression analyses of LEA genes under ABA or cold/drought stresses support the idea of a crosstalk between cold/drought and ABA transduction pathways (Hundertmark & Hincha, 2008) but further analysis of promoter elements is required to establish the general regulatory pathways.

4. Desiccation tolerance within the Linderniaceae

Most of the resurrection plants are found in tropical and subtropical regions particularly on rock outcrops or inselbergs which are characterised by scarce water availability. In dicotyledonous, desiccation tolerance was reported in the order of Gunnerales and Lamiales (Porembski, 2011). Originally, desiccation tolerant plants, such as *Craterostigma plantagineum*, were grouped in the Scrophulariaceae family in the order Lamiales. Phylogenetic studies determined that this family was polyphyletic and this led to a revision of the original classification of the genera within this family (Rahmanzadeh et al., 2005). According to the new classification introduced by Rahmanzadeh *et al.* (2005) *Craterostigma* and *Lindernia* spp. belong to the Linderniaceae family. Linderniaceae as a separate lineage was confirmed by following studies with the inclusion of new taxa (Albach *et al.*, 2005; Oxelman *et al.*, 2005; Schäferhoff *et al.*, 2010). Recently, the first detailed phylogenetic analysis of the Linderniaceae family was reported (Fischer *et al.*, 2013).

Although *Craterostigma* and *Lindernia* spp. are phylogenetically closely related, they differ for their ability to tolerate desiccation. All *Craterostigma* spp. and several

Lindernia spp. from rock outcrops have been shown to be desiccation tolerant; however, the majority of *Lindernia* spp. are desiccation sensitive (Fischer, 1995; Seine *et al.*, 1995). Recently, *Lindernia brevidens*, a plant of the montane rain forests of East coastal Africa, was shown to tolerate extreme desiccation (Phillips *et al.*, 2008), although it never experiences dry periods in its natural habitat. *Lindernia subracemosa* is a close relative of *Lindernia brevidens* and *Craterostigma plantagineum* but it is not desiccation tolerant. Phylogenetic relationships of selected members of the Linderniaceae family including *C. plantagineum*, *L. brevidens* and *L. subracemosa* are shown in Fig. 1.

Diversity centres of Linderniaceae are tropical Africa and Southeast Asia (Fischer, 1992). Asian representatives mainly occur in rain forest areas whereas African Linderniaceae are mainly found in specialised habitats such as seasonally water-filled rock pools, inselbergs and heavy metal (copper or cobalt) soils (Rahmanzadeh *et al.*, 2005).
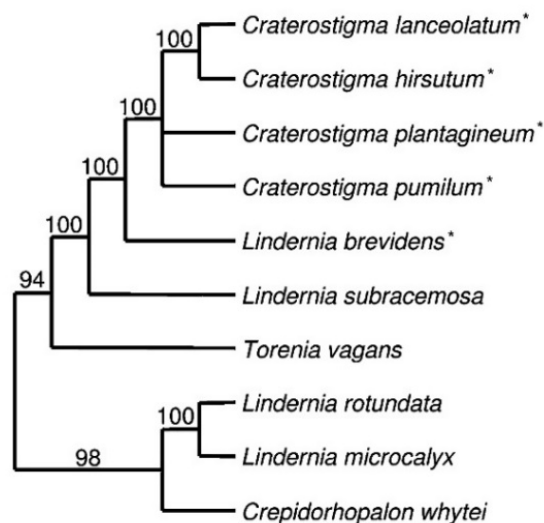
Fig. 1 Phylogenetic relationships of selected members within the Linderniaceae family. Bootstrap values are indicated above the branches. The desiccation tolerant species are indicated with an asterisk (from Phillips *et al.*, 2008).

5. The experimental system *Craterostigma plantagineum*

The Southern African plant *C. plantagineum* has been extensively studied at the molecular level to understand mechanisms underlying desiccation tolerance. The dehydration process in *C. plantagineum* is characterised by the accumulation of many dehydration-induced genes, which rapidly disappear upon rehydration (Bernacchia *et*

al., 1996). A representation of a typical dehydration/rehydration cycle and of the most frequent dehydration and rehydration stages which are compared in molecular studies is shown in Fig. 2. One advantage of using *C. plantagineum* as experimental system resides in the possibility to perform molecular analysis using both differentiated plant tissues and undifferentiated callus and thus compare gene expression in two genetically identical systems (Bartels, 2005). A remarkable feature of the *C. plantagineum* callus is that it can be switched between a desiccation sensitive state (normal condition) to a desiccation tolerant state by exogenous ABA treatments (Bartels *et al.*, 1990). This observation together with the fact that ABA level increases in leaves during dehydration and many dehydration-induced genes can also be induced by ABA supports the essential role of ABA in the acquisition of desiccation tolerance in *C. plantagineum* (Bartels, 2005). The same genes were found to be induced in leaves and callus by ABA and thus the use of the callus system may help to identify genes involved in the acquisition of desiccation tolerance in this plant.
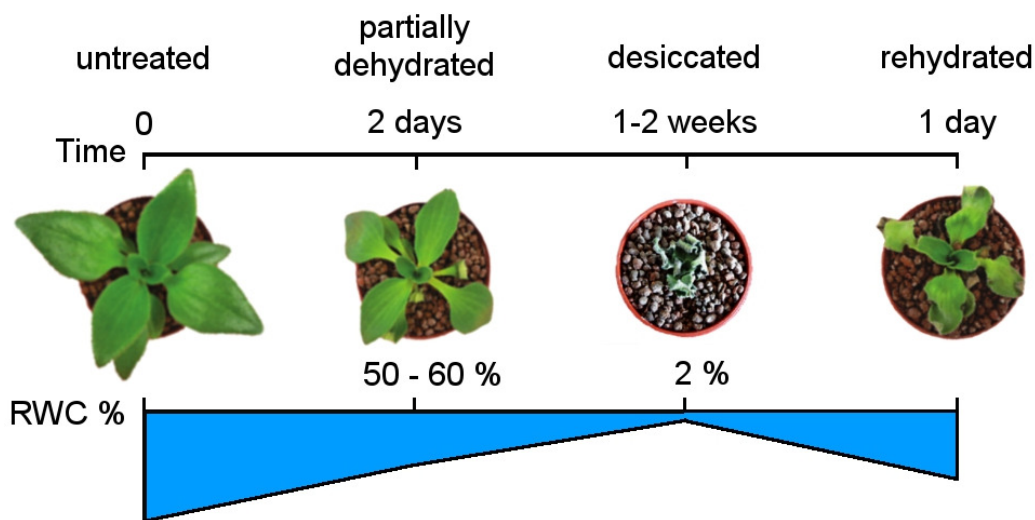
Fig. 2 Dehydration/rehydration cycle in *C. plantagineum*. The plant suspends all its activities upon desiccation and rapidly resumes them after rehydration. Early responses to dehydration are often studied in plants dehydrated to 50-60 % RWC (about 2 days) whereas late responses are analysed in plants dehydrated to 2 % RWC (1-2 weeks dehydration). Plant rehydration is achieved by submerging desiccated plants in water for 1 day. Reported dehydration times refer to plant grown and dehydrated in pots under conditions described by Bartels *et al.* (1990). RWC stands for relative water content a common measure of the hydration status of *C. plantagineum* leaves (Bernacchia *et al.*, 1996).

6. New sequencing and gene quantification technologies: use of RNA-seq and RT-qPCR to study plant desiccation tolerance

Transcripts are the functional elements of the genome. The complete set of transcripts in a cell for a specific developmental stage or physiological condition is defined as the cell transcriptome. The study of the transcriptome is the key to understand how organisms grow and respond to different conditions. In the past years, the analysis of gene expression was mainly performed by microarray or reverse transcription quantitative PCR (RT-qPCR) techniques. The main problems and limitations associated with the microarray technique are the requirement of genome sequence information, the high background levels and the limited dynamic range of detection (Okoniewski & Miller, 2006; Royce *et al.*, 2007). Currently, with the development of next generation sequencing technologies a new technique called RNA-sequencing (RNA-seq) was introduced and has become the method of choice for genome-wide expression studies. (Wang *et al.*, 2010). RNA-seq enables the identification and absolute quantification of any single transcript down to cell level and thus allows the detection of changes in gene expression under different biological conditions (Zhou *et al.*, 2010; Cullum *et al.*, 2011). RNA-seq overcomes all the limitations observed by microarray and RT-qPCR especially concerning the requirement of sequence information. Additionally, RNA-seq permits to detect novel transcripts and isoforms, to map exon/intron boundaries, to reveal sequence variations (e.g. SNPs) and splice variants and to identify non protein coding RNAs (Mutz *et al.*, 2013).

All the desiccation tolerant plants have unknown genome sequences and the lack of sequence information has hampered the study of this trait. RNA-seq permits to analyse desiccation tolerance on a genome-wide scale and thus could be of great use to understand the underlying mechanisms. Recently, the transcriptome from different dehydration and rehydration conditions was obtained for two resurrection plants, i.e., *Craterostigma plantagineum* (Rodriguez *et al.*, 2010) and *Haberlea rhodopensis* (Gechev *et al.*, 2013). The two studies suggest that most of the mechanisms involved in vegetative desiccation tolerance are similar to those conferring desiccation tolerance to seed and pollen. Besides the analysis of transcript profiles of known genes, a large fraction of putative TRGs were identified. This fraction may contain unexplored functions which may be important for plant desiccation tolerance. It was reported that quantitative differences detected by RNA-seq not always correlate with RT-qPCR

expression analysis (Nagalakshmi *et al.*, 2008; Bloom *et al.*, 2009). Thus, the validation of quantitative RNA-seq data by different methods including RT-qPCR is mandatory before drawing conclusions on gene expression.

RT-qPCR is one of the most sensitive, accurate and rapid technique to study gene expression. In the past years, the lack of a general consensus on common guidelines describing how to best perform RT-qPCR experiments could have determined the publication of potential wrong and misleading results. Recently, general guidelines for the publication of quantitative real-time PCR experiments were introduced to help the scientific community to correctly perform and evaluate qPCR experiments (Bustin *et al.*, 2009). Normalization is the key for a reliable qPCR assay because it permits to monitor variations in RNA extraction yield, reverse transcription and amplification efficiencies and thus allows the accurate comparisons of mRNA levels across different samples (Bustin *et al.*, 2009). Usually, stably expressed genes defined as reference genes are used to normalize the expression data. It has now become evident that gene stability is linked to the particular organism, tissue type and experimental conditions which are analysed (Schmittgen & Zakrajsek, 2000; Suzuki *et al.*, 2000; Lee *et al.*, 2002; Dheda *et al.*, 2004; Czechowski *et al.*, 2005) and thus it is required to systematically check the stability of genes which are used as references within each experimental condition (Guénin *et al.*, 2009). Statistical methods such as GeNorm (Vandesompele *et al.* 2002), NormFinder (Anderson *et al.*, 2004) and BestKeeper (Pfaffl *et al.*, 2004) were developed to facilitate the selection of stable genes. Using these methods reference genes for a variety of experimental conditions were reported in different plant species including *Arabidopsis thaliana* (Czechowski *et al.*, 2005), rice (Kim *et al.*, 2003; Jain *et al.*, 2006), poplar (Brunner *et al.*, 2004), potato (Nicot *et al.*, 2005), soybean (Jian *et al.*, 2008), coffee (Barsalobres-Cavallari *et al.*, 2009), chicory (Maroufi *et al.*, 2010) etc. In *C. plantagineum* and other resurrection plants there are no reports about validated stable genes to use for RT-qPCR studies.

7. Objectives of the study

Desiccation tolerance in *C. plantagineum* is acquired through the induction of protection mechanisms during dehydration. Several dehydration-induced genes were identified and characterised using both the *C. plantagineum* callus and the plant tissues as experimental systems. Expression studies in this plant have mainly been

performed by RNA blots. RT-qPCR is a very sensitive and accurate technique to analyse gene expression but its use in *C. plantagineum* is still hampered due to the lack of validated reference genes to normalise dehydration experiments of both plant tissues and callus. One objective of this thesis was to identify some stable genes in *C. plantagineum* leaves, roots and callus to allow the correct normalisation of RT-qPCR expression data in these tissues (CHAPTER 2).

The molecular characterisation of the desiccation tolerance trait in resurrection plants has been hampered by the lack of genome sequence information. Currently, RNA-sequencing allows transcriptomic studies in any organism without the requirement of their genome sequences. For this reason, RNA-seq represents the best method to identify dehydration-induced gene expression changes in resurrection plants on a genome-wide level. Recently, the transcriptome of *C. plantagineum* was sequenced revealing a large fraction of putative taxonomically restricted genes (TRGs). TRGs are known to participate in adaptation of organisms to particular environmental conditions. TRGs linked to desiccation tolerance, e.g., *CDT-1* were previously identified in *C. plantagineum* but underlying mechanisms remain unknown. It is hypothesised that TGRs may contribute to the desiccation tolerance trait in this plant. In order to validate this hypothesys, three putative TRGs were identified in the transcriptome data of *C. plantagineum* and were molecularly characterised (CHAPTER 3 and 4). These genes are predicted to take part in three different aspects of the desiccation tolerance phenomenon of this plant, i.e., leaf folding (*CpGRP1*; CHAPTER 3), chloroplast protection (*CpEDR1*; CHAPTER 4) and rehydration responses (*CpCRP1*; CHAPTER 4). Although these aspects are essential for the survival strategy of resurrection plants, the underlying molecular mechanisms are poorly understood. Here, the molecular characterisation of these genes contributes to advance our understanding of the molecular basis of these important processes.

CHAPTER 2

Identification of reference genes in the resurrection plant *Craterostigma plantagineum* for accurate gene expression analysis under different experimental conditions using RT-qPCR

Valentino Giarola

Institute of Molecular Physiology and Biotechnology of Plants (IMBIO), University of Bonn, Kirschallee 1, D-53115 Bonn, Germany

Abstract

Craterostigma plantagineum is a desiccation tolerant resurrection plant. The mechanisms that enable desiccation tolerance are not fully understood, however many of the genes that are induced during desiccation have been reported. RT-qPCR is considered the most sensitive method to study gene expression but inaccurate normalization can lead to incorrect conclusions. In our previous study we reported that RNA from leaves and roots of C*. plantagineum* are stable during desiccation and subsequent rehydration thus allowing us to use the RT-qPCR to test the stability of some candidate reference genes in leaves, roots and callus tissues during complete desiccation and rehydration. The reference genes which are commonly used for other plant species were investigated to identify genes displaying highly uniform expression patterns in leaves, roots and callus of *C. plantagineum* during desiccation. The expression stability of eight candidate reference genes (*ACT*, *CAC*, *EF1α*, *EIF5A*, *GAPDH*, *TKT3*, *PP2AA3*, *YLS8*) were tested on leaves and roots during desiccation and rehydration and on callus upon dehydration and ABA treatment. These eight genes were ranked according to their stability of gene expression using GeNorm[PLUS] and RefFinder and the most consistently expressed reference genes in each tissue is/were identified. *YLS8*, *EIF5A* and *TKT3* were identified as optimal reference genes for normalizing gene expression data in leaves whereas *EF1α*, *EIF5A* and *GAPDH* for roots and *YLS8*, *CAC* and *GAPDH* for callus.

Introduction

Response of plants to dehydration is a complex process mainly dependent on the modulation of transcriptional activity of stress-related genes. Gene expression analysis is important to understand the metabolic pathways and complex regulatory networks that underlie development and survival under stress conditions in plants. Several of the dehydration-induced genes encode proteins that play an important role in stress protection. For gene expression studies, quantitative polymerase chain reaction (RT-qPCR) is the most reliable and common technique with wide uses. For genes with conserved segments it is possible to design primers in the non-conserved regions and amplify a shorter fragment and thus check the gene expression. The prerequisite for this method is an appropriate internal reference gene to correctly normalize the expression data (Gutierrez *et al.*, 2008). Since variations exist between different samples in terms of RNA stability, quantity, purity, efficiency of reverse transcription and PCR, it is important to use a suitable reference gene and set of standardized experimental conditions to get accurate and reproducible results. Reference genes are key genes that are supposed to be constitutively expressed under different experimental conditions. Some examples of such genes include *18S rRNA* (18S ribosomal RNA), *ACTB* (β-actin), *TUBA* (β-tubulin), *TBP* (TATA sequence binding protein), *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase), *RPII* (RNA polymerase II), and *EF1α* (elongation factor 1α) (Kozera & Rapacz, 2013). The stability of reference genes can also significantly vary within different experimental conditions and tissue types (Schmittgen & Zakrajsek, 2000; Suzuki *et al.*, 2000; Lee *et al.*, 2002; Dheda *et al.*, 2004; Czechowski *et al.*, 2005). Thus it is mandatory to systematically check the stability of potential reference genes prior to their use (Guénin *et al.*, 2009). The use of a single reference gene for normalization may lead to relatively large errors, therefore, the use of multiple internal control genes is an essential approach for accurate normalization of gene expression data. Several statistical methods such as GeNorm (Vandesompele *et al.,* 2002), NormFinder (Andersen *et al.*, 2004) and BestKeeper (Pfaffl *et al.*, 2004), have been developed to facilitate the identification of the most stable gene(s) in a set of potential reference genes under different experimental conditions. These methods involve statistical analysis of the variation in each gene, by either single-factor analysis of variance (Brunner *et al.*, 2004) or by comparison of the mean variation in each gene relative to the mean variation of other

genes in and/or between data sets (Andersen *et al.*, 2004; Pfaffl *et al.*, 2004; Exposito-Rodriguez *et al.*, 2008). Although reference genes for many plant species are available, so far reference genes have not been defined for desiccation tolerant plants.

Evaluation of RNA quality is very much important for gene expression analysis using RT-qPCR studies. The RNA quality in different tissue samples can be evaluated by measuring its stability. RNA stability means the equal distribution of stable housekeeping genes despite different heterogeneous experimental conditions. Prolonged dehydration stress in plants may lead to RNA degradation and thus prevent the analysis of transcript level changes. In the current study, several of the house keeping genes were tested for their stability during desiccation in leaves, roots and callus. The identified genes now permit the use of RT-qPCR to study gene expression in *C. plantagineum* leaves and roots during desiccation and rehydration and in callus with and without abscisic acid (ABA) treatment.

Materials and Methods

Plant materials and treatments

Craterostigma plantagineum Hochst plants and callus were grown and propagated as described by Bartels *et al.* (1990). For the dehydration/desiccation and rehydration treatments, fully grown six- to eight-week-old pot-grown plants were subjected to gradual dehydration by withholding water. For early dehydration treatment (approximately 50 % relative water content [RWC]) plants were collected after 48 h of dehydration. The remaining plants were further dried over a period of 15 days until desiccation and re-watered by submerging them in water for 24 h to obtain rehydration samples. The plants reached RWCs of 1-2 % during desiccation and recovered fully upon rehydration. RWC measurements were done according to Phillips *et al.* (2008). The tissue samples were ground in liquid nitrogen and stored at -80 ℃. Dehydration treatments were also performed with the callus of *C. plantagineum.* For ABA treatment, callus was cultured on medium supplemented with 20 µM ABA for 6 days. Untreated or ABA-treated callus was placed on filter paper and desiccated in the air stream of a laminar flow cabinet.

RNA isolation and cDNA synthesis

Frozen plant tissue samples were ground in liquid nitrogen using a mortar and pestle. Total RNA was extracted according to Valenzuela-Avendaño *et al.* (2005). RNA concentration and purity were determined by measuring $OD_{260,280,230}$ using a Biospec-nano spectrophotometer (Shimadzu Biotech, Japan). After assessing the integrity of the RNA samples on a 2.0 % (w/v) agarose gel, four µg of total RNA were treated with DNase I (Thermo Fisher Scientific, St Leon-Rot, Germany) to remove any DNA. Two µg of DNase I-treated RNA were reverse transcribed using the RevertAid Reverse Transcriptase (Thermo Fisher Scientific, St Leon-Rot, Germany) following the manufacturer's instructions. The remaining two µg of RNA were used as negative control for the reverse transcription reaction. Each cDNA was diluted 15 times with diethylpyrocarbonate treated water and five µl were used as a template for a 20 µl PCR reaction.

Selection of candidate genes, design of primers and RT-qPCR conditions

To identify the most stably expressed reference genes prior to use in RT-qPCR studies, eight candidate reference genes were selected based on reports of known stable genes from diverse plant species (Czechowski *et al.*, 2005). The primers for amplifying reference genes and dehydrin genes were designed based on the sequence information from the *C. plantagineum* transcriptome database (Rodriguez *et al.*, 2010). Primer3 software (version 0.4.0) (Koressaar & Remm, 2007; Untergasser *et al.*, 2012) was used to find the best primer pair able to amplify a 70 to 149 bp fragment in the 3' untranslated region (3'UTR) region of the transcript sequence. The resulting target sequences were checked for secondary structures using the M-fold web server (http://mfold.rit.albany.edu/?q=mfold/DNA-Folding-Form) (Zuker, 2003). All primer pairs were further tested by RT-PCR and the amplicon size was confirmed by running the PCR product on a 2 % agarose gel. Target amplicons were sequenced to confirm the specificity of the PCR product. The primer specificities were confirmed with the melting-curve analysis after amplification during the subsequent RT-qPCR analysis. Melting curve analysis for each amplicon showed a single peak confirming the homogeneity and specificity of amplicons produced in RT-qPCR for all target genes. The standard curve was repeated at least four times using different dilutions of mixed cDNAs from all the samples differing with five orders of magnitude to calculate the gene-specific PCR amplification efficiency (E) and correlation coefficient (R^2) for each gene. The primer sequences and amplicon characteristics of the eight candidate genes are listed in Table 1.

RT-qPCR was performed in a 96 well plate using Mastercycler® ep realplex^2 (Eppendorf AG, Hamburg, Germany) and Maxima SYBR Green qPCR Master Mix (Thermo Fisher Scientific, St Leon-Rot, Germany). The PCR was performed in a final volume of 20 µl. Each reaction contained 5 µl of diluted cDNAs, 10 µl of Maxima SYBR Green qPCR mix (2x) and from 0.3 to 0.4 µM of each primer. The following conditions were used for the PCR: 95 ℃ for 10 min, followed by 40 cycles of 95 ℃ for 15 sec, 60 ℃ for 30 sec and 72 ℃ for 30 sec followed by 95 ℃ for 15 sec. The melting curves were analysed at 60 ℃ – 95 ℃ after 40 cycles in sequential steps by increasing the temperature in 0.5 ℃ steps for 20 min. Controls for reverse transcription were included in the PCR reactions to reveal DNA contaminations. Each PCR analysis was

performed with at least three biological replicates and three technical replicates each and the mean values were used for the analysis.

Data analysis

The quantification cycle (Cq) for each amplification curve was determined using the CalQplex™ algorithm embedded in the realplex 2.2 software (Eppendorf AG, Hamburg, Germany). These Cq values were imported in qbase^{PLUS} software V.1.3.5 (Biogazelle NV, Zwijnaarde, BE) and all the samples which passed the software quality control checks were used to calculate the expression stability of reference genes under different stress conditions with the GeNorm^{PLUS} algorithm included in the qbase^{PLUS} software. GeNorm indicates a stability measure (M), which is the average pair wise variation of a gene compared to the other control genes included in the same analysis and via a stepwise exclusion of the least stable gene. This generates a stability ranking among the reference gene candidates and estimates the number of genes required to calculate a robust normalization factor. GeNorm suggests M = 1.5 as a cut-off value. The gene with the lowest M value is considered to be the most stable one. Apart from GeNorm^{PLUS}, RefFinder, (http://www.leonxie.com/referencegene.php) was used as additional tool to get a ranking of reference genes according to GeNorm (Vandesompele *et al.*, 2002), NormFinder (Andersen *et al.*, 2004), BestKeeper (Pfaffl *et al.*, 2004) and the comparative ΔCt method (Silver *et al.*, 2006). Based on the rankings from each program, RefFinder assigns an appropriate weight to each gene and after calculating the geometric mean of these weights provides a final comprehensive ranking. All other statistical analyses were performed using SIGMA (v12, SPSS Inc., Chicago, IL) and SigmaStat 3.1 software (San Jose, CA, USA).

Results

Primer optimization and RT-qPCR

For the reliable quantification of dehydrin gene expression, RT-qPCR is the preferred method which requires the use of a set of stably expressed reference genes to normalize expression data. The *C. plantagineum* transcriptome data was searched for homologs of known stably expressed genes to identify suitable reference genes. The stability of eight genes i.e., actin (*ACT*), clathrin adaptor complex (*CAC*), eukaryotic elongation factor 1α (*EF1α*), eukaryotic translation initiation factor 5A (*EIF5A*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), transketolase 3 (*TKT3*), protein phosphatase 2A subunit A3 (*PP2AA3*), and yellow leaf specific gene 8 (*YLS8)* which were known to be involved in different aspects of cellular functions were tested. Some of these genes are already known and used as internal control genes for other plant species while some have been shown to be constitutively expressed during different stress conditions by either RNA blots or by RT-PCR in studies with *C. plantagineum* (Bernacchia *et al.*, 1995; Rodriguez *et al.*, 2010; Dinakar & Bartels, 2012). Primers were designed using the sequences obtained from the *C. plantagineum* transcriptome (Rodriguez *et al.*, 2010). *In silico* analysis of target gene secondary structures and primer/template accessibility were assessed using the M-fold web server (http://mfold.rit.albany.edu/?q=mfold/DNA-Folding-Form) using corrections for ionic conditions of 50 mM Na^+ and 25 mM Mg^{+2} (Fig. 1). Standard curves using dilutions of the cDNA spanning five orders of magnitude were made to calculate gene specific PCR efficiencies. The correlation coefficient (R^2) of the slope of the standard curve was used to calculate gene specific amplification efficiency (E). The specificity of each primer pair was also assessed by amplification followed by dissociation curve analysis. Different primer pair combinations were tested for the same target gene and those that showed the best amplification efficiency and specificity were used for expression analyses. The optimized primer pairs for each gene produced a single amplicon as demonstrated by dissociation curve analysis (Fig. 2a) and agarose gel electrophoresis (Fig. 2b). The primer sequences and amplicon characteristics such as Tm, length, amplification efficiency and correlation coefficient of the eight reference genes are shown in Table 1.
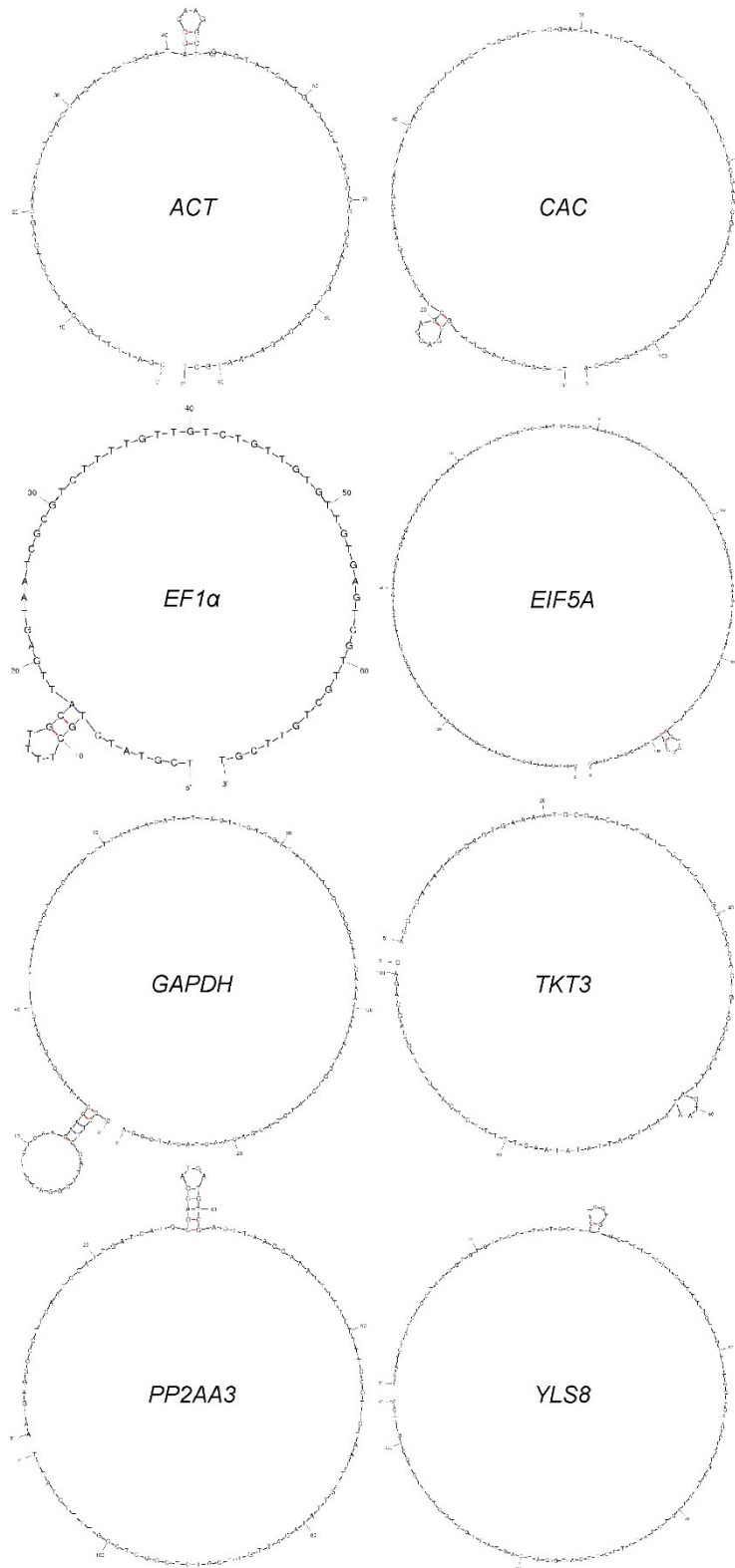
Fig. 1 The M-fold predicted secondary structures of all the amplicons which were tested for their stability during desiccation in leaves, roots and callus tissues of *C. plantagineum*. Actin (*ACT*), clathrin adaptor complex (*CAC*), eukaryotic elongation factor 1α (*EF1α*), eukaryotic translation initiation factor 5A (*EIF5A*), glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*), transketolase 3 (*TKT3*) protein phosphatase 2A subunit A3 (*PP2AA3*) and yellow leaf specific gene 8 (*YLS8*) were used in the current study.

(a)
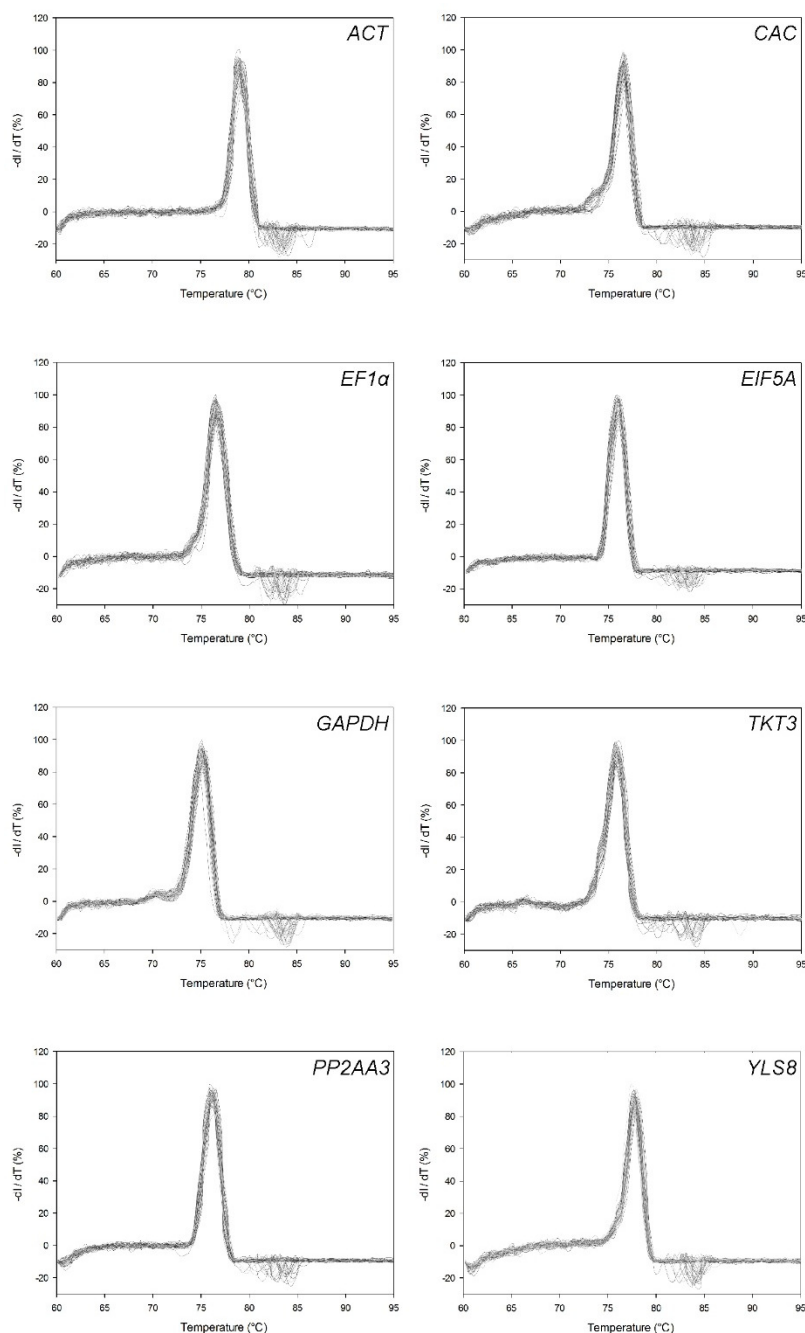
Fig. 2 Specificity of the primer pairs used for RT-qPCR amplifications. (a) The dissociation curves of eight potential reference genes derived from at least two biological repeats with 3 technical repeats showing the specificity of the RT-qPCR by amplifying a single amplicon. (b) Agarose gel electrophoresis of the RT-qPCR products showing the single amplicon with particular molecular weight.

(b)
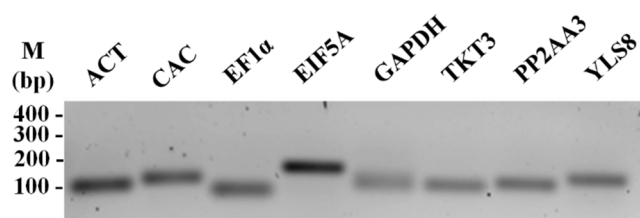
Table 1 The selected reference genes, primers and amplicon characteristics.

| Gene	Description	Primer sequences (forward/reverse)	Amplicon length (bp)	Amplicon Tm (°C)	Amplific. Efficiency (%)	R^2
ACT	Actin family protein	CGATTTTGGCATCTCTCAGC AGCATTTTCTGTGAACAATCGAC	93	79.1	97	0.998
CAC	Clathrin adaptor complexes medium subunit	TTGAGGAAGTTTGCGACGAG TGGCCTTGTAATAGAATGGTACG	107	76.5	104	0.995
EF1α	Elongation factor 1α	TCGTATCTGCTTTTGCATTGAGT ACGAACAGCAACGACTCACA	70	76.5	100	0.998
EIF5A	Eukaryotic initiation factor 5A	CATCAATCGTGAAGGAGAGGA CTAAACCTTGGCAGCAGCAC	149	75.9	96	0.998
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	GGCTTCGATTCGGATCTTCT AGCCCACAAATATATCAAGAACTCA	94	75.0	97	0.999
TKT3	Transketolase 3	AGGTGAAAAAGGAGTGAAAATGG GTCTGCTGCAAACATGAGAAAA	101	75.8	98	0.999
PP2AA3	Protein phosphatase 2A subunit A3	AATGAGGCCCTTCAATCCAT AAATAGAAAACCGAGCCGAGA	111	76.1	106	0.995
YLS8	Yellow-Leaf-Specific gene 8	TCATCCGTCTCCTCAGGTGT TCAACAGCACAAAAGCATCAA	125	77.8	104	0.999

Expression profiles and stability of reference genes

Expression of reference genes may vary in different tissues under different experimental conditions. To evaluate the stability of the reference genes in tissue/treatment combinations, mRNA levels of candidate reference genes were analysed in leaves and roots under untreated, partially dehydrated, desiccated and rehydrated conditions and in callus upon ABA treatment with and without desiccation. The quantification cycle (Cq) representing the cycle at which a significant increase in the PCR product is observed differed between the selected reference genes. The Cq values ranged between 18.6 to 32.9 in different genes expressed in leaves, roots and callus tissues. Amplicons with high Cq value represent low abundant genes and those with low Cq value represent abundant genes. In all samples tested (leaves, roots and callus) the *PP2AA3* gene was expressed at a low level with a high Cq value (from 27.3 to 32.9) while the *TKT3* gene was abundantly expressed in both leaves and roots (from 19.2 to 23.0) and the *EF1α* gene was abundantly expressed in callus samples (from 18.6 to 20.2) (Fig. 3).

The average expression levels of all the eight genes varied widely with different Cq values in different samples (Fig. 3).
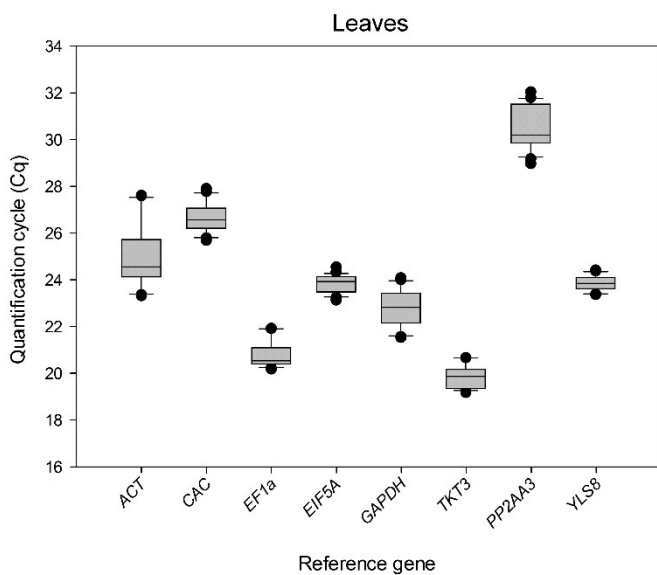
Fig. 3 Cq values of the potential reference genes from leaves, roots and callus under different stress conditions. Genes with low Cq values are abundantly expressed whereas genes with high Cq values are low expressed. While the box represents the 50% of the distribution, the upper and lower whiskers represent the entire spread of the data. The central horizontal line represents the median. Outlier which is an observation that is numerically distant from the rest of the data is represented by a dark circle.
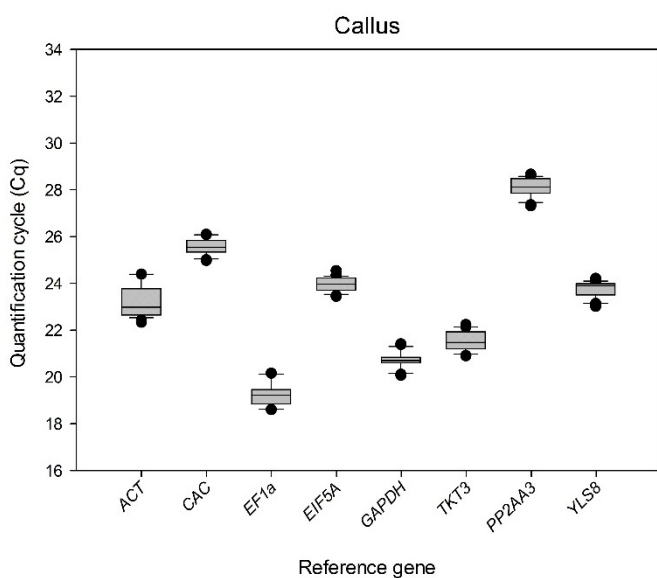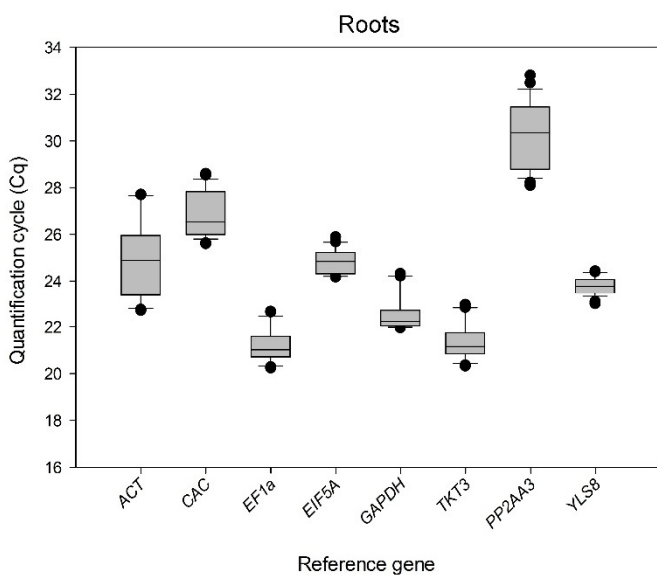
To select the best reference genes for accurate gene expression profiling, different statistical methods were used to rank the genes according to their stabilities. GeNorm[PLUS] tool available from qbase[PLUS] software and web-based tool RefFinder (http://www.leonxie.com/referencegene.php) were used to analyse the transcript expression stabilities under different stress conditions. Using the GeNorm[PLUS] software the eight genes were ranked according to their stability of transcript expression (M) (Fig. 4a).
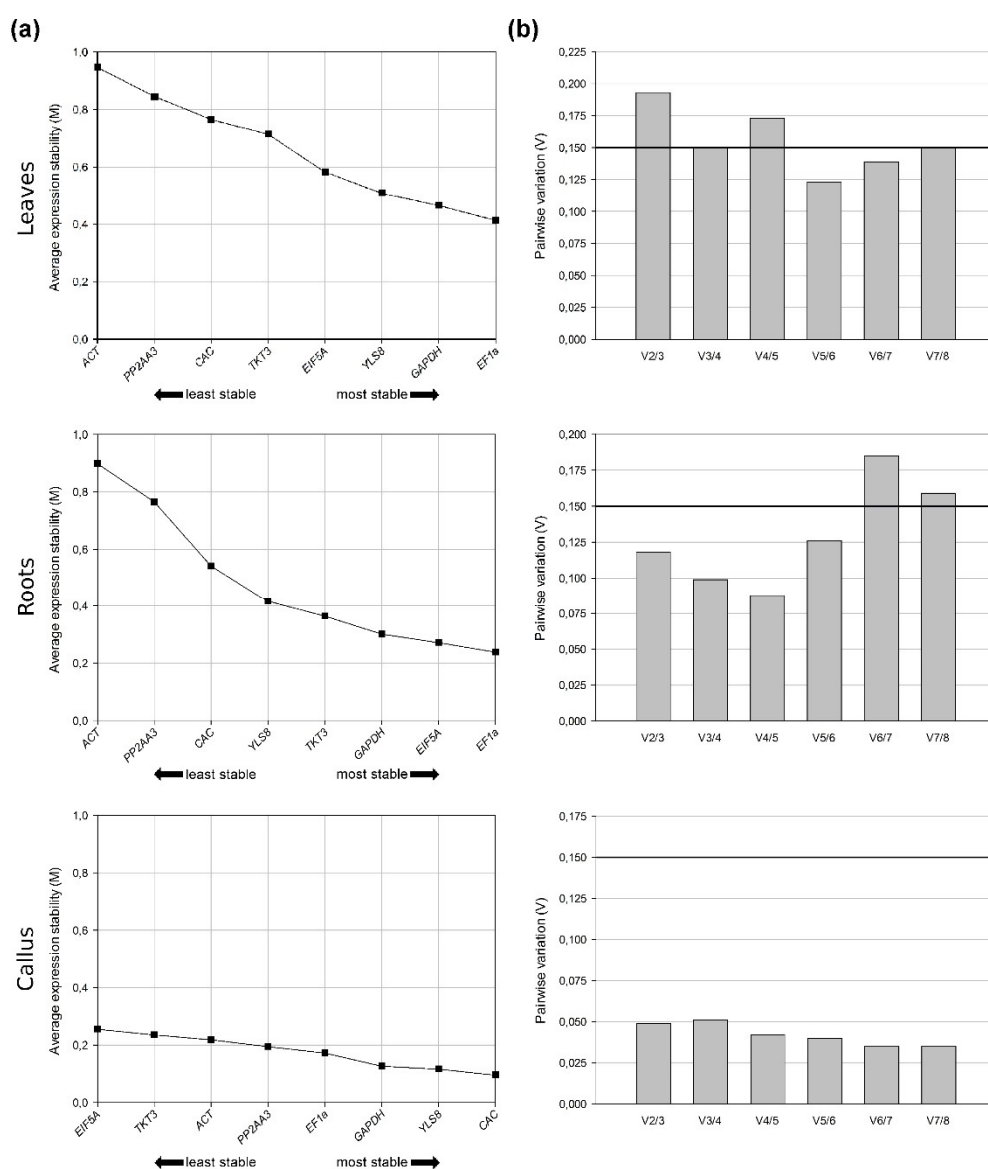
Fig. 4 GeNorm[PLUS] analysis. (a) The expression stability values (M) and the ranking of the potential reference genes as calculated by GeNorm from cDNAs of leaves, roots and callus tissues of *C. plantagineum.* (b) Pairwise variation values (V) calculated from the "M" values of different reference genes which will be useful for selecting optimal number of reference genes for accurate normalisation.

The GeNorm algorithm first calculates an expression stability value (M) for each gene and then the pairwise variation (V) with the other genes compared to the gene of interest. This algorithm is based on the average pairwise variation between all the genes tested with the lowest M values representing greater stability (Table 2). The ranking of the eight reference genes according to their M values is shown in Fig. 4a. While the genes *EF1α*, *GAPDH* and *YLS8* represent the most stable reference genes in leaf tissues of *C. plantagineum*, *ACT*, *PP2AA3* and *CAC* are the least stable ones (Fig. 4a). Similarly *GAPDH, EIF5A and EF1α* represent the most stable reference genes in root samples and GAPDH, YLS8 and CAC in callus tissues. Since normalization with a single reference gene can produce significant errors (Vandesompele *et al.*, 2002), we identified the minimum number of genes required to minimize variation in normalization using GeNorm algorithm in leaf, root and callus tissues by performing pairwise variation analysis (Fig. 4b).

Table 2 Ranking of candidate reference genes in order of their expression stability as calculated by GeNorm[PLUS]. Higher "M" values represent lower stability of gene expression whereas lower "M" value represents higher gene expression stability. The three most stable genes in leaves, roots and callus are indicated in bold. Please refer to Table 1 for full names of the abbreviated genes.

Leaves		Roots		Callus	
Gene	M	Gene	M	Gene	M
ACT	0.946	*ACT*	0.898	*EIF5A*	0.255
PP2AA3	0.845	*PP2AA3*	0.765	*TKT3*	0.235
CAC	0.765	*CAC*	0.541	*ACT*	0.218
TKT3	0.713	*YLS8*	0.416	*PP2AA3*	0.194
EIF5A	0.582	*TKT3*	0.364	*EF1α*	0.172
YLS8	0.508	**GAPDH**	0.301	**GAPDH**	0.126
GAPDH	0.465	**EIF5A**	0.271	**YLS8**	0.116
EF1α	0.413	**EF1α**	0.238	**CAC**	0.095

The RefFinder tool integrates the major computational programs such as GeNorm, NormFinder, BestKeeper and the comparative delta CT method to compare and rank the tested reference genes. From the ranking of the genes from each program the appropriate weight is assigned to each individual gene. The geometric mean of their weights is calculated for the comprehensive ranking indicating the most stable and least stable genes. For analysing the data by the RefFinder program, the Cq values were copied from the excel file to the programme file directly and the ranking of the genes by four programs and comprehensive ranking of the genes can be calculated. The results obtained by RefFinder analysis are shown in Table 3.

Table 3 Relative stability ranking of reference genes in leaves, roots and callus of *C. plantagineum*. Delta CT, BestKeeper, NormFinder and GeNorm rankings were calculated using RefFinder (http://www.leonxie.com/referencegene.php). Please refer to Table 1 for full names of the abbreviated gene names.

Leaves

Method	Ranking order (Least stable → Most stable)							
	8	7	6	5	4	3	2	1
Delta CT	*ACT*	*GAPDH*	*PP2AA3*	*EF1α*	*CAC*	*YLS8*	*EIF5A*	*TKT3*
BestKeeper	*ACT*	*PP2AA3*	*GAPDH*	*CAC*	*TKT3*	*EF1α*	*EIF5A*	*YLS8*
NormFinder	*ACT*	*GAPDH*	*PP2AA3*	*EF1α*	*YLS8*	*CAC*	*EIF5A*	*TKT3*
GeNorm	*ACT*	*PP2AA3*	*TKT3*	*CAC*	*EIF5A*	*YLS8*	*EF1α	GAPDH*	
Comp. ranking	***ACT***	***PP2AA3***	***CAC***	***GAPDH***	***EF1α***	***YLS8***	***EIF5A***	***TKT3***

Roots

Method	Ranking order (Least stable → Most stable)							
	8	7	6	5	4	3	2	1
Delta CT	*ACT*	*PP2AA3*	*YLS8*	*TKT3*	*CAC*	*GAPDH*	*EIF5A*	*EF1α*
BestKeeper	*ACT*	*PP2AA3*	*CAC*	*TKT3*	*GAPDH*	*EF1α*	*EIF5A*	*YLS8*
NormFinder	*ACT*	*PP2AA3*	*YLS8*	*TKT3*	*GAPDH*	*EIF5A*	*CAC*	*EF1α*
GeNorm	*ACT*	*PP2AA3*	*CAC*	*YLS8*	*TKT3*	*GAPDH*	*EF1α	EIF5A*	
Comp. ranking	***ACT***	***PP2AA3***	***TKT3***	***CAC***	***YLS8***	***GAPDH***	***EIF5A***	***EF1α***

Callus

Method	Ranking order (Least stable → Most stable)							
	8	7	6	5	4	3	2	1
Delta CT	*EIF5A*	*ACT*	*TKT3*	*PP2AA3*	*EF1α*	*GAPDH*	*CAC*	*YLS8*
BestKeeper	*ACT*	*EF1α*	*PP2AA3*	*YLS8*	*TKT3*	*CAC*	*GAPDH*	*EIF5A*
NormFinder	*EIF5A*	*ACT*	*TKT3*	*PP2AA3*	*EF1α*	*GAPDH*	*CAC*	*YLS8*
GeNorm	*EIF5A*	*ACT*	*TKT3*	*PP2AA3*	*EF1α*	*GAPDH*	*CAC	YLS8*	
Comp. ranking	***ACT***	***TKT3***	***PP2AA3***	***EIF5A***	***EF1α***	***GAPDH***	***CAC***	***YLS8***

The comprehensive stability ranking of the reference genes estimated by RefFinder in leaves, roots and callus tissues is shown in Fig. 5. From the analysis, *TKT3*, *EIF5A* and *YLS8* were identified as the three most stable reference genes for normalizing gene expression data in leaves and *EF1α*, *EIF5A* and *GAPDH* were found to be the three optimal reference genes for roots. *YLS8*, *CAC* and *GAPDH* are the most stable reference genes for callus tissue (Fig. 5). From the results of the comprehensive ranking in leaves *ACT* and *PP2AA3* were found to be the least stable genes, whereas all other genes are stable and can be used for normalizing gene expression data (Fig. 5).
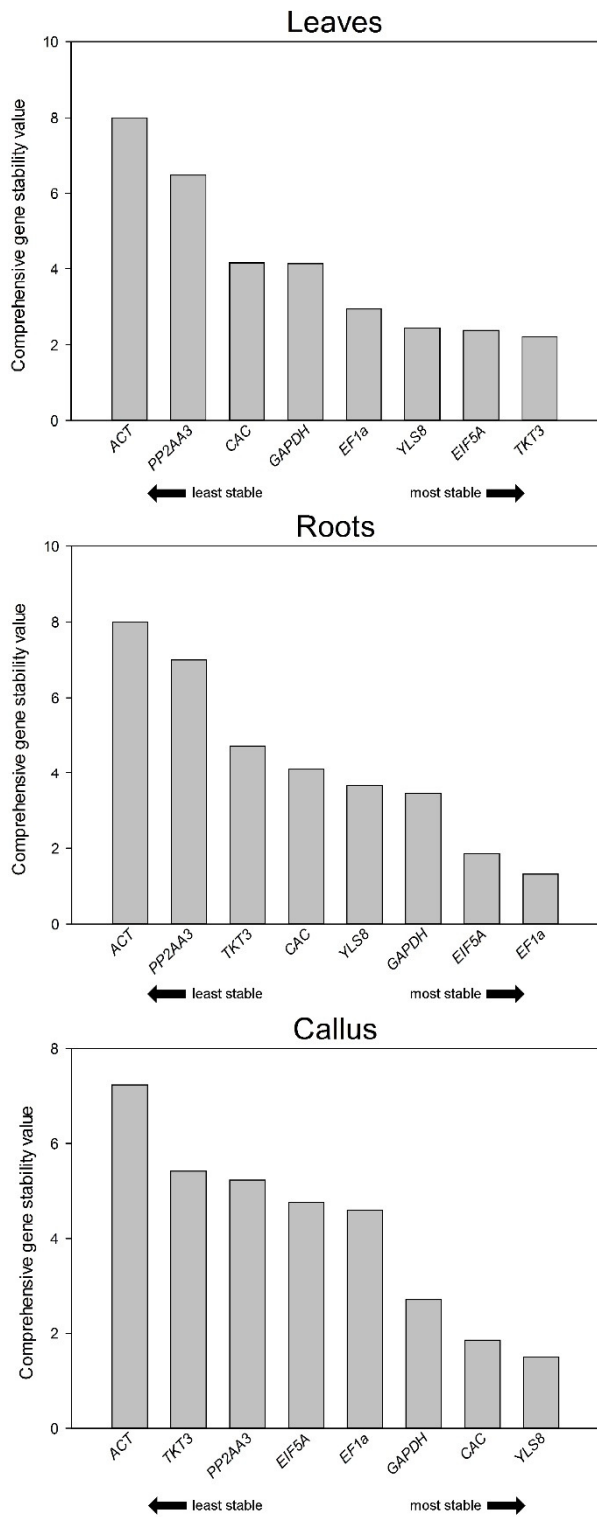
Fig. 5 The comprehensive gene stability values obtained by RefFinder analysis showing the most stable and the least stable genes in leaves, roots and callus tissues of *C. plantagineum*.

Discussion

Previously, transcript analysis of constitutively expressed genes suggested that mRNAs are stable in *C. plantagineum* during desiccation and rehydration (Dinakar & Bartels, 2012). However, since the presence of degraded RNAs may influence gene expression data, the stability of the reference genes was checked by RT-qPCR. This analysis is also necessary for identifying the best reference genes for gene expression studies in *C. plantagineum*. Constitutive expression of the stable reference genes without major differences in their Cq values at different stress conditions confirms the stability of RNA in desiccated tissues of *C. plantagineum* (Fig. 3). How *C. plantagineum* protects the RNA during desiccation and rehydration is still an open question. Although, microarray data from *Arabidopsis* resulted in the identification of many potential reference genes that show exceptional expression stability throughout development and under wide range of environmental conditions (Czechowski *et al.*, 2005), this analysis cannot be directly extrapolated to *C. plantagineum*. Recent studies have indicated that the traditionally used reference genes are not always stably expressed when tested in other plant species at different experimental conditions (Artico *et al.*, 2010; Chen *et al.*, 2011). Thus it is mandatory to do the reference gene analysis under different experimental conditions in different tissue types. Reference genes for *C. plantagineum* were selected based on their sequence similarity with genes already reported from *Arabidopsis thaliana*, *Vitis vinifera* or *Populus trichocarpa*. *TKT3* is selected based on the constitutive expression pattern shown previously by RNA blot analysis (Bernacchia *et al.*, 1995). The expression analysis of the eight candidate reference genes using geNorm[PLUS] and RefFinder resulted in the identification of the best suitable reference genes. Based on the stability of gene expression, these reference genes were ranked using GeNorm[PLUS] and RefFinder. The most stable optimal reference genes were identified for leaves, roots and callus. The ranking of the reference genes by RefFinder was similar to that generated by GeNorm[PLUS] for roots and callus, however differences were observed for leaves. Since GeNorm[PLUS] and RefFinder are two different mathematical approaches to calculate stability, it is not surprising to observe differences in the rankings (Paolacci *et al.*, 2009; Zhu *et al.*, 2012). However, all the genes except actin and protein phosphatase 2A subunit A3 were found to be stable and can potentially be used for normalizing the gene

expression data in leaves. In all the tissues, actin was found to be the least stable gene under different stress conditions.

Several isoforms of dehydration-induced genes were found in the *C. plantagineum* transcriptome (Rodriguez *et al.*, 2010). These isoforms may play different roles during the acquisition of desiccation tolerance in *C. plantagineum* but the large sequence similarity of these genes limits the use of hybridisation-based techniques to study the expression of the single isoforms. The reference genes identified in this study now permit the use of RT-qPCR for the accurate quantification of transcript levels of these isoforms and other desiccation-related genes in *C. plantagineum* leaves, roots and callus*.*

CHAPTER 3

The *Craterostigma plantagineum* glycine-rich protein CpGRP1 interacts with a cell wall associated protein kinase 1 (CpWAK1) and accumulates in leaf cell walls during dehydration

Valentino Giarola[1], Stephanie Krey[2], Barbara von den Driesch[1] and Dorothea Bartels[1]

[1] Institute of Molecular Physiology and Biotechnology of Plants (IMBIO), University of Bonn, Kirschallee 1, D-53115 Bonn, Germany

[2] Present address: Botanical Institute, Cologne Biocenter, University of Cologne, Zülpicherstraße 47b, D-50674, Cologne, Germany

Abstract

Craterostigma plantagineum tolerates extreme desiccation. Leaves of this plant shrink and extensively fold during dehydration and expand again during rehydration preserving their structural integrity but the underlying mechanisms remain largely unknown. Among transcripts abundantly expressed in desiccated leaves we identified a gene putatively coding for an apoplastic glycine-rich protein (CpGRP1) and we investigated the role of this protein in the cell wall during dehydration. The CpGRP1 protein accumulates in the apoplast of desiccated leaves. Analysis of the promoter revealed that the gene expression is mainly regulated on the transcriptional level, is independent of ABA and involves a drought responsive *cis*-element (DRE). We demonstrated that CpGRP1 interacts with a cell wall associated protein kinase (WAK1) from *C. plantagineum* which is down-regulated in response to dehydration. WAKs are known to be required for cell expansion. Our data suggest a role of the CpGRP1-CpWAK1 complex in the dehydration-induced morphological changes of the cell wall during dehydration in *C. plantagineum*. Cell wall pectins and dehydration-induced pectin modifications are predicted to be involved in the activity of the CpGRP1-CpWAK1 complex.

Introduction

Desiccation tolerance is the ability of an organism to equilibrate its internal water content to the air-dry state and suspend all metabolic activities until water becomes available again. Although this trait is very common in lower plants, in higher plants it is restricted to seeds and to the vegetative tissue of a small group of species termed resurrection plants (Gaff, 1971). Among resurrection plants, *Craterostigma plantagineum* has been extensively studied to understand the molecular basis of desiccation tolerance (Bartels & Salamini, 2001). It is thought that *C. plantagineum* and other desiccation tolerant angiosperms mainly acquire their tolerance through the induction of protection mechanisms during dehydration (Gaff, 1989; Oliver *et al.*, 1998). Cell walls of *C. plantagineum* and other resurrection plants extensively shrink and fold during dehydration (Gaff, 1989; Farrant, 2000; Farrant *et al.*, 2003; Willigen *et al.*, 2003; Farrant *et al.*, 2007). These extreme morphological changes appear to be essential for resurrection plants to cope with the mechanical stress due to the shrinkage of the cytoplasm (Iljin, 1957; Farrant, 2000). It has been proposed for *Craterostigma* ssp. that dehydration-induced modifications of existing cell wall components might play a major role to increase cell wall flexibility during dehydration (Moore *et al.*, 2013). Expansin activity, xyloglucan remodelling and calcium ion accumulation have been proposed to take part in the cell wall folding process (Vicré *et al.*, 1999; Jones & McQueen-Mason, 2004; Vicré *et al.*, 2004) but the underlying molecular mechanisms remain largely unknown. The study of genes important for desiccation tolerance in *C. plantagineum* is hampered by factors such as the poor ability to transform this plant and the lack of the genome sequence. Expression analysis by deep RNA sequencing (RNA-seq) permits to quantify changes in the whole transcriptome of an organism under different physiological conditions and thus can be used to identify candidate genes involved in the acquisition of desiccation tolerance in *C. plantagineum*. Recently, an RNA-seq analysis of the transcriptomes from dehydrated/rehydrated *C. plantagineum* leaves revealed the presence of uncharacterised genes potentially involved in the dehydration process (Rodriguez *et al.*, 2010).

Here we report the characterisation of one of these genes which is coding for an apoplastic glycine-rich protein (CpGRP1). The plant glycine-rich protein (GRP) superfamily is a large group of proteins sharing a high glycine-content region arranged as $(Gly)_n$-X repeats. Other distinct protein domains are often found in GRPs, e.g., the

cysteine-rich region, the oleosin domain, the RNA-recognition motif, the cold-shock domains and the zinc-finger motif. GRPs are classified in five groups according to the arrangements of glycine-rich repeats and the presence of additional protein motifs (Sachetto-Martins *et al.*, 2000; Fusaro *et al.*, 2001; Nora Bocca *et al.*, 2005; Mangeon *et al.*, 2010). Class I GRPs have a high glycine-content region with $(GGX)_n$ repeats. Class II GRPs contain a characteristic C-terminal cysteine-rich region. Lower glycine content and the possible presence of an oleosin domain are characteristics of the class III GRPs. Class IV GRPs or RNA-binding GRPs have either an RNA-recognition motif or a cold-shock domain with additional zinc-finger motifs in some members. The presence of mixed patterns of glycine repeats distinguish the class V from the class I GRPs. GRPs may contain a target peptide which directs them to specific subcellular compartments. The variable structure of these proteins together with diverse expression patterns and different subcellular localisations indicate that GRPs have many different roles *in planta* (Mangeon *et al.*, 2010).

GRPs are known to be modulated by biotic and abiotic stresses including dehydration but only a few have been studied and functionally characterized (Sachetto-Martins *et al.*, 2000; Mousavi & Hotta, 2005; Mangeon *et al.*, 2010). Many GRPs contain a signal peptide which might indicate an apoplastic localisation and thus a role for these proteins in the apoplasm (Sachetto-Martins *et al.*, 2000). Since the isolation of the first apoplastic GRP from petunia, apoplastic GRPs have been proposed to be a main component of the cell wall structure (Condit & Meagher, 1986). Recently, a cell wall dehydration-responsive class V GRP from the resurrection plant *Boea hygrometrica* (BhGRP1) was proposed to be required for cell wall integrity during dehydration (Wang, L *et al.*, 2009). Different functions have been demonstrated for class II GRPs. In *Nicotiana tabacum* a class II GRP protein which is specifically induced by low cadmium concentrations was shown to inhibit turnip vein-clearing tobamovirus movement by increasing callose deposition in phloem cell walls (Ueki & Citovsky, 2002). The cysteine-rich region of a cell wall GRP from *A. thaliana* (AtGRP-3) was found to be required for interaction with a cell wall-associated protein kinase 1 (AtWAK1) and AtGRP-3 was proposed to be involved in pathogenesis-related responses (Park *et al.*, 2001).

Cell wall-associated protein kinases are transmembrane spanning proteins with a cytoplasmic serine/threonine kinase domain and a variable extracellular region tightly bound to pectins (He *et al.*, 1999; Anderson *et al.*, 2001). Down-regulation of WAK

protein levels *via* antisense constructs led to reduction of cell elongation suggesting the involvement of *WAK* genes in the control of cell volume (Anderson *et al.*, 2001; Lally *et al.*, 2001; Wagner & Kohorn, 2001). It was demonstrated that pectins are able to activate signalling pathways downstream of WAKs (Kohorn *et al.*, 2006; Kohorn *et al.*, 2009; Brutus *et al.*, 2010). Currently, WAKs are thought to sense the state of cell wall pectin and mediate different signalling pathways accordingly (Kohorn & Kohorn, 2012).

Here the expression and subcellular localization of CpGRP1 were investigated and the analysis proposes a role for CpGRP1 in the apoplasm of dehydrating *C. plantagineum* leaves. To understand gene regulation promoter deletions were analysed and a region containing a putative drought responsive element (DRE) was found to be mainly responsible for gene expression during dehydration. Because of the high structural similarity between CpGRP1 and AtGRP-3 an interaction between CpGRP1 and a *C. plantagineum* wall associated protein kinase 1 (CpWAK1) was demonstrated. These data suggest an involvement of the CpGRP1/CpWAK1 complex in cell wall folding and unfolding mechanisms during dehydration and rehydration in *C. plantagineum* leaf and callus cells.

Materials and Methods

Plant material and treatments

Craterostigma plantagineum Hochst. plants and callus were grown according to Bartels *et al.* (1990). Dehydration and rehydration treatments were conducted according to Rodriguez *et al.* (2010). Callus was grown on medium supplemented with 20 µM ABA for 6 days for ABA treatment. Detached leaves were incubated in 100 µM ABA or 0.5 M or 0.8 M mannitol or sorbitol solutions for the indicated time. Plant material was ground in liquid nitrogen and stored at -80 °C. Relative water content (RWC) measurements were done according to Bernacchia *et al.* (1996).

Molecular techniques and sequence analysis

Standard molecular techniques were performed as described by Sambrook *et al.* (1989). Sequencing of DNA and primer synthesis were carried out by Eurofins MWG Operon (http://www.eurofinsgenomics.eu). Protein localization was predicted using WoLF PSORT (http://wolfpsort.org) (Horton *et al.*, 2007). All primers used are listed in Table S1.

Amplification of 5' and 3' cDNA fragments

The 5' region of the *CpGRP1* gene was obtained by PCR from a λ ZAP II cDNA library prepared from 2 h dried *C. plantagineum* leaves (Bockel *et al.*, 1998) using nested gene specific (CpGRP1_R1/2; Table S1) and vector specific (ZAP_F1/2; Table S1) primers. The following primer combinations based on two *C. plantagineum* cDNA sequences were used to obtain the full-length *WAK* gene sequence by sequential PCR reactions: CpWAK5p_R1/2 and ZAP_F1/2 to obtain the 5' end; CpWAK3p_F1/2 and ZAP_R1/2 to obtain the 3' end; CpWAK3p_F1 and CpWAK5p_R1 to obtain the middle part (for all primers see Table S1). The 5' and 3' ends of a second CpWAK isoform were obtained using CpWAK5p2_R and ZAP_F1 or CpWAK3p2_F and CpWAK_R primer combinations, respectively (Table S1). CpWAK1_F and CpWAK2_F primers were used in combination with the CpWAK_R primer (Table S1) to isolate the full length coding sequences of *CpWAK1* and *CpWAK2*. PCR products were cloned into pJET

44

1.2 vectors using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific, St Leon-Rot, Germany) and sequence identities were verified by DNA sequencing.

Transcript analysis

Total RNA from plant samples was isolated according to Valenzuela-Avendaño *et al.* (2005). RNA concentrations were determined using a Biospec-nano spectrophotometer (Shimadzu Biotech, Japan). One µg of each RNA sample was loaded on a 2 % agarose gel to monitor RNA quality. Preparation of cDNA for RT-PCR or RT-qPCR analysis was performed as described in Giarola *et al.* (unpublished). RNA blot analysis was performed as described by Bartels *et al.* (1990). The template sequence used to generate radioactive probes was obtained by PCR using coding sequence specific primers (CpGRP1_CDS_F/R; Table S1). Hybridization probes were obtained with ^{32}P-labelled dCTP using the DecaLabel DNA labelling kit (Thermo Fisher Scientific).

CpGRP1 antibody generation and protein analysis

A cDNA fragment coding for amino acids 23-156 was amplified by PCR from *CpGRP1* cDNA using a forward primer to add an *Nco*I site and a reverse primer to introduce an *Xho*I site at the 5' and 3' end respectively (CpGRP1_CDS_F/R; Table S1). The *Nco*I/*Xho*I fragment was cloned into the expression vector pET28a(+) (Novagen, Darmstadt, Germany) to obtain the C-terminal 6His-tag translational fusion protein CpGRP1His. The resulting pET28 CpGRP1His plasmid was introduced into BL21 (DE3) *E. coli* cells (Amersham Pharmacia Biotech; Piscataway, NJ, USA) and protein overexpression was induced by the addition of 1 mM isopropyl-1-thio-β-d-galactopyranoside (IPTG). The recombinant protein was purified from bacteria 5 hours after IPTG induction according to Kirch & Röhrig (2010), freeze-dried and sent to Seqlab (Sequence Laboratories Göttingen GmbH, Göttingen, Germany; www.seqlab.de) to raise a polyclonal antiserum in rabbit. The specificity of the antibody was confirmed with both recombinant proteins and plant leaf extracts. Total proteins were extracted from *C. plantagineum* samples according to Wang *et al.* (2003). Ten µg of total proteins were separated by 15 % (w/v) SDS-polyacrylamide gel electrophoresis (PAGE) (Laemmli, 1970). Twenty µg proteins were separated by

isoelectric focusing using 7 cm IPG strips pH 3–10 (GE Healthcare, Freiburg, Germany) for two-dimensional SDS-PAGE (Röhrig *et al.*, 2006). Separated proteins were transferred to nitrocellulose membranes (Towbin *et al.*, 1979). A 1:5000 dilution of the polyclonal antiserum was used to detect CpGRP1 protein.

Promoter isolation and analysis

Genomic DNA was extracted from leaves (Murray & Thompson, 1980) and prepared for the PCR-based gene walking procedure according to the GenomeWalker™ Universal kit (Clontech, Heidelberg, Germany) protocol. An amplification product of about 1.2 kb was obtained using a primer specific for the intron present in the *CpGRP1* genome sequence (CpGRP1_intron_R; Table S1) and the library adaptor primer (AP1; Table S1) and cloned into the pJET 1.2 vector. The vector insert was sequenced and putative promoter *cis*-acting elements were identified using the plant *cis*-acting regulatory DNA elements (PLACE) database scan tool (http://www.dna.affrc.go.jp/PLACE/signalscan.html) (Higo *et al.*, 1999). To generate the full promoter::GUS fusion construct (CpGRP1pFULL::GUS), the *CpGRP1* upstream region (promoter + 5' UTR) was amplified using the library adaptor primer (AP1; Table S1) and a primer introducing an *Nco*I site at the first ATG of the coding sequence (CpGRP1p_R; Table S1). The amplified fragment was inserted into the pBT10GUS vector (Sprenger-Haussels & Weisshaar, 2000) using *Sal*I and *Nco*I restriction sites. Promoter deletion fragments were generated by PCR using the same reverse primer but different forward primers to generate an *Xba*I site at different positions along the promoter (CpGRP1p605_F, CpGRP1p407_F, CpGRP1p280_F, CpGRP1p185_F; Table S1). Promoter::GUS deletion constructs were obtained by cloning promoter fragments into pBT10GUS using *Xba*I and *Nco*I restriction sites. Promoter analysis was performed using *C. plantagineum* leaves according to van den Dries *et al.* (2011) with the following modifications. All promoter::GUS constructs were co-bombarded with a CaMV35S::LUC vector (pTSLUC obtained substituting the *GUS* coding sequence in PBT10GUS with the luciferase coding sequence). The CaMV35S::LUC construct was used as efficiency normalizator of each bombardment. Relative promoter activities were calculated using the method of Schenk *et al.* (1998). Luciferin solution (1 mM luciferin in 100 mM phosphate buffer pH 8, 0.01 % (v/v) Triton X-100) was sprayed on leaves 16 h after the bombardment and luminescence emitted

from transformed cells was measured using a luminescent image analyzer LAS-1000plus (Fujifilm Co., Tokyo, JP). The number of luciferase (LUC) expressing cells was estimated by counting the number of spots on the acquired pictures. After detection of the LUC spots leaves were incubated in 100 µM ABA, 0.8 M mannitol, water or left to dehydrate for 24h. The number of GUS spots was detected after histochemically staining the leaves with 5-bromo-4-chloro-3-indolyl-beta-D-glucuronic acid (X-Gluc, Apollo Scientific, Bredbury, UK). Promoter activities were expressed as percentage of the CaMV35S promoter activity (% GUS/LUC) and were determined from at least four independent repetitions.

Protein localization analysis

To study protein localization different fragments of the *CpGRP1* coding sequence or the N-terminal coding sequence of *CpWAK1* and *CpWAK2* were fused to the 5' end of the *GFP* gene in CaMV35S::GFP vector (pGJ280, Willige *et al.,* 2009). The *CpGRP1* full length sequence (CpGRP1-FULL), the signal peptide (CpGRP1-SP, amino acids 1-21) and the C-terminal fragment (CpGRP1-CTERM, amino acids 23-156) were amplified by PCR using CpGRP1_GFP_F1/R1, CpGRP1_GFP_F1/R2 or CpGRP1_GFP_F2/R1 primer combinations (Table S1) to generate *Nco*I sites at both ends. *CpWAKs* already contain an *Nco*I site close to the predicted signal peptide cleavage site. Thus, the fragments encoding the signal peptides were obtained by PCR from corresponding clones using the CpWAKs_GFP_F primer in combination with either the CpWAK1_RT_R or the CpWAK2_RT_R primer (Table S1) to add *Nco*I sites at the 5' end. The *Nco*I/*Nco*I fragments were cloned into the pGJ280 vector to obtain the corresponding translational fusions. Onion cells were transiently transformed *via* particle bombardment (van den Dries *et al.*, 2011). Protein fluorescence was observed using an inverted confocal laser scanning microscope (Nikon Eclipse TE2000-U/D-Eclipse C1, Nikon, Düsseldorf, Germany). The excitation wavelengths were 488 nm for GFP and 543 nm for chloroplast auto-fluorescence and emitted light was detected at 515-530 nm and 570 nm, respectively. Images were captured and processed with EZ-C1 software version 3.20 (Nikon).

Targeted yeast two-hybrid assays

For yeast two-hybrid assays, the HybriZAP^TM Vector kit (Stratagene) was used. The pBD-GAL4 Cam (bait) and pAD-GAL4 (prey) vectors were used to generate translational fusions with the GAL4 DNA binding or activation domain. All the GAL4 fused coding sequences except for *AtWAK1* were amplified with primers to introduce *Eco*RI and *Sal*I sites at the 5' and 3' end, respectively (CpGRP1_Y2H_F/R, AtGRP-3_Y2H_F/R, AtWAK2_Y2H_F/R, AtWAK3_Y2H_F/R, AtWAK4_Y2H_F/R, AtWAK5_Y2H_F/R, CpWAK1_Y2H_F/R primers; Table S1). *AtWAK1* was amplified with the AtWAK1_Y2H_F forward primer to generate an *Eco*RI site and with the AtWAK1_Y2H_R reverse primer to generate an *Xho*I site (Table S1). *CpGRP1* or *AtGRP-3* were cloned between in the pBD-GAL4 Cam vector, *AtWAK2-5* and *CpWAK1* between *Eco*RI and *Sal*I sites of the pAD-GAL4 vector and *AtWAK1* between *Eco*RI and *Xho*I of the pAD-GAL4 vector. Bait and prey plasmids were sequentially transformed into yeast YRG-2 cells using the TRAFO lithium acetate method (Gietz *et al.*, 2002). Yeast cotransformed with the p53-Gal4BD vector consisting of murine p53 fused to the GAL4 DNA binding domain and the pSV40-Gal4AD vector consisting of the SV40 large T-antigen fused to the GAL4 activation domain were used as positive control. Interactions were tested on synthetic defined (SD) minimal medium lacking tryptophan, leucine and histidine with increasing concentrations of 3-amino-1,2,4-triazole (3-AT).

Results

Isolation of the glycine-rich *CpGRP1* cDNA and sequence analysis

The transcriptome of the resurrection plant *C. plantagineum* contains several transcripts with low similarity to described genes (Rodriguez *et al.*, 2010). Some of these transcripts accumulate in desiccated leaves and thus they are predicted to be involved in desiccation tolerance. Among them, a contig sequence of 580 bp coding for the C-terminal of a putative glycine-rich protein (GRP) was identified and termed *CpGRP1*. To characterize the gene, the 5'-end of the *CpGRP1* was isolated. The analysis of the transcript sequence revealed an open reading frame which codes for a protein of 156 amino acids (Fig. 1a). The deduced protein sequence contains a predicted N-terminal signal peptide, a central semi-repetitive glycine-rich motif and a C-terminus rich in cysteines (Fig. 1a). The mature protein has an overall glycine content of 39 %, a predicted molecular weight of 13.0 kDa and an isoelectric point of 8.7. Based on the protein motifs CpGRP1 can be classified as a class II glycine-rich protein (Sachetto-Martins *et al.*, 2000; Fusaro *et al.*, 2001; Nora Bocca *et al.*, 2005). CpGRP1 shows 52 % sequence similarity to the *A. thaliana* protein AtGRP-3 (Q9SL15) (de Oliveira *et al.*, 1990; Park *et al.*, 2001). Both proteins have a conserved domain organization, a similar glycine content and a similar molecular weight (Fig. 1b).

CpGRP1 accumulates in response to dehydration in leaves

The transcriptome data indicated that *CpGRP1* was more abundant in *C. plantagineum* desiccated leaves than in hydrated and rehydrated leaves. To analyse the gene expression in detail, RNA blots were performed of plant samples subjected to a dehydration-rehydration cycle (Fig. 2b). The *CpGRP1* transcript level was higher in partially dried and desiccated leaves than in untreated and rehydrated leaves (Fig. 2b). The transcript was not detectable in roots. *CpGRP1* accumulation was quantified by RT-qPCR analysis (Fig. 2a). The relative transcript abundance was two to three times higher in partially dried and in desiccated leaf tissues than in untreated leaves (Fig. 2a).

(a)

(b)

```
CpGRP1   MMASKVFLLIVLCMAVALFLSSEYVVAREMAETSNAVDTSEKPNTSG-ELKDAKYP-GGGYGGNPGGGYGGNPGGGYGGN 78
AtGRP-3  -MASKA--LVLLGLFAVLLVVSEVAAASSATVNSESKET-VKPDQRGYGDNGGNYNNGGGYQGG-GGNYQGG-GGNYQGG 74

CpGRP1   PGGGYGGGGRGGYGGGGRGGYGGGGRGGYGGGGRGGYGGRGR--CRYGCCGRGGYYGGCRCCDYADQKPDTGFTAAVNNP 156
AtGRP-3  -GGNYQGGG-GRYQGGG-GRYQGGGGRYQGGGGRQGGGGSGGSYCRHGCCYRG-YNGCSRCCSYAGEAVQTQPGH----- 145
```

Fig. 1 *C. plantagineum* CpGRP1 sequence analysis. (a) Nucleotide sequence and predicted protein sequence of CpGRP1. The translation start (ATG) and stop (TGA) codons are underlined and in bold. The different protein domains are indicated by coloured boxes above the sequence. Red: predicted signal peptide; blue: Semi-repetitive Gly-rich region; the two different amino acid tandem repeat motifs are shown with solid or dashed arrows respectively. Yellow: Cys-rich region; the six Cys residues are marked by black triangles. (b) Sequence alignment between CpGRP1 and the *A. thaliana* AtGRP-3 (Q9SL15). Identical (black) and conserved (grey) amino acids are indicated. The Cys-rich region is marked with a black line.

To investigate the protein accumulation antibodies were raised against the recombinant CpGRP1 protein. The specificity of the antibodies is shown in Fig. S1. The antibodies detected CpGRP1 proteins with an apparent molecular weight of 27 kDa which is twice that of the expected value (Fig. 2c). CpGRP1 protein levels correlate with transcript accumulation in dehydrated leaves (Fig. 2c). Two-dimensional separation of proteins revealed at least three distinct protein spots with a similar molecular weight and an isoelectric point of about 5 (Fig. 3). The three protein spots represent probably isoforms of CpGRP1. This is in agreement with the transcriptome data which showed other dehydration-induced transcripts with high similarity to *CpGRP1*.

Fig. 2 Transcript and protein expression analysis of *C. plantagineum* CpGRP1. *CpGRP1* transcript expression was analysed in untreated, partially dried (RWC 50 %), desiccated (RWC 2 %) and 24 h rehydrated leaf and root tissue samples (a,b), ABA, mannitol (osmotic stress) or water (control) 48 h incubated leaf samples (d,e) or untreated, ABA treated, dried and ABA-treated, dried callus samples (g,h). Transcript expression was additionally investigated in leaf samples incubated for 4 h, 24 h and 48 h in 0.5 M or 0.8 M mannitol or sorbitol solutions (i). Real-time quantitative PCR analysis of the *CpGRP1* expression is shown in a,d,g. cDNA was prepared from total RNA isolated after the treatments and amplified using CpGRP1 specific primers. Relative expression values in (a) and (d) were calculated from three different biological repetitions (mean ± SE). Different letters above bars denote statistically significant differences within each group of samples (One-way ANOVA, $P < 0.05$). Expression values in (g) were calculated from three technical repetitions (mean ± SE). RNA blot analysis of *CpGRP1* transcript abundance is shown in b,e,h. Total RNA samples were separated by denaturing agarose gel electrophoresis, blotted on a nylon membrane and probed with a ^{32}P-labelled *CpGRP1* gene specific fragment. Ethidium bromide (EtBr)-stained denaturing agarose gels are shown to indicate equal loading of RNA. RT-PCR analysis of the *CpGRP1* expression in leaves subjected to different osmotic stress treatments is shown in i. CpGRP1 protein analysis is shown in c,f,i. Total protein samples were separated by 15 % (w/v) SDS-PAGE, transferred onto a nitrocellulose membrane and polyclonal antibodies were used to detect the CpGRP1 protein. In (c) and (f) antibodies against the transketolase 3 (TKT3) protein (Bernacchia *et al.*, 1995) were used to show equal loading of proteins. The protein levels of the ABA-induced *CDeT11-24* gene (van den Dries *et al.*, 2011) were analysed as controls for correct ABA treatment. In (i) protein levels of the 48 h treated samples were reported. CpGRP1 protein bands in (c) are indicated by black triangles.

CpGRP1 accumulation is independent of ABA

Many drought responsive genes characterized from *C. plantagineum* were shown to be up-regulated in response to ABA and/or osmotic stress. To investigate whether CpGRP1 is responsive to ABA and/or osmotic stress, RNA was isolated from leaves treated either with 100 µM ABA or incubated in 0.8 M mannitol solution for 48 h. Gene expression was analysed by RT-qPCR (Fig. 2d) and RNA blot (Fig. 2e). Neither ABA

nor mannitol led to induction of the *CpGRP1* transcript after 48 h. The osmotic stress even reduced the transcript level to the level observed in untreated leaves (Fig. 2e). Surprisingly, control leaves which were incubated in water only showed an increased transcript level compared to untreated leaves. *CpGRP1* expression was also analysed in callus by RT-qPCR (Fig. 2g) and RNA blot (Fig. 2h). Similarly to leaves, dehydration but not ABA induced the transcript (Fig. 2g,h). The dehydration of ABA-treated callus failed to induce *CpGRP1* (Fig. 2g,h) thus supporting the role of ABA as negative regulator.

Immunological analysis was conducted using protein samples from ABA and mannitol treated leaves (Fig. 2f) to investigate whether transcript levels correlated with protein levels. Like the transcripts, CpGRP1 proteins are reduced in ABA-treated leaves (Fig. 2f). On the other hand, incubation of leaves in mannitol for 48h caused higher accumulation of the CpGRP1 protein compared to untreated leaves. The level of the ABA-responsive LEA-like CDeT11-24 protein (van den Dries *et al.*, 2011) was determined parallel to CpGRP1 to confirm that ABA signalling pathways were active (Fig. 2f). Further, the effect of osmotic stress on the *CpGRP1* expression was investigated using leaves incubated for 4 h, 24 h and 48 h in 0.5 M or 0.8 M mannitol or sorbitol solutions. *CpGRP1* transcripts were transiently induced during the osmotic stress conditions but they were reduced in all samples subjected to osmotic stress for 48 h (Fig. 2i). The immunological analysis of the 48h-treated samples showed that the CpGRP1 protein is still present in these samples suggesting high protein stability (Fig. 2i).
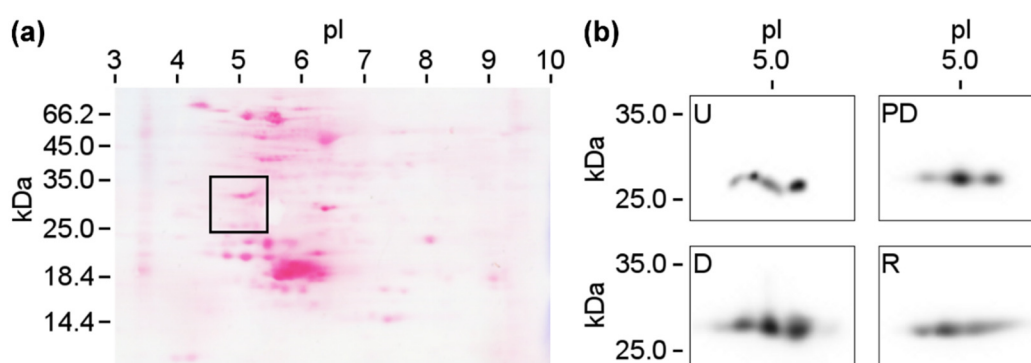
Fig. 3 Two-dimensional immunoblot analysis of the CpGRP1 protein. Total proteins (20 µg) were first separated by isoelectric focusing over a pH range of 3-10 and then by 15 % (w/v) SDS-PAGE. Separated proteins were transferred to nylon membranes and polyclonal antibodies were used to detect the CpGRP1 protein. (a) Ponceau S stained representative membrane showing transferred total proteins. The section of the blot which is shown in (b) is indicated with a black square in (a). (b) Sections of the immunoblots showing detected CpGRP1 protein spots in untreated (U), partially dehydrated (PD; RWC 50 %), desiccated (D; RWC 2 %) and 24 h rehydrated (R) leaf samples from *C. plantagineum*.

Cis-elements conferring dehydration responsive promoter activity are located between -605 to -407 of the *CpGRP1* promoter

To identify factors regulating transcript expression, the 5' upstream sequence of the *CpGRP1* gene was isolated and analysed. Initially, a genomic clone of the *CpGRP1* gene was obtained which permitted to identify the presence of a single intron in the gene (data not shown). A primer designed on the intron sequence was used to walk upstream of the *CpGRP1* gene coding sequence and led to the isolation of 921 bp 5' from the ATG translational start codon. The promoter sequence was analysed using the signal scan tool from the PLACE database (Higo *et al.*, 1999) to identify putative regulatory *cis*-acting elements. Important elements such as ABA-responsive elements (ABREs) and drought-responsive elements (DREs) are present among the putative *cis*-acting elements associated with dehydration-induced gene expression (Table 1).
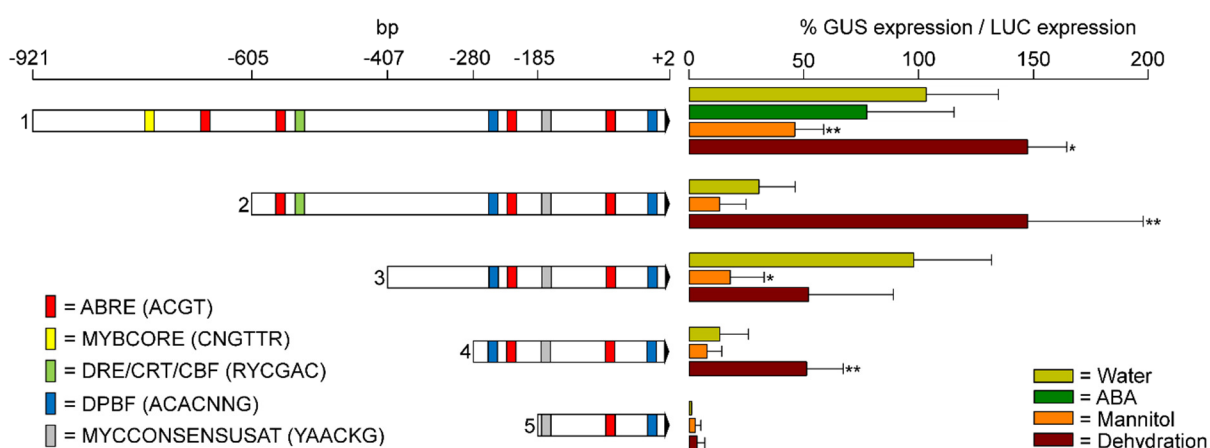
Table 1 Putative *cis*-acting regulatory elements associated with ABA- and dehydration-responsive gene expression identified within the *CpGRP1* promoter region. Sequence motifs and positions in the *CpGRP1* promoter (N is any nucleotide, R is adenine or guanine, Y is cytosine or thymine and K is guanine or thymine) are shown. The c*is*-acting regulatory elements were identified using the PLACE database (Higo *et al.*, 1999).

Cis-element	Sequence	Position	Reference
ABRELATERD1	ACGTG	-224	(Simpson *et al.*, 2003)
ACGTATERD1	ACGT	-670, -561, -80	(Simpson *et al.*, 2003)
DPBFCOREDCDC3	ACACNNG	-252, -21	(Kim *et al.*, 1997)
CBF DRE/CRT	RYCGAC RCCGAC	-533	(Xue, 2002) (Dubouzet *et al.*, 2003)
MYBCORE	CNGTTR	-753	(Lüscher & Eisenman, 1990)
MYCCONSENSUSAT	YAACKG	-174	(Abe *et al.*, 2003)

Several putative ABA-responsive elements (ABREs) were observed, although the gene is not responsive to ABA under the conditions used. A unique DRE element (GCCGAC) was found at position -533. To investigate whether the observed *CpGRP1* transcript levels were the direct result of transcriptional regulation, the *CpGRP1* promoter fragment (921 bp) was fused to the *GUS* reporter gene and the promoter activity was analysed in response to dehydration, ABA and osmotic stress using *C. plantagineum* leaves (Fig. 4). The activity in dehydrating leaves was significantly higher

53

after 24 h which demonstrates that the isolated promoter fragment contains all the *cis*-elements required for drought responsiveness. ABA had no significant effect on the promoter activity compared to the water control which confirms that ABA does not induce transcription and that the putative ABREs are not functional. The mannitol treatment caused a strong reduction in the promoter activity. The promoter activities observed correlated directly with *CpGRP1* transcript levels indicating a transcriptional control of gene expression.

To identify the region responsible for osmotic and dehydration-specific promoter activity, promoter deletions were tested in *C. plantagineum* leaves for responsiveness to osmotic stress and dehydration (Fig. 4). Most of the dehydration-induced promoter activities seem to be due to the region located between nucleotides -605 and -407 (constructs 2,3) which supports an active DRE motif (Fig. 4). The GUS activity detected with the deletion construct 4 must be due to the region located between nucleotides -280 and -185. *Cis*-elements located upstream of the position -605 seem to be partially responsible for the activity in leaves incubated in water and mannitol because the deletion of the nucleotides between -921 and -605 caused a 50 % reduction of the promoter activity. The deletion of the region essential for most of the dehydration-responsiveness of the promoter restored the promoter activity in response to water to the same level as observed for the full length promoter (Fig. 4; constructs 3,1).

Fig. 4 *CpGRP1* promoter activity analysis. Putative *cis*-acting regulatory elements associated with ABA- and dehydration-responsive gene expression are indicated with coloured boxes (left panel). The full length *CpGRP1* promoter (1) and progressive 5' deletions (2, 3, 4, 5) were tested for responsiveness to ABA, mannitol and dehydration in *C. plantagineum* leaves using the transient expression protocol described by van den Dries *et al.* (2010). Bars indicate the relative promoter activities, expressed as percentage of *CpGRP1* promoter fragment activity compared with the CamV35S promoter (% beta-glucuronidase (GUS) / luciferase (LUC)). The values are calculated from 4 independent determinations (mean ± SD) for each treatment. Statistical significant differences from control mean of each promoter fragment are indicated with asterisks above error bars (*$P < 0.05$; **$P < 0.01$).

CpGRP1 is targeted to the apoplast

The presence of an N-terminal signal peptide was predicted for the CpGRP1 protein sequence using WoLF PSORT (http://wolfpsort.org) (Horton *et al.*, 2007). The putative signal peptide is cleaved off at amino acid position 21-22, generating a 13 kDa mature protein. No other signals were predicted in the protein sequence suggesting an apoplasmic localisation for the mature CpGRP1 protein. To validate this prediction *in vivo*, onion cells were biolistically transformed with constructs overexpressing either the full length CpGRP1-GFP (CpGRP1-**FULL**-GFP) or the signal peptide CpGRP1-GFP (CpGRP1-**SP**-GFP) or the C-terminal fragment CpGRP1-GFP (CpGRP1-**CTERM**-GFP) translational fusion proteins. The microscopic analysis of onion cells expressing either the CpGRP1-**FULL**-GFP or the CpGRP1-**SP**-GFP construct revealed the presence of a dot-like fluorescence pattern at the cell boundaries (Fig. 5a,b). This pattern was not observed in cells expressing the CpGRP1-**CTERM**-GFP (Fig. 5c) construct or the GFP alone (data not shown). To demonstrate the apoplastic localization of CpGRP1, GFP fluorescence of transformed onion cells expressing the CpGRP1-**FULL**-GFP or GFP only were analysed during plasmolysis. Plasmolysed cells expressing the CpGRP1-**FULL**-GFP (Fig. 5d) showed GFP fluorescence in the apoplasmic space between the plasma membrane and the cell wall as well as at the cell wall, while cells expressing only GFP retained the GFP signal in the cytoplasm within the plasma membrane (Fig. 5e).

CpGRP1 interacts with the *A. thaliana* wall associated protein kinase 2 (AtWAK2) and with a newly identified *C. plantagineum* WAK (CpWAK1)

It was previously reported that the CpGRP1 *Arabidopsis* homologue AtGRP-3 interacts with the extracellular domain of the wall associated protein kinase 1 (AtWAK1) (Park *et al.*, 2001). Therefore it was tested in yeast two-hybrid assays whether CpGRP1 also interacts with the extracellular domain of AtWAK1 or any other WAK protein from *A. thaliana*. First, the *CpGRP1* coding sequence without the signal peptide was introduced into the bait vector and the resulting construct was used to transform yeast cells. The lack of growth of the transformed yeast on medium lacking tryptophan and histidine excluded any reporter auto-activation (data not shown). Next, the extracellular domain of the five *A. thaliana WAK* genes (*AtWAK1-5*) (He *et al.*, 1999) were cloned

into prey vectors and used to transform yeast cells carrying the bait vector. Yeast cells were not able to grow on interaction medium (SD-WLH) when the CpGRP1-AtWAK1 bait-prey combination was co-expressed (Fig. 6a). Among the other tested combinations, only CpGRP1-AtWAK2 showed an interaction (Fig. 6a). In contrast to the *in vitro* binding assay reported by Park *et al.* (2001), the extracellular domain of AtWAK2 also interacted with the AtGRP-3 in yeast (Fig. 6b). These results suggest that protein motifs capable of interacting with AtWAK2 are conserved between *A. thaliana* AtGRP-3 and *C. plantagineum* CpGRP1.
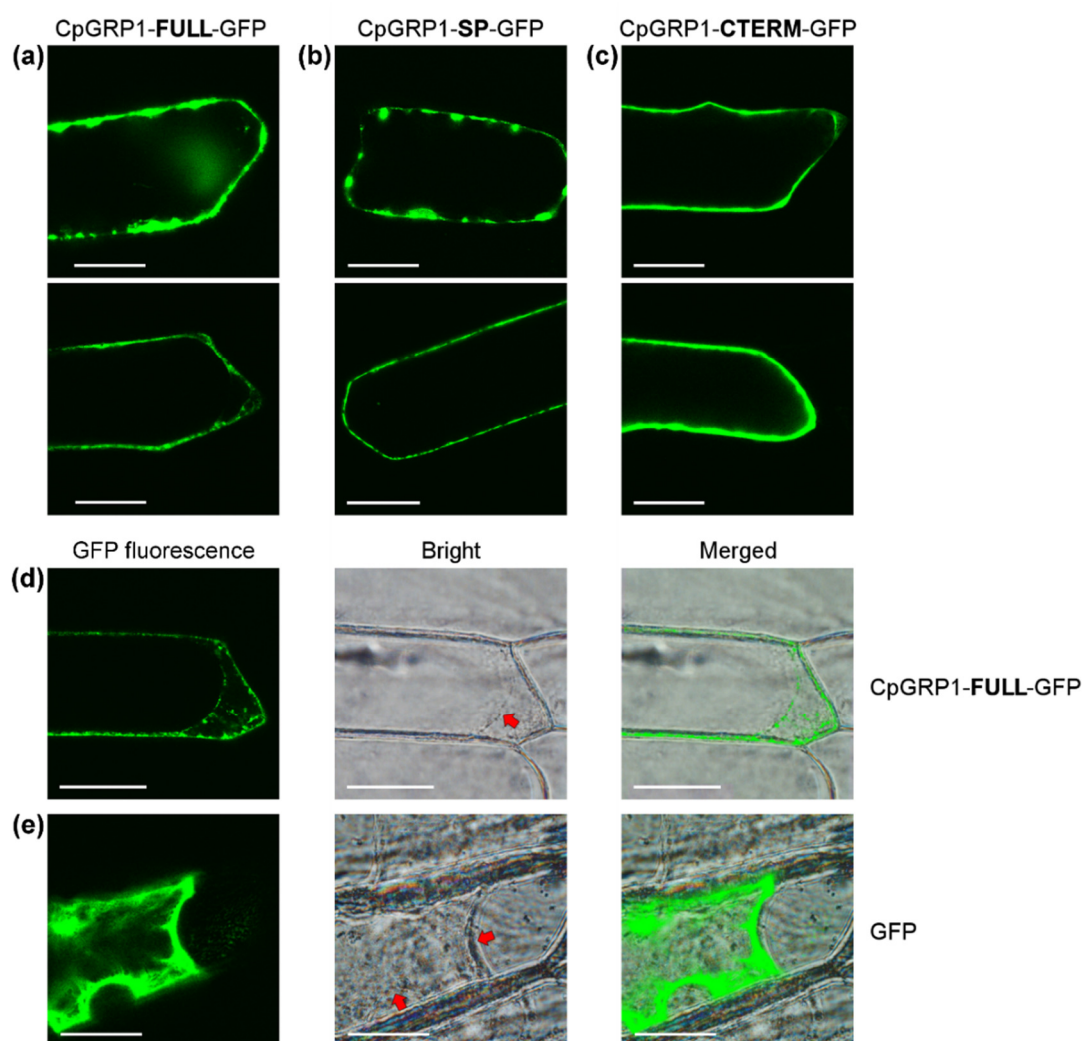
Fig. 5 Cellular localization of the CpGRP1-GFP translational fusions. The full-length *CpGRP1* coding sequence (CpGRP1-**FULL**-GFP) (a), the signal peptide only (CpGRP1-**SP**-GFP) (b) or the *CpGRP1* coding sequence without the signal peptide (CpGRP1-**CTERM**-GFP) (c) were fused to GFP and transiently expressed in onion cells *via* particle bombardment. Images of transformed cells were taken with a confocal microscope after 16 h. Two representative images from different transformations are shown for each construct. Cells transformed with the full length protein coding sequence-GFP fusion (d) or GFP alone (e) were incubated in a 0.5 M sucrose solution for 5 minutes before analysing GFP fluorescence. In (d) and (e) GFP fluorescence only, bright field only or merged images are shown. Red arrows indicate plasma membranes of cells undergoing plasmolysis. The white bar represents 50 µm.
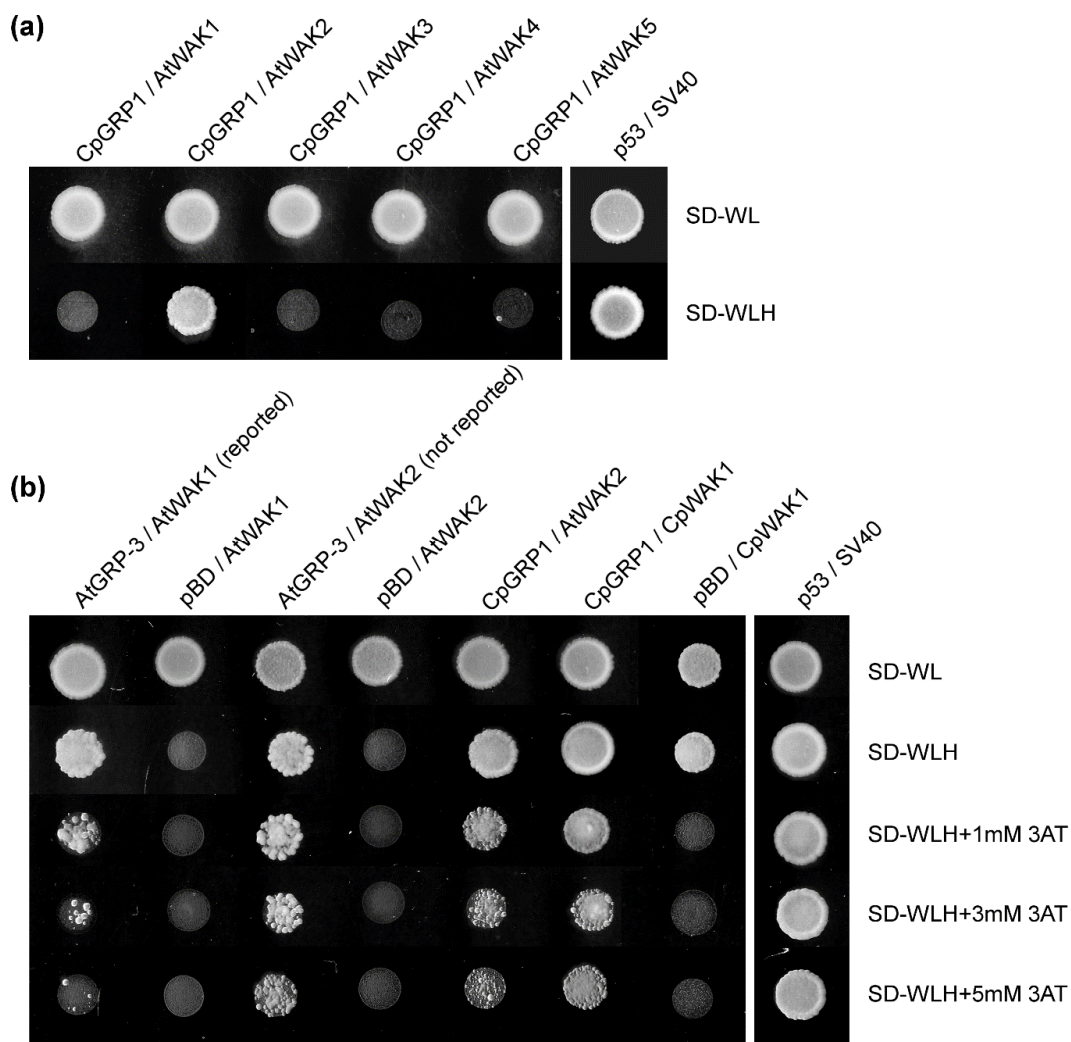
Fig. 6 Yeast two-hybrid interaction assays of selected glycine-rich proteins (GRPs) and cell wall associated protein kinases (WAKs) from *A. thaliana* and *C. plantagineum*. Transformed yeast cells carrying the different bait and prey combinations were resuspended in water and pipetted on synthetic minimal medium lacking tryptophan and leucine (SD-WL) or tryptophan, leucine and histidine (SD-WLH). Yeast carrying interacting protein pairs were additionally pipetted on SD-WLH plates supplemented with increasing concentrations of 3-amino-1,2,4-triazole (3-AT). Images were taken after 5 days using a scanner. (a) Interaction assays of *C. plantagineum* glycine-rich protein 1 (CpGRP1) with *A. thaliana* WAK1-5 extracellular protein domains (AtWAK1-5). Yeast carrying the p53/SV40 bait/prey combination was used as positive control for interactions. (b) Interaction of the *A. thaliana* glycine-rich protein 3 (AtGRP-3) with AtWAK1 and 2 and CpGRP1 with AtWAK2 and the cloned extracellular domain of the *C. plantagineum* wall associated protein kinase 1 (CpWAK1). The vector expressing the GAL4 binding domain only (pBD) was used as a replacement for AtGRP-3 or CpGRP1 to check for WAK-GAL4AD autoactivation.

To test whether CpGRP1 also interacts with *C. plantagineum* WAKs, two *WAK* genes were isolated by 5' and 3' PCR walking. *CpWAK1* codes for a protein of 708 amino acids with a predicted molecular weight of 78.7 kDa and *CpWAK2* for a 726 amino acid protein corresponding to 81.6 kDa. CpWAK1 and 2 show 91 % overall similarity. The analysis of the sequences confirmed the presence of the domains characteristic for WAK proteins (Fig. 7a). CpWAK1 was chosen for further experiments. The sequence

with the highest similarity to CpWAK1 within the UniProtKB databank is a putative kinase from *Ricinus communis* (B9RE26) whereas among the *A. thaliana* sequences, the WAKL8 (Q9SA25) had the highest similarity. All three WAK sequences showed conservation of the kinase domain (Fig. 7b), whereas the extracellular part of these proteins shows high variation. The overall amino acid similarity between CpWAK1 and B9RE26 was 58 % and between CpWAK1 and *Arabidopsis* WAKL8 53 %. Even less similarity was found between CpWAK1 and AtWAK2.

We used the yeast two-hybrid assay to test whether CpGRP1 also interacts with *Craterostigma* WAK1. Although the yeast cells expressing the CpWAK1 extracellular domain-pBD combination (control) were able to grow on interaction medium, this growth was abolished in the presence of 1 mM 3-amino-1,2,4-triazole (3-AT) (Fig. 6b). However, in yeast cells expressing the CpGRP1-CpWAK1 bait-prey combination growth was not arrested with increasing 3-AT concentrations. (Fig. 6b). This result indicates an interaction between the two proteins in yeast and suggests a functional link between the CpGRP1 and CpWAK1 in the plant.

**CpWAKs are targeted to the apoplast; *CpWAKs* are mainly expressed in leaves and down-regulated upon drought**

CpWAK1 and 2 were predicted to carry an N-terminal signal peptide (WoLF PSORT; http://wolfpsort.org). Although only few residues are conserved between CpWAKs and CpGRP1 signal peptides (Fig. 8a), they show very similar hydropathy profiles (Kyte & Doolittle, 1982) (Fig. 8b) suggesting a conserved signalling function. The signal peptide sequence of both CpWAK1 and CpWAK2 was used to study the protein localization *in vivo*. Onion cells overexpressing CpWAK1 or CpWAK2 signal peptide-GFP fusion proteins showed uniform fluorescence signal at the cell boundaries (Fig. 8c). After cell plasmolysis, the fluorescence signal appeared to be distributed in the region between the plasma membrane and the cell wall (Fig. 8c) indicating protein secretion into the apoplastic space like it was shown for CpGRP1 (Fig. 5d).

Gene expression of *CpWAK1* and *CpWAK2* was investigated in the same tissue samples which were used to analyse *CpGRP1* expression (Fig. 9). The *CpWAK1* and *CpWAK2* transcripts were mainly expressed in leaves and down-regulated during dehydration (Fig. 9a,b). Although both ABA and mannitol repressed *CpWAKs* expression, mannitol treatment had a stronger effect than ABA (Fig. 9c).
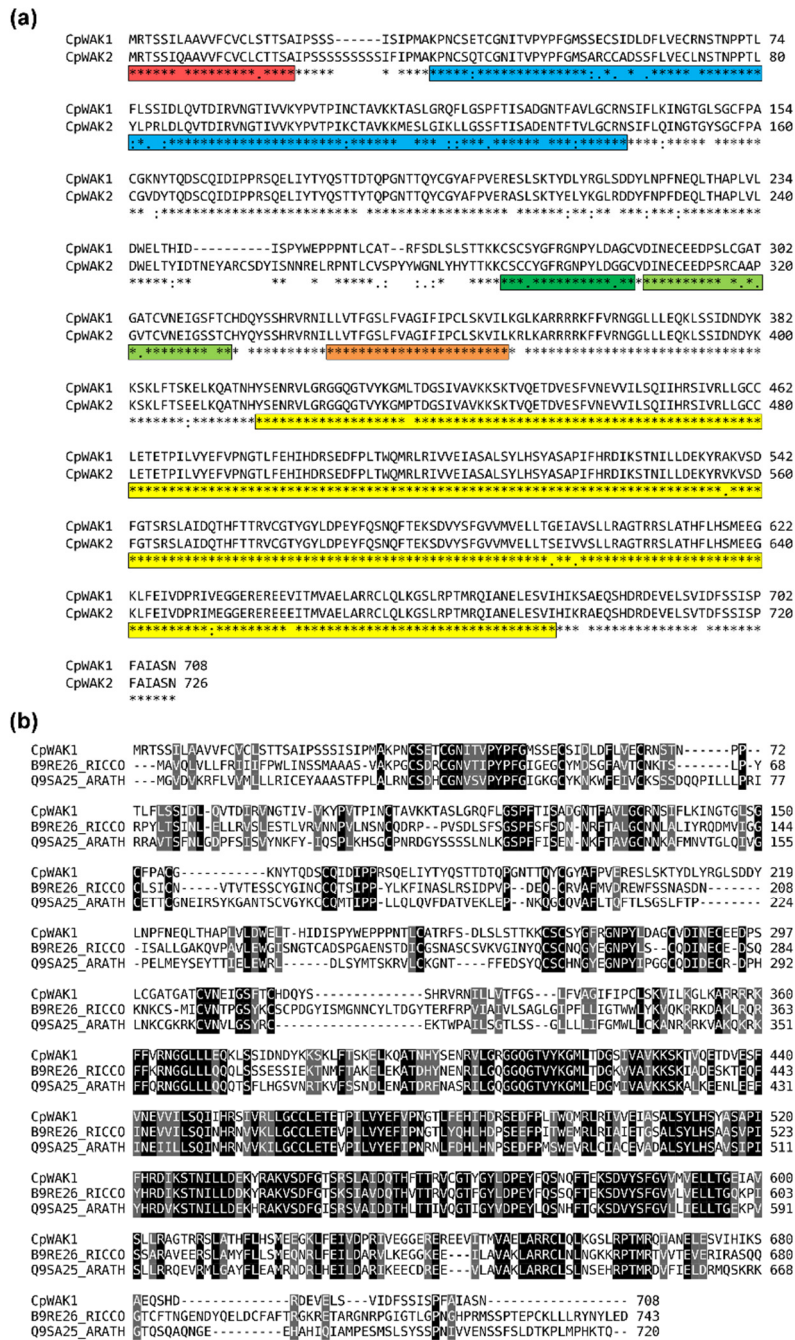
Fig. 7 Analysis of the CpWAK amino acid sequences. (a) Alignment of the CpWAK1 and CpWAK2 amino acid sequences deduced from the cDNA clones. Identical amino acids are indicated by asterisks, and conserved amino acids are indicated by dots or colons. Protein domains were identified using the NCBI CD-Search tool (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer *et al.*, 2011) and the ExPASy ScanProsite tool (http://prosite.expasy.org/scanprosite/) (de Castro *et al.*, 2006). The N-terminal signal peptide and transmembrane domain were predicted using WoLF PSORT (http://wolfpsort.org) (Horton *et al.*, 2007) and TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/), respectively. Different domains are indicated by coloured boxes below the sequence. Red: predicted signal peptide; blue: wall-associated receptor kinase galacturonan-binding domain (pfam13947); dark green: EGF-like domain signature 2 (PS01186); light green: Calcium-binding EGF-like domain signature (PS01187); orange: predicted transmembrane domain; yellow: predicted protein kinase domain (PS50011). (b) Overall amino acid conservation among CpWAK1, a putative kinase from *Ricinus communis* (B9RE26) and the wall-associated protein kinase-like 8 (WAKL8) from *A. thaliana* (Q9SA25). Identical and conserved amino acids are indicated in black and grey, respectively.
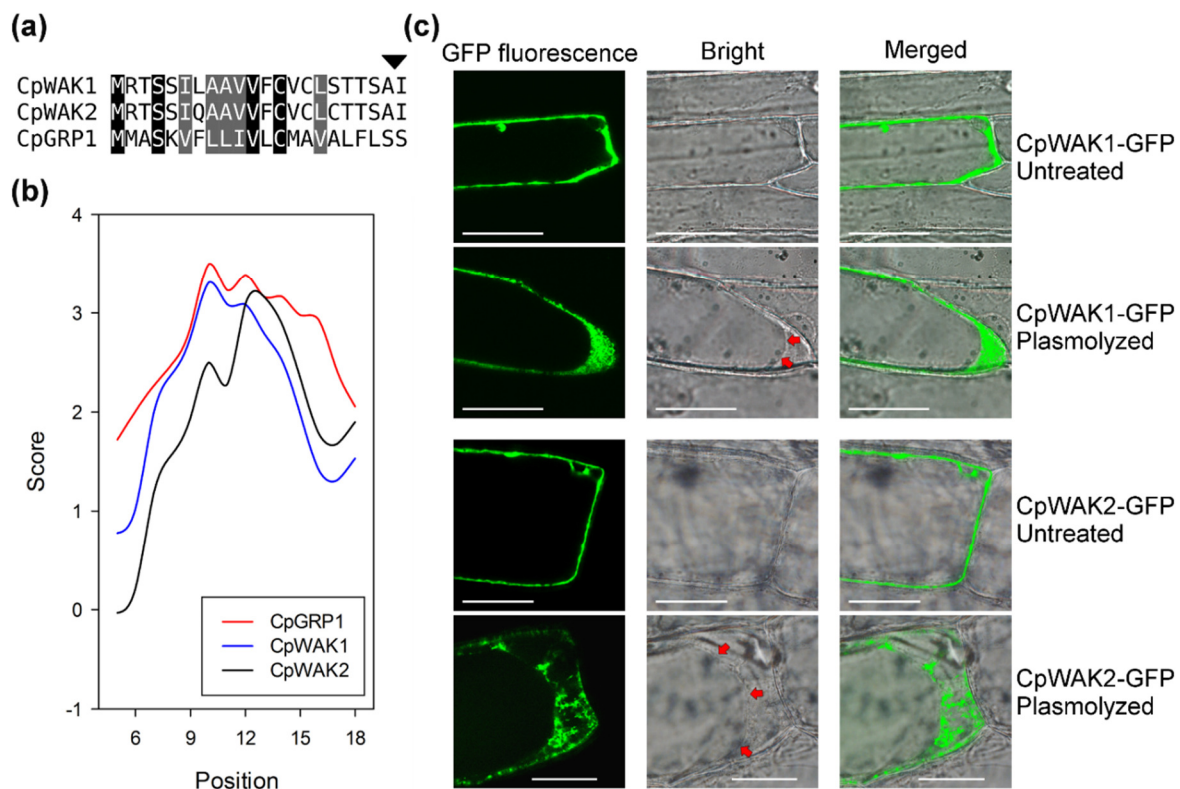
Fig. 8 Analysis of the CpWAK1 and CpWAK2 subcellular localisation. Hydropathy plot (c) and sequence conservation (d) of CpGRP1, CpWAK1 and CpWAK2 signal peptide sequences. Curves in (c) were generated according to Kyte and Doolittle (1982) using the ExPASy ProtScale tool (http://web.expasy.org/cgi-bin/protscale/protscale.pl). The signal peptide cleavage site was predicted using WoLF PSORT (http://wolfpsort.org) (Horton *et al.*, 2007) and is marked by a black triangle (d). *In vivo* localization of the CpWAKs (e). The predicted signal peptide was fused to GFP and then transiently expressed in onion cells *via* particle bombardment. The GFP fluorescence was analysed 16 h after bombardment using a confocal microscope. Pictures of plasmolysed cells were taken 5 minutes after incubation in a 0.5 M sucrose solution. GFP fluorescence only, bright field only or merged images are shown (e). Red arrows indicate the cell plasma membrane undergoing plasmolysis. The white bar represents 50 µm.
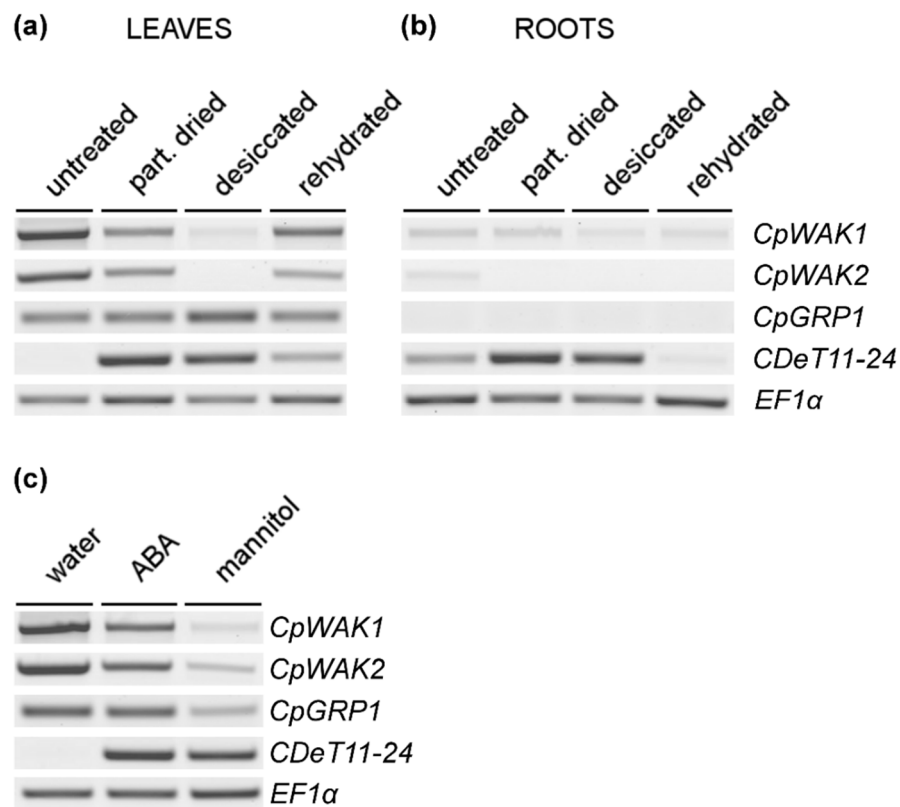
Fig. 9 Analysis of the *CpWAK1* and *CpWAK2* expression. cDNA prepared from different tissue samples was amplified using primers specific for the *CpWAK1* and *CpWAK2* genes. Transcript expression was also analysed for *CpGRP1*, the LEA-like gene *CDeT11-24* (van den Dries *et al.*, 2011) and elongation factor 1α (*EF1α*) for comparison and controls. (a,b) *CpWAK1* and *CpWAK2* expression in leaves and roots from *C. plantagineum* subjected to dehydration and rehydration. (c) *CpWAK1* and *CpWAK2* transcript levels in detached leaves incubated in 100 µM ABA, 0.8 M mannitol solutions or water for 48 h.

Discussion

The analysis of the transcriptome revealed the presence of unknown genes potentially involved in desiccation tolerance mechanisms of *C. plantagineum* (Rodriguez *et al.*, 2010). Among these genes the dehydration-induced glycine-rich protein *CpGRP1* was identified. CpGRP1 was classified as member of the class II secreted glycine-rich proteins (Fig. 1a). The analysis of the localization via CpGRP1-GFP fusions (Fig. 5) confirmed the extracellular localization of this protein suggesting an apoplastic role for CpGRP1. Only two other secreted class II GRPs, the *A. thaliana* glycine-rich protein 3 (AtGRP-3) and the *Nicotiana tabacum* cadmium-induced glycine-rich protein (cdiGRP) have been studied (Park *et al.*, 2001; Ueki & Citovsky, 2002). AtGRP-3 was proposed to be involved in plant-pathogen related signal transduction (Park *et al.*, 2001) whereas cdiGRP was suggested to be an enhancer for callose deposition leading to plant virus defence (Ueki & Citovsky, 2002). These findings show that members of the class II GRPs have different roles *in planta*. A dehydration-induced glycine-rich protein from the resurrection plant *Boea hygrometrica* (BhGRP1) was correlated with cell wall integrity during dehydration (Wang, L *et al.*, 2009). BhGRP1 does not contain a conserved C-terminal Cys-rich domain found in CpGRP1, AtGRP-3 and other GRP and nodulins (Sandal *et al.*, 1992) and thus it might have a different role than CpGRP1.

The dehydration specific accumulation of *CpGRP1* transcript (Fig. 2a,b) and protein (Fig. 2c) in leaves suggests a role in the dehydration process in *C. plantagineum* leaf tissue. The CpGRP1 protein has an apparent molecular weight of 27 kDa which is twice that of the expected value for the mature protein (Fig. 2c). This may be due to abnormal migration related to the CpGRP1 amino acid sequence composition and to the formation of multimers. Indeed, the CpGRP1 recombinant protein has an expected molecular weight of 14.1 kDa but it was detected as a 17.6 kDa protein (Fig. S1). Further analyses are required to understand whether CpGRP1 might form protein multimers and whether this is connected to the function. BLAST analysis revealed the presence of CpGRP1-related sequences in the *C. plantagineum* transcriptomic data (data not shown). Our 2D SDS-PAGE analysis results support the existence of CpGRP1 isoforms (Fig. 3b) suggesting they might have a common role in the dehydration process.

CpGRP1 promoter analysis directly reflected the transcript expression data which indicates that the differential gene expression is mainly regulated at the

transcriptional level. The stress hormone ABA is a known mediator of abiotic stress responses including dehydration (Leung & Giraudat, 1998). Many genes which are thought to be required for the acquisition of desiccation tolerance in *C. plantagineum,* accumulate in response to exogenous ABA treatment (Bartels *et al.*, 1990). Exogenous ABA application however, reduced *CpGRP1* promoter activity as well as transcript and protein levels. ABA treatment of callus prevented desiccation-induced *CpGRP1* transcript accumulation (Fig. 2g,h). These results indicate that ABA acts as negative regulator of *CpGRP1* gene expression and the putative ABA-responsive elements (ABREs) might not be functional as ABRE. The presence of non-functional ABREs has already been reported, e.g., for the LEA *CDeT6-19* promoter from *C. plantagineum* (Michel *et al.*, 1994). To mimic dehydration stress leaves were incubated with high concentrations of mannitol. Promoters of different LEA-like genes have been reported to be highly active in response to this treatment (e.g. Fig. 9c) (Phillips *et al.*, 2008; van den Dries *et al.*, 2011). We found that the mannitol treatment reduced *CpGRP1* transcript abundance after 48 h (Fig. 2d,e). Conversely, the protein levels detected were higher than the control (Fig. 2f). Transcript expression analyses of leaves incubated for different times with mannitol and sorbitol solutions (Fig. 2g) suggest that the opposite effect shown on the transcript and protein levels by the mannitol treatment after 48 h is likely due to a transient upregulation of the *CpGRP1* transcript.

Many dehydration-inducible genes are regulated by DREBs (dehydration responsive element binding) transcription factors which bind to DRE/CRT *cis*-elements in the promoter of different drought-, high salinity- and cold-responsive genes (Lata & Prasad, 2011). A putative DRE/CRT element was identified in the promoter of *CpGRP1* at position -533 (Table 1). The promoter deletion analysis suggested that this element could be responsible for most of the dehydration-induced promoter activities because deletion of the region between the nucleotides -605 and -407 (Fig. 4; constructs 2,3) caused loss of the dehydration-induced promoter activity. Nevertheless, the existence of additional other *cis*-elements within the same region cannot be ruled out. In addition, also the region from -280 to -185 contributed to induction by dehydration and additional elements contained in this region must be involved in the activation by dehydration (Fig. 4; construct 4). Further analysis is required to identify these elements and their binding proteins.

It has been reported by Park *et al.* (2001) that AtGRP-3 binds to the extracellular domain of different cell wall associated protein kinases (AtWAK1,3,5). Also CpGRP1

interacted with the extracellular domain of an *A. thaliana* WAK protein (AtWAK2) (Fig. 6a), which suggests that elements which are important for interaction are conserved between *Arabidopsis* and *Craterostigma* proteins. We therefore investigated the functional link between CpGRP1 and WAKs from *C. plantagineum*. Two genes putatively coding for wall associated protein kinases from *C. plantagineum* were cloned and termed *CpWAK1* and *CpWAK2*. CpWAKs showed similarities to other WAKs (Fig. 7b) and contained the domains commonly found in the family of the *A. thaliana* WAKs and WAK-like proteins (He *et al.*, 1996; He *et al.*, 1999; Verica & He, 2002) (Fig. 7a). We showed that the predicted N-terminal signal peptides of CpWAKs were almost identical (Fig. 8a), had hydropathy profiles similar to CpGRP1 (Fig. 8b) and we further confirmed their apoplastic localization *in vivo* (Fig. 8c).

Recently, *A. thaliana* WAKs were proposed as pectin receptors for both pathogen or wound-generated pectin fragments (oligogalacturonic acids; OGs) and pectin in the cell wall (Kohorn & Kohorn, 2012). It was shown that WAKs bind to pectins in the cell wall through their galacturonan-binding domain (GUB) (Decreux & Messiaen, 2005; Decreux *et al.*, 2006) and that this binding activates downstream signalling pathways (Kohorn *et al.*, 2006; Kohorn *et al.*, 2009; Brutus *et al.*, 2010). Pectins are the major structural component of dicot cell walls. Communication between cell wall and the cytoplasm is essential for cell growth, perceiving environmental changes or plant defence reactions. The ability of WAKs to bind pectins makes WAK proteins good candidates to sense the cell wall status and to transduce the information into the cell. Similarly to other WAKs, CpWAK1 and 2 contain a putative GUB in the extracellular domain and a kinase domain in the intracellular part (Fig. 7a), thus CpWAKs are potentially able to phosphorylate intracellular targets in response to pectin perceived stimuli which could be related to the water status of the cell.

Our expression studies suggest that both CpWAK1 and CpWAK2 are related to the dehydration process in *C. plantagineum* and the corresponding transcripts are down-regulated during desiccation in contrast to CpGRP1 (Fig. 9a). Moreover, it is likely that ABA plays a role in the down-regulation of the transcript during dehydration (Fig. 9c). The expression of the CpWAK1 and CpWAK2 genes fits with a role in desiccation. A decrease of WAK levels led to a reduction of cell expansion in *Arabidopsis* (Anderson *et al.*, 2001; Lally *et al.*, 2001; Wagner & Kohorn, 2001) and WAK protein levels have been correlated with growth inhibition (Lally *et al.*, 2001). Leaves of *C. plantagineum* and other resurrection plants extensively fold during

dehydration to avoid the mechanical stress due to the massive water loss. This phenomenon is only observed in leaves and is a distinctive characteristic of desiccation tolerant plants (Gaff, 1989; Farrant, 2000; Farrant *et al.*, 2003; Willigen *et al.*, 2003; Farrant *et al.*, 2007). It was demonstrated that leaf curling or folding is essential for surviving desiccation in *C. wilmsii* (Farrant *et al.*, 2003). Cell walls of *Craterostigma* showed a pectin content and composition very similar to other dicot plants and no significant differences were observed between untreated and dehydrated leaf cell walls (Vicré *et al.*, 2004; Moore *et al.*, 2013). It is hypothesised that in *C. plantagineum* small but still undefined differences might be involved in preparing plant cell walls to desiccation (Moore *et al.*, 2013). Dehydration-induced cell wall modifications which may enhance cell wall flexibility have already been reported for *Craterostigma* (Vicré *et al.*, 1999; Jones & McQueen-Mason, 2004; Vicré *et al.*, 2004) but the molecular basis of these modifications is unknown. CpGRP1 and CpWAKs are selectively expressed in leaves and they must exert their function in this tissue. We hypothesize that CpWAKs might be working similarly to *A. thaliana* WAKs in the control of the cell volume and that protein levels need to be reduced during dehydration in order to reduce cell volume so that leaves can fold. Cell folding during dehydration and expansion during rehydration may require an interaction of CpWAKs and CpGRP1. Both CpGRP1 and CpWAKs localise to the apoplastic space (Fig. 5 and 8c) and interaction between CpGRP1 and CpWAK1 was demonstrated in yeast (Fig. 6b). Considering these results, CpWAK1 could be essential in regulating leaf cell growth under non-stress conditions in *C. plantagineum* and CpGRP1 might be part of a receptor complex together with CpWAK1. CpGRP1 may mediate interaction between WAK proteins and other cell wall components. This hypothesis is supported by the structure of the CpGRP1. The repetitive glycine rich motif is able to form a β-pleated sheet and mediates interactions with hydrophobic components (Condit & Meagher, 1986) and the Cys-rich domain might form disulphide bonds with other proteins. During dehydration the CpGRP1-CpWAK1 complex might sense the water deficit and activate dehydration-specific signalling pathways. How the stress is sensed by the complex is unclear but dehydration-induced modification of pectins could be involved. It has been shown that WAK binding to pectins was modulated by calcium ions and occurred in conditions promoting the formation of calcium-bridged pectin dimers and/or multimers (also known as 'egg-boxes') indicating that a particular pectin conformation is required for the interaction (Decreux & Messiaen, 2005). Ca^{2+} levels were reported to increase

in *C. wilmsii* cell walls upon dehydration (Vicré *et al.*, 1999) suggesting that calcium-induced pectin conformations might form upon dehydration and be involved in this hypothetical water deficit sensing mechanism. However, CpWAK1 binding to pectins needs to be experimentally shown.

AtGRP-3 was shown to regulate the signalling activity of WAK1 in *A. thaliana* protoplasts (Park *et al.*, 2001). Similarly, CpGRP1 could also have a direct role in stress signalling. This is supported by its increased expression upon dehydration. Under this scenario CpGRP1 might be secreted in the cell wall as inactive protein or prevented to bind to CpWAK1 under non-stress conditions. Dehydration would induce both CpGRP1 protein activation and pectin modifications which are eventually responsible for the formation of the dehydration-specific signalling complex.

Starting from the recent model on function of WAKs as proposed by Kohorn and Kohorn (2012) and considering the observations reported here we propose the model presented in Fig. 10. In non-stress conditions, CpWAK1 is bound to cell wall pectins and regulates cell expansion. CpGRP1 is secreted in the cell wall and can potentially bind to CpWAK1 and participate in CpWAK1 growth regulation. During dehydration, CpGRP1 accumulates and might contribute to pectin modifications together with Ca^{2+} or form with Ca^{2+}-induced pectin structures a complex which bind with high affinity to CpWAK1 leading to the activation of dehydration-specific signalling pathways (Fig. 10, right side). With the progression of dehydration, CpWAK1 levels are reduced in order to decrease cell volume and finally achieve leaf folding.

Further experiments are required to demonstrate the biological relevance of the CpGRP1/CpWAK1 interaction and to elucidate whether CpWAK2 is also able to interact with CpGRP1. The role of CpGRP1 and its relation to pectins in downstream signalling is still puzzling and the identification of additional CpGRP1 interaction partners is required. The identification of intracellular targets of CpGRP1/CpWAK1 will help to clarify whether *WAK* genes are involved in the hypothesized control of morphological adaptations which are essential for the acquisition of desiccation tolerance in *C. plantagineum*.
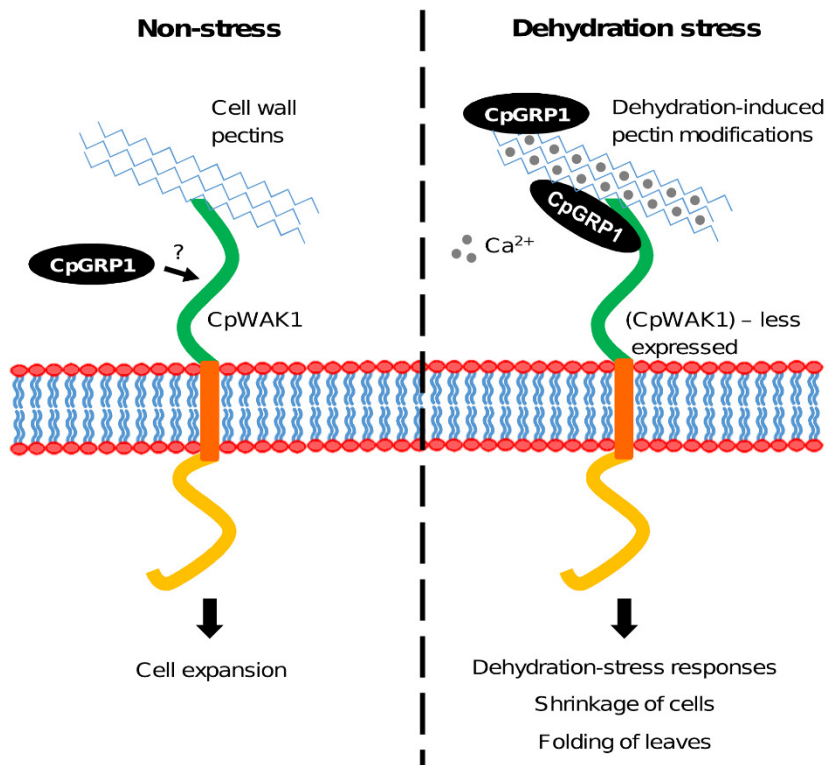
Fig. 10 Model explaining how CpGRP1, CpWAK1 and cell wall pectins regulate growth and dehydration-stress responses. Cell wall pectins (jagged blue lines) are bound to CpWAK1 and regulate growth under non-stress conditions. CpGRP1 is secreted into the cell wall and could bind to CpWAK1 and participate in growth regulation. During dehydration, CpGRP1 is accumulated and might play a role in the modifications of the existing pectin matrix together with Ca^{2+}. Alternatively, a complex between CpGRP1 and Ca^{2+}-induced pectin structures might form and bind with high affinity to CpWAK1. Perturbations of the cell wall pectins are eventually sensed by CpWAK1 and would lead to the activation of dehydration signalling pathways.

Acknowledgements

Authors contributions

Giarola V. mainly conceived, designed and carried out the experiments reported in this study and wrote the manuscript. Krey S. prepared the antibody against CpGRP1 and performed CpGRP1 protein expression and localisation analyses. von den Driesch B.

studied the CpGRP1 promoter activity. Bartels D. supervised the work, contributed to the discussion of the results and corrected the manuscript.

Supporting Data

Table S1 Primers used in this study

5' and 3'-end cloning	Sequence (5' to 3')	Restriction site
ZAP_F1	AGCGGATAACAATTTCACACAGG	-
ZAP_F2	CAGGAAACAGCTATGAACC	-
ZAP_R1	CCCAGTCACGACGTTGTAAAACG	-
ZAP_R2	TGTAAAACGACGGCCAGT	-
CpGRP1_R1	CCTCTTCCACCATAACCACC	-
CpGRP1_R2	CATAACCACCACCGGGATT	-
CpWAK5p_R1	CCTTGTAAACGGTCCCTTGG	-
CpWAK5p_R2	TGAGTTTGCACTGCACCTTG	-
CpWAK3p_F1	TACCGTCCCATATCCATTCG	-
CpWAK3p_F2	CACCAATCATTACAGCGAGAATC	-
CpWAK5p2_R	GGGAAAGATCGCTAAATCGAG	-
CpWAK3p2_F	GAGCTCAGAGTGCAGTATCGAC	-
CpWAK1_F	AAATGAGGACTTCTTCAATACTTGCT	-
CpWAK2_F	CACGGAAATGAGGACTTCTTC	-
CpWAK_R	TCAATATCCATAACCGATAATCCA	-
5' genome walking	**Sequence (5' to 3')**	**Restriction site**
CpGRP1_intron_R	TTGGGCCTCCATATGCTCAG	-
AP1	GTAATACGACTCACTATAGGGC	-
Protein overexpression	**Sequence (5' to 3')**	**Restriction site**
CpGRP1_CDS_F	<u>CCATGG</u>AGTACGTTGTGGCTCGTG	*Nco*I
CpGRP1_CDS_R	<u>CTCGAG</u>AGGATTATTGACGGCCG	*Xho*I
Promoter::GUS fusions	**Sequence (5' to 3')**	**Restriction site**
CpGRP1p605_F	GGAGGTAAC<u>TCTAGA</u>ACATGGA	*Xba*I
CpGRP1p407_F	GTAAATTCATT<u>TCTAGA</u>CCACGCAAA	*Xba*I
CpGRP1p280_F	CTCCA<u>TCTAGA</u>CTTAAAGATATT	*Xba*I
CpGRP1p185_F	CAATCATT<u>TCTAGA</u>CACCAACG	*Xba*I
CpGRP1p_R	GGAAG<u>CCATGG</u>TTAGAATGTGT	*Nco*I
Protein localization	**Sequence (5' to 3')**	**Restriction site**
CpGRP1_GFP_F1	<u>CCATGG</u>TGGCTTCCAAAGTATTT	*Nco*I
CpGRP1_GFP_F2	<u>CCATGG</u>AGTACGTTGTGGCTCGTG	*Nco*I
CpGRP1_GFP_R1	<u>CCATGG</u>GATTATTGACGGCCG	*Nco*I
CpGRP1_GFP_R2	CACGAGCCACAACGTACT<u>CCATGG</u>	*Nco*I
CpWAKs_GFP_F	CACGG<u>CCATGG</u>GGACTTCTTC	*Nco*I
Yeast two-hybrid	**Sequence (5' to 3')**	**Restriction site**
CpGRP1_Y2H_F	CGCCGAATTCACATCTGAGAAGCCT	*Eco*RI
CpGRP1_Y2H_R	TCATCGTCGACATAATCTTCAAG	*Sal*I
AtGRP-3_Y2H_F	<u>GAATTC</u>GAGAGTAAGGAAACTGTGAAACC	*Eco*RI
AtGRP-3_Y2H_R	CACCATG<u>GTCGAC</u>TTAAAATC	*Sal*I
AtWAK1_Y2H_F	<u>GAATTC</u>CAACATCAACCTGGTGAGA	*Eco*RI
AtWAK1_Y2H_R	<u>CTCGAG</u>TTAACGCTTGCAGCTCATAG	*Xho*I
AtWAK2_Y2H_F	<u>GAATTC</u>CAACCTCGCAAGGAGTG	*Eco*RI
AtWAK2_Y2H_R	<u>GTCGAC</u>TTATTCAGGCCTGACTTTACG	*Sal*I
AtWAK3_Y2H_F	<u>GAATTC</u>CAACATCAACCTCGCGAA	*Eco*RI

Yeast two-hybrid	Sequence (5' to 3')	Restriction site
AtWAK3_Y2H_R	GTCGACTACGTGCAGCTCATACTGGA	*Sal*I
AtWAK4_Y2H_F	GAATTCCAAACCTTGCCTCGTTG	*Eco*RI
AtWAK4_Y2H_R	GTCGACTATTCAGGATTGCCTTTAGGTT	*Sal*I
AtWAK5_Y2H_F	GAATTCTGCCAAACTAGATGTGGTG	*Eco*RI
AtWAK5_Y2H_R	GTCGACTAAGGCTCTTCTTTAGGTGTGT	*Sal*I
CpWAK1_Y2H_F	GAATTCACCGTCCCATATCCATTC	*Eco*RI
CpWAK1_Y2H_R	GTCGACTACTGATCATGACACGTGAA	*Sal*I
RT-PCR	**Sequence (5' to 3')**	**Restriction site**
CpWAK1_RT_F	GAGCTCAGAGTGCAGTATCGAC	-
CpWAK1_RT_R	GGGAAAGATCGCTAAATCGAG	-
CpWAK2_RT_F	AAGGTGCTGTGCAGACTCATC	-
CpWAK2_RT_R	AAGATTGCCCCAGTAATATGGAC	-
CpGRP1_RT_F	GGTGGTTATGGTGGAAGAGG	-
CpGRP1_RT_R	TTGCATGAATGAAACGGAGAT	-
CDeT11-24_RT_F	TCGGAAGACGAGCCTAAGAA	-
CDeT11-24_RT_R	ACAGCGCCTTGTCTTCATCT	-
CpEF1α_RT_F	AGTCAAGTCCGTCGAAATGC	-
CpEF1α_RT_R	CACTTGGCACCCTTCTTAGC	-
RT-qPCR	**Sequence (5' to 3')**	**Restriction site**
CpGRP1_QRT_F	TTGTCCCATTCCTCCTATCCTA	-
CpGRP1_QRT_R	TTGCATGAATGAAACGGAGAT	-
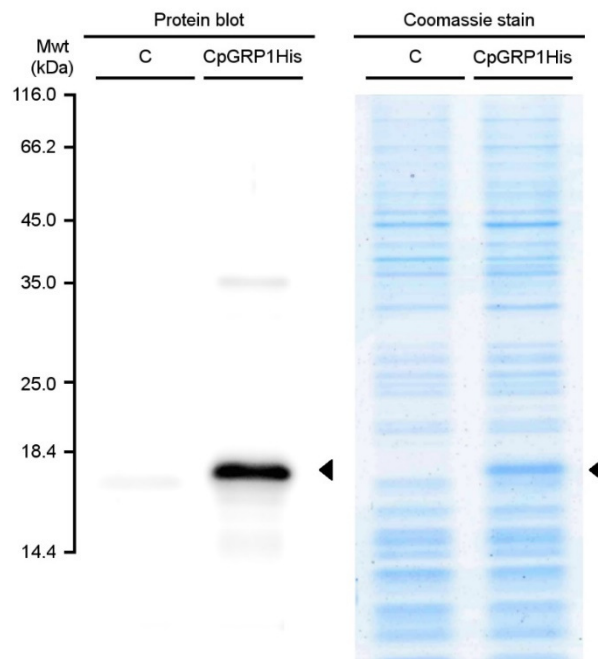
Fig. S1 Protein blot analysis showing the specificity of the polyclonal antiserum used to detect CpGRP1. BL21 *E. coli* cells carrying the empty pET28a(+) vector(C) or the CpGRP1 (CpGRP1His) overexpressing vector were induced with 1 mM IPTG for 5 hours. Total proteins were separated by 15 % (w/v) SDS-PAGE. The gel was either stained with Coomassie blue or blotted to a nitrocellulose membrane for protein immunodetection. The CpGRP1 protein is indicated by black triangles.

CHAPTER 4

Taxonomically restricted genes of *Craterostigma plantagineum* are modulated in their expression during dehydration and rehydration

Valentino Giarola[1], Stephanie Krey[2], Anneke Frerichs[3] and Dorothea Bartels[1]

[1] Institute of Molecular Physiology and Biotechnology of Plants (IMBIO), University of Bonn, Kirschallee 1, D-53115 Bonn, Germany

[2] Present address: Botanical Institute, Cologne Biocenter, University of Cologne, Zülpicherstraße 47b, D-50674, Cologne, Germany

[3] Present address: Institute of Developmental Biology, Cologne Biocenter, University of Cologne, Zülpicherstraße 47b, D-50674, Cologne, Germany

Abstract

The resurrection plant *Craterostigma plantagineum* can tolerate extreme water loss. Desiccation tolerance in this plant is mainly acquired through the induction of protection mechanisms but these mechanisms remain unknown. Transcriptome analysis revealed that a fraction of transcripts with scarce similarity to genes in other organisms are modulated in expression during dehydration/rehydration. Genes which occur only in some species are defined as orphan or taxonomically restricted genes (TRGs) and may be important for the evolution of new traits. Here we report the characterisation of two of these TRGs encoding a cysteine-rich rehydration-responsive protein 1 (CpCRP1) and an early dehydration-responsive protein 1 (CpEDR1). The involvement of CpCRP1 and CpEDR1 in different phases of the dehydration/rehydration cycle is shown by transcript and protein expression analysis. *In silico* sequence analyses predicted that both genes are likely to interact with other cellular components and are localised in two different cellular compartments. We demonstrated by using GFP fusion proteins that CpCRP1 is secreted into the apoplasm whereas CpEDR1 is imported into chloroplasts. We identified orthologs of *CpCRP1* and *CpEDR1* in *Lindernia brevidens* and *Lindernia subracemosa* which belong to the same family as *C. plantagineum* thus suggesting a recent evolution of the genes in this family. Our data support the notion that CpCRP1 and CpEDR1 are TRGs which play a role in the acquisition of desiccation tolerance in *C. plantagineum*. This study highlights the importance of the use of deep transcriptome sequencing data to identify new genes involved in desiccation tolerance mechanisms.

Introduction

Response of plants to dehydration stress is a complex phenomenon involving physiological, biochemical and molecular mechanisms. Most of the flowering plants do not tolerate complete dehydration of vegetative tissues and die when their water content drops below 59–30 % (Höfler *et al.*, 1941). Conversely, a restricted number of plants defined as resurrection plants can survive dehydration of less than 2 % of cellular water. Phylogenetic studies suggest that the desiccation tolerance trait appeared early during evolution of plants and was subsequently lost in the order of higher plants and only lately re-evolved in selected clades (Oliver *et al.*, 2000). Increasing evidence supports the idea that each plant re-evolved specific mechanisms that collectively confer plant desiccation tolerance (Farrant, 2000; Farrant *et al.*, 2007; Moore *et al.*, 2013). Within the angiosperm resurrection plants these mechanisms are inductive rather than constitutive and desiccation tolerance is mainly acquired through the accumulation of protective molecules during dehydration (Gaff, 1989; Oliver *et al.*, 1998).

The resurrection plant *Craterostigma plantagineum* is widely studied to understand the mechanisms underlying desiccation tolerance and many of the genes which are induced during dehydration have been reported (Bartels & Salamini, 2001). Many of the characterised genes belong to the class of the late embryogenesis abundant (*LEA*) and are genes which are normally accumulated in seeds of higher plants and are supposed to contribute to seed desiccation tolerance (Ingram & Bartels, 1996). Recently, the analysis of the transcriptome of *C. plantagineum* at four different stages of dehydration and rehydration revealed that most of the desiccation tolerance mechanisms observed in vegetative tissues of *C. plantagineum* are shared with seeds and pollen (Rodriguez *et al.*, 2010). This suggests that the differential use of similar components rather than specific mechanisms are responsible for the evolution of vegetative desiccation tolerance in this plant. The "rewiring" of the gene regulatory networks to generate new traits is a widely accepted explanation for the large physiological and morphological diversity observed in the animal kingdom (Prud'homme *et al.*, 2007). Well-studied sources of morphological and physiological evolution are changes in *cis*-regulatory elements (Prud'homme *et al.*, 2007) and gene duplication (Ohno, 1970). *C. plantagineum* belongs to the Linderniaceae family. Members of this family differ in their ability to tolerate desiccation and thus comparative

analysis could help to understand the molecular mechanisms underlying this trait. Comparative analysis of regulatory sequences of the *CDeT11-24* gene from the desiccation tolerant *C. plantagineum* and *Lindernia brevidens* plants and the desiccation sensitive *Lindernia subracemosa* plant demonstrated a link between the variability in desiccation tolerance and the presence and distribution of *cis*-elements (van den Dries *et al.*, 2011).

Apart from gene duplication and regulatory changes, an additional and less-studied source of diversity is represented by orphan or taxonomically restricted genes (TRGs) (Khalturin *et al.*, 2009). TRGs do not share any similarity with genes in other species and account for 10-20 % of genes in every eukaryotic genome (Wilson *et al.*, 2005). Increasing evidence suggests that TRGs together with the variation of regulatory networks permit organisms to evolve and adapt to different environmental conditions (Khalturin *et al.*, 2009). TRGs linked to desiccation tolerance were previously identified in *C. plantagineum*. One such gene is *CDT-1* which was shown to confer constitutive desiccation tolerance to *C. plantagineum* callus (Furini *et al.*, 1997). Another example is the *CpEdi-9* gene which was found to be particularly expressed in seeds and in leaf phloem tissues in response to dehydration stress and is hypothesised to protect the cells against desiccation mediated damage (Rodrigo *et al.*, 2004). A significant amount of the transcripts which were identified in *C. plantagineum* did not show any similarity to previously identified genes and thus are putative TRGs (Rodriguez *et al.*, 2010). Transcriptome analysis of the resurrection plant *Haberlea rhodopensis* also revealed the presence of a large fraction of TRGs (Gechev *et al.*, 2013). Some of these genes whose expression vary upon dehydration may be related to desiccation tolerance in these plants. It is very important to analyse the functions of TRGs in desiccation tolerant plants not only to understand the molecular mechanisms but also as source for genetic engineering of crop plants to improve stress tolerance (Dinakar & Bartels, 2013).

Resurrection plants mainly face two different kind of stresses which are associated with the water loss, i.e. mechanical stress and oxidative stress. As a consequence, cell walls, membranes, macromolecules (DNA and proteins) and other cellular components must be protected or repaired in order to survive desiccation. Leaves of *C. plantagineum* and other resurrection plants undergo extensive shrinking and folding to minimize the mechanical stress (Gaff, 1989; Farrant, 2000; Farrant *et al.*, 2003; Willigen *et al.*, 2003; Farrant *et al.*, 2007). Leaf folding in *Craterostigma*

wilmsii and *Myrothamnus flabellifolius* was also proposed as mechanism to shade chlorophyll and avoid ROS formation in water-limiting conditions (Farrant, 2000). Molecular mechanisms underlying leaf folding and genes which are involved remain mainly unknown. The photosynthetic machinery works with high efficiency at a particular light intensity and light above the optimal level causes ROS formation that eventually leads to damage of the photosynthetic apparatus. Resurrection plants evolved different mechanisms to cope with ROS formation (Sherwin & Farrant, 1998; Moore *et al.*, 2009; Scheibe & Beck, 2011). In *C. plantagineum* chlorophyll is maintained during desiccation and compared to plants that degrade chlorophyll, *C. plantagineum* is more exposed to oxidative stress due to ROS generation with the scarcity of water and light. This suggests that this plant must utilize specific mechanisms to protect the photosynthetic apparatus during water-limiting conditions. It was previously shown that during dehydration photosynthesis is inhibited and chloroplasts undergo reversible ultrastructural changes (Schneider *et al.*, 1993; Bockel *et al.*, 1998). Several dehydration- and ABA-induced proteins which accumulate in chloroplasts were proposed to protect photosynthetic structures, e.g. an early light-inducible protein (ELIP) (Bartels *et al.*, 1992), a LEA D29 similar protein (dsp 21), a protein of unknown function (dsp 34) (Schneider *et al.*, 1993) and the plastid targeted proteins (CpPTPs) (Phillips *et al.*, 2002).

Here we have investigated the role of two TRGs encoding the cysteine-rich rehydration-responsive protein 1 (CpCRP1) and the early dehydration-responsive protein 1 (CpEDR1) in *C. plantagineum*. Their importance in desiccation tolerance is discussed.

Materials and Methods

Plant material and treatments

Craterostigma plantagineum Hochst. plants were grown according to Bartels *et al.* (1990) and *Lindernia brevidens* Skan and *Lindernia subracemosa* De Wild according to Dinakar & Bartels (2012). For dehydration treatments, plants were gradually dried in pots with or without illumination. Desiccated plants were rehydrated by submerging the plants in water for the indicated period. After 24 h submersion *C. plantagineum* plants were removed from water and allowed to grow for additional 24 h before collecting the 48 h rehydrated sample. The relative water content (RWC) was determined according to Bernacchia *et al.* (1996). Detached leaves were incubated in either 100 µM ABA or 0.8 M mannitol for 48 h. Plant material was ground in liquid nitrogen and stored at -80 °C until use.

Molecular techniques and sequence analysis

Standard molecular techniques were performed as described by Sambrook *et al.* (1989). Sequencing of DNA and primer synthesis were carried out by Eurofins MWG Operon (http://www.eurofinsgenomics.eu). Protein localizations were predicted using WoLF PSORT (http://wolfpsort.org) (Horton *et al.*, 2007) and ChloroP 1.1 (http://www.cbs.dtu.dk/services/ChloroP/) (Emanuelsson *et al.*, 1999). Protein disorder was analysed using PONDR-FIT™ meta-predictor (http://www.disprot.org/pondr-fit.php) (Xue *et al.*, 2010) and the structure was predicted using the I-TASSER web server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) (Roy *et al.*, 2010). Similarity search analyses were performed with the BLAST algorithm using GenBank® (http://blast.ncbi.nlm.nih.gov/Blast.cgi) (Benson *et al.*, 2012), PLAZA (http://bioinformatics.psb.ugent.be/plaza/) (Proost *et al.*, 2009) and phytozome (http://www.phytozome.net/) (Goodstein *et al.*, 2012) databases. ScanProsite tool (http://www.expasy.org/tools/scanprosite/) (de Castro *et al.*, 2006) was used to search for patterns within the UniProtKB database. Coiled-coil prediction was performed using COILS (http://embnet.vital-it.ch/software/COILS_form.html) (Lupas *et al.*, 1991) and the helical wheel projection of predicted coiled-coil region was obtained using HELIQUEST (Gautier *et al.*, 2008). Putative orthologs of *CpCRP1* and *CpEDR1* in *L.*

brevidens and *L. subracemosa* were retrieved from our transcriptomic databanks (unpublished data) using BLAST+ software (Camacho *et al.*, 2009). Sequence alignments were obtained using ClustalX 2.1 (Larkin *et al.*, 2007). All primers used are listed in Table S1.

Identification of new genes in the *C. plantagineum* transcriptome

The 10119 *C. plantagineum* contigs which were not assigned to UniProtKB entries (Rodriguez *et al.*, 2010) and which show a normalized expression in the dehydrated or desiccated stages higher than 100 or have a covariation coefficient (standard deviation/mean) higher than 1.4 were selected and ranked according to the contig lengths. BLASTX analysis was performed with standard parameters and without filter for low complexity regions and the best match was recorded.

Amplification of 5' and 3' cDNA fragments

The 5' region of the *CpCRP1* and *CpEDR1* genes were obtained by PCR from a λ ZAP II cDNA library prepared from dehydrated *C. plantagineum* leaves (Bockel *et al.*, 1998) using gene and vector specific primers. CpCRP1_R and ZAP_F2 primers (Table S1) were used to amplify the *CpCRP1* 5' region. The 5' region of *CpEDR1* was obtained in two subsequent PCR reactions. In the first PCR, CpEDR1_R1 and ZAP_F2 primers (Table S1) were used to generate the template which was reamplified in a second PCR using CpEDR1_R2 and ZAP_F2 primers (Table S1). PCR products were cloned into pJET 1.2 vectors using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific, St Leon-Rot, Germany) and sequence identities were verified by DNA sequencing.

Transcript analysis

Total RNA from plant samples was isolated according to Valenzuela-Avendaño *et al.* (2005). RNA concentrations were determined using a Biospec-nano spectrophotometer (Shimadzu Biotech, Kyoto, Japan). One μg of each RNA sample was loaded on a 2 % agarose gel to monitor RNA quality. Preparation of cDNA for RT-PCR and RT-qPCR analyses was performed as described in Giarola *et al.* (unpublished) and primers used are listed in Table S1. RNA blot analysis was

performed as described by Bartels *et al.* (1990). The template sequences used to generate radioactive probes were obtained by PCR from *CpCRP1* and *CpEDR1* cloned coding sequences using either CpCRP1_GFP_F1 and CpCRP1_CDS_R or CpEDR1_GFP_F1 and CpEDR1_GFP_R1 primers (Table S1). Hybridization probes were obtained with ^{32}P-labelled dCTP using the DecaLabel DNA labelling kit (Thermo Fisher Scientific, St Leon-Rot, Germany).

Protein analysis

For protein analysis antibodies were raised against the CpCRP1 and CpEDR1 proteins. A *CpCRP1* cDNA fragment (coding for amino acids 26-165) and a *CpEDR1* cDNA fragment (coding for amino acids 71-186) were amplified by PCR using CpCRP1_CDS_F and CpCRP1_CDS_R or CpEDR1_CDS_F and CpEDR1_CDS_R primers to add 5' *Nco*I sites and 3' *Xho*I sites (Table S1). The *Nco*I/*Xho*I fragments were cloned into the expression vector pET28a(+) (Novagen; Darmstadt, Germany) to obtain the corresponding C-terminal 6His-tag translational fusion proteins CpCRP1His and CpEDR1His. The resulting pET28 CpCRP1His and CpEDR1His plasmids were introduced into BL21 (DE3) *E. coli* cells (Amersham Pharmacia Biotech; Piscataway, NJ, USA) and proteins were overexpressed by the addition of 1 mM isopropyl-1-thio-β-d-galactopyranoside (IPTG) for 5 hours. The CpEDR1His recombinant protein was purified using the HIS-Select® Nickel Affinity Gel (Sigma-Aldrich; Saint Louis, MO, USA) under native conditions according to Kirch and Röhrig (2010). To obtain the CpCRP1His recombinant protein, bacterial inclusion bodies were isolated (Schmidt *et al.*, 1986) and dissolved in 8 M urea / 0.1 M sodium phosphate buffer (pH 8.0). Resuspended proteins were filtered through a 0.2 μm filter (Filtropur S; SARSTEDT, Nümbrecht, Germany) and then purified using the HIS-Select® Nickel Affinity Gel (Sigma-Aldrich) under denaturing conditions following manufacturer's instructions. Purified proteins were eluted with extraction buffer (0.1 M Hepes [pH 7.9], 6 M urea, 0.5 M imidazole) and then dialysed against 0.1 M ammonium hydrogen carbonate buffer using a Slide-A-Lyzer™ 10K Dialysis Cassette (Thermo Fisher Scientific, St Leon-Rot, Germany). Purified CpCRP1 and CpEDR1 proteins were freeze-dried and sent to Seqlab (Sequence Laboratories Göttingen GmbH, Göttingen, Germany; www.seqlab.de) to raise polyclonal antisera in rabbits. The specificity of the antibody was confirmed using recombinant proteins in protein blots (Fig. S1). Total proteins

were extracted from *C. plantagineum* samples according to Wang *et al.* (2003). The cell wall fraction was isolated according to Feiz *et al.* (2006) and dissolved in SDS sample buffer to extract cell wall proteins. Total proteins (10 µg) and cell wall proteins were separated by 15 % (w/v) SDS-polyacrylamide gel electrophoresis (PAGE) according to Laemmli (1970). Separated proteins were transferred to nitrocellulose membranes (Towbin *et al.*, 1979). A 1:5000 dilution of the polyclonal antisera was used to detect CpCRP1 and CpEDR1 proteins.

Protein localization analysis

To study the subcellular localization of proteins different fragments of the *CpCRP1* and *CpEDR1* coding sequences were fused to the 5' end of the *GFP* gene in the CaMV35S::GFP vector (pGJ280, Willige *et al.* 2009). The near full length *CpCRP1* sequence (CpCRP1, amino acids 1-144), the signal peptide (CpCRP1-SP, amino acids 1-24) and the C-terminal fragment (CpCRP1-CTERM, amino acids 26-144) were amplified by PCR using CpCRP1_GFP_F1/R1, CpCRP1_GFP_F1/R2 or CpCRP1_GFP_F2/R1 primer combinations (Table S1) to generate *Nco*I sites at both ends. The *CpEDR1* sequence (CpEDR1, amino acids 1-186), the signal peptide (CpEDR1-SP, amino acids 1-73) and the C-terminal fragment (CpEDR1-CTERM, amino acids 75-186) were amplified by PCR using CpEDR1_GFP_F1/R1, CpEDR1_GFP_F1/R2 or CpEDR1_GFP_F2/R1 primer combinations (Table S1) to generate *Nco*I sites at both ends. The different *Nco*I/*Nco*I fragments were cloned into the pGJ280 vector to obtain the corresponding translational fusions. *C. plantagineum* leaf or onion cells were transiently transformed by particle bombardment (van den Dries *et al.*, 2011). Protein fluorescence was observed using an inverted confocal laser scanning microscope (Nikon Eclipse TE2000-U/D-Eclipse C1, Nikon, Düsseldorf, Germany). The excitation wavelengths were 488 nm for GFP and 543 nm for chloroplast auto-fluorescence and emitted light was detected at 515-530 nm and 570 nm, respectively. Leaves were incubated in a 0.5 M sucrose / 0.05 M Tris buffer (pH 8.0) solution for 5 minutes for plasmolysis before analysing GFP fluorescence. Images were captured and processed with EZ-C1 software version 3.20 (Nikon, Düsseldorf, Germany).

Results

CpCRP1 and CpEDR1 transcripts have no sequence similarities to described genes and participate in desiccation tolerance mechanisms in *C. plantagineum*.

BLASTX analysis of the transcriptome of *C. plantagineum* revealed that part of the transcripts (33 %) have low similarity to described genes (Rodriguez *et al.*, 2010). For each transcript, normalized expression values during dehydration and rehydration were obtained and used to calculate the covariation coefficient which is a measure of the variation in expression level (Rodriguez *et al.*, 2010). Among the transcripts with low similarity to described genes, eight transcripts were identified with high expression levels in dehydrated or desiccated tissues or with high covariation coefficients. These transcripts may be specific for the desiccation tolerant plant *C. plantagineum* and thus be important for desiccation tolerance (Table 1). BLASTX analysis performed without filtering for low complexity regions revealed that some of the identified transcripts contain motifs which are similar to previously identified genes, e.g. to a glycine rich protein or dehydrins (Table 1).

Table 1 Most abundant transcripts in dehydrated or desiccated leaves or with the highest covariation coefficient lacking UniProt annotation from *C. plantagineum* transcriptome. The table shows the total read count for each gene after normalization across the four samples: (U) untreated, (PD) partially dried, (D) desiccated, (R) rehydrated. The expectation value (E) of the best BLASTX hit is reported. The *06401* and *23731* transcripts were selected for further analyses and termed *CpCRP1* and *CpEDR1*, respectively.

Transcript number	Length (bp)	U	PD	D	R	Covar	BLASTX hit	E
23731	592	5.34	121.77	8.35	13	1.52	transmembrane channel-like protein 2-like [*Xenopus tropicalis*]	5.1
28056	580	33.97	75.82	104.57	57.42	0.44	glycine-rich protein-like [*Citrus sinensis*]	1.1E-2
27641	562	5.34	158.66	14.64	21.49	1.45	hypothetical protein [*Vibrio campbellii*]	3.2
06401	513	31.76	7.14	8.35	166.08	1.43	no similarity	-
25497	358	3.70	67.58	8.35	7.90	1.40	group 3 LEA protein [*Medicago sativa*]	2.0E-7
00250	303	17.65	230.13	293.60	66.47	0.86	dehydrin Rab18-like [*Cucumis sativus*]	2.0E-15
08958	283	227.76	122.25	41.30	43.26	0.81	metallothionin 3 [*Salvia miltiorrhiza*]	1.0E-4
28852	266	3.7	4.55	101.62	28.15	1.34	no similarity	-

The transcripts *23731* and *27641* showed 87 % identity on the nucleotide level and thus are probably isoforms of the same gene. The transcripts *23731* and *06401* showed the lowest similarities to described genes. According to the normalized expression values both genes were transiently induced in a phase of the dehydration and rehydration cycle. The *23731* transcript was highly expressed in partially dehydrated leaf tissue whereas the *06401* transcript was transiently expressed in rehydrated leaf tissue. The 5' ends of *06401* and *23731* transcripts were cloned and the complete transcripts were further analysed. The *06401* transcript codes for a protein of 165 amino acids with an N-terminal signal peptide which likely targets the protein to the endoplasmic reticulum and eventually to the cell apoplasm (Fig. 1a). The mature protein has an unusual high cysteine content of 10 %, a predicted molecular weight of 15.4 kDa and an isoelectric point of 8.6. Due to the high cysteine content of the *06401* encoded protein and the high abundance in rehydrated leaves, the gene was termed cysteine-rich rehydration-responsive protein 1 (*CpCRP1*). In the CpCRP1 sequence we identified two repeat units (Fig. 1c). The position of cysteine residues in these motifs is almost completely conserved (Fig. 1c). When a protein pattern representing the distribution of the cysteine residues in these two stretches was generated and used to search proteins in UniProtKB (http://www.expasy.org/tools/scanprosite/), the same pattern was retrieved only twice in the pesticidal crystal protein cry5Ac (P56955) from *Bacillus thuringiensis*. Except for the cysteine position, no other sequence similarities were found to CpCRP1. The *23731* gene codes for a protein of 188 amino acids, with a predicted plastid transit peptide of about 73 amino acids (Fig. 1b). The deduced mature protein has a molecular weight of 12.8 kDa and an isoelectric point of 5.6 and is predicted to be mainly unfolded using PONDR-FIT™ meta-predictor (http://www.disprot.org/pondr-fit.php) (Xue *et al.*, 2010). The major structural feature which is predicted for 23731 is a coiled-coil structure (Fig. 1b). Coiled coil structures typically consist of two or more amphiphatic α-helices. Alpha-helical wheel projection using HELIQUEST shows the potential of the predicted region to form an amphiphatic α-helix. Indeed, positively and negatively charged motifs mainly composed of lysine and glutamic acid residues and a stretch of hydrophobic amino acids are distributed at opposite faces of the helical model (Fig. 1d). Because of the high abundance in partially dehydrated leaves, the *23731* transcript was termed early dehydration-responsive 1 (*CpEDR1*).
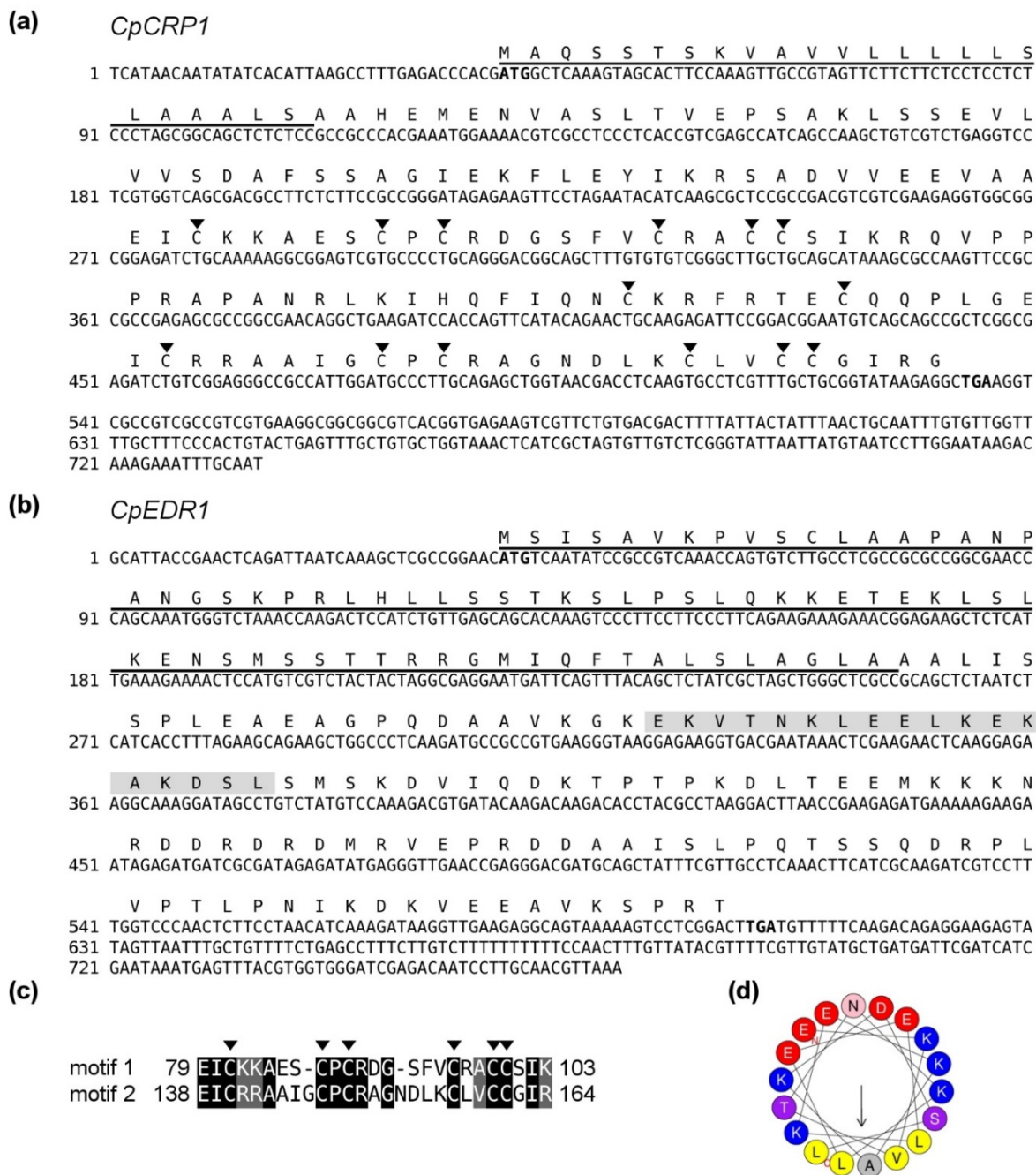
Fig. 1 *C. plantagineum* CpCRP1 and CpEDR1 sequence analyses. The translation start (ATG) and stop (TGA) codons are in bold. Predicted signal peptides are indicated by solid black lines. (a) Nucleotide sequence and predicted protein sequence of the *CpCRP1* gene. The cysteine residues are indicated by black triangles. (b) Nucleotide sequence and predicted protein sequence of the *CpEDR1* gene. The grey box indicates predicted coiled-coil region. (c) Alignment of two conserved amino acid motifs contained in CpCRP1. Identical (black) and conserved (grey) amino acids are indicated. The cysteine residues are indicated by black triangles. (d) Helical wheel projection of the predicted coiled-coil region of CpEDR1. Yellow: hydrophobic residues; purple: serine and threonine; blue: basic residues; red: acidic residues; grey: alanine; pink: asparagine. The hydrophobic region is indicated by an arrow.

Similarity search analysis in PLAZA (http://bioinformatics.psb.ugent.be/plaza/) (Proost *et al.*, 2009) and phytozome (http://www.phytozome.net/) (Goodstein *et al.*, 2012) genomic databases as well as in the transcriptome of the resurrection plants *Haberlea*

rhodopensis and *Boea hygrometrica* failed to identify any gene similar to CpCRP1 or CpEDR1 (X. Deng and T. S. Gechev personal communication). The analysis of the CpCRP1 and CpEDR1 protein structure and function using the I-TASSER web server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) (Roy *et al.*, 2010) did also not show any significant result.

CpCRP1 transcript transiently accumulates during rehydration in leaves and roots; ABA reduces transcript accumulation whereas osmotic stress induces the transcript.

Expression of the *CpCRP1* gene was analysed in leaf and root samples subjected to a dehydration-rehydration cycle by RT-qPCR (Fig. 2a,b) and RNA blot (Fig. 2d,e). RT-qPCR analysis showed that the transcript levels were about two times higher in rehydrated leaves and five times higher in rehydrated roots than in the corresponding untreated tissues (Fig. 2a,b). Transcript levels in desiccated leaf and root samples were about ten times lower than in untreated samples. RNA blot analysis confirmed the RT-qPCR result (Fig. 2d,e).

To investigate whether *CpCRP1* is responsive to ABA and/or osmotic stress, RNA was isolated from leaves treated either with 100 µM ABA or incubated in 0.8 M mannitol solution for 48 h and gene expression was analysed by RT-qPCR (Fig. 2c) and RNA blot (Fig. 2f). RT-qPCR analysis showed that ABA reduced the *CpCRP1* transcript levels by about half whereas *CpCRP1* levels were more than two times higher in osmotically stressed leaves than in control leaves. RNA blot analysis (Fig. 2f) confirmed the RT-qPCR results and showed that leaves which were incubated in water had an increased transcript level compared to untreated leaves.

Antibodies were raised against the recombinant CpCRP1 protein to investigate protein expression. A protein with a molecular weight of 15 kDa was detected in cell wall protein samples from untreated leaves (Fig. 2g). No proteins were detected in the corresponding total protein samples from leaves (Fig. 2g) and roots (data not shown) as well as in total protein samples from water-, ABA- and mannitol-treated leaves (data not shown).
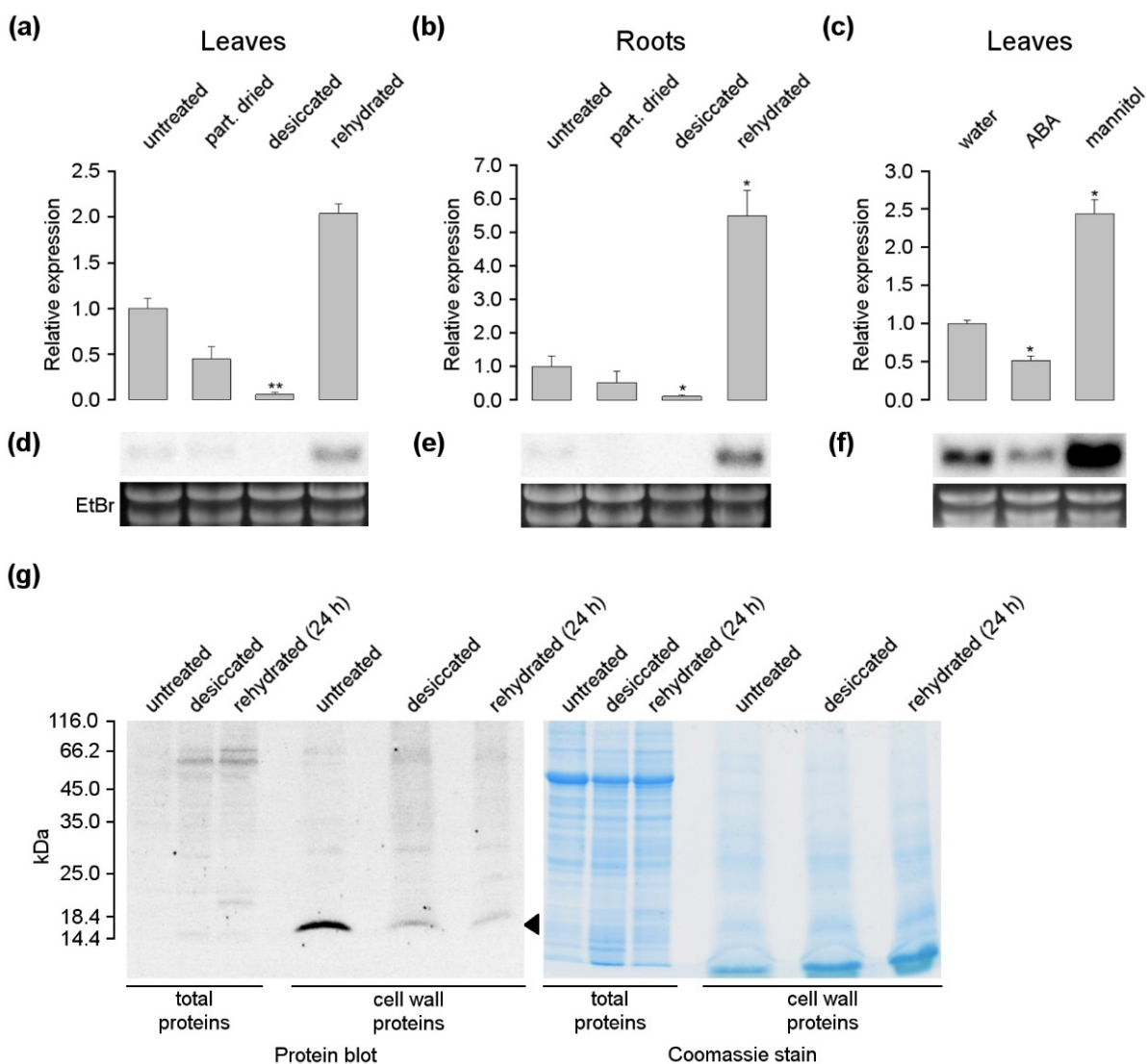
Fig. 2 Transcript and protein expression analyses of CpCRP1. *CpCRP1* transcript abundance was analysed in untreated, partially dried (RWC 50 %), desiccated (RWC 2 %) and 24 h rehydrated leaf (a,d) and root (b,e) tissue samples or ABA, mannitol (osmotic stress) or water (control) incubated leaf samples (c,f). Real-time quantitative PCR analysis of the *CpCRP1* expression is shown in a,b,c. cDNA was prepared from total RNA and amplified using *CpCRP1* specific primers. Relative expression values in (a) and (b) were calculated from three different biological repetitions and in (c) from two different biological repetitions (mean ± SE). Asterisks above bars denote statistically significant differences compared to untreated or control samples (One-way ANOVA, $*P < 0.05$; $**P < 0.01$). RNA blot analysis of *CpCRP1* transcript abundance is shown in d,f. Total RNA samples were separated by denaturing agarose gel electrophoresis, blotted on a nylon membrane and probed with a ^{32}P-labelled *CpCRP1* specific fragment. Ethidium bromide (EtBr)-stained denaturing agarose gels are shown to indicate equal loading of RNA. CpCRP1 protein expression was analysed in total protein and cell wall protein samples prepared from untreated, desiccated (RWC 2 %) and 24 h rehydrated leaves (g). Protein samples were separated by 15 % (w/v) SDS-PAGE, transferred onto a nitrocellulose membrane and polyclonal antibodies were used to detect the CpCRP1 protein. The coomassie-stained gel of protein samples is shown to indicate protein levels. CpCRP1 protein band is indicated by a black triangle.

CpCRP1 is targeted to the apoplast

The presence of an N-terminal signal peptide was predicted for the CpCRP1 protein sequence (WoLF PSORT; http://wolfpsort.org) (Horton *et al.*, 2007). The putative signal peptide is cleaved at amino acid position 24-25, generating a 15.4 kDa mature protein. No other signals were predicted in the protein sequence suggesting an apoplasmic localisation for the mature CpCRP1 protein. To study the localisation *in vivo*, *C. plantagineum* and onion cells were biolistically transformed with constructs expressing either the near full length CpCRP1 protein sequence fused to the GFP (CpCRP1-GFP) or only the signal peptide fused to GFP (CpCRP1-**SP**-GFP) or only the C-terminal part fused to GFP (CpCRP1-**CTERM**-GFP). The microscopic analysis of *C. plantagineum* cells expressing the CpCRP1-GFP construct revealed the presence of dot-like fluorescence patterns at the cell boundaries (Fig. 3a). This pattern was not observed in *C. plantagineum* cells expressing the CpCRP1-**SP**-GFP construct but only in cells expressing the CpCRP1-**CTERM**-GFP construct indicating that the C-terminal part of CpCRP1 is responsible for the formation of the fluorescence dots (Fig. 3a). *C. plantagineum* cells expressing the CpCRP1-**SP**-GFP construct showed uniform fluorescence signals. Similar fluorescence patterns were observed in onion cells expressing the corresponding constructs. To demonstrate the apoplastic localization of CpCRP1, GFP fluorescence was analysed in transformed onion cells expressing the CpCRP1-GFP during plasmolysis. Plasmolysed cells showed GFP fluorescence between the plasma membrane and the cell wall thus proving the apoplastic localisation of CpCRP1 (Fig. 3b).

CpEDR1 transcript is transiently induced and the encoded protein accumulates during dehydration in leaves. ABA and osmotic stress induce transcript and protein accumulation.

RT-qPCR and RNA blot analyses of the *CpEDR1* gene expression in leaf and root samples subjected to dehydration and rehydration are reported in Fig. 4a,b and 4d,e, respectively. *CpEDR1* levels were about 14 times higher in partially dehydrated leaves than in untreated leaves (Fig. 4a) whereas in partially dehydrated roots *CpEDR1* levels were 20 times higher than in untreated roots (Fig. 4b). Although a higher induction of the transcript was measured by qPCR in roots than in leaves the transcript abundance

was lower in root samples than in leaf samples as indicated by RNA blot analysis (Fig. 4d,e) and RT-qPCR Cq values (data not shown).
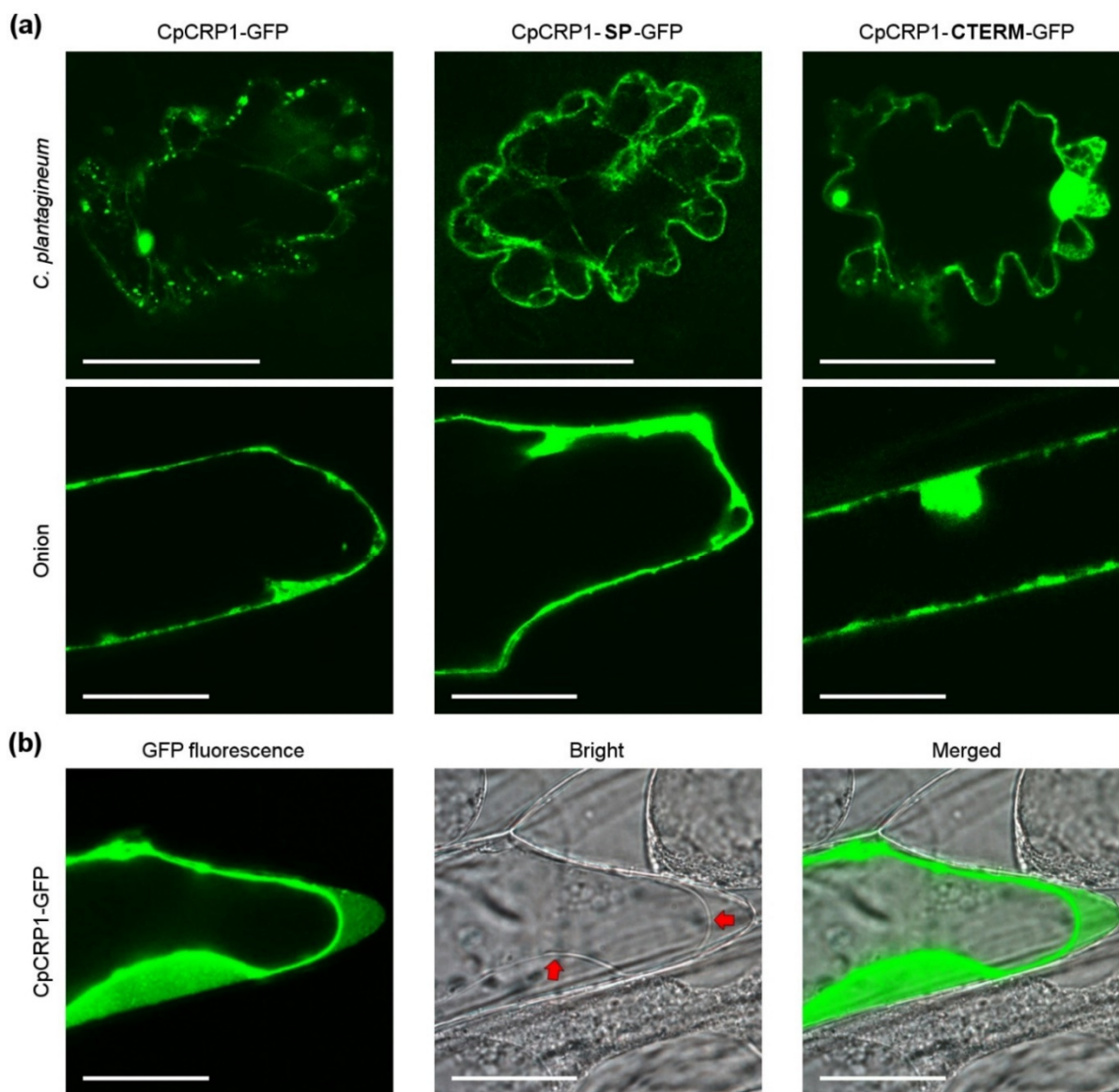
Fig. 3 Cellular localization of the CpCRP1-GFP translational fusions. (a) CpCRP1-GFP subcellular localisation. The *CpCRP1* coding sequence (CpCRP1-GFP), the signal peptide only (CpCRP1-**SP**-GFP) or the *CpCRP1* coding sequence without the signal peptide (CpCRP1-**CTERM**-GFP) were fused to GFP and transiently expressed in *C. plantagineum* and onion cells *via* particle bombardment. A representative image from different transformations is shown for each construct. (b) Onion cells transformed with the full length protein coding sequence-GFP fusion were incubated in a 0.5 M sucrose / 0.05 M Tris buffer (pH 8.0) solution for 5 minutes before analysing GFP fluorescence. In (b) GFP fluorescence only, bright field only or merged images are shown. Red arrows indicate plasma membranes of cells undergoing plasmolysis. Images of transformed cells were taken after 16 h using a confocal microscope. The white bar represents 50 µm.

Importance of light on transcript expression was already shown for the *C. plantagineum* early light-inducible protein (ELIP) (Bartels *et al.*, 1992). We analysed the influence of

light on *CpEDR1* transcript levels using RNA samples prepared from leaves dehydrated under dark or light (Fig. 4g). Accumulation of the *CpEDR1* transcript was also induced in samples subjected to dehydration in the dark for 48 hours or 5 days (90 % and 54 % RWC) (Fig. 4g) but transcript levels were reduced compared to leaves dehydrated to 50 % RWC in the light (Fig. 4d). Similarly to *CpCRP1*, the effect of ABA and osmotic stress was investigated on the expression of *CpEDR1*. RT-qPCR and RNA blot analyses showed that both ABA and mannitol led to increased *CpEDR1* transcript levels (Fig. 4c,f), but ABA was much more effective than mannitol in inducing transcript accumulation (Fig. 4c,f).

To investigate protein accumulation antibodies were raised against the recombinant CpEDR1 protein. A protein band with the expected molecular weight of 13 kDa was detected in partially dried, desiccated and 24h rehydrated leaf samples indicating that the protein like the transcript accumulated early during dehydration (Fig. 4h). Although the transcript levels decreased, protein levels remained stable up to 24 h after rehydration (Fig. 4a,d and 4h). In roots, no protein bands were detected in the corresponding dehydrated and rehydrated samples (Fig. 4i). Lack of protein detection in this tissue correlates with the low *CpEDR1* transcript levels observed (Fig. 4e). The level of the dehydration-responsive LEA-like CDeT11-24 protein (van den Dries *et al.*, 2011) was determined parallel to CpEDR1 to confirm that dehydration-induced signalling pathways were active in root samples (Fig. 4i). Similarly to transcripts, proteins accumulated in ABA- and mannitol-treated leaves (Fig. 4j). Additional protein bands of 22 and 24 kDa with the same expression pattern of the 13 kDa protein were also detected.
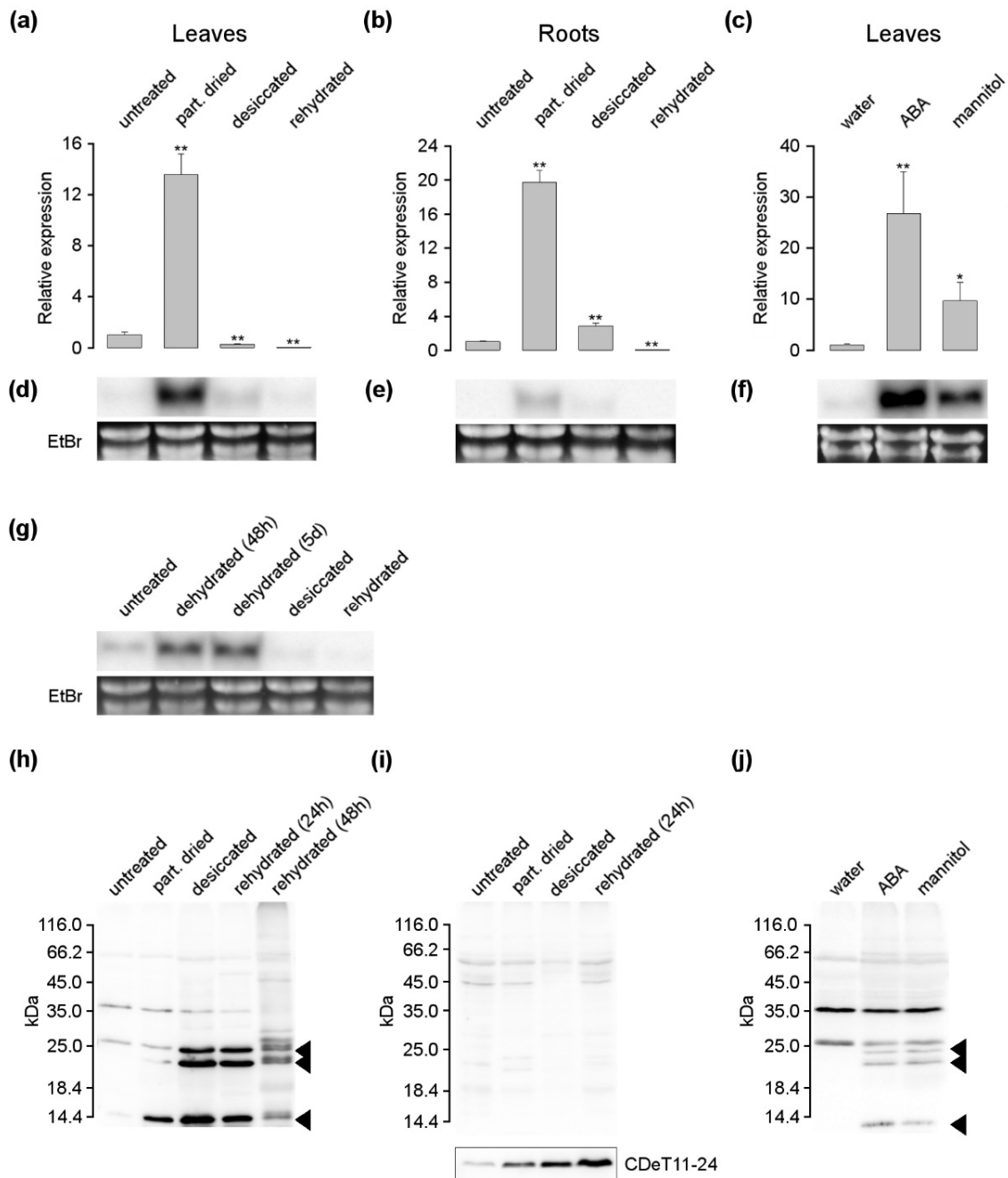
Fig. 4 CpEDR1 transcript and protein expression analyses. *CpEDR1* transcript abundance was analysed in untreated, partially dried (RWC 50 %), desiccated (RWC 2 %) and 24 h rehydrated leaf (a,b) and root (d,e) tissue samples or ABA, mannitol (osmotic stress) or water (control) incubated leaf samples (c,f). The *CpEDR1* transcript accumulation in absence of light was monitored using untreated, 48 h dehydrated (RWC 90 %), five days dehydrated (RWC 54 %), desiccated (RWC 2 %) and 24 h rehydrated leaf samples (g). Real-time quantitative PCR analysis of the *CpEDR1* expression is shown in a,b,c. cDNA was prepared from total RNA and amplified using *CpEDR1* specific primers. In a,b,c relative expression values and statistical significance were calculated as described in Fig. 2. RNA blot analysis (see Fig. 2 for details) is shown in d,e,f,g. CpEDR1 protein expression was analysed in untreated, partially dried (RWC 50 %), desiccated (RWC 2 %), 24 h rehydrated and 48 h rehydrated leaf samples (h) or untreated, partially dried, desiccated and 24 h rehydrated root samples (i) or ABA, mannitol (osmotic stress) or water (control) incubated leaf samples (j). Protein samples were analysed as described for CpCRP1 (Fig. 2). In (i) the protein levels of the dehydration-induced *CDeT11-24* gene (van den Dries *et al.*, 2011) were additionally analysed as controls for correct dehydration of roots. CpEDR1 protein bands are indicated by black triangles.

CpEDR1 is targeted to the chloroplast

To validate the predicted CpEDR1 subcellular localisation *in vivo*, *C. plantagineum* cells were biolistically transformed with constructs overexpressing either the full length CpEDR1-GFP or the signal peptide CpEDR1-GFP (CpEDR1-**SP**-GFP) translational fusion proteins. *C. plantagineum* cells expressing either CpEDR1-GFP or CpEDR1-**SP**-GFP showed that GFP and chloroplastic fluorescence overlapped thus confirming the localisation of CpEDR1 in the chloroplast (Fig. 5). A complementary experiment was performed using a CpEDR1-GFP translational fusion construct where the first 74 N-terminal amino acids of CpEDR1 were deleted (CpEDR1-**CTERM**-GFP). Transformed cells expressing CpEDR1-**CTERM**-GFP did not show GFP fluorescence in chloroplasts but in the cytoplasm (Fig. 5). This confirms the correct identification of the chloroplast signal peptide.

***CpCRP1* and *CpEDR1* orthologs are present in *L. brevidens* and *L. subracemosa*, close relatives of *C. plantagineum*.**

BLAST analysis failed to identify genes similar to *CpCRP1* and *CpEDR1* in public databanks. Recently, we obtained transcriptome data from two species very closely related to *C. plantagineum*, i.e., the desiccation tolerant *L. brevidens* (Lb) and the desiccation sensitive *L. subracemosa* (Ls) (unpublished data) and it was possible to identify putative orthologs of *CpCRP1* and *CpEDR1*. The overall amino acid similarity between CpCRP1 and LbCRP1 was 73 % and between CpCRP1 and LsCRP1 77 % (Fig. 6a). The C-terminal region of CRP1 proteins show the highest conservation. CRP1 proteins have the same number of cysteine residues and the position of these residues is well conserved among the proteins from the three species. CpEDR1 and LbEDR1 show 60 % similarity between each other whereas the similarity between CpEDR1 and LsEDR1 was 59 % (Fig. 6b). High amino acid conservation is observed in the region predicted to form a coiled-coil structure.

To understand the importance of *CRP1* and *EDR1* genes in desiccation tolerance, the expression of *CpCRP1* and *CpEDR1* was compared with the expression of the *L. brevidens* and *L. subracemosa* orthologue *CRP1* and *EDR1* genes. *CpCRP1*, *LbCRP1* and *LsCRP1* transcript levels decreased upon dehydration suggesting a common gene regulation. In both tolerant plants, the transcript accumulated during

rehydration. Both *CpEDR1* and *LbEDR1* accumulated early during dehydration and then transcript abundance decreased to levels similar to untreated leaves in desiccated leaves (2% RWC) (Fig. 6c). The *LsEDR1* transcript was also induced during dehydration but transcript levels are lower compared to levels observed in the two tolerant plants (Fig. 6c).
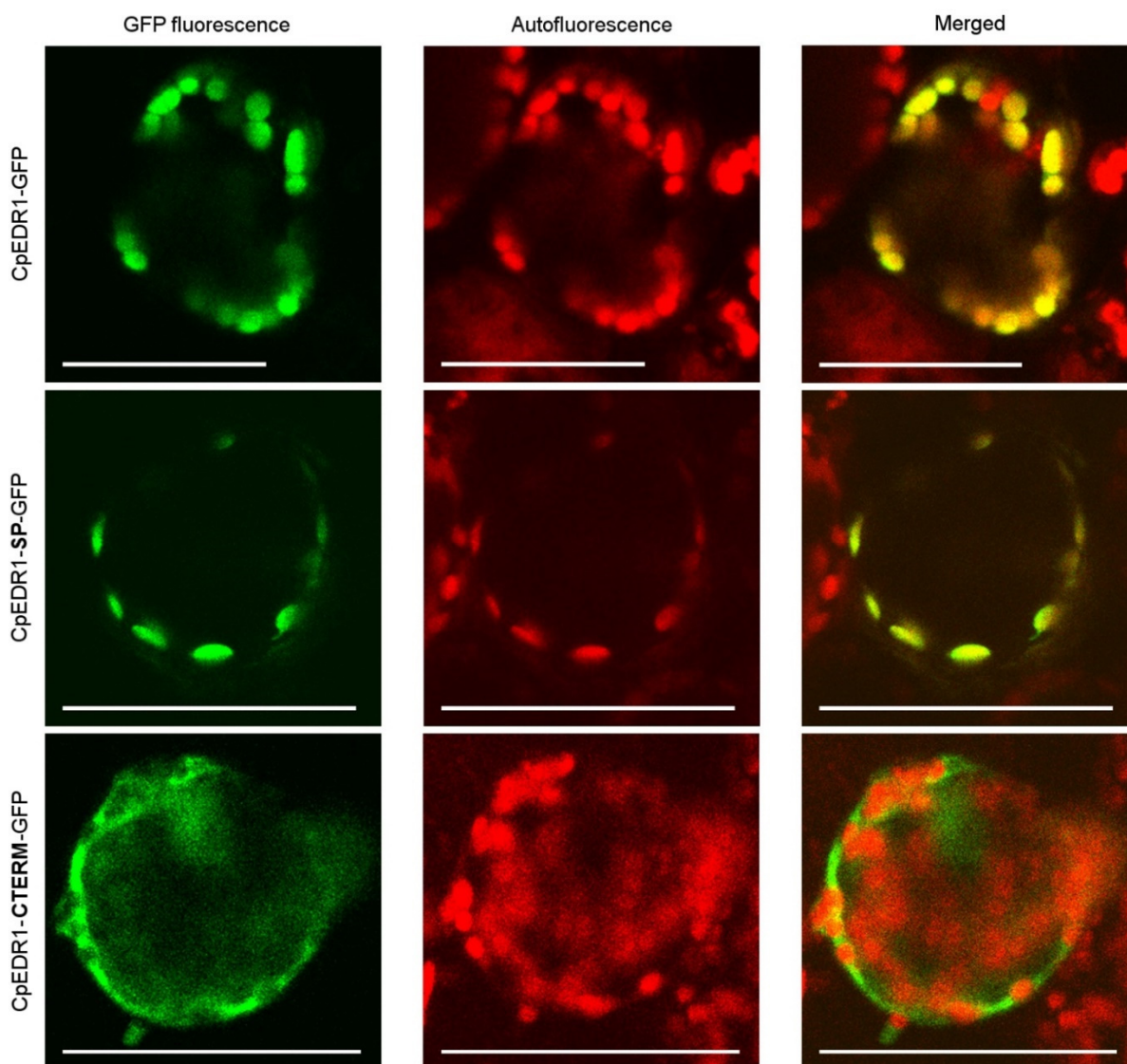
Fig. 5 Cellular localization of CpEDR1-GFP translational fusions. The *CpEDR1* coding sequence (CpEDR1-GFP), the predicted signal peptide (CpEDR1-**SP**-GFP) or the *CpEDR1* coding sequence without the signal peptide (CpEDR1-**CTERM**-GFP) were fused to GFP and transiently expressed in *C. plantagineum* cells *via* particle bombardment. GFP fluorescence only, chlorophyll autofluorescence only and merged images of GFP and chlorophyll fluorescence are shown. Images of transformed cells were taken after 16 h using a confocal microscope. The white bar represents 50 μm.
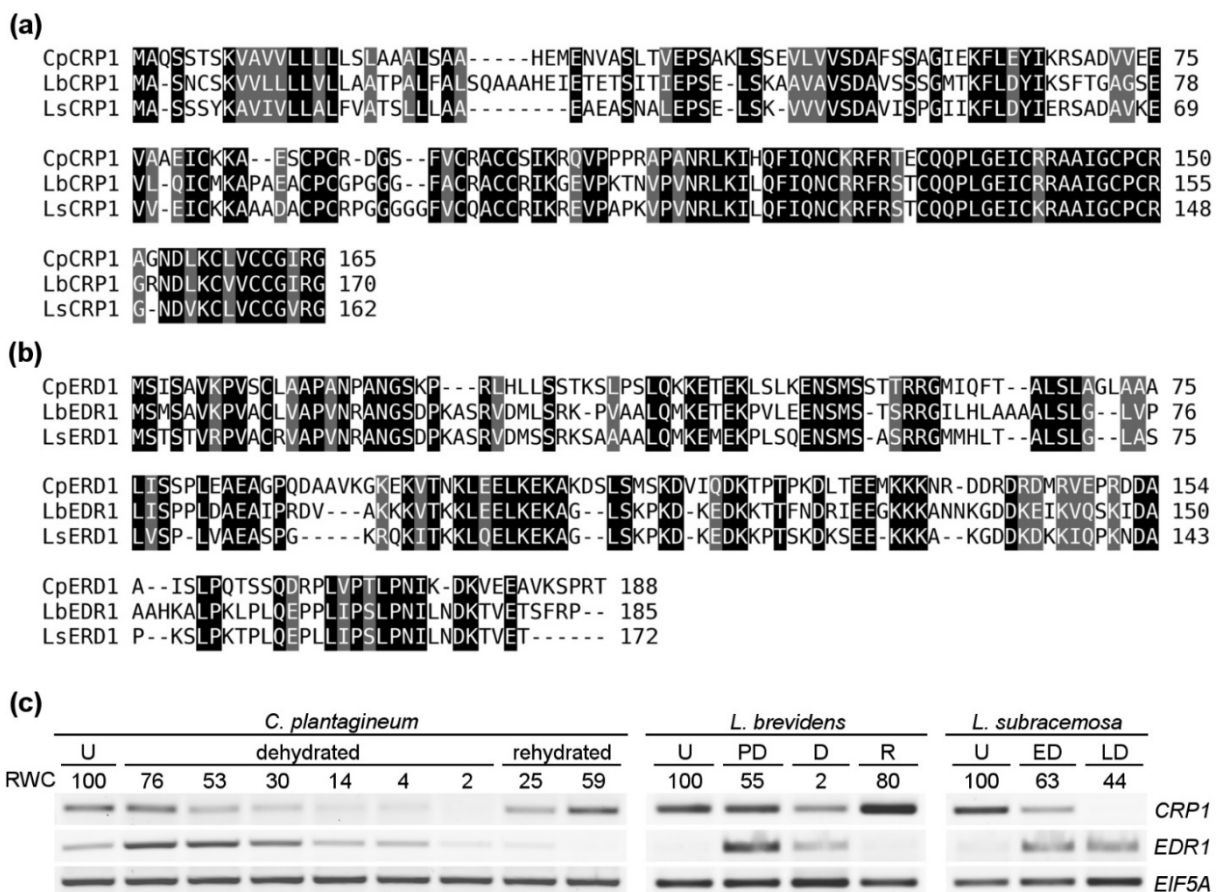
Fig. 6 Amino acid conservation and expression analysis of CpCRP1 and CpEDR1 putative orthologs in *L. brevidens* and *L. subracemosa*. Putative orthologs of CpCRP1 and CpEDR1 in *L. brevidens* and *L. subracemosa* were retrieved from our transcriptome databanks (unpublished data) using BLAST+ software (Camacho *et al.*, 2009) and aligned with CpCRP1 and CpEDR1 using ClustalX 2.1 (Larkin *et al.*, 2007). (a) Sequence alignment between *C. plantagineum* CpCRP1 and *L. brevidens* LbCRP1 and *L. subracemosa* LsCRP1 proteins. (b) Sequence alignment between *C. plantagineum* CpEDR1 and *L. brevidens* LbEDR1 and *L. subracemosa* LsEDR1 proteins. Identical (black) and conserved (grey) amino acids are indicated. (c) RT-PCR analysis showing the expression of *CRP1* and *EDR1* genes in *C. plantagineum*, *L. brevidens* and *L. subracemosa*. Gene expression was analysed in untreated (U), dehydrated (RWC 76, 53, 30, 14, 4, 2 %) and 12 h and 24 h rehydrated (RWC 25 and 59 %) *C. plantagineum* leaf tissue samples, untreated (U), partially dried (PD; RWC 55 %), desiccated (D; RWC 2 %) and 24 h rehydrated (R; RWC 80 %) *L. brevidens* leaf tissue samples and in untreated (U), early dehydrated (ED; RWC 63 %) and late dehydrated (LD; RWC 44 %) *L. subracemosa* leaf tissue samples. cDNA was prepared from total RNA and amplified using *CRP1* and *EDR1* specific primers. The expression of the *EIF5A* gene was analysed as a control for cDNA preparations.

Discussion

Transcriptome analysis and gene discovery in *C. plantagineum*.

The *C. plantagineum* transcriptome databank contains a large fraction of putative TRGs with unknown function (Rodriguez *et al.*, 2010). We used the expression data of these transcripts to identify TRGs important for the desiccation tolerance mechanisms. Among transcripts which showed high expression in dehydration or high covariation coefficients only eight were longer than 250 bp (Table 1). Some transcripts showed significant matches with other previously identified genes like late embryogenesis abundant or a glycine rich proteins when were analysed by BLAST excluding filtering for low complexity regions. This indicates that BLAST setting used for the computer-based annotation of genes by Rodriguez *et al.* (2010) could have failed to identify putative homologs in some cases. The *CpCRP1* (*06401*) and *CpEDR1* (*23731*) transcripts showed scarce similarities by BLAST analysis and transient regulation during the dehydration-rehydration cycle and thus are potentially interesting genes that are predicted to participate in desiccation tolerance mechanisms.

Functional characterisation of CpCRP1 and CpEDR1

CpCRP1 contains a predicted signal peptide for the endoplasmic reticulum (ER). Proteins which do not contain any additional sorting signal migrate non-specifically from the lumen of the ER towards the cell surface through the so-called default pathway (Denecke *et al.*, 1990). Antibodies raised against a recombinant CpCRP1 protein detected the corresponding protein only in protein samples obtained from cell wall fractions (Fig. 2g) thus supporting an extracellular localisation for CpCRP1. This localisation was further confirmed using CpCRP1-GFP translational fusion proteins (Fig. 3b). The *CpCRP1* transcript transiently accumulates in leaves and roots during rehydration (Fig. 2a,b,d,e). Conversely, the corresponding protein was found to be highly accumulated only in untreated leaves (Fig. 2g). This data point to a role for CpCRP1 in preparing/protecting leaf cell walls to dehydration stress. The high transcript accumulation after osmotic stress treatments suggests that the gene may be also functioning in the osmotic stress management in the cell wall (Fig. 2c,f). CpCRP1 proteins were not detected in total leaf (Fig. 2g) and root (data not shown) extracts

from different dehydration-rehydration stages and from osmotically stressed leaves (data not shown). This suggests that the presence of the CpCRP1 protein must be tightly linked to the cell wall. The analysis of the encoded sequence revealed a high abundance of cysteines which are known to form intra- and intermolecular disulphide bonds. Previous attempts to detect cysteine-rich cell wall proteins by Western blot from tomato (ToTLRP) or tobacco (NtTLRP) were negative (Domingo *et al.*, 1994; Domingo *et al.*, 1999). Although the NtTLRP protein was not detected in crude extracts, it was immunodetected in plant sections indicating that the protein was present. It was shown by translational fusions that the cysteine-rich domain of NtTLRP was affecting the detection of the highly soluble secreted protein PR1 and thus it was hypothesized that the protein remained cross-linked to the cell wall after secretion (Domingo *et al.*, 1999). Except for the presence of several cysteine residues, CpCRP1 does not share any sequence similarities with TLRPs but the specific occurrence of the CpCRP1 protein in cell wall protein extracts suggests a similar cysteine-mediated cross-linking to cell wall components (Fig. 2g).

CpEDR1 codes for a chloroplast-targeted protein (Fig. 5) which accumulates during dehydration and in response to ABA treatments in leaves. The dehydration specific accumulation of the transcript (Fig. 4a,d) and protein (Fig. 4h) in leaves suggests a role in the dehydration process in *C. plantagineum* leaf tissue. The transcript is upregulated very early during dehydration, when leaf relative water content is still 76 % (Fig. 6c) suggesting that early dehydration events affect transcript levels. Plants respond to dehydration by changes in the level of the stress hormone ABA. *C. plantagineum* showed a six- to seven-fold increase of ABA in desiccated leaves compared to untreated leaves (Bartels *et al.*, 1990). ABA treated leaves showed high expression of *CpEDR1* (Fig. 4c,f) indicating that ABA is involved in the upregulation of *CpEDR1*.

The CpEDR1 protein levels did not perfectly correlate with the transcript levels. The protein but not the transcript was still detected in completely desiccated and 24 h rehydrated leaf tissue (Fig. 4a,h). This indicates that the CpEDR1 protein is stable and persists until the plant has completely recovered from desiccation. Chloroplasts need more than eight hours of rehydration to restore their ultrastructure organisation (Schneider *et al.*, 1993). The observed partial recovery of the chloroplast structures after eight hours of rehydration together with the presence of the dehydration-induced CpEDR1 protein points to a role for this protein in rehydrating leaves. The predicted

molecular weight for the CpEDR1 protein was 13 kDa but protein bands of 22 and 24 kDa with similar expression patterns as the 13 kDa protein were co-detected using specific antiserum (Fig. 4h,j). These additional protein bands are likely protein dimers of CpEDR1 or dimers of CpEDR1 isoforms. According to transcriptome data, at least one additional dehydration-induced isoform (*27641*) is predicted to exist and ranked with the 8 longest uncharacterized genes reported in Table 1. The inability of SDS to completely disrupt non-covalent interactions between subunits of particular proteins was already reported, e.g., for the oligomeric SP1 protein from aspen (Wang *et al.*, 2002). The other protein bands which were detected (25 and 35 kDa) are likely unspecific.

Beyond a common localisation and induction by dehydration and ABA other aspects of the CpEDR1 protein structure and gene regulation are similar to previously identified dehydration-induced chloroplastic proteins. The analysis of the expression of *CpEDR1* transcript under dehydration and dark conditions showed that light is not essential but strongly increased gene accumulation suggesting a role in the regulation of *CpEDR1* gene expression (Fig. 4g). *In silico* analyses showed that the CpEDR1 protein sequence is mainly unstructured and only a short stretch is predicted to form a coiled coil structure (Fig. 1b). The mostly unstructured nature of CpEDR1 with only limited structural elements is a characteristic observed for most of the LEA proteins. Coiled coil structures were already predicted for CpPTPs (Phillips *et al.*, 2002) and for the cytosolic LEA-like CDeT11-24 protein (Röhrig *et al.*, 2006). In the CDeT11-24 protein one predicted coiled coil structure lies within the N-terminal lysine-rich sequence which was demonstrated to be required for phosphatidic acid binding and enzyme protection against desiccation (Petersen *et al.*, 2012). Intriguingly, the stretch of amino acids predicted to form a coiled coil structure in CpEDR1 is also rich in lysine (Fig. 1b) and showed a lysine distribution on a helical wheel projection (Fig. 1d) similar to the one observed for CDeT11-24 (Petersen *et al.*, 2012). Coiled coil domains typically consist of two or more amphiphatic α-helices which are wrapped together to bury their hydrophobic faces and form a supercoil. Coiled-coil-forming proteins are involved in many biological processes and play a structural role which requires protein-protein interaction (Lupas, 1996). Amphiphatic helices are also known to bind through their hydrophobic face to membranes and participate in membrane remodelling (Drin & Antonny, 2010). Taken together these results suggest that CpEDR1 plays a role in

the chloroplast during dehydration which requires molecular interactions with either coiled coil-forming proteins or membranes or both.

Taxonomically restricted genes may be important for desiccation tolerant mechanisms in *C. plantagineum*

The lack of any similarity between *CpCRP1* and *CpEDR1* and previously identified genes was confirmed using different public available resources and the transcriptome data from two other desiccation tolerant plants. The I-TASSER web server also did not predict any significant match with previously reported proteins and protein structures. The lack of any similarity together with the regulation during dehydration and rehydration suggests that *CpCRP1* and *CpEDR1* are new TRGs important for plant desiccation tolerance. *CpEDR1* shares features with some of the *C. plantagineum* dehydration-induced chloroplast genes (Bartels *et al.*, 2006). These genes are likely TRGs which may be specifically required to protect the photosynthetic apparatus during desiccation. *CpCRP1* likely evolved to protect the cell wall from damage associated with water loss. The only presence of *CpEDR1* and *CpCRP1* orthologs in two plants belonging to the same family as *C. plantagineum* suggests that these genes are restricted to the Linderniaceae and thus may have been recently evolved in this family. Comparative analysis of gene expression revealed that *EDR1* genes were upregulated in response to dehydration but gene expression amplitudes were reduced in the desiccation sensitive *L. subracemosa* (Fig. 6c). The lower accumulation of the *EDR1* gene in *L. subracemosa* may be linked to the lack of the desiccation tolerant trait in this plant.

Only the *CRP1* response to dehydration could be analysed in *L. subracemosa*. *CRP1* genes are expressed in untreated leaves in the three species and are similarly downregulated upon dehydration. According to the protein structure, CRP1 may be cross-linked to cell walls and exert a structural role which is related to the water status. The common expression pattern in tolerant and sensitive plants suggests that CRP1 may have an additional physiological function unrelated from desiccation tolerance.

Acknowledgements

Authors contributions

Giarola V. mainly conceived, designed and carried out the experiments reported in this study and wrote the manuscript. Krey S. prepared and tested the antibody against CpCRP1. Frerichs A. prepared the antibody against CpEDR1 and performed some CpEDR1 protein expression and localisation analyses. Bartels D. supervised the work, contributed to the discussion of the results and corrected the manuscript.

Supporting Data

Table S1 Primers used in this study

5' and 3'-end cloning	Sequence (5' to 3')	Restriction site
ZAP_F2	CAGGAAACAGCTATGAACC	-
CpCRP1_R	GCGATGAGTTTACCAGCACA	-
CpEDR1_R1	GGACGATCTTGCGATGAAGT	-
CpEDR1_R2	TCCTTTGCCTTCTCCTTGAG	-
Protein overexpression	**Sequence (5' to 3')**	**Restriction site**
CpCRP1_CDS_F	<u>CCATGG</u>CCCACGAAATGGAAAAC	*Nco*I
CpCRP1_CDS_R	<u>CTCGAG</u>GCCTCTTATACCGCAGCAAAC	*Xho*I
CpEDR1_CDS_F	<u>CCATGG</u>GGCTCGCCGCAG	*Nco*I
CpEDR1_CDS_R	<u>CTCGAG</u>AGGACTTTTTACTGCCT	*Xho*I
Protein localization	**Sequence (5' to 3')**	**Restriction site**
CpCRP1_GFP_F1	ACCCA<u>CCATGG</u>CTCAAAG	*Nco*I
CpCRP1_ GFP_F2	<u>CCATGG</u>CCCACGAAATGGAAAAC	*Nco*I
CpCRP1_GFP_R1	AGGGCATC<u>CCATGG</u>CG	*Nco*I
CpCRP1_GFP_R2	GTTTTCCATTTCGTGGG<u>CCATGG</u>	*Nco*I
CpEDR1_GFP_F1	CGGA<u>CCATGG</u>CAATATCCGCCGTC	*Nco*I
CpEDR1_GFP_F2	GCTCG<u>CCATGG</u>CTCTAATCTCATC	*Nco*I
CpEDR1_GFP_R1	CAAG<u>CCATGG</u>GACTTTTTACTGCCTC	*Nco*I
CpEDR1_GFP_R2	AGAG<u>CCATGG</u>CGAGCCCAGCTAGC	*Nco*I
RT-PCR	**Sequence (5' to 3')**	**Restriction site**
CpCRP1_RT_F	CAGCATAAAGCGCCAAGTTC	-
CpCRP1_RT_R	GCGATGAGTTTACCAGCACAG	-
LbCRP1_RT_F	AAGACCAATGTTCCCGTCAA	-
LbCRP1_RT_R	CTCGTTTTCAGTGGATGTGC	-
LsCRP1_RT_F	CAAGGTCCCCGTCAATAGG	-
LsCRP1_RT_R	TGGCATTATTATGTTATGGTGAGTG	-
CpEDR1_RT_F	TTGAACCGAGGGACGATG	-
CpEDR1_RT_R	GTTTAACGTTGCAAGGATTGTCTC	-
LbEDR1_RT_F	ACAAGGCGTTGCCTAAATTG	-
LbEDR1_RT_R	TTTCTCGATCACATCCCACA	-
LsEDR1_RT_F	AAGAACGATGCACCCAAATC	-
LsEDR1_RT_R	AGGATGTTCTCGATCCCAAC	-
CpEIF5A_RT_F	CATCAATCGTGAAGGAGAGGA	-
CpEIF5A_RT_R	CTAAACCTTGGCAGCAGCAC	-
LbEIF5A_RT_F	GGGAAAGGACCTTGTTGTGA	-
LbEIF5A_RT_R	TGGGCTCATTACTCCACTGA	-
LsEIF5A_RT_F	TGGGTGGTTTTTATCAATGTCC	-
LsEIF5A_RT_R	GGTTGAGTGATAACACAGTAGCAAA	-
RT-qPCR	**Sequence (5' to 3')**	**Restriction site**
CpCRP1_QRT_F	CGTCACGGTGAGAAGTCGT	-
CpCRP1_QRT_R	GCGATGAGTTTACCAGCACAG	-
CpEDR1_QRT_F	AGGCAGTAAAAAGTCCTCGGAC	-
CpEDR1_QRT_R	GTTTAACGTTGCAAGGATTGTCTC	-
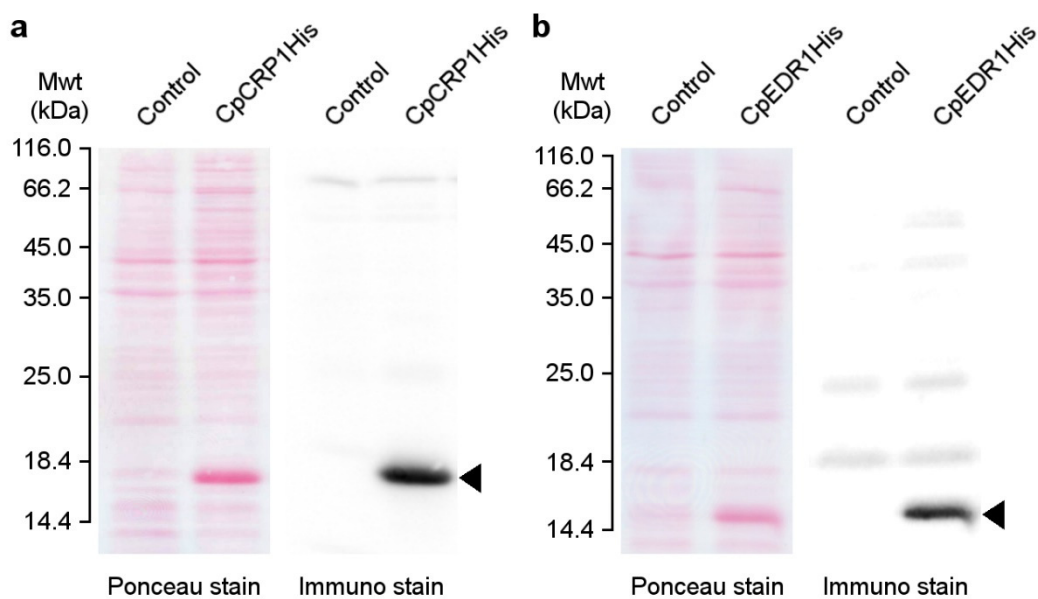
Fig. S1 Protein blot analysis showing the specificity of the antibodies used to detect CpCRP1 and CpEDR1 proteins. BL21 *E. coli* cells carrying the empty pET28a(+) vector (Control) or the CpCRP1 (CpCRP1His) or the CpEDR1 (CpEDR1His) overexpressing vectors were induced with 1 mM IPTG for 5 hours. Total proteins were separated by 15 % (w/v) SDS-PAGE, transferred onto a nitrocellulose membrane, stained with Ponceau S and polyclonal antibodies were used to detect the CpCRP1 and CpEDR1 proteins. CpCRP1 and CpEDR1 Ponceau-stained and immuno-stained membranes are shown in (a) and (b), respectively. The CpCRP1 and CpEDR1 proteins are indicated by black triangles.

CHAPTER 5

General Discussion

Resurrection plants show extreme adaptation to water scarcity (Gaff, 1971) and thus represent the best resource to study mechanisms evolved by plants to cope with water stress. Among resurrection plants, *C. plantagineum* has been extensively characterised at the molecular level and many genes linked to desiccation tolerance were reported (Bartels & Salamini, 2001). Major changes in vegetative tissues of this plant are similar to those occurring in reproductive plant tissues, i.e., seeds and pollen (Ingram & Bartels, 1996; Rodriguez *et al.*, 2010). Additionally, general responses like the accumulation of non-reducing sugars and LEA proteins are mechanisms which are shared also with desiccation sensitive plants subjected to water, cold or salt stress (Bartels & Sunkar, 2005). The identification and the analysis of the genes and gene isoforms involved in desiccation tolerance has been hampered by the lack of genome sequence information for *C. plantagineum*. Nowadays, high throughput RNA sequencing technologies, which do not require any previous knowledge of genome sequence, have been developed thus allowing the study of desiccation tolerance in resurrection plants on a genome-wide level.

Use of RT-qPCR and validated genes for gene expression studies in *C. plantagineum*

Presently, RT-qPCR is the best method to study gene expression. To achieve high sensitivity and accuracy with RT-qPCR, stable genes must be used as reference to remove technical variation between samples. It has been demonstrated that gene stability is linked to the particular organism, tissues and experimental conditions and thus it is not possible to directly use genes as references without their previous validation within particular experimental settings (Guénin *et al.*, 2009). Currently, the transketolase 3 (*TKT3*) isoform is the only gene with reported stability during dehydration in *C. plantagineum* leaves. In the past years, the lack of sequence information has precluded any systematic validation of reference genes in this plant.

The availability of *C. plantagineum* transcriptome data permitted to identify homologs of genes with demonstrated stability in other species and thus to test their stability (CHAPTER 2). In agreement with previous reports in other plants, different genes were found to be stable in leaves and roots during dehydration/desiccation and in the callus with desiccation and ABA treatments (Fig. 5 and Table 3 in CHAPTER 2). The reliability of the selected genes was provided by the comparison of expression levels of genes analysed by RT-qPCR and RNA-blot in CHAPTER 3 (Fig. 2) and 4 (Fig. 2 and Fig. 4). Validated genes now permit the use of RT-qPCR to analyse the expression of desiccation-induced gene in *C. plantagineum*. Particularly, the high specificity of this technique is required in those studies where accurate quantification of gene isoforms is needed. Recently, several genes encoding dehydrin isoforms were identified in the transcriptome of *C. plantagineum* (data not published) but the large sequence similarity of these genes limits the use of hybridisation-based techniques to study the expression of the single isoforms. Although, dehydrins are supposed to have essential roles in desiccation tolerance, a comprehensive analysis of these genes has never been performed in *C. plantagineum*. Currently, RT-qPCR is used to investigate differences in the expression of these isoforms in leaves, roots and callus in order to study the roles of the individual isoforms.

Isolation and characterisation of new desiccation-related genes in *C. plantagineum*

RNA-seq revealed the presence of several transcripts in *C. plantagineum* with scarce similarity to previously characterised genes (Rodriguez *et al*., 2010). Some of these transcripts are abundant in the partially dehydrated or desiccated transcriptome or show high expression variation (covariation) during a dehydration/rehydration cycle and thus are likely important so far unknown genes involved in desiccation tolerance. To restrict the number of putative interesting genes, only the most abundant and/or most covaried genes were chosen and among them three were further characterised (CHAPTER 3 and 4). Most of the identified transcripts were shorter than 250 bp (data not shown). The reduced size observed is likely due to incomplete assembly of sequencing reads of original transcripts which is mainly related to scarce sequence coverage. Indeed, one of the major drawbacks in RNA-seq analysis is due to the initial fragmentation of nucleic acids which are going to be sequenced. The two methods

which are currently used, i.e., cDNA- or RNA-based fragmentation, do not permit to get high representation of each part of the original transcripts but some parts, especially transcript ends, are generally less represented (Wang, Z *et al.*, 2009). The *C. plantagineum* transcriptome was obtained after cDNA-based fragmentation and this is known to bias the sequencing towards the transcript 3′-ends. This resulted in the loss of 5'-ends of transcripts for several genes. A required step in the characterisation of these new genes was the extension of transcript sequences towards the 5'-end. The use of lambda cDNA libraries prepared from *C. plantagineum* leaves as a template for the PCR based amplification of transcript ends has revealed to be a rapid and specific approach to overcome the RNA-seq limitations. Among top ranking sequences a glycine-rich class II protein (CpGRP1; CHAPTER 3), a Cys-rich rehydration responsive protein (CpCRP1; CHAPTER 4) and an early dehydration-responsive protein (CpEDR1; CHAPTER 4) were functionally characterised in this thesis. Among the three genes, *CpGRP1* was the only one which showed some similarities to already characterised genes but only when BLASTX analysis was performed without filtering the low complexity regions. This indicates that the BLAST setting used to annotate transcripts in the *C. plantagineum* transcriptome could have failed to identify putative homologs especially in the case of genes mainly possessing unusual composition, e.g., Gly-rich regions. The link of these genes to desiccation tolerance mechanisms in *C. plantagineum* was shown through analysis of transcript and protein levels during dehydration and rehydration. Two genes, i.e., CpGRP1 (Fig. 2 in CHAPTER 3) and CpEDR1 (Fig. 4 in CHAPTER 4), were upregulated in response to dehydration whereas CpCRP1 (Fig. 2 in CHAPTER 4) showed downregulation with dehydration and rapid re-induction upon rehydration.

It is believed that rehydration-induced repair mechanisms do not play a major role in vegetative desiccation tolerance in angiosperms (Cushman & Oliver, 2011). The rehydration process of resurrection plants has been scarcely investigated and most of the studies in these plants have focused on the characterisation of protection mechanisms in the "dehydration phase". However, several events occur leading to re-establish the original cellular organization, e.g., leaf unfolding and metabolic activities during rehydration. These mechanisms are part of the survival strategy of resurrection plants and must be at least in part regulated and coordinated to permit successful recovery from desiccation. The rehydration process in *C. plantagineum* is a relative long process requiring several hours to be accomplished (Bartels *et al.*, 1990).

Comparative analysis of *in vitro* translated mRNA samples from *C. plantagineum* suggested the specific accumulation of some genes in the rehydration phase but only two of such genes encoding a chlorophyll a/b binding protein and the transketolase 7 (TKT7) were eventually identified (Bernacchia *et al.*, 1996). This led to the conclusion that in *C. plantagineum* the events occurring during rehydration are not essential for desiccation tolerance but only contribute to restore the normal metabolic functions (Bernacchia *et al.*, 1996). However, the requirement of rehydration-induced mechanisms for the recovery from rapid drying was demonstrated in *Craterostigma wilmsii* thus suggesting the importance of such mechanisms to protect plant vitality when dehydration-induced protection is inadequate (Cooper & Farrant, 2002). In this thesis, it was shown that CpCRP1 is transiently upregulated upon rehydration (Fig. 2a,b,d,e in CHAPTER 4) thus suggesting its involvement in the rehydration process of *C. plantagineum*. This result represents the first characterisation of an unknown gene which participates in the so far largely hypothetical rehydration-specific mechanisms in *C. plantagineum*.

CpGRP1, CpCRP1 and CpEDR1 seem to play different functions in the plant. CpGRP1 and CpCRP1 are both apoplastic-localised proteins (Fig. 5 in CHAPTER 3 and Fig. 3 in CHAPTER 4, respectively). It is hypothesised that both proteins may take part in the mechanisms controlling the leaf folding which appear to be essential to counteract both mechanical and oxidative stress generated by the extensive water loss during dehydration. Although, other components such as expansin proteins and calcium ions have already been hypothesised to take part in this process (Vicré et al., 1999; Jones & McQueen-Mason, 2004; Vicré et al., 2004), underlying molecular mechanisms remain largely uncharacterised. In this thesis, it was reported that the CpCRP1 protein is tightly linked to cell wall components in untreated conditions (Fig. 2g in CHAPTER 4) and thus it appears to have a structural role in preparing or protecting the cell walls in the early stages of dehydration. The upregulation of the *CpCRP1* transcript (Fig. 2a,d in CHAPTER 4) and the absence of protein accumulation after 24 h of rehydration (Fig. 2g in CHAPTER 4) suggest that CpCRP1 may be required late during rehydration in leaves. On the other hand, CpGRP1 may play a role in controlling the cell turgor during dehydration, through the interaction with cell wall-associated protein kinases (WAKs) (Fig. 10 in CHAPTER 3). Interaction between CpGRP1 and CpWAK1 was demonstrated in yeast thus supporting the proposed role (Fig. 6 in CHAPTER 3). *In planta* interaction of CpGRP1-CpWAK1 by bimolecular

fluorescence complementation (BiFC) assays need to be performed to support the results obtained by yeast two-hybrid analysis. Pectins are known ligands for WAKs (Decreux & Messiaen, 2005). Dehydration-induced modifications of pectin structures may also be involved in the proposed mechanism. It would be interesting to establish whether CpGRP1 can directly bind to pectins and determine structural changes or indirectly modify pectin structure through the interaction with other cell wall modifying enzymes. Currently, yeast two-hybrid screenings and pectin binding assays are being performed to identify putative interactors of CpGRP1 and to determine CpGRP1 pectin-binding activity, respectively.

CpEDR1 accumulates in chloroplasts of desiccated *C. plantagineum* leaves (Fig. 4h and 5 in CHAPTER 4) and together with other dehydration-induced chloroplastic proteins it is supposed to protect these organelles upon dehydration. Accumulation of dehydration-inducible proteins seem to be one of the main strategies evolved in the homoiochlorophyllous plant *C. plantagineum* to protect from oxidative stress while maintaining the photosynthetic structures intact. This gives the advantage to *C. plantagineum* to recover and rapidly restart photosynthesis after rehydration with respect to poikilochlorophyllous desiccation tolerance plants which prior need to resynthesize chlorophyll and thylakoids. How CpEDR1 protect the chloroplast remains to be established but the involvement of protein-protein and membrane-protein interactions is hypothesised. Lipid binding assays may provide some evidence on the ability of CpEDR1 to bind to membranes.

Overall, the results presented in CHAPTER 3 and 4 provide evidence of the existence of uncharacterised genes which may play important roles in the complex mechanisms conferring desiccation tolerance to *C. plantagineum*. The study also highlights the utility of using RNA-seq data to identify such genes.

Importance of comparative studies

It was proposed that desiccation tolerance in *C. plantagineum* may be at least in part a quantitative characteristic and reside in the expression pattern of common plant genes (Bartels & Salamini, 2001). To validate this hypothesis, comparative expression studies of dehydration-induced genes in *C. plantagineum* and in the close relative desiccation tolerant *L. brevidens* and desiccation sensitive *L. subracemosa* were initiated. Evidence of quantitative differences was provided by the analysis of the

expression of the dehydration- and ABA-induced *LEA-like 11-24* genes from these Linderniaceae species (van den Dries *et al.*, 2011). It was demonstrated that the reduced expression of the *11-24* gene was correlated with changes in promoter *cis-*elements and to the desiccation sensitive phenotype. Further support to the "quantitative" hypothesys may be provided by the use of RT-qPCR for the analysis of expression levels of dehydration-responsive genes. Recently, transcriptomes from *L. brevidens* and *L. subracemosa* were obtained and used to identify homologs of *C. plantagineum* dehydration-induced genes (data not published). Preliminary RT-PCR experiments indicated that large differences in the expression of these genes exist between tolerant and sensitive species (data not published). Currently, RT-qPCR is used to precisely quantify transcript levels of selected dehydration-induced genes in *C. plantagineum*, *L. brevidens* and *L. subracemosa* in order to confirm previous observations and provide support to the "quantitative" hypothesis.

Analysis of the *CpEDR1* and *CpCRP1* sequences by several publically available databank services failed to identify similar genes in other species. Homologs of *CpEDR1* and *CpCRP1* were identified in the transcriptome data from *L. brevidens* and *L. subracemosa* (data not published). The exclusive occurrence of *EDR1* and *CRP1* genes in three members of the Linderniaceae family suggests that these genes may be restricted to this taxon and may have evolved after this taxon group was formed. Taxonomically restricted genes (TRGs) are known to participate together with the variation of regulatory networks in the physiological adaptation of organisms to different environmental conditions (Khalturin *et al.*, 2009) but the importance of these genes for desiccation tolerance in *C. plantagineum* needs to be evaluated. Further characterisation of CpCRP1, CpEDR1 and other TRG homologs in *L. brevidens* and *L. subracemosa* may help to understand whether, beside quantitative expression differences, TRGs are important for the extreme tolerance to dehydration observed in *C. plantagineum*.

To summarise, the work of this thesis contributed to our better understanding of three important aspects which are essential for the desiccation tolerance phenomenon in *C. plantagineum*, i.e., leaf folding, chloroplast protection and rehydration specific responses. In this thesis, molecular components involved in the cell wall folding were isolated and characterised. The identified genes, i.e. *CpGRP1* and *CpWAKs* are together with expansin (Jones & McQueen-Mason, 2004), the only genes reported so

far in *C. plantagineum*. Additionally, a new gene, *CpEDR1*, was found to be taxonomically restricted to the Linderniaceae and linked to protection mechanisms in chloroplasts thus suggesting the existence of specific protection genes exclusively evolved in this plant taxon. In this work for the first time, an unknown gene, *CpCRP1*, was found to be associated with the rehydration process suggesting that rehydration-specific mechanisms may also occur in resurrection plants. Finally yet importantly, reference genes were identified thus allowing the accurate quantification of gene expression in *C. plantagineum* and the future comparison of expression levels between tolerant and non-tolerant species. As indicated by the results of this thesis, desiccation tolerance is a complex trait. It is known that water scarcity has a huge economic impact on productivity of crops and due to the predicted global environmental changes in the forthcoming years, the breeding of crops able to maintain high yields with less water has become a critical requirement. Thus, the characterisation of more and more genes involved in different aspects of desiccation tolerance in resurrection plants may contribute in the next future to improve the yield of crops under drought conditions.

REFERENCES

Abba S, Ghignone S, Bonfante P. 2006. A dehydration-inducible gene in the truffle *Tuber borchii* identifies a novel group of dehydrins. *BMC Genomics* **7**(1): 1-15.

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2003. *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *The Plant Cell* **15**(1): 63-78.

Alamillo JM, Bartels D. 1996. Light and stage of development influence the expression of desiccation-induced genes in the resurrection plant *Craterostigma plantagineum*. *Plant, Cell & Environment* **19**(3): 300-310.

Albach DC, Meudt HM, Oxelman B. 2005. Piecing together the "new" Plantaginaceae. *American Journal of Botany* **92**(2): 297-315.

Alpert P. 2005. The limits and frontiers of desiccation-tolerant life. *Integrative and Comparative Biology* **45**(5): 685-695.

Alpert P. 2006. Constraints of tolerance: why are desiccation-tolerant organisms so small or rare? *Journal of Experimental Biology* **209**(9): 1575-1584.

Andersen CL, Jensen JL, Orntoft TF. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research* **64**(15): 5245-5250.

Anderson CM, Wagner TA, Perret M, He ZH, He D, Kohorn BD. 2001. WAKs: cell wall-associated kinases linking the cytoplasm to the extracellular matrix. *Plant Molecular Biology* **47**(1-2): 197-206.

Artico S, Nardeli S, Brilhante O, Grossi-de-Sa M, Alves-Ferreira M. 2010. Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data. *BMC Plant Biology* **10**(1): 1-12.

Artus NN, Uemura M, Steponkus PL, Gilmour SJ, Lin C, Thomashow MF. 1996. Constitutive expression of the cold-regulated *Arabidopsis thaliana COR15a* gene affects both chloroplast and protoplast freezing tolerance. *Proceedings of the National Academy of Sciences of the United States of America* **93**(23): 13404-13409.

Barsalobres-Cavallari CF, Severino FE, Maluf MP, Maia IG. 2009. Identification of suitable internal control genes for expression studies in *Coffea arabica* under different experimental conditions. *BMC Molecular Biology* **10**(1): 1.

Bartels D. 2005. Desiccation tolerance studied in the resurrection plant *Craterostigma plantagineum*. *Integrative and Comparative Biology* **45**(5): 696-701.

Bartels D, Ditzer A, Furini A. 2006. What can we learn from resurrection plants? In: Ribaut J ed. *Drought adaptation in cereals*: Hayworth Publisher, 599-622.

Bartels D, Hanke C, Schneider K, Michel D, Salamini F. 1992. A desiccation-related *Elip*-like gene from the resurrection plant *Craterostigma plantagineum* is regulated by light and ABA. *The EMBO Journal* **11**(8): 2771-2778.

Bartels D, Salamini F. 2001. Desiccation tolerance in the resurrection plant *Craterostigma plantagineum*. A contribution to the study of drought tolerance at the molecular level. *Plant Physiology* **127**(4): 1346-1353.

Bartels D, Schneider K, Terstappen G, Piatkowski D, Salamini F. 1990. Molecular cloning of abscisic acid-modulated genes which are induced during desiccation of the resurrection plant *Craterostigma plantagineum*. *Planta* **181**(1): 27-34.

Bartels D, Sunkar R. 2005. Drought and salt tolerance in plants. *Critical Reviews in Plant Sciences* **24**(1): 23-58.

Battaglia M, Olvera-Carrillo Y, Garciarrubio A, Campos F, Covarrubias AA. 2008. The enigmatic LEA proteins and other hydrophilins. *Plant Physiology* **148**(1): 6-24.

Battista JR, Park MJ, McLemore AE. 2001. Inactivation of two homologues of proteins presumed to be involved in the desiccation tolerance of plants sensitizes *Deinococcus radiodurans* R1 to desiccation. *Cryobiology* **43**(2): 133-139.

Benson DA, Karsch-Mizrachi I, Clark K, Lipman DJ, Ostell J, Sayers EW. 2012. GenBank. *Nucleic Acids Research* **40**(D1): D48-D53.

Bernacchia G, Salamini F, Bartels D. 1996. Molecular characterization of the rehydration process in the resurrection plant *Craterostigma plantagineum*. *Plant Physiology* **111**(4): 1043-1050.

Bernacchia G, Schwall G, Lottspeich F, Salamini F, Bartels D. 1995. The transketolase gene family of the resurrection plant *Craterostigma plantagineum*: differential expression during the rehydration phase. *The EMBO Journal* **14**(3): 610-618.

Bewley JD. 1979. Physiological aspects of desiccation tolerance. *Annual Review of Plant Physiology and Plant Molecular Biology* **30**: 195-238.

Bewley JD, Krochko JE. 1982. Desiccation-tolerance. In: Lange OL, Nobel PS, Osmond CB, Ziegler H eds. *Physiological plant ecology II*: Springer Berlin Heidelberg, 325-378.

Bewley JD, Reynolds TL, Oliver MJ. 1993. Evolving strategies in the adaptation to desiccation. In: Close TJ, Bray EA eds. *Plant responses to cellular dehydration during environmental stress*. Rockville, MD: American Society of Plant Physiologists, 193-201.

Bianchi G, Gamba A, Murelli C, Salamini F, Bartels D. 1991. Novel carbohydrate metabolism in the resurrection plant *Craterostigma plantagineum*. *The Plant Journal* **1**(3): 355-359.

Bloom J, Khan Z, Kruglyak L, Singh M, Caudy A. 2009. Measuring differential gene expression by short read sequencing: quantitative comparison to 2-channel gene expression microarrays. *BMC Genomics* **10**(1): 221.

Bockel C, Salamini F, Bartels D. 1998. Isolation and characterization of genes expressed during early events of the dehydration process in the resurrection plant *Craterostigma plantagineum*. *Journal of Plant Physiology* **152**(2-3): 158-166.

Borovskii GB, Stupnikova IV, Antipina AI, Vladimirova SV, Voinikov VK. 2002. Accumulation of dehydrin-like proteins in the mitochondria of cereals in response to cold, freezing, drought and ABA treatment. *BMC Plant Biology* **2**(1): 5.

Boucher V, Buitink J, Lin X, Boudet J, Hoekstra FA, Hundertmark M, Renard D, Leprince O. 2010. MtPM25 is an atypical hydrophobic late embryogenesis-abundant protein that dissociates cold and desiccation-aggregated proteins. *Plant, Cell & Environment* **33**(3): 418-430.

Bravo LA, Gallardo J, Navarrete A, Olave N, Martinez J, Alberdi M, Close TJ, Corcuera LJ. 2003. Cryoprotective activity of a cold-induced dehydrin purified from barley. *Physiologia Plantarum* **118**(2): 262-269.

Browne J, Tunnacliffe A, Burnell A. 2002. Anhydrobiosis: plant desiccation gene found in a nematode. *Nature* **416**(6876): 38-38.

Browne JA, Dolan KM, Tyson T, Goyal K, Tunnacliffe A, Burnell AM. 2004. Dehydration-specific induction of hydrophilic protein genes in the anhydrobiotic nematode *Aphelenchus avenae*. *Eukaryotic Cell* **3**(4): 966-975.

Brunner AM, Yakovlev IA, Strauss SH. 2004. Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biology* **4**(1): 14.

Brutus A, Sicilia F, Macone A, Cervone F, De Lorenzo G. 2010. A domain swap approach reveals a role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. *Proceedings of the National Academy of Sciences of the United States of America* **107**(20): 9452-9457.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL*, et al.* 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* **55**(4): 611-622.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden T. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**(1): 421.

Campbell SA, Close TJ. 1997. Dehydrins: genes, proteins, and associations with phenotypic traits. *New Phytologist* **137**(1): 61-74.

Chakrabortee S, Boschetti C, Walton LJ, Sarkar S, Rubinsztein DC, Tunnacliffe A. 2007. Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. *Proceedings of the National Academy of Sciences of the United States of America* **104**(46): 18073-18078.

Chakrabortee S, Tripathi R, Watson M, Kaminski Schierle GS, Kurniawan DP, Kaminski CF, Wise MJ, Tunnacliffe A. 2012. Intrinsically disordered proteins as molecular shields. *Molecular BioSystems* **8**(1): 210-219.

Chen D, Pan X, Xiao P, Farwell MA, Zhang B. 2011. Evaluation and identification of reliable reference genes for pharmacogenomics, toxicogenomics, and small RNA expression analysis. *Journal of Cellular Physiology* **226**(10): 2469-2477.

Clark MS, Thorne MAS, Purac J, Grubor-Lajsic G, Kube M, Reinhardt R, Worland MR. 2007. Surviving extreme polar winters by desiccation: clues from Arctic springtail (*Onychiurus arcticus*) EST libraries. *BMC Genomics* **8**: 12.

Clegg JS, Seitz P, Seitz W, Hazlewood CF. 1982. Cellular responses to extreme water loss: the water-replacement hypothesis. *Cryobiology* **19**(3): 306-316.

Close T, Lammers P. 1993. An osmotic stress protein of cyanobacteria is immunologically related to plant dehydrins. *Plant Physiology* **101**: 773 - 779.

Close TJ, Fenton RD, Yang A, Asghar R, DeMason DA, Crione D, Meyer NC, Moonan F. 1993. Dehydrin: the protein. In: Close TJ, Bray EA eds. *Plant responses to cellular dehydration during environmental stress*. Rockville, MD: American Society of Plant Physiologists, 104-118.

Close TJ, Kortt AA, Chandler PM. 1989. A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley and corn. *Plant Molecular Biology* **13**(1): 95-108.

Colmenero-Flores J, Moreno L, Smith C, Covarrubias A. 1999. *Pvlea-18*, a member of a new late-embryogenesis-abundant protein family that accumulates during water stress and in the growing region of well-irrigated bean seedlings. *Plant Physiology* **120**: 93 - 103.

Condit CM, Meagher RB. 1986. A gene encoding a novel glycine-rich structural protein of petunia. *Nature* **323**(6084): 178-181.

Cooper K, Farrant JM. 2002. Recovery of the resurrection plant *Craterostigma wilmsii* from desiccation: protection versus repair. *Journal of Experimental Botany* **53**(375): 1805-1813.

Crowe JH, Carpenter JF, Crowe LM. 1998. The role of vitrification in anhydrobiosis. *Annual Review of Physiology* **60**(1): 73-103.

Cullum R, Alder O, Hoodless PA. 2011. The next generation: using new sequencing technologies to analyse gene regulation. *Respirology* **16**(2): 210-222.

Cuming AC. 1999. LEA proteins. In: Shewry PR, Casey R eds. *Seed proteins*. Dordrecht, NL: Springer Netherlands, 753-780.

Cushman JC, Oliver MJ. 2011. Understanding vegetative desiccation tolerance using integrated functional genomics approaches within a comparative evolutionary framework. In: Lüttge U, Beck E, Bartels D eds. *Plant desiccation tolerance*: Springer Berlin Heidelberg, 307–338.

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR. 2005. Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis. Plant Physiology* **139**(1): 5-17.

Dang NX, Popova AV, Hundertmark M, Hincha DK. 2014. Functional characterization of selected LEA proteins from *Arabidopsis thaliana* in yeast and in vitro. *Planta*: 1-12.

Danyluk J, Perron A, Houde M, Limin A, Fowler B, Benhamou N, Sarhan F. 1998. Accumulation of an acidic dehydrin in the vicinity of the plasma membrane during cold acclimation of wheat. *Plant Cell* **10**(4): 623-638.

de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N. 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Research* **34**(Web Server issue): W362-365.

de Oliveira DE, Seurinck J, Inzé D, Van Montagu M, Botterman J. 1990. Differential expression of five *Arabidopsis* genes encoding glycine-rich proteins. *The Plant Cell* **2**(5): 427-436.

Decreux A, Messiaen J. 2005. Wall-associated kinase WAK1 interacts with cell wall pectins in a calcium-induced conformation. *Plant Cell Physiology* **46**(2): 268-278.

Decreux A, Thomas A, Spies B, Brasseur R, Van Cutsem P, Messiaen J. 2006. In vitro characterization of the homogalacturonan-binding domain of the wall-associated kinase WAK1 using site-directed mutagenesis. *Phytochemistry* **67**(11): 1068-1079.

Denecke J, Botterman J, Deblaere R. 1990. Protein secretion in plant cells can occur via a default pathway. *The Plant Cell Online* **2**(1): 51-59.

Denekamp NY, Reinhardt R, Kube M, Lubzens E. 2010. Late embryogenesis abundant (LEA) proteins in nondesiccated, encysted, and diapausing embryos of rotifers. *Biology of Reproduction* **82**(4): 714-724.

Denekamp NY, Thorne MAS, Clark MS, Kube M, Reinhardt R, Lubzens E. 2009. Discovering genes associated with dormancy in the monogonont rotifer *Brachionus plicatilis*. *BMC Genomics* **10**: 17.

Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A. 2004. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *BioTechniques* **37**(1): 112-119.

Dinakar C, Bartels D. 2012. Light response, oxidative stress management and nucleic acid stability in closely related Linderniaceae species differing in desiccation tolerance. *Planta* **236**(2): 541-555.

Dinakar C, Bartels D. 2013. Desiccation tolerance in resurrection plants: new insights from transcriptome, proteome and metabolome analysis. *Frontiers in Plant Science* **4**: 482.

Ditzer A, Bartels D. 2006. Identification of a dehydration and ABA-responsive promoter regulon and isolation of corresponding DNA binding proteins for the group 4 *LEA* gene *CpC2* from *C. plantagineum*. *Plant Molecular Biology* **61**(4-5): 643-663.

Domingo C, Gomez MD, Canas L, Hernandez-Yago J, Conejero V, Vera P. 1994. A novel extracellular matrix protein from tomato associated with lignified secondary cell walls. *Plant Cell* **6**(8): 1035-1047.

Domingo C, Saurí A, Mansilla E, Conejero V, Vera P. 1999. Identification of a novel peptide motif that mediates cross-linking of proteins to cell walls. *The Plant Journal* **20**(5): 563-570.

Drin G, Antonny B. 2010. Amphipathic helices and membrane curvature. *FEBS Letters* **584**(9): 1840-1847.

Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2003. *OsDREB* genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *The Plant Journal* **33**(4): 751-763.

Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW*, et al.* 2001. Intrinsically disordered protein. *Journal of Molecular Graphics and Modelling* **19**(1): 26-59.

Dure L, III, Crouch M, Harada J, Ho THD, Mundy J, Quatrano R, Thomas T, Sung ZR. 1989. Common amino acid sequence domains among the LEA proteins of higher plants. *Plant Molecular Biology* **12**(5): 475-486.

Dure L, III, Greenway S, Galau G. 1981. Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by in vitro and in vivo protein synthesis. *Biochemistry* **20**: 4162 - 4168.

Dure L, III. 1993. Structural motifs in LEA proteins. In: Close TJ, Bray EA eds. *Plant responses to cellular dehydration during environmental stress*. Rockville, MD: American Society of Plant Physiologists, 91–103.

Dure L, III. 2001. Occurrence of a repeating 11-mer amino acid sequence motif in diverse organisms. *Protein and Peptide Letters* **8**(2): 115-122.

Eichinger L, Pachebat JA, Glockner G, Rajandream MA, Sucgang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q*, et al.* 2005. The genome of the social amoeba *Dictyostelium discoideum. Nature* **435**(7038): 43-57.

Emanuelsson O, Nielsen H, Heijne GV. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Science* **8**(5): 978-984.

Exposito-Rodriguez M, Borges A, Borges-Perez A, Perez J. 2008. Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biology* **8**(1): 131.

Farrant JM. 2000. A comparison of mechanisms of desiccation tolerance among three angiosperm resurrection plant species. *Plant Ecology* **151**(1): 29-39.

Farrant JM, Brandt W, Lindsey GG. 2007. An overview of mechanisms of desiccation tolerance in selected angiosperm resurrection plants. *Plant Stress* **1**(1): 72-84.

Farrant JM, Cooper K, Kruger LA, Sherwin HW. 1999. The effect of drying rate on the survival of three desiccation-tolerant angiosperm species. *Annals of Botany* **84**(3): 371-379.

Farrant JM, Vander Willigen C, Loffell DA, Bartsch S, Whittaker A. 2003. An investigation into the role of light during desiccation of three angiosperm resurrection plants. *Plant Cell & Environment* **26**(8): 1275-1286.

Feiz L, Irshad M, Pont-Lezica RF, Canut H, Jamet E. 2006. Evaluation of cell wall preparations for proteomics: a new procedure for purifying cell walls from *Arabidopsis* hypocotyls. *Plant Methods* **2**(1): 10.

Fischer E. 1992. Systematik der afrikanischen Lindernieae (Scrophulariaceae). *Tropische und subtropische Pflanzenwelt* **81**: 1-365.

Fischer E. 1995. Revision of the Lindernieae (Scrophulariaceae) in Madagascar. 1. The genera *Lindernia Allioni* and *Crepidorhopalon* E. Fischer. *Bulletin du Muséum National d'Histoire Naturelle, section B, Adansonia: Botanique Phytochemie ser. 4,* **17**: 227–257.

Fischer E, Schäferhoff B, Müller K. 2013. The phylogeny of Linderniaceae – The new genus *Linderniella*, and new combinations within *Bonnaya*, *Craterostigma*, *Lindernia*, *Micranthemum*, *Torenia* and *Vandellia*. *Willdenowia - Annals of the Botanic Garden and Botanical Museum Berlin-Dahlem* **43**(2): 209-238.

Forster F, Liang CG, Shkumatov A, Beisser D, Engelmann JC, Schnolzer M, Frohme M, Muller T, Schill RO, Dandekar T. 2009. Tardigrade workbench: comparing stress-related proteins, sequence-similar and functional protein clusters as well as RNA elements in tardigrades. *BMC Genomics* **10**: 10.

Frank W, Phillips J, Salamini F, Bartels D. 1998. Two dehydration-inducible transcripts from the resurrection plant *Craterostigma plantagineum* encode interacting homeodomain-leucine zipper proteins. *The Plant Journal* **15**(3): 413-421.

Furini A, Koncz C, Salamini F, Bartels D. 1997. High level transcription of a member of a repeated gene family confers dehydration tolerance to callus tissue of *Craterostigma plantagineum. The EMBO Journal* **16**(12): 3599-3608.

Fusaro A, Mangeon A, Magrani Junqueira R, Andréa Benício Rocha C, Cardoso Coutinho T, Margis R, Sachetto-Martins G. 2001. Classification, expression pattern and comparative analysis of sugarcane expressed sequences tags (ESTs) encoding glycine-rich proteins (GRPs). *Genetics and Molecular Biology* **24**.

Gaff DF. 1971. Desiccation-tolerant flowering plants in southern Africa. *Science* **174**(4013): 1033-1034.

Gaff DF. 1989. Responses of desiccation-tolerant "resurrection" plants to water stress. In: Kreeb KH, Richter H, Hinckley TM eds. *Structural and functional responses to environmental stresses: water shortages*. The Hague, NL: SPB Academic, 264-311.

Gal TZ, Glazer I, Koltai H. 2004. An LEA group 3 family member is involved in survival of *C. elegans* during exposure to stress. *FEBS Letters* **577**(1-2): 21-26.

Garay-Arroyo A, Colmenero-Flores JM, Garciarrubio A, Covarrubias AA. 2000. Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit. *Journal of Biological Chemistry* **275**(8): 5668-5674.

Gasulla F, vom Dorp K, Dombrink I, Zähringer U, Gisch N, Dörmann P, Bartels D. 2013. The role of lipid metabolism in the acquisition of desiccation tolerance in *Craterostigma plantagineum*: a comparative approach. *The Plant Journal* **75**(5): 726-741.

Gautier R, Douguet D, Antonny B, Drin G. 2008. HELIQUEST: a web server to screen sequences with specific α-helical properties. *Bioinformatics* **24**(18): 2101-2102.

Gechev T, Benina M, Obata T, Tohge T, Sujeeth N, Minkov I, Hille J, Temanni M-R, Marriott A, Bergström E, *et al.* 2013. Molecular mechanisms of desiccation tolerance in the resurrection glacial relic *Haberlea rhodopensis*. *Cellular and Molecular Life Sciences* **70**(4): 689-709.

Gietz DR, Woods RA, Christine G, Gerald RF. 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. In: Guthrie C, Fink GR eds. *Methods in Enzymology*. New York, USA: Academic Press, 87-96.

Giraudat J, Parcy F, Bertauche N, Gosti F, Leung J, Morris P-C, Bouvier-Durand M, Vartanian N. 1994. Current advances in abscisic acid action and signalling. *Plant Molecular Biology* **26**(5): 1557-1577.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, *et al.* 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* **40**(D1): D1178-D1186.

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

Guénin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L. 2009. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *Journal of Experimental Botany* **60**(2): 487-493.

Gutierrez L, Mauriat M, Guénin S, Pelloux J, Lefebvre JF, Louvet R, Rusterucci C, Moritz T, Guerineau F, Bellini C*, et al.* 2008. The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription-polymerase chain reaction (RT-PCR) analysis in plants. *Plant Biotechnology Journal* **6**(6): 609-618.

Hand SC, Menze MA, Toner M, Boswell L, Moore D. 2011. LEA proteins during water stress: not just for plants anymore. *Annual Review of Physiology* **73**(1): 115-134.

Hara M, Terashima S, Kuboi T. 2001. Characterization and cryoprotective activity of cold-responsive dehydrin from *Citrus unshiu*. *Journal of Plant Physiology* **158**(10): 1333-1339.

He ZH, Cheeseman I, He D, Kohorn BD. 1999. A cluster of five cell wall-associated receptor kinase genes, *Wak1-5*, are expressed in specific organs of *Arabidopsis*. *Plant Molecular Biology* **39**(6): 1189-1196.

He ZH, Fujiki M, Kohorn BD. 1996. A cell wall-associated, receptor-like protein kinase. *The Journal of Biological Chemistry* **271**(33): 19789-19793.

Heyen BJ, Alsheikh MK, Smith EA, Torvik CF, Seals DF, Randall SK. 2002. The calcium-binding activity of a vacuole-associated, dehydrin-like protein is regulated by phosphorylation. *Plant Physiology* **130**(2): 675-687.

Higo K, Ugawa Y, Iwamoto M, Korenaga T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research* **27**(1): 297-300.

Hincha DK, Thalhammer A. 2012. LEA proteins: IDPs with versatile functions in cellular dehydration tolerance. *Biochemical Society Transactions* **40**(5): 1000-1003.

Hoekstra FA, Golovina EA, Buitink J. 2001. Mechanisms of plant desiccation tolerance. *Trends in Plant Science* **6**(9): 431-438.

Höfler K, Migsch H, Rottenburg W. 1941. Über die Austrocknungresistenz landwirtschaftlicher Kulturpflanzen. *Forschungsdienst* **12**: 50–61.

Honjoh KI, Matsumoto H, Shimizu H, Ooyama K, Tanaka K, Oda Y, Takata R, Joh T, Suga K, Miyamoto T*, et al.* **2000.** Cryoprotective activities of group 3 late embryogenesis abundant proteins from *Chlorella vulgaris* C-27. *Bioscience, Biotechnology, and Biochemistry* **64**(8): 1656-1663.

Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. 2007. WoLF PSORT: protein localization predictor. *Nucleic Acids Research* **35**(Web Server issue): W585-587.

Houde M, Dallaire S, N'Dong D, Sarhan F. 2004. Overexpression of the acidic dehydrin WCOR410 improves freezing tolerance in transgenic strawberry leaves. *Plant Biotechnology Journal* **2**(5): 381-387.

Hundertmark M, Hincha DK. 2008. LEA (Late Embryogenesis Abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics* **9**(1): 118.

Hutin C, Nussaume L, Moise N, Moya I, Kloppstech K, Havaux M. 2003. Early light-induced proteins protect *Arabidopsis* from photooxidative stress. *Proceedings of the National Academy of Sciences of the United States of America* **100**(8): 4921-4926.

Iljin WS. 1957. Drought resistance in plants and physiological processes. *Annual Review of Plant Physiology and Plant Molecular Biology* **8**(1): 257-274.

Imai R, Chang L, Ohta A, Bray E, Takagi M. 1996. A lea-class gene of tomato confers salt and freezing tolerance when expressed in *Saccharomyces cerevisiae*. *Gene* **170**: 243 - 248.

Ingram J, Bartels D. 1996. The molecular basis of dehydration tolerance in plants. *Annual Review of Plant Biology* **47**(1): 377-403.

Ismail AM, Hall AE, Close TJ. 1999a. Allelic variation of a dehydrin gene cosegregates with chilling tolerance during seedling emergence. *Proceedings of the National Academy of Sciences of the United States of America* **96**(23): 13566-13570.

Ismail AM, Hall AE, Close TJ. 1999b. Purification and partial characterization of a dehydrin involved in chilling tolerance during seedling emergence of cowpea. *Plant Physiology* **120**(1): 237-244.

Iturriaga G, Schneider K, Salamini F, Bartels D. 1992. Expression of desiccation-related proteins from the resurrection plant *Craterostigma plantagineum* in transgenic tobacco. *Plant Molecular Biology* **20**(3): 555-558.

Jain M, Nijhawan A, Tyagi AK, Khurana JP. 2006. Validation of housekeeping genes as internal control for studying gene expression in rice by quantitative real-time PCR. *Biochemical and Biophysical Research Communications* **345**(2): 646-651.

Jian B, Liu B, Bi Y, Hou W, Wu C, Han T. 2008. Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Molecular Biology* **9**(1): 59.

Jones L, McQueen-Mason S. 2004. A role for expansins in dehydration and rehydration of the resurrection plant *Craterostigma plantagineum. FEBS Letters* **559**(1-3): 61-65.

Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P*, et al.* 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi. Nature* **414**(6862): 450-453.

Kaye C, Neven L, Hofig A, Li QB, Haskell D, Guy C. 1998. Characterization of a gene for spinach CAP160 and expression of two spinach cold-acclimation proteins in tobacco. *Plant Physiology* **116**(4): 1367-1377.

Khalturin K, Hemmrich G, Fraune S, Augustin R, Bosch TCG. 2009. More than just orphans: are taxonomically-restricted genes important in evolution? *Trends in Genetics* **25**(9): 404-413.

Kikawada T, Nakahara Y, Kanamori Y, Iwata K, Watanabe M, McGee B, Tunnacliffe A, Okuda T. 2006. Dehydration-induced expression of LEA proteins in an anhydrobiotic chironomid. *Biochemical and Biophysical Research Communications* **348**: 56 - 61.

Kim BR, Nam HY, Kim SU, Kim SI, Chang YJ. 2003. Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice. *Biotechnology Letters* **25**(21): 1869-1872.

Kim SY, Chung HJ, Thomas TL. 1997. Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the *Dc3* promoter using a modified yeast one-hybrid system. *The Plant Journal* **11**(6): 1237-1251.

Kirch H-H, Nair A, Bartels D. 2001. Novel ABA- and dehydration-inducible aldehyde dehydrogenase genes isolated from the resurrection plant *Craterostigma plantagineum* and *Arabidopsis thaliana*. *The Plant Journal* **28**(5): 555-567.

Kirch H-H, Röhrig H. 2010. Affinity purification and determination of enzymatic activity of recombinantly expressed aldehyde dehydrogenases. In: Sunkar R ed. *Plant stress tolerance*. New York, USA: Humana Press, 281-290.

Kohorn BD, Johansen S, Shishido A, Todorova T, Martinez R, Defeo E, Obregon P. 2009. Pectin activation of MAP kinase and gene expression is WAK2 dependent. *The Plant Journal* **60**(6): 974-982.

Kohorn BD, Kobayashi M, Johansen S, Friedman HP, Fischer A, Byers N. 2006. Wall-associated kinase 1 (WAK1) is crosslinked in endomembranes, and transport to the cell surface requires correct cell-wall synthesis. *Journal of Cell Science* **119**(11): 2282-2290.

Kohorn BD, Kohorn SL. 2012. The cell wall-associated kinases, WAKs, as pectin receptors. *Frontiers in Plant Science* **3**: 88.

Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**(10): 1289-1291.

Kovacs D, Kalmar E, Torok Z, Tompa P. 2008. Chaperone activity of ERD10 and ERD14, two disordered stress-related plant proteins. *Plant Physiology* **147**(1): 381-390.

Kozera B, Rapacz M. 2013. Reference genes in real-time PCR. *Journal of Applied Genetics* **54**(4): 391-406.

Kramer PJ, Boyer JS. 1995. Functions and properties of water. In: Kramer P, Boyer J eds. *Water relations of plants and soils*. San Diego: Academic press, 16-41.

Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**(1): 105-132.

Laemmli UKU. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**(5259): 680-685.

Lally D, Ingmire P, Tong HY, He ZH. 2001. Antisense expression of a cell wall-associated protein kinase, WAK4, inhibits cell elongation and alters morphology. *The Plant Cell* **13**(6): 1317-1331.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R*, et al.* 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**(21): 2947-2948.

Lata C, Prasad M. 2011. Role of DREBs in regulation of abiotic stress responses in plants. *Journal of Experimental Botany* **62**(14): 4731-4748.

Le TN, McQueen-Mason SJ. 2006. Desiccation-tolerant plants in dry environments. *Reviews in Environmental Science and Bio/Technology* **5**(2-3): 269-279.

Lee PD, Sladek R, Greenwood CM, Hudson TJ. 2002. Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies. *Genome Research* **12**(2): 292-297.

Leung J, Giraudat J. 1998. Abscisic acid signal transduction. *Annual Review of Plant Physiology and Plant Molecular Biology* **49**: 199-222.

Lin CT, Thomashow MF. 1992. A cold-regulated *Arabidopsis* gene encodes a polypeptide having potent cryoprotective activity. *Biochemical and Biophysical Research Communications* **183**(3): 1103-1108.

Lupas A. 1996. Coiled coils: new structures and new functions. *Trends in Biochemical Sciences* **21**(10): 375-382.

Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. *Science* **252**(5009): 1162-1164.

Lüscher B, Eisenman RN. 1990. New light on Myc and Myb. Part II. Myb. *Genes & Development* **4**(12B): 2235-2241.

Mangeon A, Junqueira RM, Sachetto-Martins G. 2010. Functional diversity of the plant glycine-rich proteins superfamily. *Plant Signaling & Behavior* **5**(2): 99-104.

Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, *et al.* 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Research* **39**(Database issue): D225-229.

Marcotte W, Russel S, Quatrano R. 1989. Abscisic acid-responsive sequences from the *Em* gene of wheat. *Plant Cell* **1**: 969 - 976.

Maroufi A, Van Bockstaele E, De Loose M. 2010. Validation of reference genes for gene expression analysis in chicory (*Cichorium intybus*) using quantitative real-time PCR. *BMC Molecular Biology* **11**(1): 15.

McCubbin WD, Kay CM, Lane BG. 1985. Hydrodynamic and optical properties of the wheat germ Em protein. *Canadian Journal of Biochemistry and Cell Biology* **63**(8): 803-811.

Mehta PA, Rebala KC, Venkataraman G, Parida A. 2009. A diurnally regulated dehydrin from *Avicennia marina* that shows nucleo-cytoplasmic localization and is phosphorylated by Casein kinase II in vitro. *Plant Physiology and Biochemistry* **47**(8): 701-709.

Michel D, Furini A, Salamini F, Bartels D. 1994. Structure and regulation of an ABA- and desiccation-responsive gene from the resurrection plant *Craterostigma plantagineum. Plant Molecular Biology* **24**(4): 549-560.

Michel D, Salamini F, Bartels D, Dale P, Baga M, Szalay A. 1993. Analysis of a desiccation and ABA-responsive promoter isolated from the resurrection plant *Craterostigma plantagineum. The Plant Journal* **4**(1): 29-40.

Moore JP, Le NT, Brandt WF, Driouich A, Farrant JM. 2009. Towards a systems-based understanding of plant desiccation tolerance. *Trends in Plant Science* **14**(2): 110-117.

Moore JP, Nguema-Ona EE, Vicré-Gibouin M, Sørensen I, Willats WG, Driouich A, Farrant JM. 2013. Arabinose-rich polymers as an evolutionary strategy to plasticize resurrection plant cell walls against desiccation. *Planta* **237**(3): 739-754.

Mousavi A, Hotta Y. 2005. Glycine-rich proteins: a class of novel proteins. *Applied Biochemistry and Biotechnology* **120**(3): 169-174.

Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8**(19): 4321-4325.

Mutz K-O, Heilkenbrinker A, Lönne M, Walter J-G, Stahl F. 2013. Transcriptome analysis using next-generation sequencing. *Current Opinion in Biotechnology* **24**(1): 22-30.

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M. 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**(5881): 1344-1349.

Nakayama K, Okawa K, Kakizaki T, Honma T, Itoh H, Inaba T. 2007. *Arabidopsis* Cor15am is a chloroplast stromal protein that has cryoprotective activity and forms oligomers. *Plant Physiology* **144**(1): 513-523.

Navari-Izzo F, Ricci F, Vazzana C, Quartacci MF. 1995. Unusual composition of thylakoid membranes of the resurrection plant *Boea hygroscopica*: changes in lipids upon dehydration and rehydration. *Physiologia Plantarum* **94**(1): 135-142.

Neale AD, Blomstedt CK, Bronson P, Le TN, Guthridge K, Evans J, Gaff DF, Hamill JD. 2000. The isolation of genes from the resurrection grass *Sporobolus stapfianus* which are induced during severe drought stress. *Plant, Cell & Environment* **23**(3): 265-277.

Nicot N, Hausman JF, Hoffmann L, Evers D. 2005. Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress. *Journal of Experimental Botany* **56**(421): 2907-2914.

Nora Bocca S, Magioli C, Mangeon A, Magrani Junqueira R, Cardeal V, Margis R, Sachetto-Martins G. 2005. Survey of glycine-rich proteins (GRPs) in the *Eucalyptus* expressed sequence tag database (ForEST). *Genetics and Molecular Biology* **28**(3).

Ohno S. 1970. *Evolution by gene duplication*. Heidelberg and New York: Springer-Verlag.

Okoniewski M, Miller C. 2006. Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. *BMC Bioinformatics* **7**(1): 276.

Oliver M. 2008. Biochemical and molecular mechanisms of desiccation tolerance in bryophytes. In: Goffinet B, Shaw AJ eds. *Bryophyte biology*. Cambridge: Cambridge University Press, 269-298.

Oliver M, Dowd S, Zaragoza J, Mauget S, Payton P. 2004. The rehydration transcriptome of the desiccation-tolerant bryophyte *Tortula ruralis*: transcript classification and analysis. *BMC Genomics* **5**(1): 89.

Oliver M, Tuba Z, Mishler B. 2000. The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology* **151**(1): 85-100.

Oliver MJ, Bewley JD. 1997. Desiccation-tolerance of plant tissues: a mechanistic overview. *Horticultural Reviews* **18**: 171-214.

Oliver MJ, O'Mahony P, Wood AJ. 1998. "To dryness and beyond" – Preparation for the dried state and rehydration in vegetative desiccation-tolerant plants. *Plant Growth Regulation* **24**(3): 193-201.

Oxelman B, Kornhall P, Olmstead RG, Bremer B. 2005. Further disintegration of Scrophulariaceae. *Taxon* **54**(2): 411-425.

Paolacci A, Tanzarella O, Porceddu E, Ciaffi M. 2009. Identification and validation of reference genes for quantitative RT-PCR normalization in wheat. *BMC Molecular Biology* **10**: 27.

Park AR, Cho SK, Yun UJ, Jin MY, Lee SH, Sachetto-Martins G, Park OK. 2001. Interaction of the *Arabidopsis* receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. *The Journal of Biological Chemistry* **276**(28): 26688-26693.

Petersen J, Eriksson SK, Harryson P, Pierog S, Colby T, Bartels D, Röhrig H. 2012. The lysine-rich motif of intrinsically disordered stress protein CDeT11-24 from *Craterostigma plantagineum* is responsible for phosphatidic acid binding and protection of enzymes from damaging effects caused by desiccation. *Journal of Experimental Botany* **63**(13): 4919-4929.

Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnology Letters* **26**(6): 509-515.

Phillips JR, Fischer E, Baron M, van den Dries N, Facchinelli F, Kutzer M, Rahmanzadeh R, Remus D, Bartels D. 2008. *Lindernia brevidens*: a novel desiccation-tolerant vascular plant, endemic to ancient tropical rainforests. *The Plant Journal* **54**(5): 938-948.

Phillips JR, Hilbricht T, Salamini F, Bartels D. 2002. A novel abscisic acid- and dehydration-responsive gene family from the resurrection plant *Craterostigma plantagineum* encodes a plastid-targeted protein with DNA-binding activity. *Planta* **215**(2): 258-266.

Popova AV, Hundertmark M, Seckler R, Hincha DK. 2011. Structural transitions in the intrinsically disordered plant dehydration stress protein LEA7 upon drying are modulated by the presence of membranes. *Biochimica et Biophysica Acta* **1808**(7): 1879-1887.

Porembski S. 2011. Evolution, diversity, and habitats of poikilohydrous vascular plants. In: Lüttge U, Beck E, Bartels D eds. *Plant desiccation tolerance*: Springer Berlin Heidelberg, 139-154.

Pouchkina-Stantcheva N, McGee B, Boschetti C, Tolleter D, Chakrabortee S, Popova A, Meersman F, Macherel D, Hincha D, Tunnacliffe A. 2007. Functional divergence of former alleles in an ancient asexual invertebrate. *Science* **318**: 268 - 271.

Proost S, Van Bel M, Sterck L, Billiau K, Van Parys T, Van de Peer Y, Vandepoele K. 2009. PLAZA: A Comparative Genomics Resource to Study Gene and Genome Evolution in Plants. *The Plant Cell Online* **21**(12): 3718-3731.

Prud'homme B, Gompel N, Carroll SB. 2007. Emerging principles of regulatory evolution. *Proceedings of the National Academy of Sciences of the United States of America* **104**(suppl 1): 8605-8612.

Puhakainen T, Hess MW, Makela P, Svensson J, Heino P, Palva ET. 2004. Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in *Arabidopsis*. *Plant Molecular Biology* **54**(5): 743-753.

Quartacci MF, Forli M, Rascio N, Vecchia FD, Bochicchio A, Navari-Izzo F. 1997. Desiccation-tolerant *Sporobolus stapfianus*: lipid composition and cellular ultrastructure during dehydration and rehydration. *Journal of Experimental Botany* **48**(6): 1269-1279.

Rahman LN, Chen L, Nazim S, Bamm VV, Yaish MW, Moffatt BA, Dutcher JR, Harauz G. 2010. Interactions of intrinsically disordered *Thellungiella salsuginea* dehydrins TsDHN-1 and TsDHN-2 with membranes - synergistic effects of lipid composition and temperature on secondary structure. *Biochemistry and Cell Biology* **88**(5): 791-807.

Rahmanzadeh R, Müller K, Fischer E, Bartels D, Borsch T. 2005. The Linderniaceae and Gratiolaceae are further lineages distinct from the Scrophulariaceae (Lamiales). *Plant Biol (Stuttg)* **7**(1): 67-78.

Ramanjulu S, Bartels D. 2002. Drought- and desiccation-induced modulation of gene expression in plants. *Plant, Cell & Environment* **25**(2): 141-151.

Reyes JL, Rodrigo MJ, Colmenero-Flores JM, Gil JV, Garay-Arroyo A, Campos F, Salamini F, Bartels D, Covarrubias AA. 2005. Hydrophilins from distant organisms can protect enzymatic activities from water limitation effects in vitro. *Plant Cell and Environment* **28**(6): 709-718.

Rinne PL, Kaikuranta PL, van der Plas LH, van der Schoot C. 1999. Dehydrins in cold-acclimated apices of birch (*Betula pubescens* ehrh.): production, localization and potential role in rescuing enzyme function during dehydration. *Planta* **209**(4): 377-388.

Rodrigo MJ, Bockel C, Blervacq AS, Bartels D. 2004. The novel gene *CpEdi-9* from the resurrection plant *C. plantagineum* encodes a hydrophilic protein and is expressed in mature seeds as well as in response to dehydration in leaf phloem tissues. *Planta* **219**(4): 579-589.

Rodriguez MC, Edsgärd D, Hussain SS, Alquezar D, Rasmussen M, Gilbert T, Nielsen BH, Bartels D, Mundy J. 2010. Transcriptomes of the desiccation-tolerant resurrection plant *Craterostigma plantagineum*. *The Plant Journal* **63**(2): 212-228.

Röhrig H, Schmidt J, Colby T, Bräutigam A, Hufnagel P, Bartels D. 2006. Desiccation of the resurrection plant *Craterostigma plantagineum* induces dynamic changes in protein phosphorylation. *Plant, Cell & Environment* **29**(8): 1606-1617.

Rorat T, Grygorowicz WJ, Irzykowski W, Rey P. 2004. Expression of KS-type dehydrins is primarily regulated by factors related to organ type and leaf developmental stage during vegetative growth. *Planta* **218**(5): 878-885.

Rossi M, Lijavetzky D, Hopp HE, Iusem N, Bernacchi D, Rossi M, Hopp HE, Iusem N. 1996. *Asr* genes belong to a gene family comprising at least three closely linked loci on chromosome 4 in tomato. *Molecular and General Genetics MGG* **252**(4): 489-492.

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols* **5**(4): 725-738.

Royce TE, Rozowsky JS, Gerstein MB. 2007. Toward a universal microarray: prediction of gene expression through nearest-neighbor probe sequence identification. *Nucleic Acids Research* **35**(15): e99.

Saavedra L, Svensson J, Carballo V, Izmendi D, Welin B, Vidal S. 2006. A dehydrin gene in *Physcomitrella patens* is required for salt and osmotic stress tolerance. *The Plant Journal* **45**(2): 237-249.

Sachetto-Martins G, Franco LO, de Oliveira DE. 2000. Plant glycine-rich proteins: a family or just proteins with a common motif? *Biochimica et Biophysica Acta* **1492**(1): 1-14.

Sales K, Brandt W, Rumbak E, Lindsey G. 2000. The LEA-like protein HSP 12 in *Saccharomyces cerevisiae* has a plasma membrane location and protects membranes against desiccation and ethanol-induced stress. *Biochimica et Biophysica Acta* **1463**(2): 267-278.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*. New York, USA: Cold Spring Harbor Laboratory Press.

Sanchez-Ballesta M, Rodrigo M, Lafuente M, Granell A, Zacarias L. 2004. Dehydrin from Citrus, which confers in vitro dehydration tolerance and freezing protection activity, is constitutive and highly expressed in the flavedo of fruit but responsive to cold and water stress in leaves. *Journal of Agricultural and Food Chemistry* **52**: 1950 - 1957.

Sandal NN, Bojsen K, Richter H, Senguptagopalan C, Marcker KA. 1992. The nodulin-24 protein family shows similarity to a family of glycine-rich plant proteins. *Plant Molecular Biology* **18**(3): 607-610.

Schäferhoff B, Fleischmann A, Fischer E, Albach DC, Borsch T, Heubl G, Müller KF. 2010. Towards resolving Lamiales relationships: insights from rapidly evolving chloroplast sequences. *BMC Evolutionary Biology* **10**(1): 352.

Scheibe R, Beck E. 2011. Drought, desiccation, and oxidative stress. In: Lüttge U, Beck E, Bartels D eds. *Plant desiccation tolerance*: Springer Berlin Heidelberg, 209-231.

Schenk PM, Elliott AR, Manners JM. 1998. Assessment of transient gene expression in plant tissues using the green fluorescent protein as a reference. *Plant Molecular Biology Reporter* **16**(4): 313-322.

Schmidt J, John M, Wieneke U, Krüssmann H-D, Schell J. 1986. Expression of the nodulation gene *nodA* in *Rhizobium meliloti* and localization of the gene product in the cytosol. *Proceedings of the National Academy of Sciences of the United States of America* **83**(24): 9581-9585.

Schmittgen TD, Zakrajsek BA. 2000. Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR. *Journal of Biochemical and Biophysical Methods* **46**(1-2): 69-81.

Schneider K, Wells B, Schmelzer E, Salamini F, Bartels D. 1993. Desiccation leads to the rapid accumulation of both cytosolic and chloroplastic proteins in the resurrection plant *Craterostigma plantagineum* Hochst. *Planta* **189**(1): 120-131.

Segrest JP, Jones MK, De Loof H, Brouillette CG, Venkatachalapathi YV, Anantharamaiah GM. 1992. The amphipathic helix in the exchangeable apolipoproteins: a review of secondary structure and function. *Journal of Lipid Research* **33**(2): 141-166.

Seine R, Fischer E, Barthlott W. 1995. Notes on the Scrophulariaceae of Zimbabwean inselsbergs, with the description of *Lindernia syncerus* sp. nov. *Feddes Repertorium* **106**(1-5): 7-12.

Sharon MA, Kozarova A, Clegg JS, Vacratsis PO, Warner AH. 2009. Characterization of a group 1 late embryogenesis abundant protein in encysted embryos of the brine shrimp *Artemia franciscana*. *Biochemistry and Cell Biology* **87**(2): 415-430.

Sherwin H, Farrant J. 1998. Protection mechanisms against excess light in the resurrection plants *Craterostigma wilmsii* and *Xerophyta viscosa*. *Plant Growth Regulation* **24**(3): 203-210.

Silhavy D, Hutvágner G, Barta E, Bánfalvi Z. 1995. Isolation and characterization of a water-stress-inducible cDNA clone from *Solanum chacoense*. *Plant Molecular Biology* **27**(3): 587-595.

Silver N, Best S, Jiang J, Thein SL. 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Molecular Biology* **7**(1): 33.

Simpson SD, Nakashima K, Narusaka Y, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. 2003. Two different novel *cis*-acting elements of *erd1*, a *clpA* homologous *Arabidopsis* gene function in induction by dehydration stress and dark-induced senescence. *The Plant Journal* **33**(2): 259-270.

Sivamani E, Bahieldin A, Wraith JM, Al-Niemi T, Dyer WE, Ho THD, Qu RD. 2000. Improved biomass productivity and water use efficiency under water deficit conditions in transgenic wheat constitutively expressing the barley *HVA1* gene. *Plant Science* **155**(1): 1-9.

Sprenger-Haussels M, Weisshaar B. 2000. Transactivation properties of parsley proline-rich bZIP transcription factors. *The Plant Journal* **22**(1): 1-8.

Stacy R, Aalen R. 1998. Identification of sequence homology between the internal hydrophilic repeated motifs of Group 1 late-embryogenesis-abundant proteins in plants and hydrophilic repeats of the general stress protein GsiB of *Bacillus subtilis*. *Planta* **206**: 476 - 478.

Suzuki T, Higgins PJ, Crawford DR. 2000. Control selection for RNA quantitation. *BioTechniques* **29**(2): 332-337.

Swire-Clark GA, Marcotte WR. 1999. The wheat LEA protein Em functions as an osmoprotective molecule in *Saccharomyces cerevisiae*. *Plant Molecular Biology* **39**(1): 117-128.

Teets NM, Peyton JT, Colinet H, Renault D, Kelley JL, Kawarasaki Y, Lee RE, Denlinger DL. 2012. Gene expression changes governing extreme dehydration tolerance in an Antarctic insect. *Proceedings of the National Academy of Sciences of the United States of America* **109**(50): 20744-20749.

Thalhammer A, Hundertmark M, Popova AV, Seckler R, Hincha DK. 2010. Interaction of two intrinsically disordered plant stress proteins (COR15A and COR15B) with lipid membranes in the dry state. *Biochimica et Biophysica Acta* **1798**(9): 1812-1820.

Thomson WW, Platt KA. 1997. Conservation of cell order in desiccated mesophyll of *Selaginella lepidophylla* ([Hook and Grev] Spring). *Annals of Botany* **79**(4): 439-447.

Tolleter D, Hincha DK, Macherel D. 2010. A mitochondrial late embryogenesis abundant protein stabilizes model membranes in the dry state. *Biochimica et Biophysica Acta* **1798**(10): 1926-1933.

Tompa P. 2002. Intrinsically unstructured proteins. *Trends in Biochemical Sciences* **27**(10): 527-533.

Towbin H, Staehelin T, Gordon J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets - Procedure and some applications. *Proceedings of the National Academy of Sciences of the United States of America* **76**(9): 4350-4354.

Tuba Z, Lichtenthaler H. 2011. Ecophysiology of homoiochlorophyllous and poikilochlorophyllous desiccation-tolerant plants and vegetations. In: Lüttge U, Beck E, Bartels D eds. *Plant desiccation tolerance*: Springer Berlin Heidelberg, 157-183.

Tuba Z, Protor CF, Csintalan Z. 1998. Ecophysiological responses of homoiochlorophyllous and poikilochlorophyllous desiccation tolerant plants: a comparison and an ecological perspective. *Plant Growth Regulation* **24**(3): 211-217.

Tunnacliffe A, Hincha D, Leprince O, Macherel D. 2010. LEA Proteins: versatility of form and function. In: Lubzens E, Cerda J, Clark M eds. *Dormancy and resistance in harsh environments*. Heidelberg: Springer, 91-108.

Tunnacliffe A, Lapinski J, McGee B. 2005. A putative LEA protein, but no trehalose, is present in anhydrobiotic bdelloid rotifers. *Hydrobiologia* **546**(1): 315-321.

Tunnacliffe A, Wise MJ. 2007. The continuing conundrum of the LEA proteins. *Naturwissenschaften* **94**(10): 791-812.

Ueki S, Citovsky V. 2002. The systemic movement of a tobamovirus is inhibited by a cadmium-ion-induced glycine-rich protein. *Nature Cell Biology* **4**(7): 478-486.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Research* **40**(15): e115.

Uversky V, Gillespie J, Fink A. 2000. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* **41**: 415 - 427.

Valenzuela-Avendaño JP, Mota IAE, Uc GL, Perera RS, Valenzuela-Soto EM, Aguilar JJZ. 2005. Use of a simple method to isolate intact RNA from partially hydrated *Selaginella lepidophylla* plants. *Plant Molecular Biology Reporter* **23**(2): 199-200.

van den Dries N, Facchinelli F, Giarola V, Phillips JR, Bartels D. 2011. Comparative analysis of *LEA-like 11-24* gene expression and regulation in related plant species within the Linderniaceae that differ in desiccation tolerance. *New Phytologist* **190**(1): 75-88.

Vandesompele J, DePreter K, Pattyn F, Poppe B, Roy N, DePaepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric avaraging of multiple internal control genes. *Genome Biology* **3**: 1 - 12.

Velasco R, Salamini F, Bartels D. 1998. Gene structure and expression analysis of the drought- and abscisic acid-responsive *CDeT11-24* gene family from the resurrection plant *Craterostigma plantagineum* Hochst. *Planta* **204**(4): 459-471.

Verica JA, He ZH. 2002. The cell wall-associated kinase (*WAK*) and *WAK*-like kinase gene family. *Plant Physiology* **129**(2): 455-459.

Vicré M, Lerouxel O, Farrant J, Lerouge P, Driouich A. 2004. Composition and desiccation-induced alterations of the cell wall in the resurrection plant *Craterostigma wilmsii. Physiologia Plantarum* **120**(2): 229-239.

Vicré M, Sherwin HW, Driouich A, Jaffer MA, Farrant JM. 1999. Cell wall characteristics and structure of hydrated and dry leaves of the resurrection plant *Craterostigma wilmsii*, a microscopical study. *Journal of Plant Physiology* **155**(6): 719-726.

Wagner TA, Kohorn BD. 2001. Wall-associated kinases are expressed throughout plant development and are required for cell expansion. *The Plant Cell* **13**(2): 303-318.

Wang L, Feng Z, Wang X, Wang X, Zhang X. 2010. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26**(1): 136-138.

Wang L, Shang H, Liu Y, Zheng M, Wu R, Phillips J, Bartels D, Deng X. 2009. A role for a cell wall localized glycine-rich protein in dehydration and rehydration of the resurrection plant *Boea hygrometrica. Plant Biology* **11**(6): 837-848.

Wang W, Scali M, Vignani R, Spadafora A, Sensi E, Mazzuca S, Cresti M. 2003. Protein extraction for two-dimensional electrophoresis from olive leaf, a plant tissue containing high levels of interfering compounds. *Electrophoresis* **24**(14): 2369-2375.

Wang WX, Pelah D, Alergand T, Shoseyov O, Altman A. 2002. Characterization of SP1, a stress-responsive, boiling-soluble, homo-oligomeric protein from aspen. *Plant Physiology* **130**(2): 865-875.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews: Genetics* **10**(1): 57-63.

Webb SJ. 1965. In: Thomas CC ed. *Bound water in biological integrity*. Springfield, Illinois, 173-181.

Willige BC, Kutzer M, Tebartz F, Bartels D. 2009. Subcellular localization and enzymatic properties of differentially expressed transketolase genes isolated from the desiccation tolerant resurrection plant *Craterostigma plantagineum*. *Planta* **229**(3): 659-666.

Willigen CV, Pammenter NW, Jaffer MA, Mundree SG, Farrant JM. 2003. An ultrastructural study using anhydrous fixation of *Eragrostis nindensis*, a resurrection grass with both desiccation-tolerant and -sensitive tissues. *Functional Plant Biology* **30**(3): 281.

Wilson GA, Bertrand N, Patel Y, Hughes JB, Feil EJ, Field D. 2005. Orphans as taxonomically restricted and ecologically important genes. *Microbiology* **151**(8): 2499-2501.

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

Wise MJ. 2002. The POPPs: clustering and searching using peptide probability profiles. *Bioinformatics* **18**(suppl 1): S38-S45.

Wolkers WF, McCready S, Brandt WF, Lindsey GG, Hoekstra FA. 2001. Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses in vitro. *Biochimica Et Biophysica Acta* **1544**(1-2): 196-206.

Wood AJ. 2005. Eco-physiological adaptations to limited water environments. In: Haegawa PM, Jenks MA eds. *Plant abiotic stress*. Oxford, UK: Blackwell Press, 1-13.

Wood AJ, Jenks MA. 2007. Plant desiccation tolerance: diversity, distribution, and real-world applications. In: Jenks MA, Wood AJ eds. *Plant desiccation tolerance*. Oxford, UK: Blackwell Publishing, 3-10.

Wood AJ, Oliver MJ. 1999. Translational control in plant stress: the formation of messenger ribonucleoprotein particles (mRNPs) in response to desiccation of *Tortula ruralis* gametophytes. *The Plant Journal* **18**(4): 359-370.

Xu D, Duan X, Wang B, Hong B, Ho T, Wu R. 1996. Expression of a late embryogenesis abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. *Plant Physiology* **110**(1): 249-257.

Xue B, Dunbrack RL, Williams RW, Dunker AK, Uversky VN. 2010. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *Biochimica et Biophysica Acta* **1804**(4): 996-1010.

Xue GP. 2002. Characterisation of the DNA-binding profile of barley HvCBF1 using an enzymatic method for rapid, quantitative and high-throughput analysis of the DNA-binding activity. *Nucleic Acids Research* **30**(15): e77.

Yamaguchi-Shinozaki K, Shinozaki K. 1994. A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *The Plant Cell Online* **6**(2): 251-264.

Yamaguchi-Shinozaki K, Shinozaki K. 2005. Organization of *cis*-acting regulatory elements in osmotic- and cold-stress-responsive promoters. *Trends in Plant Science* **10**(2): 88-94.

Zhang L, Ohta A, Takagi M, Imai R. 2000. Expression of plant group 2 and group 3 *lea* genes in *Saccharomyces cerevisiae* revealed functional divergence among LEA proteins. *J Biochem* **127**: 611 - 616.

Zhou X, Ren L, Meng Q, Li Y, Yu Y, Yu J. 2010. The next-generation sequencing technology and application. *Protein & Cell* **1**(6): 520-536.

Zhu X, Li X, Chen W, Chen J, Lu W, Chen L, Fu D. 2012. Evaluation of new reference genes in papaya for accurate transcript normalization under different experimental conditions. *PLoS ONE* **7**(8): e44405.

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**(13): 3406-3415.

ACKNOWLEDGEMENTS

Many people either directly or indirectly contributed to the realization of this PhD thesis during these years. Among them, mainly two people deserve my sincere gratitude.

Firstly, I want to thank Professor Dr. Dorothea Bartels for offering me the opportunity to do a PhD in her group. The many challenges which I had to face during these years permitted me to greatly improve my research skills. I thank you very much for your constant moral support during this time especially in the most difficult moments. Your valuable advices and recommendations throughout my PhD contributed very much to the obtainment of this result. I want also to thank you for the time you spent to read the PhD thesis and for the valuable comments you made to it.

Secondly but yet importantly, I want to thank Linda Lonardi, the person with whom I shared any single moment of the past seven years of my life. Despite all the privations you made for me in these years you never stopped supporting me. Your constant presence by my side importantly contributed to keep me motivated and permitted me to overcome the most critical moments and to successfully complete this thesis. I am very lucky to have you in my life and I will never forget what you did for me.

I want also to thank all the current and former group members who helped me during my PhD and contributed to keep a pleasant atmosphere in the laboratory. Special thanks go to Christine Marikar for the kind help and assistance to solve any administrative problem I encountered during these years.

Lastly, I want to thank Michele Lauritano, Martina Nogara, Francisco Gasulla, Dinakar Challabathula, Alessandro Zucchelli, Naim Stiti and Saeedeh Ataei which significantly made my stay and work in Bonn much more pleasant and nice.

Ich versichere, dass ich diese Arbeit selbstständig verfasst, keine anderen Quellen und Hilfsmittel als die angegebenen benutzt und die Stellen der Arbeit, die anderen Werken dem Wortlaut oder Sinn nach entnommen sind, kenntlich gemacht habe.

Diese Arbeit hat in dieser oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegen.

Bonn, den 7.07.2014

Valentino Giarola