

**Characterization of *ELIP* genes in the desiccation tolerant *Lindernia brevidens* and
its desiccation sensitive counterpart *Lindernia subracemosa***

Thesis

submitted in partial fulfillment of the requirements for the degree

Master of Science

Master program in

Plant Sciences

Faculty of Mathematics and Natural Sciences

Rheinische Friedrich-Wilhelms-Universität Bonn

Presented by

Ahmad Sabri Mahmoud Ammar

From *Amman, Jordan*

Bonn, April 2020

This work has been performed at Instituts für Molekulare Physiologie und Biotechnologie der Pflanzen (IMBIO) of Rheinischen Friedrich-Wilhelms-Universität Bonn

In the team of Prof. Dr. Dorothea Bartels

1. Referee 1: Prof. Dr. Dorothea Bartels

Institut für Molekulare Physiologie und Biotechnologie der Pflanzen

2. Referee 2: Dr. Valentino Giarola

Fondazione Edmund Mach - Istituto Agrario San Michele All'Adige

Affirmation for the Master's Thesis

I herewith declare, that I have written this thesis independently and myself. I have used no other sources than those listed. I have indicated all places where the exact words or analogous text were taken from sources. I assure that this thesis has not been submitted for examination elsewhere.

Bonn, 14.04.2020

(Signature)

Table of contents

1. Introduction	1
1.1. Drought stress.....	5
1.2. Desiccation tolerance	6
1.3. Light stress	7
1.4. Photoinhibition	8
1.5. Plant photoacclamatory defense processes	9
1.5.1. Dissipation of excess absorbed light energy.....	9
1.5.2. ROS detoxification by antioxidants.....	11
1.5.3. Light-induced Stress Proteins.....	12
1.6. Linderniaceae	17
1.6.1. Resurrection plants of the family Linderniaceae	17
1.6.2. Evolution of desiccation tolerance in Linderniaceae.....	18
1.7. Objectives of the study.....	20
2. Materials and Methods.....	21
2.1. Materials	21
2.1.1. Plant material	21
2.1.2. Main stock Buffers and Media solutions	21
2.1.3. Primers.....	21
2.1.4. Devices	23
2.1.5. Enzymes and Vectors	23
2.2. Methods	24
2.2.1. Standard molecular biology techniques and nucleotide sequence analysis.....	24
2.2.2. Phylogenetic analysis	24
2.2.3. Amino acid sequence analysis	27
2.2.4. Plant treatment.....	28
2.2.5. Preparation of DEPC treated water	28
2.2.6. Preparation of colony-selection plates.....	29
2.2.7. Total RNA isolation (Valenzuela-Avendano et al., 2005)	29
2.2.8. cDNA synthesis using ReverseAid Reverse Transcriptase	29
2.2.9. Determination of nucleic acid concentration.....	30
2.2.10. RT-PCR and Realtime RT-qPCR analysis	31

2.2.11. Agarose gel electrophoresis.....	32
2.2.12. DNA extraction from agarose gels purification (MACHEREY-NAGEL GmbH & Co. KG)	32
2.2.13. Ligation.....	33
2.2.14. Preparation of electrocompetent <i>Escherichia coli</i> cells (RbCl method)	33
2.2.15. Transformation of electrocompetent <i>E. coli</i> cells.....	34
2.2.16. Colony PCR.....	34
2.2.17. Plasmid Isolation	35
2.2.18. Purification of Isolated Plasmids (MACHEREY-NAGEL GmbH & Co. KG)	37
2.2.19. Sending purified plasmids for sequencing.....	37
4. Results.....	38
4.1. Characterization of ELIPs	38
4.2. Phylogenetic analysis	49
4.3. RNA stability.....	51
4.4. Transcript expression of Lbr and Lsu <i>ELIPs</i> in response to dehydration treatments	51
4.5. SANGER sequencing results	53
5. Discussion	56
5.1. Analysis of ELIP amino acid sequences	60
5.2. ELIPs role as pigment binding proteins	63
5.3. Expansion and expression of <i>ELIPs</i> in <i>L. brevidens</i> and <i>L. subracemosa</i>	65
6. Conclusions	70
7. References.....	71
8. Acknowledgments.....	89
9. Supplementary material	90

List of abbreviations

At	<i>Arabidopsis thaliana</i>
ALA	5-aminolevulinate biosynthetic pathway
ATP	Adenosine triphosphate
<i>Ca. sinensis</i>	<i>Camellia sinensis</i>
CAB	Chlorophyll a/b binding
CaCl ₂ 2H ₂ O	Calcium chloride dihydrate
Cbr	Carotene biosynthesis-related proteins
cDNA	Complementary DNA
Chl	Chlorophyll
Chl a	Chlorophyll a
Chl b	Chlorophyll b
<i>Ci. sinensis</i>	<i>Citrus sinensis</i>
Cp	<i>Craterostigma plantagineum</i>
cpSRP43	Chloroplast signal recognition particle 43
C-t	C terminal
DA	Drought avoidance
DAG1	DOF AFFECTING GERMINATION 1
DE	Drought escape
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
DR	Drought recovery
dsp	desiccation stress protein
DT	Drought tolerance
EDTA	Ethylenediaminetetraacetic acid
EF5 α	Elongation factor 5-alpha
ELIP	Early light-induced proteins
<i>E. coli</i>	<i>Escherichia coli</i>
FDT	Full desiccation tolerance
FG	Full gene amplification
Gel Electr.	Gel electrophoresis
GluTR	Glu tRNA reductase biosynthetic pathway
GTE	Glucose/Tris/EDTA
GTE	Grams
h	Hour
H ₂ O	Water
HL	High light
Hlips	High light-induced proteins

Kac	Potassium acetate
KCl	Potassium chloride
kDa	Kilodalton
KOH	Potassium hydroxide
LB agar	Lysogeny broth agar
Lbr	<i>Lindernia brevidens</i>
LHC	Light harvesting complex
LHCB1	LHCII type I CAB-8
LiCl	Lithium chloride
LL	Low light
Lsu	<i>Lindernia subracemosa</i>
MDT	Modified desiccation tolerance
mg	milligram
MgCl ₂	Magnesium chloride
MgSO ₄	Magnesium sulfate
min	Minutes
ML	Maximum likelihood
ml	Milliliter
MnCl ₂ 4H ₂ O	Manganese(II) chloride tetrahydrate
MOPS	3-(N-morpholino) propane sulfonic acid
MP	Maximum parsimony
NaCl	Sodium chloride
NADPH	nicotinamide adenine dinucleotide phosphate hydrogen
NaOH	Sodium hydroxide
NPQ	Nonphotochemical chlorophyll fluorescence quenching
N-t	N terminal
OD	Optical density
ORFs	Open reading frames
PCR	Polymerase chain reaction
PG	Partial gene amplification
pI	Isoelectric point
PSII	Photosystem II
qE	High-energy quenching
qI	Quenching resulting from photoinhibition
qT	State transition
RbCl	Rubidium chloride
RNA	Ribonucleic acid
RNase	Ribonuclease
ROS	Reactive oxygen species
rpm	Rounds per minute
RT	Room temperature

RT-PCR	Reverse transcription PCR
RWC	Relative water content
Scps	Small CAB-like protein
SDS	Sodium dodecyl sulfate
sec	Seconds
Seps	Stress-enhanced proteins
SOC	Super Optimal broth
TAE buffer	Tris-Acetate-EDTA buffer
UV	Ultraviolet
w/v	Weight/volume
wcr12	Wheat cold-regulated 12 protein
µg	Microgram
µl	Microliter

1. Introduction

Plants are sessile organisms which makes them exposed and vulnerable to a wide range of environmental abnormalities (Al-Whaibi, 2011; Smékalová et al., 2014; Pan et al., 2016). Their sessile nature imply that they are incapable of migrating from sub-optimal environment to an optimal one (Kumar et al., 2012). In the light of this, organisms have been under strong selection to evolve adaptive responses to environmental stressors (Bartels and Salamini, 2001; Gassmann et al., 2016). Owing to their immobility, plants have selected for the evolution of traits capable of conferring stress tolerance to both biotic and abiotic stressors (Mafakheri et al., 2010). Frequent exposure to stressful conditions is unfavorable for plant growth and development (Zhu, 2016). Their presence represent major constraints that negatively affects plant health (Rajkumar et al., 2017), limits productivity (Kissoudis et al., 2015), impacts their geographical distribution, and threatens food security (Zhu, 2016).

The adverse effects of these stressors are exacerbated by intensifying global warming (Ramegowda and Senthil-Kumar, 2015) and climate change (Zhu, 2016). The spread and intensity of biotic and abiotic stressors (Kissoudis et al., 2015) as well as the abnormalities in weather conditions associated with them (Ramegowda and Senthil-Kumar, 2015) are expected to increase in frequency which will put a strain on agricultural crop production and undermine plant conservation efforts. Furthermore, plants inhabiting the affected areas are predicted to bear the brunt of an increased number of abiotic and biotic stress combinations (Kissoudis et al., 2015; Ramegowda and Senthil-Kumar, 2015) which even puts an additional strain on global food security and the preservation of ecological diversity. The coup de grâce of these events culminates in the unsustainable anthropoid activities of human populations which compels them to compete with plants for the finite environmental resources on Earth (Wallace et al., 2003). To maintain the environmental balance between ecological conservation and human need for food, it is of paramount importance to invest in the research and development of climate smart crops which are resilient to climate change (Wheeler and Von Braun, 2013; Cohen and Leach, 2019) in addition to the promotion of new sustainable agricultural practices and technologies capable of slowing down or reversing the damage wrought upon the ecologically distressed areas.

Biotic stressors are represented by animate threats such as cold, heat, light, radiation (UV-B and UV-A), drought, salinity, flooding (hypoxia and anoxia), nutrient deficiency, and the presence of toxic metals like aluminum, arsenate, cadmium, ozone, and sulfur dioxide in the soil (Suzuki et al., 2014; Zhu, 2016)

whereas abiotic stressors are denoted by animate ones such as bacteria, fungi, oomycetes (Saijo and Poiani Loo, 2019), and herbivore attacks (Zhu, 2016).

Plants have evolved sophisticated signaling and protective mechanisms to survive sub-lethal stress situations (Kumar et al., 2012). Their adaptations to environmental stressors vary depending on the level of at which the response is reacting to those stressors. Responses could be initiated at either the tissue, multicellular, cellular, or subcellular levels (Chaves et al., 2008) and could involve the activation of multiple responses in all levels at the same time. Amongst the most common responses to avoid or tolerate abiotic or biotic stresses include the closure of stomata, reduction in the rate of photosynthesis, upregulation reactive oxygen species (ROS) scavenging genes, reduction of leaf growth, and increased root length (Melotto et al., 2006; Bilgin et al., 2010; Maiti and Satya, 2014). These alterations in response to stress are of paramount importance particularly because stress alters the timing of transition from vegetative to flowering phase, which is vital for both survival and reproductive success (Kumar et al., 2012).

Plant responses to stresses are parametric and dynamic in their nature (Skirycz and Inze, 2010) that is, they are up- or down-regulated as function of several parameters which could be both elastic (reversible) and plastic (irreversible) (Cramer et al., 2011). Furthermore, the level-dependent response is controlled by the intensity (Soliman and Kostandi, 1998), duration (acute vs chronic), and rate of progression of stress. These parametric variables are deterministic factors which affect the complexity and outcome of the stress response (Pinheiro and Chaves, 2011). In the example of drought stress, erroneous results could be avoided by monitoring the impact of drought using several parameters including monitoring water availability in the substrate, photosynthetic activity, plant water status, and the radiation to which plants are subjected. This will be followed by integrating the data to determine the nature of the elicited response. Fine-tuning experimental methods to follow this parametric approach is important for obtaining high-confidence results that are accurate and reproducible (Dehyolos, 2010). For example, it has been shown that the intensity and/or duration of stress are two factors that influence the velocity and the extent of recovery of photosynthesis (Miyashita et al., 2005).

The complexity of stress tolerance adaptations is compounded by the intricate crosstalk networks which overlap in many ways and give rise to cross-tolerance phenomena (Foyer et al., 2016). Indications for such crosstalk between multiple regulatory networks can be found at the phenotypic as well at the gene expression level (Kissoudis et al., 2014). Very less is known about these crosstalk networks. Moreover,

their integrative signals and convergence points are not well understood (Fraire-Velázquez et al., 2011). More research is ought to be carried out to unravel the different elements of these fine-tuned regulatory networks before the development of any reliable genetically modified crop can be realized.

In natural environments, stress responses are either activated in response to a combination of environmental stressors simultaneously or in rapid succession. Nonetheless, environmental stress conditions and their effects are typically studied in isolation (Suzuki et al., 2014; Gassmann et al., 2016) and under controlled growth conditions in the confined areas of laboratories (Suzuki et al., 2014) which represent a major constraint to biological studies. In the field, plants are frequently exposed simultaneously to a combination of abiotic and/or biotic stress condition such as: double threats which include drought and heat, drought and cold, salinity and heat (Suzuki et al., 2014), triple threats such as drought and heat along with light stress, or any the abiotic stressors combined with a biotic stress such as pathogenic or fungal infections.

In spite of the prospects of obtaining a fine picture of single stress responses using current scientific approaches, little is known about how exposure to additional stressors might impact plant growth and development. Due to the lack of studies of more complex systems of interacting stresses, the impact of environmental conditions on phenotypes in nature is not completely understood (Gassmann et al., 2016). Based on the limited data available for plant responses under combinatorial stress, the research points to predominantly negative interactions at the phenotypic level (Kissoudis et al., 2014). It has been suggested that when plant cells encounter different stresses, their responsiveness to stress stimuli depends on the functional specialization of gene families to each stress (Li et al., 2017).

The genes modulating stress responses might proceed in cumulative or redundant fashions as well as activate common or unique downstream targets (Kumar et al., 2012). Nonetheless, there is increasing evidence in the scientific literature which supports the uniqueness of responses to multiple environmental stresses and shows that their simultaneous activation does not always impart an additive-like response (Atkinson and Urwin, 2012) because the nature of interactions between stress factors dictates the outcome of the acclimation process (Atkinson et al., 2013). Therefore, it has been suggested that each stress combination is ought to be studied exclusively as an entirely new stress (Mittler and Blumwald, 2010). This has been demonstrated in transcriptome Studies carried out on *Nicotiana tabacum* and *Arabidopsis thaliana*. It has been shown that a combination of drought and heat stress induces a novel program of gene expression thereby activating transcripts that are not induced by either stress individually (Rizhsky et al.,

2004). Moreover, when carrying out an experiment to study the combinatorial effect of two stresses, drought and salt stress, it will be helpful to differentiate between primary and secondary stress signals. In the case of a concurrent drought-salt tolerance, it is helpful to determine whether the initial stress response was caused by a deficiency in water or abundance of salt (Zhu, 2002).

Breeding for single stress type (e.g., abiotic or biotic) or a single stress (e.g., drought or pathogens) may be risky because increasing tolerance to one stress or stress type may be established at the expense of tolerance to the other (Atkinson and Urwin, 2012). Abiotic stress conditions such as high incident light (Barczak-Brzyżek et al., 2017), drought, high and low temperature, and salinity are capable of influencing the development and pervasiveness of pathogens, insects, and weeds (Peters et al., 2014).

On one hand abiotic stress could be detrimental to plant health. For instance, plants exposed to high salinity and drought stress are more prone to display reduced immune activity (Bostock et al., 2014). Another study has linked high light stress to plant defense against the two-spotted spider mite (TSSM). An *A. thaliana* mutant is uncapable of performing non-photochemical quenching (NPQ) when exposed to excess light results in a phenotype susceptible to TSSM infestation (Barczak-Brzyżek et al., 2017). Thus, high light stress could make plants more vulnerable to infestations. Under water deficit a similar trend could be observed (Jordan et al., 1984).

However, the preponderance of evidence does not suggest that a combination of either abiotic–abiotic or abiotic–biotic stresses necessarily lead to negative impact on plants. Some stress combinations could result in a net neutral impact or elicit positive responses in plants. (Pandey et al., 2017). It has been shown that aphid attacks may provide the plant with some protective measures. This notion may hold true for other abiotic–biotic stress interactions (Gassmann et al., 2016). The colonization of plant species with certain beneficial species of *Pseudomonas* can alleviate the impact of environmental stresses on plant by enhancing plant nutrition acquisition, modulating plant hormone levels, inducing accumulation of osmolytes and antioxidants, and up or down regulating the growth-related genes in plants. This shows that plant-microbe interactions are much more complex than previously thought and that our knowledge concerning both plant-microbe interactions and microbe-driven stress tolerance in plants is still in its infantile stages. Similarly, increased soil salt concentration results in enhanced susceptibility to the soil borne plant pathogen *Pythium aphanidermatum* in cucumber (Al-Sadi et al. 2010).

Therefore, to optimize targeted crop improvement strategies (Cohen and Leach, 2019) to ultimately shield crops from the combinatorial effects of stressors hinges upon advancing our understanding of the

concurrent occurrence of multiple stress responses along with the genetic architecture responsible for regulating it (Kissoudis et al., 2015). That is, the development of plants with enhanced tolerance to combined abiotic and biotic stresses involves the identification of each element of the crosstalk network, how these elements interact with each other (Li et al., 2017), and characterizing the physio-morphological traits that are affected by combined stresses (Pandey et al., 2017). Plant breeding strategies should also be accompanied with the identification or development of complementary bioinoculants to maximize the efficiency of breeding and optimize stress tolerance responses (Rajkumar et al., 2017).

1.1. Drought stress

One of the most potent and pervasive form of environmental stresses affecting the growth and productivity of the world's agricultural production is drought stress (Araus et al., 2002). It constitutes a threat for sustainable crop production especially in a world that is experiencing abrupt temperature changes due to climate change (Tubiello et al., 2000) and in which the frequency of severe drought conditions is expected to increase drastically (Dai, 2012) particularly in arid and semiarid areas (Ahmadizadeh et al., 2012). In meteorological terms, drought is defined as a sufficiently long period of dry weather (Kneebone et al., 1992) caused by high temperatures (Tester and Bacic, 2005) or low precipitation rates (Ahmadizadeh et al., 2012). By the year 2025, the International Water Management Institute portends that one-third of the world's population will be living in regions that experience severe water scarcity (Bartels and Salamini, 2001).

The negative physiological and biochemical effects of drought stress encompass loss of plant-water potential and turgor pressure (Hsiao, 2000), loss of water use efficacy, induction of heat stress, reduction stem extension (Farooq et al., 2009), dwindling rates of photosynthesis (Ghosh and Xu, 2014), hindrance of stomatal conductance to conserve water yielding a substantial reduction in CO₂ fixation and lessening assimilate production for growth and yield of plants (Boyer, 1970), the induction of stomatal closure (Flexas et al., 2004), and the accumulation of ROS which damage chloroplasts (Smirnoff, 1995). In addition, drought stress decreases the diffusion of nutrients from the soil and into the roots. Since nutrients are carried to the roots by water, transport of water-soluble soil nutrients such as nitrate, sulfate, Ca, Mg, and Si is impeded (Vurukonda et al., 2016). Subsequently, these stresses collectively cause irreversible mechanical damage to the cellular membranes and organelles (Levitt, 1980; Bohnert et al., 1995) culminating in the termination of growth and cessation of general metabolism (Hsiao, 1973).

In the light of this, to avoid the harmful impact of changing environments, which are detrimental to plant health, growth, and productivity, it has become imperative for plant biologists to understand the mechanisms by which plants can adapt to water deficit while retaining their capacity to serve as sources of food and other raw materials (Bartels and Salamini, 2001). Plant responses and adaptations to dehydration stress is displayed by complex and multifaceted changes that are evident at the biochemical, molecular, physiological, and morphological levels (Czolpinska and Rurek, 2018). The acclimation and deacclimation processes are meticulously mediated by regulating the expression of many genes. Fluctuations in homeostatic conditions that favor plant growth bring about changes in the transcript levels of specific genes thereby altering gene expression and protein production (Gaff, 1971).

1.2. Desiccation tolerance

The four primary mechanisms by which plants acclimate to drought involve drought avoidance (DA), drought tolerance (DT), drought escape (DE), and drought recovery (DR) (Lawlor, 2013). However, the two major mechanisms that govern the processes conferring plant tolerance are DA and DT (Yue et al., 2006). Morgan (1984) defines drought tolerance in plants as their ability to endure and survive low tissue water content by activating adaptive traits such as the maintenance of cell turgor and increasing protoplasmic resistance. In contrast, Levitt (1980) defined DA as the ability of plants to maintain a relatively high tissue water content even though the water content of the soil has been reduced. DA is achieved by various mechanisms that differ between water spenders and water savers. For instance, water spenders maintain high tissue water content by conserving their uptake of water through increased rooting and hydraulic conductance. On the other hand, water savers utilize the uptaken water effectively and reduce water loss by reducing transpiration and transpiration area (Morgan, 1984).

In lower plants, DT mechanisms are commonly found. On the other hand, higher plants do not enjoy the same drought-resistance privileges of lower plants particularly when the relative water content of their cells decreases below 60% (Giarola et al., 2017), with two notable exceptions. First, seeds and pollen grains can withstand desiccation to an air-dried state with the former surviving for long periods of time in the desiccated state while the latter losing its tolerance rather quickly (Walters et al., 2005). Second, resurrection plants, a minor group of vascular angiosperm plants, can lose most of their cellular water content (Gaff, 1971) up to 98% of the cellular water content in their vegetative tissue (Bartels and Hussein, 2011) and revive from an air-dried state within few hours of rewetting (Bartels, 2005). Among the few

other primitive species which possess desiccation tolerance in their vegetative tissues are ferns, algae, lichens, and bryophytes (Bartels and Salamini, 2001).

Resurrection plants can be classified under three different classifications:

1. Based on the number of cotyledons in the seed: resurrection plants could either be monocotyledonous such as *Sporobulus stapfianus* (Neale et al., 2000) or dicotyledonous such as the South African *Craterostigma plantagineum* (Bartels et al., 1990).
2. Based on whether the plant retains or degrades chlorophyll: resurrection plants could either be homoiochlorophyllous (retains chlorophyll) such as *L. brevidens* (VanBuren et al., 2018) or poikilochlorophyllous (degrades chlorophyll) such as the South African *Xerophyta viscosa* (Mundree et al. 2000).
3. Based on the time required for the repair and recovery: resurrection plants could either exhibit full desiccation tolerance (FDT) or modified desiccation tolerance (MDT). On one hand, FDT prevails in lower order plants such as the moss *Tortula ruralis*. FDT plants desiccate rather quickly and rely on constitutive damage and repair mechanisms during rehydration. One key characteristic of this group is that the progression of FDT is not affected by the rate of drying. On the other hand, MDT prevails in higher plants such as *L. brevidens*. MDT plants desiccate rather slowly and rely on the induction of protective mechanisms to preserve cellular integrity during dehydration. Unlike FDT the induction of DT is directly affected by the rate of drying and needs a certain amount of time before tolerance is fully manifested (Oliver and Bewley, 2010).

1.3. Light stress

Another major factor that diminishes the productivity of higher plants is high light (HL) stress. Light is the primary source of energy that does not only power plant growth through photosynthesis but also serves as a developmental signal that modulates morphogenesis (i.e., de-etiolation and the transition to reproductive development) (Harari-Steinberg et al., 2001). In addition to its regulatory roles, it modulates the photosynthetic rate and accumulate-assimilation (Yang et al., 2019). Throughout the day, light intensity varies temporally (because of the diurnal cycle) and spatially (because of shading by clouds, objects, and other organisms and objects) (Ruban et al., 2012). In order to utilize light energy for photosynthesis, plants possess molecules that act as light-harvesting antennas. First, light quanta are collected through chlorophyll (Chl) excitation and the collected energy is subsequently transferred to the

photosynthetic reaction centers (Heddad et al., 2012). Nonetheless, light could be deleterious for plants (Harari-Steinberg et al., 2001) due to seasonal fluctuations caused by the varying intensity and quality of the incoming light (Brugnoli and Björkman, 1992). Overall, HL and low light (LL) stresses induce contrasting responses in plants (Yang et al., 2019). Under HL stress, when excess photochemical energy can't be used in photochemical reactions, photoinhibition of Photosystem II (PSII) occurs (Adamska, 1997).

1.4. Photoinhibition

Photoinhibition is the light-induced loss of PSII electron-transfer activity (Tyystjärvi, 2013). The balance between the rate of photodamage and repair determines the extent of photoinhibition (Takahashi and Badge, 2011). Even though the impact of photoinhibition could be alleviated by repairing damaged PSII at the early stages of HL stress, prolonged exposure to HL is accompanied by the photooxidation of pigments, degradation of carotenoids, bleaching of chlorophyll, destruction of chloroplasts structure (Adamska, 1997) and the obstruction of chlorophyll biosynthesis, the latter being a process which involves many potent photosensitizing tetrapyrrole intermediate molecules (Aarti et al., 2006; Tzvetková-Chevolleau et al., 2007). Photoinhibition of PSII, if left uncurbed, will ultimately induce photooxidative stress (Krieger-Liszkay, 2004) which promotes the formation of ROS (Heddad et al., 2012). Likewise, other biotic and abiotic environmental stresses (e.g., pathogen attacks, high or low temperature, salinity, drought, and nutrient deficiency) could trigger the generation of ROS in photosynthetic organisms (Apel and Hirt, 2004; Zhao et al., 2017). Cellular ROS take four basic forms including superoxide radical (O_2^-), hydrogen peroxide (H_2O_2), and hydroxyl radical ($\cdot OH$) which forms via the partial reduction of atmospheric oxygen (Mittler, 2002) whereas singlet oxygen (1O_2) is the resultant of the electronic excitation of oxygen (Aarti et al., 2006). These ROS carry the potential to oxidize multiple cellular biomolecules such as pigments, proteins, lipids, carbohydrates, DNA, and RNA which ultimately culminate in plant cellular death (Halliwell and Gutteridge, 1999; Mittler, 2002; Tripathy and Oelmüller, 2012; Das and Roychoudhury, 2014).

Light stress is one of the least characterized and studied abiotic stress. Despite the recent upsurge in publications aimed at deciphering the significance of light stress and how high and low light affect plants (Zhu et al., 2017) more research should be carried out to discover new potential candidate genes and

signalling pathways which can be used for biotechnological applications and development of transgenic plants (Yang et al., 2019).

1.5. Plant photoacclamatory defense processes

To preserve their physiological functions, plants have evolved a multilevel network of responses and adaptations to cope with conditions of high irradiance. In general, plant photo-acclamatory responses proceed via two modes of actions: the control of light absorption capacity and/or the management of the already captured light energy (Ruban, 2009). Three primary mechanisms, either concurrently or separately, could operate at the physiological, cellular (Brugnoli and Björkmann, 1992), or subcellular levels (Heddad et al., 2012) to establish short-term (occurring within hours) and long-term (occurring after days/weeks) adaptations.

At the physiological level, plants regulate their light absorption capacity by altering leaf orientations, leaf movements, and leaf reflectance as well as by building up inorganic deposits on the leaf surface or developing air filled hairs (Ruban, 2009) whereas at the cellular level chloroplast avoidance is manifested by the movement of chloroplasts to the cell wall perpendicular to the light source (Brugnoli and Björkmann, 1992; Kasahara et al., 2002). However, to protect the photosynthetic apparatus from the hazardous effects of photoinhibition at the subcellular level, numerous biochemical and molecular photoprotective mechanisms such as altering thylakoid protein profiles, quenching of high energy, scavenging of ROS by the antioxidant machinery (Nama et al., 2018; Yang et al., 2019), modulating the activities of xanthophyll cycle pigments, in addition to upregulation of light-induced stress proteins and ROS scavengers (Heddad et al., 2012; Yang et al., 2019) are employed.

1.5.1. Dissipation of excess absorbed light energy

At the subcellular level, long-term (acclimation) and short-term (regulatory mechanisms) adaptations control the input of light energy into reaction centers. Acclimation is the outcome of light-dependent regulation of complex gene expression taking place at three regulatory levels: transcriptional, translational, and posttranslational levels (Anderson et al., 1988). Nonetheless, acclimation on its own is not capable of completely achieving photoprotection (Ruban et al., 2012). This does not allow the plants to quickly intervene in protecting themselves from photodamage which can happen within minutes of exposure to high incident light (Tyystjärvi and Aro, 1996).

The primary line of defense against high irradiance is the stress-related nonphotochemical chlorophyll fluorescence quenching (NPQ). It is an adaptation which evolved to provide a faster response to protect photosynthetic membranes from excess light (Demmig-Adams et al., 2014). Briefly, NPQ is a process in which the light-harvesting antennae are quickly and reversibly converted into a photo-protected quenched state to dissipate excess absorbed light energy into heat (Ruban et al., 2007). This process is induced in PSII of photosynthetic membranes (Horton et al., 1996; Demmig-Adams et al., 2014) and integrates several cellular processes including the chlorophyll binding quenching process and zeaxanthin formation via the VAZ cycle as well as state transition (qT), high-energy quenching (qE), and quenching resulting from photoinhibition (qI) (Ruban, 2016). Changes in the relative antenna sizes of PS II and I are represented by qT whereas changes in the thermal dissipation of light energy triggered by lumen acidification are triggered by qE (Allorent et al., 2013). The energy released by these cellular processes does not only damage the reaction center of PSII which drives the splitting of water and evolution of oxygen (Barber, 1995) but also harm the antenna pigments (Fleming et al., 2012) which can result in the intermittent decline in photosynthetic efficiency (Ruban, 2016). Sensitivity to photoinhibition is more prominent in mutants deficient in their ability to induce NPQ (Allorent et al., 2013). Similarly, resistance to biotic stressors such as insects (Barczak-Brzyżek et al., 2017) and abiotic stressors such as heat (Tang et al., 2007), drought (Cousins et al., 2002), low temperature (Xu et al., 1999), and salt (Neto et al., 2014) is greatly diminished in NPQ mutants.

NPQ resides in the antenna (Fleming et al., 2012) and is triggered either by changes in the proton gradient of the antenna components or indirectly by the changes in the activity of the xanthophyll cycle(s). Accumulation of protons which occur in the inter-thylakoid membrane space (lumen) are the result of ATP (NADPH) overproduction. Changes in the proton gradient subsequently triggers the expression of PsbS which is hypothesized to act as a switch and be closely localized to the light harvesting complex II (LHCII). The switch makes LHCII sensitive to protonation (Ruban et al., 2012) and prompts the aggregation of LHCII which is required to establish the NPQ state (Horton et al., 1991). Moreover, changes in the activity of some xanthophyll cycle pigments is associated with the activation of NPQ. High-energy quenching (qE) also called thermal energy dissipation is associated with the conversion of violaxanthin to zeaxanthin via antheraxanthin either by the protonation of PsbS or the catalyst violaxanthin de-epoxidase (VDE) (Niyogi, 1999). Experiments on plant leaves have been demonstrated that NPQ function is drastically enhanced with the conversion of violaxanthin into zeaxanthin, a process induced by

the proton gradient (Demmig-Adams et al., 1989a; Demmig-Adams et al., 1989b). In diatoms, dinophytes and haptophytes, qE is associated with the conversion of dianinoxanthin to diatoxanthin (Lohr and Wilhelm, 2001). Ultimately, qE via the xanthophyll cycle pigments suppresses the transfer of energy from antenna proteins to photosystems (Wilson et al., 2016).

1.5.2. ROS detoxification by antioxidants

ROS are constantly being generated as the byproduct of aerobic metabolism and their production negatively influence membrane-linked electron transport processes, redox-cascades, and metabolism (Foyer and Noctor, 2005; Bhattacharjee, 2010). Their impact is primarily confined to plant cellular compartments including: chloroplasts, mitochondria, peroxisomes, and apoplasts (Jubany-Marí et al., 2009; Roychoudhury and Basu, 2012; Tripathy and Oelmüller, 2012). Since ROS are continuously being generated and ROS imbalances are frequently encountered due to environmental stressors, all aerobic organisms have evolved discreet pathways to deal with ROS imbalances if they occurred. Maintaining the balance between production of ROS and antioxidants to scavenge or quench them is a prerequisite for redox homeostasis (Choudhury et al., 2013) and any imbalances in the generation and metabolism of ROS will lead to oxidative stress (Bhattacharjee, 2010; Tripathy and Oelmüller, 2012). Therefore, the availability of ROS in cellular compartments is tightly controlled by ROS-scavenging pathways that metabolize them. The antioxidant system includes enzymes (e.g., Superoxide Dismutase, Catalase, Ascorbate Peroxidase, Monodehydroascorbate reductase, Dehydroascorbate reductase, Glutathione Reductase, and Guaiacol Peroxidase) and lipophilic antioxidants which scavenges lipid radicals as well as ROS (e.g., α -tocopherol and carotenoids). Secondary ROS scavenging could be carried out by flavonoids (flavonols, flavones, isoflavones, and anthocyanins), osmolytes such as proline which is considered to be a very powerful non-enzymatic antioxidant (Das and Roychoudhury, 2014), and non-enzymatic phytochemicals (e.g., glutathione, polyphenols, bioflavonoids, hydroxycinnamates, and vitamins) (Engwa, 2018).

Eventually, ROS are harmful agents that could hamper plant growth and development, yet they positively regulate gene expression. ROS act as signaling molecules and secondary messengers which inherently activate molecular responses and activate signal transduction processes to proceed with the acclimation process (Bhattacharjee, 2005; Bhattacharjee, 2012; Tripathy and Oelmüller, 2012). Surprisingly, under favorable conditions, ROS are considered to be important metabolites which participate in metabolism,

growth, and morphogenesis of plant cells (Bhattacharjee, 2012). For example, *A. thaliana* responds to oxidative stress by concurrently activating and upregulating the expression of antioxidative and stress-inducible genes (Santos et al., 1996). Aslund et al. (1999) and Goyer et al. (2002) have linked ROS-mediated gene expression to some oxidative stress responsive elements (e.g., promoters and transcription factors). However, the knowledge surrounding elements of the signalling transduction pathways involved in redox-sensing and gene expression upregulation are still obscure (Bhattacharjee, 2012). Advances in functional genomics, proteomics, and metabolomics will offer detailed insights into the function of every individual element of ROS networks and shed the light on their respective downstream responses. Since ROS generation and scavenging are the result of various environmental stressors, unravelling their stress-induced responses will represent a key development in the search for molecular mechanisms that counter combinations of stress at the same time (Oztetik, 2012).

1.5.3. Light-induced Stress Proteins

The disruption of chloroplast photosynthetic equilibrium leads to photoinhibition and the generation of ROS, resulting in severe photooxidative damage. These photosynthetic imbalances do not only elicit responses to decrease light absorption, remove excess excitation, and scavenge reactive oxygen species but also lead to the up- and down-regulation of photosynthesis-related genes (Demmig-Adams and Adams, 1992). At the epicenter of up-regulated genes lies the Early Light-induced Stress Proteins (*ELIPs*).

1.5.3.1. Early Light-induced Stress Proteins

ELIP was initially identified in pea as a transiently expressed protein present at the early stage of greening after an etiolated-to-light transition (Grimm and Kloppstech, 1987; Kolanus et al., 1987; Grimm et al., 1989; Meyer and Kloppstech, 1984). They are CAB (Chlorophyll a/b binding) -related proteins that are targeted to the non-appressed region of the thylakoid membranes (Grimm et al., 1989; Adamska and Kloppstech, 1991) and encoded by nuclear genes (Meyer and Kloppstech, 1984; Adamska, 2001). The non-appressed regions of thylakoids are known to be the site for both the assembly and disassembly of pigment-protein complexes (Adamska, 2001). When ELIPs reach the non-appressed regions of the thylakoid membranes they are inserted in such a manner that the carboxy terminus faces the luminal side of the membrane whereas the amino-terminus faces the stromal side (Grimm et al., 1989). However, unlike other CAB proteins which are constitutively expressed in thylakoids, ELIPs show transient accumulation

patterns in response to a wide range of environmental stresses (Meyer and Kloppstech, 1984; Grimm and Kloppstech, 1987; Grimm et al., 1989; Adamska, 1997; Adamska, 2001).

After the translation of ELIPs is completed in the cytoplasm, they are imported into the chloroplast and subsequently inserted into the thylakoid membranes via a pathway involving chloroplast signal recognition particle 43 (cpSRP43) (Hutin et al., 2002). Furthermore, it has been shown that the binding of ELIPs to chlorophyll (Chl) a is an essential requirement that determines the stable insertion of the proteins into etioplast membranes (Adamska et al., 2001). On a further note, ELIP transcripts and proteins are expressed much faster than the other light-induced genes during de-etiolation, yet, they disappear before chloroplast development is completed (Jansson, 1999).

A characteristic property of ELIPs is the presence of two internally duplicated transmembrane helices (I and III) (Adamska, 2001) located at an angle of 56° to each other and a second helix (II) which is almost perpendicular to the thylakoid membrane plane (Green and Kühlbrandt, 1995; Adamska, 1997) and appears to contain several polymorphic regions but overall is conserved. Despite the diversity of its amino acid content, the second helix shows significant conservation within monocot or dicot plant species (Adamska, 2001). It has been found that ELIPs are rapidly and transiently expressed in green tissues, during the development of photosynthetic units in the first hours of the greening process (Barrero et al., 2014), the early stages of seedling development (Adamska, 2001) and seed germination (Rizza et al., 2011) in response to high light stress.

1.5.3.2. Structure and evolution of ELIPs

ELIPs belong to the CAB/ELIP/HLIP superfamily of proteins (Montane and Kloppstech, 2000) and consist of more than 100 different stress proteins found in prokaryotic and eukaryotic photosynthetic organisms. Photosynthetic eukaryotes are classified based on their antenna systems: (1) chlorophyll a/b antennas occur in Chlorophytes (green algae and higher plants) (2) chlorophyll a/c antennas occur in Chromophytes (3) chlorophyll a and phycobilisomes occur in Rhodophytes (red algae) (Wolfe et al., 1994; Dunford et al., 1999).

Nonetheless, at the amino acid level, chlorophyll a/b-binding and fucoxanthin-chlorophyll a/c binding antenna proteins of photosystem I and II share sequence homology with many members of the CAB/ELIP/HLIP superfamily (Heddad and Adamska, 2002). ELIPs belong to the LHCs family of proteins

which contain than 30 different members including several distant relatives (Adamska, 1997; Jansson, 1999; Montané and Kloppstech, 2000). ELIPs were discovered as three helix proteins in pea but in recent years additional members have been discovered in other species. Despite having conserved amino acid sequences their secondary structures vary based on the number of helices present. Thus, ELIPs can be classified based on their predicted secondary structure: one-helix proteins (Ohps) (Heddad and Adamska, 2002; Andersson et al., 2003) also called Hlips (high light-induced proteins) or Scps (small CAB-like proteins) (Adamska 2001; Heddad and Adamska, 2002), two-helix Seps (stress-enhanced proteins) (Heddad and Adamska 2000), three-helix Elips and related proteins (Heddad and Adamska, 2002), and the four-helix PsbS protein of PSII (Funk, 2001). The discovery of cyanobacterial Hlips and the SCPs which share the same conserved residues with the eukaryotic CAB proteins clearly points to common ancestor (Wolfe et al., 1994; Green and Kuhlbrandt, 1995). Sequence homologies suggest that Hlips are the distant ancestors of the entire CAB/ELIP/HLIP superfamily (Green and Kühlbrandt, 1995; Montane and Kloppstech, 2000). Over the year, multiple theories postulating the origin of members of CAB/ELIP/HLIP superfamily have emerged, but the conservation of amino acid composition in helices I and III as well as II and IV of the PsbS protein (Funk, 2001) favors the following scenario: During evolution, the two-helix ancestor of CAB proteins arose from the fusion of two Hlip/Scp-type genes. Then, the PsbS protein arose from an internal gene duplication of a two helix Sep-like protein, thereby creating the four-helix intermediate. Finally, the three-helix ancestor of ELIP and CAB proteins arose from the deletion of the fourth helix of PsbS (Green and Pichersky, 1994). This scenario postulates that the dissipation of absorbed energy in the form of heat or fluorescence was the primary function of the ancient antenna not light-harvesting (Montane and Kloppstech, 2000).

Traditionally, ELIPs and CAB proteins have been thought to contain only three predicted α -helices with a conserved consensus sequence present in the first and the third helices (Grimm et al. 1989; Green and Pichersky 1994; Green and Kühlbrandt 1995; Jansson, 1999). Based on the structural analysis conducted by Kühlbrandt et al. (1994) on CAB proteins and due to the structural similarity between CABs and ELIPs, it could be assumed that helices 1 and 3 in ELIPs are kept together by ion pairs formed by charged residues (Adamska 1997, 2001). Similarly, helices I and III of ELIPs show very high similarity with the corresponding regions of the LHCIs and LHCIIIs (Grimm et al., 1989; Green and Kühlbrandt, 1995; Jansson, 1999). The structural similarity of ELIPs with the LHCs further supports a transient pigment-binding function to the xanthophyll lutein (Adamska et al., 1999) to at least four Chl a molecules (Green and Kühlbrandt, 1995; Heddad and Adamska, 2000).

1.5.3.3. Functions of ELIP

It has been thought that ELIPs function as photoprotectants when only exposed to light stress but recent research has shown that their upregulation is possible in a light-independent fashion under dehydration stress (Alamillo and Bartels, 2001; Zeng et al., 2002; Lee et al., 2006; Liu et al., 2019), cold stress (Adamska and Klaus Kloppstech, 1994; Montané et al., 1996; Montané et al., 1997; Lee et al., 2006; Hayami et al., 2015), heat stress (Harari-Steinberg et al., 2001; Lee et al., 2006; Martins et al., 2016), and salinity stress (Zeng et al., 2002; Lee et al., 2006). Moreover, in all surveyed resurrection plants, ELIPs are universally up-regulated in response to desiccation. Since photosynthesis is sensitive to dehydration, the up-regulation of ELIPs may help facilitate the rapid recovery of most resurrection plants upon rehydration (VanBuren et al., 2019). Similarly, antioxidants and xanthophyll cycle pigments are not only abundant when light stress responses are induced but also when exposed to other stresses such as cold heat stress (Martins et al., 2016). Despite having extensive research on ELIPs, there is no consensus detailing their physiological role because the mechanism by which it elicits the response has not been experimentally elucidated as is the transcription regulation of ELIPs (Harari-Steinberg et al., 2001). Nonetheless, there have been several suggestions (Krol et al., 1995; Braun et al., 1996; Bruno and Wetzel, 2004; Tzvetkova-Chevolleau et al., 2007).

The primary function of LHCs, which are members of the CAB/ELIP/HLIP superfamily, is to intercept solar photons and deliver the excitation energy to their core photosystem reaction centers. That is, they possess a primary light-harvesting function. Nevertheless, the LHCs differentially bind to xanthophyll cycle pigments and carotenoids which include lutein, neoxanthin, and vioxanthin. Three main functions have been assigned to these carotenoids which include aiding in light harvesting, structurally stabilizing the LHCs, and protecting the photosynthetic apparatus against photooxidative damage. The latter process proceeds directly or indirectly and promotes the dissipation of excitation energy (Barros and Kühlbrandt, 2009). A substantial decrease in the abundance of D1 protein, photoinactivation of PSII (Adamska et al., 1996), and pigment disorganization (Tzvetkova-Chevolleau et al., 2007) are detected upon HL stress. The decrease in abundance of D1 protein is the result of degradation of the D1/D2 heterodimer of the PSII reaction center and LHC II. Subsequently, the degradation of the heterodimers is associated with the release of triplet chlorophyll which is easily sensitized and forms triplet chlorophyll, which in turn reacts with molecular oxygen to generate the photodamaging singlet oxygen species (Asada, 1994).

Concurrent with decrease in abundance of D1 proteins and photosynthetic activity, the inhibition of transcription of LHCs and activation of synthesis of ELIPs, is induced upon high light treatment (Pötter and Kloppstech, 1993). However, unlike LHCs, multiple research papers have proposed a non-photosynthetic photoprotective function for ELIPs. These proteins have been implicated in the transient binding to free Chl a thereby preventing the formation triplet chlorophyll and its downstream byproduct singlet oxygen (Bartels et al., 1992; Adamska et al., 1992; Funk et al., 1995; Adamska, 1997; Montane and Kloppstech, 2000; Alamillo and Bartels, 2001; Hutin et al., 2003) as well as binding xanthophyll cycle pigments thereby acting as a potent non-photochemical quencher to dissipate excess light energy (Krol et al., 1995; Krol et al. 1999; Montané and Kloppstech, 2000). As a non-photochemical quencher, ELIPs would function at two levels: controlling the amount of captured light energy and the dissipation of energy (Montane et al., 1997).

Mutants lacking Lhcb1 (type I LHC II) and Lhcb6 (CP 29) demonstrated an increase in the expression of *ELIPs* and xanthophyll cycle pigments within the first hour of high light treatment. Whereas the expression levels of PSBS proteins were not affected, thereby prompting the hypothesis which states that: the pool of photo-transformable xanthophyll cycle pigments must be associated with proteins other LHCs and PSBS or persist freely. Thus, it has been postulated that carotenoids do not exist freely in the thylakoid membranes and suggested that they are bound to ELIPs. The findings were also intriguing especially because under conditions of light stress and in the absence of LHCs, ELIPs substituted the normal LHCs and bound directly to the core antenna of PSI and PSII (Krol et al., 1995). But their assumption is not necessarily true because free carotenes and xanthophylls have been found to occupy the hydrophobic regions of membrane bilayers where they aid in rigidifying the membranes (Gruszecki and Strzalka, 2005).

Thylakoid dynamic states modulated by the ubiquity of ELIPs points to the involvement of these proteins in normal developmental processes in addition to prevention of photooxidative-induced damage (Bruno and Wetzel, 2004). For instance, the early expression of ELIPs upon illumination of etiolated seedlings and during chloroplast development lead researchers to hypothesize a role for them in plastid development. Namely, as a last step in photosystem assembly, ELIPs could possibly acts as a structural proxy for LHCs (Adamska, 2001). Additionally, ELIPs have been implicated in modulating chlorophyll content and disrupting the chlorophyll biosynthesis pathway (Tzvetkova-Cheolleau et al., 2007) as well as thylakoid biogenesis during de-etiolation (Bruno and Wetzel, 2004). Moreover, the accumulation of *ELIP* mRNA during the earliest stages of chloroplast-to-chromoplast transition in tomato have led to the

proposal of a newly-recognized function for ELIPs in fruit development and ripening (Bruno and Wetzel, 2004). Furthermore, the loss of DOF AFFECTING GERMINATION 1 (DAG1) protein in *dag1* mutants resulted in deregulation of *ELIPs*. The transcription factor DAG1 appears to influence the development of chloroplasts during seed maturation, namely by affecting the transcription of chloroplast proteins such as *ELIP1* and *ELIP2*. Thus, it is possible that DAG1-mediated *ELIP* expression regulates seed germination under abiotic stress conditions (Rizza et al., 2011). During the repair of photo-damaged PSII reaction centers, ELIPs could be involved in the stabilization and insertion of newly synthesized D1 proteins (Potter and Kloppstech, 1993). The multigenicity of *ELIPs* represents a major impediment to the systematic analysis of their functions *in vivo* (Hutin et al., 2003). *ELIP* genes within the same plant might vary in their transcriptional and/or posttranslational regulation and could be differentially expressed during different stages of dehydration and rehydration. Differential expression of *ELIP* isoforms upon exposure to the same stimuli has been demonstrated in *Arabidopsis thaliana*, *Tortula ruralis*, *Crocus sativus*, and *Syntrichia caninervis* (Harari-Steinberg et al., 2001; Zeng et al., 2002, Bruno and Wetzel, 2004; Casazza et al., 2005; Heddad et al., 2006; Rossini et al., 2006; Tzvetkova-Cheolleau et al., 2007; Rizza et al., 2011; Ahrazem et al., 2016; Liu et al., 2020).

1.6. Linderniaceae

Desiccation tolerance is trait that has been reported in many organisms including mosses (Zeng et al., 2002), green algae (VanBuren et al., 2019), ferns (Liu et al., 2020), in most bryophytes, lichens (Bartels and Salamini, 2001; Oliver et al., 2005; Porembski, 2011), seeds of flowering plants (Oliver et al., 2000), and in the vegetative tissues of around 300 angiosperm species termed resurrection plants (Porembski and Barthlott, 2000; Alpert, 2006). Although desiccation tolerance is an infrequently occurring trait, desiccation tolerant plants are found across many taxa except gymnosperms (Alpert, 2000).

1.6.1. Resurrection plants of the family Linderniaceae

Resurrection plants can be found in regions which suffer from seasonal and sporadic rainfall and weather. They favor growing at low to moderate elevations on rocky outcrops in tropical and subtropical climates (Porembski and Barthlott, 2000). Nonetheless, the most painstakingly studied resurrection plant is the dicotyledonous South African *Craterostigma plantagineum* (Bartels et al., 1990). It is a member of the genus Linderniae, a tribus of the Scrophulariaceae which is a large heterogenous family of polyphyletic

origin. It has been estimated that 170 Linderniae species have been discovered with most of them being endemic to Africa (Fischer, 2004).

Resurrection plants identified within the angiosperms could either be monocotyledonous or dicotyledonous plants. On one hand, monocotyledonous plants are distributed among various different families whereas dicotyledonous plants are represented mainly in the families of Scrophulariaceae and Myrothamnaceae. Phylogenetic analysis of Scrophulariaceae suggests a clustering of desiccation-tolerant plants represented by the genera *Craterostigma* and *Lindernia* (Rahmanzadeh et al., 2005). Microsynteny between desiccation-related genes support this clustering (Phillips et al., 2008). It has been long thought that the aforementioned two genera and their relatives are members of the family Scrophulariaceae in the order Lamiales (Takhtajan, 1997). However, large-scale parsimony and Bayesian analysis largely support the monophyly of the lineage encompassing the genera *Craterostigma*, *Lindernia*, *Artanema*, *Torenia*, and *Crepidorhopalon*. This monophyletic lineage has been annotated as Linderniaceae (Rahmanzadeh et al., 2004; Rahmanzadeh et al., 2005).

1.6.2. Evolution of desiccation tolerance in Linderniaceae

Desiccation tolerance is a trait found even in the most primitive autotrophs such as cyanobacteria, indicating that it evolved very early in evolution (Bartels and Hussein, 2011). It has been postulated that the evolutionary origins of desiccation tolerance in angiosperms lie in their distant ancestors, an assortment of poikilohydrous unicells, that contributed to the formation of the eukaryotic plant cell. Subsequently, cyanobacteria-related endosymbionts which evolved to chloroplasts or other bacteria-related endosymbionts which evolved to mitochondria could have carried the trait into land plants. Desiccation tolerance became a crucial trait allowing primitive fresh water plants (Mishler and Churchill 1985; Oliver et al., 2000) to colonize relatively dry land environments (Kappen and Valladares, 1999).

The reproductive and vegetative tissues of many species of algae, bryophytes, ferns and fern allies are desiccation tolerant (Gaff and Oliver, 2013) and phylogenetic evidence suggests that facultative adaptation to vegetative desiccation existed in the most basal clades of bryophytes (liverworts, hornworts, and mosses) (Oliver et al., 2000; Gaff and Oliver, 2013). Primitive tolerance to vegetative desiccation is very similar to modern-day desiccation tolerant bryophytes and involves constitutive cellular protection paired with active cellular repair. Nevertheless, vegetative desiccation tolerance came at a cost. This is manifested in the low metabolic in desiccation tolerant plants compared to desiccation-sensitive plants

(Oliver et al., 2000). For example, desiccation tolerant angiosperms are metabolically primed for dehydration but major metabolite changes occur upon dehydration (Dinakar and Bartels, 2013). Nonetheless, with the evolution of tracheophytes, vegetative desiccation was lost but reproductive desiccation was retained. This is attributed to the noticeable increase in morphological, developmental, and physiological complexity of the evolving land plants. Thus, selection favored vascular plants capable of internalizing water relationships and maintaining efficient carbon fixation. Vascular plants retained their desiccation tolerant genes in their seeds while genes of vegetative desiccation involved in cellular protection and repair were recruited for different but related processes (Oliver et al., 2000). Furthermore, it has been suggested that desiccation tolerance has independently evolved or re-evolved in Selaginellales and leptosporangiate ferns (Bartels and Hussein, 2011).

Similarly, vegetative desiccation tolerance occurs in 13 mainly unrelated angiosperm families, implying that it has evolved separately at least 13 times either as a result of independent evolution or re-evolution (Bartels and Hussein, 2011; Gaff and Oliver, 2013). Each of the angiosperm lineages exhibit unique and different responses to dehydration, desiccation, and rehydration. It is hypothesized that desiccation tolerance in seeds evolved from a primitive form of vegetative desiccation tolerance (Oliver et al., 2000). Once established in seeds, resurrection plants have redirected the expression of desiccation tolerant-related genes from the seeds to vegetative tissues (Illing et al., 2005) by environmental cues related to drying.

This has been supported by comparative genomic studies conducted on desiccation tolerant and sensitive species such as *Lindernia brevidens* and *Lindernia subracemosa*. VanBuren et al. (2018a) have shown that desiccation tolerance evolved through a combination of gene duplications and network-level rewiring of existing seed desiccation pathways. Another research paper revealed that massive tandem proliferation of desiccation-induced genes in *L. brevidens* supports the convergent evolution of desiccation tolerance in land plants through tandem gene duplication (VanBuren et al., 2019). Surprisingly, despite being endemic to the rainforests of coastal Africa which does not experience any drought, *L. brevidens* retained its desiccation tolerance. It has been hypothesized that *L. brevidens* preserved vegetative desiccation tolerance through genome stability (Phillips et al., 2008). In their quest to provide reliable data on stress tolerance and elucidate the role of differential stress responses, researchers have been carrying out comparative studies on the genotypes of tolerant and sensitive organisms within a species. Since differential stress responses can arise due to differences in sets of genes, a systems level study of the differences between genotypes could shed the light on how such sets of genes elicit their responses.

Interestingly, some members of the Linderniaceae are desiccation tolerant (e.g., *L. brevidens*) while others are desiccation sensitive (e.g., *L. subracemosa*). This observation is being utilized to understand the evolution of desiccation tolerance (Bartels and Hussein, 2011). The availability of a dual comparative system coupled with the fact that the desiccation results in severe damage to the photosynthetic apparatus make the genera Lindernia the perfect model system for studying desiccation tolerance.

As a consequence of the above-mentioned studies, ELIPs which are responsible for protecting the photosynthetic apparatus from degradation upon dehydration in the two homiochlorophyllous Linderniaceae species, *L. brevidens* and *L. subracemosa*, were analyzed and reviewed.

1.7. Objectives of the study

Among the many studied abiotic stresses, extreme light stress (high or low) represents a huge challenge for food production, namely due to the prevailing combinatorial effects of extreme light and dehydration during the summer season. This work will provide more insights into the stress-responsive genes of the newly sequenced genomes of the homiochlorophyllous desiccation tolerant *L. brevidens* and its desiccation sensitive counterpart *L. subracemosa*. The Master thesis aims at identifying several light-responsive genes, namely the ones coding for the Early Light-induced Proteins (*ELIPs*) in addition to analyzing both their transcript expression profiles and amino acid sequences. *ELIPs* are expressed as a result of the combinatorial effect of illumination with high incident light and desiccation treatment. They are involved in protecting the photosynthetic apparatus by binding to Chl a and preventing its degradation. Thereby indirectly protecting the plant from photooxidative stress by preventing the formation of triplet chlorophyll, a potent photosensitizer that causes oxidative damage by converting harmless molecular oxygen into harmful singlet oxygen. The analysis will include comprehensive comparisons between the regulation of *ELIP* gene expression in the resurrection plant *L. brevidens*, *L. subracemosa*, other *ELIPs* belonging to the family of Linderniaceae, and some homologous protein members of the CAB/ELIP/HLIP superfamily of distantly related polypeptides. Ultimately, the road will be paved for other researchers to use the provided data to explore the crosstalk between dehydration and light stress as well as explore new frontiers to exploit this crosstalk. This research can provide leads to more efficiently breed concurrent tolerance to crops without damaging the performance of plants nor curtailing their yields.

2. Materials and Methods

2.1. Materials

2.1.1. Plant material

Fresh and dehydrated plant material for *L. brevidens* (Lbr) Skan and *L. subracemosa* (Lsu) De Wild were provided directly by prof. Dorothea Bartels.

2.1.2. Main stock Buffers and Media solutions

MgCl ₂ solution	100 ml (1 M) was prepared then filtrated
MgSO ₄ solution	100 ml (2 M) was prepared then filtrated
Glucose solution	100 ml (2 M) was prepared then filtrated
10 X TAE buffer	0.4 M Tris acetate; 10 mM EDTA; pH 8
SOC Media	2 g tryptone; 0.5 g yeast extract; 1 ml 1 M NaCl; 0.25 ml 1 M KCl; 1 ml 1 M MgCl ₂ ; 1 ml 1 M MgSO ₄ ; 1 ml 2 M glucose; to 100 ml water. Into 97 ml of water the following ingredients: tryptone, yeast extract, NaCl and KCl were added, stirred, autoclaved, and cooled to room temperature. Next, MgCl ₂ , MgSO ₄ , and glucose solutions were added to the cooled medium. The obtained solution was filtered again then stored at room temperature.
LB Agar	35 g/l LB agar powder (10 g tryptone, 5 g yeast extract, 5 g NaCl, 15 g agar)
LB liquid	20 g/l LB powder (10 g tryptone, 5 g yeast extract, 5 g NaCl)

2.1.3. Primers

Primers designed for Lbr elongation factor 5-alpha (*EF5α*) and Lsu *EF5α* will act as positive controls.

Table 1. Primers designed to amplify the conserved gene *EF5α*.

Organism	Gene of interest	Primer	Sequence
<i>Lindernia brevidens</i>	<i>LbrEF5α</i>	Forward	GGGAAAGGACCTTGTGTGA
		Reverse	TGGGCTCATTACTCCACTGA
<i>Lindernia subracemosa</i>	<i>LsuEF5α</i>	Forward	TGGGTGGTTTTATCAATGTCC
		Reverse	GGTTGAGTGATAAACACAGTAGCAA

Two primer design approaches were used to test amplifying genes of interest:

2.1.3.1. First primer design approach

Table 2. Primers designed to partially amplify genes of interest.			
Organism	Gene of interest	Primer	Sequence
<i>Lindernia brevidens</i>	<i>ELIP 016551</i>	Forward	GAGAACGACGGCGTGATG
		Reverse	TTCTCATGTAGGCACACCATT
	<i>ELIP 016696</i>	Forward	ATGTCTTCTTGTGGAGGCCG
		Reverse	ATCAAAAGGGCGTGAGTGA
<i>Lindernia subracemosa</i>	<i>ELIP 008078</i>	Forward	ACCGATTGTGTCAGGCGA
		Reverse	AGACAATATTGCAGGCCATT
	<i>ELIP 030118</i>	Forward	GTCAAGTGGTTCCTGCTGAC
		Reverse	ACAAACCTCGACTTCCAACCA
	<i>ELIP 024731</i>	Forward	ATTCTCCCAGGTGCTAACG
		Reverse	TCGGCGGCTCACTAACATTA

2.1.3.2. Second primer design approach

Table 3. Primers designed to amplify the full length of the gene of interest.			
Organism	Gene of interest	Primer	Sequence
<i>Lindernia brevidens</i>	<i>ELIP 016551</i>	Forward	GAGCATCGACACAAACACACC
		Reverse	TTCTCATGTAGGCACACCATT
	<i>ELIP 025593.27</i>	Forward	ATCGGAAGTGCTACAACGGA
		Reverse	CGAAAGGGGTGTGACACTTT
	<i>ELIP 025593.16</i>	Forward	ATCGGAAGTGCTACAACGGA
		Reverse	AATGAGGGTTGCGAGTGAGA
<i>Lindernia subracemosa</i>	<i>ELIP 030118</i>	Forward	TCGCGCACTACATATGTCCT
		Reverse	ACGACAACAACCACCACTA

2.1.4. Devices

Type of device	Industrial name	Manufacturer
Vortex	Vortex Genie gene name 2	Scientific Industries Inc., New York, USA
UV illuminator	Intas UV systems series CONCEPT	Intas Pharmaceuticals Ltd., Gujarat, IN
Horizontal Gel Electr.	Compact S/M	Biometra GmbH, Göttingen, DE
pH-meter	-----	SCHOTT GLAS, Mainz, Germany
Power Supply	Electrophoresis power supply	Gibco BRL, Carlsbad, CA
Balance	BL 1500 S Balance	Sartorius, Göttingen, Germany
Nanodrop machine	Biospec – Nano	Shimadzu Biotech, Japan
Heating block	QBT digital block heater	Grant Instruments Ltd, Cambridgeshire, UK
Centrifuge I	Microcentrifuge 5415D	Eppendorf, Hamburg, Germany
Centrifuge II	Microcentrifuge 5417R	Eppendorf, Hamburg, Germany
Centrifuge III	Multi-purpose centrifuge 5810R	Eppendorf, Hamburg, Germany
PCR machine	T3 Thermocycler	Biometra, Göttingen, Germany
Realtime qPCR machine	RealPlex 2 qPCR Real Time PCR ThermoCycler ep Gradient S Silver 96	Eppendorf, Hamburg, Germany

2.1.5. Enzymes and Vectors

- DNA ladder (Fermentas GmbH, St. Leon-Rot, Germany)
- DNase I (Thermo Fisher Scientific, St Leon-Rot, Germany)
- RevertAid Reverse Transcriptase (Thermo Fisher Scientific, St Leon-Rot, Germany)
- Phusion High-Fidelity DNA Polymerase kit (Thermo Fisher Scientific, St Leon-Rot, Germany)
- CloneJET PCR Cloning Kit which contains pJET1.2/blunt cloning vector (Thermo Fisher Scientific, St Leon-Rot, Germany)

2.2. Methods

2.2.1. Standard molecular biology techniques and nucleotide sequence analysis

Illumina RNAseq data were downloaded from the National Center for Biotechnology Information sequence read archive under the following BioProjects: *L. brevidens* (PRJNA488068), *L. subracemosa* (PRJNA488068). *ELIP* sequences for *L. brevidens* and *L. subracemosa* were retrieved by BLAST from local RNA-Seq data by using the *C. plantagineum* desiccation stress protein DSP-22 (also called *CpELIP*) nucleotide sequence as query. Similarly, putative homologs from the CAB/ELIP/HLIP superfamily were identified on Blastn (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch) (McGinnis et al., 2004) using the *DSP-22* nucleotide sequence as query. Primer synthesis and plasmid sequencing were carried out by Eurofins MWG Operon (<http://www.eurofinsgenomics.eu>).

2.2.2. Phylogenetic analysis

Putative homologs from the CAB/ELIP/HLIP superfamily as well as some homologous members of the Carotene biosynthesis-related proteins (CBR) were identified on Blastp (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>) with the amino acid sequence of DSP-22 (Uniprot accession number: Q01931 DS22_CRAPL) (The Uniprot Consortium, 2019) as the query. The obtained amino acid sequences were used to construct phylogenetic trees. In total 57 accessions were used in the phylogenetic analysis:

Table 4. Uniprot accession numbers for proteins used in phylogenetic analysis.			
Clade	Organism	Protein	Uniprot Accession ID
Eudicots	<i>Craterostigma plantagineum</i>	CpELIP (dsp-22)	Q01931
	<i>Camellia sinensis</i>	ELIP2	C9E439
	<i>Capsicum annuum</i>	ELIP	F8RJ00
	<i>Citrus sinensis</i>	ELIP	H9A1Z7
	<i>Arabidopsis thaliana</i>	ELIP1	P93735
		ELIP2	Q94K66
	<i>Lindernia subracemosa</i>	ELIP 030118	Not annotated
		ELIP 024731	Not annotated

		ELIP 008078	Not annotated
		ELIP 021038	Not annotated
	<i>Lindernia brevidens</i>	ELIP 025540	Not annotated
		ELIP 025541	Not annotated
		ELIP 025542	Not annotated
		ELIP 025543	Not annotated
		ELIP 025544	Not annotated
		ELIP 025545	Not annotated
		ELIP 025546	Not annotated
		ELIP 025547	Not annotated
		ELIP 025548	Not annotated
		ELIP 025549	Not annotated
		ELIP 025584	Not annotated
		ELIP 025585	Not annotated
		ELIP 025587	Not annotated
		ELIP 025589	Not annotated
		ELIP 025591	Not annotated
		ELIP 025593	Not annotated
		ELIP 017192	Not annotated
		ELIP 016696	Not annotated
		ELIP 016551	Not annotated
		ELIP 025581	Not annotated
		ELIP 025579	Not annotated
		ELIP 025583	Not annotated
Monocots	<i>Triticum aestivum</i>	wcr12	Q9SXP3
	<i>Zea mays</i>	ELIP	Q94KE7
	<i>Oryza sativa subsp. japonica</i>	PSBS1	Q943K1
		PSBS2	Q0J8R9
	<i>Oryza sativa subsp. indica</i>	PSBS1	A2WXD9
		PSBS2	A2XZ94

	<i>Hordeum vulgare</i>	ELIP9_ HV90	P14897
		ELIP6_ HV60	P14896
		ELIP5_ HV58	P14895
	<i>Crocus sativus</i>	ELIPa	A0A1S5T4W8
		ELIPb	A0A1S5T4X4
		ELIPc	A0A1S5T4X9
		ELIPd	A0A1S5T4X8
Lycophytes	<i>Selaginella moellendorffii</i>	ELIP2_1	D8R6E9
		ELIP2_2	D8S4K5
Bryophytes	<i>Physcomitrella patens</i>	ELIP12	A9TAX2
		ELIP13	A9TLZ1
Chlorophytes	<i>Dunaliella salina</i>	CBR	P27516
	<i>Micromonas commoda</i>	CBR	C1EEI3
	<i>Micromonas pusilla</i>	CBR/ELIP1	C1N708
	<i>Chlamydomonas reinhardtii</i>	PSBS1	A8HPM2
		PSBS2	A8HPM5
Spermatophytes	<i>Pinus densata</i>	ELIP	J7H2Q2
	<i>Pinus yunnanensis</i>	ELIP	J7H8G1
	<i>Pinus tabuliformis</i>	ELIP	J7H9L8

The amino acid sequences of the aforementioned accessions were aligned in silico by the MEGA X program (<http://www.megasoftware.net>) (Kumar et al., 2018) using ClustalW (a built-in option in MEGA X). To optimize amino acid alignments, the settings recommended by Hall (2013) were used: Multiple Alignment Gap Opening penalty was changed to 3 and the Multiple Alignment Gap Extension penalty was changed to 1.8. Alignments were saved as .mas and .meg files. The .meg file will be used for carrying out phylogenetic analysis. The generation of phylogenetic trees was performed using the maximum likelihood (ML) and maximum parsimony (MP) methods. The ML tree was based on the JTT matrix-based model (Jones et al., 1992) with Gamma-Distributed rates. The tree with the highest log likelihood was chosen (Figure 8). The MP tree was based on the Tree-bisection–reconnection branch swapping model (Nei and Kumar, 2000) with the number of initial trees set to 10 and the maximum number of trees to retain set to 100. The tree is shown in Figure 9. The reliability of the two phylogenetic trees was

evaluated using the bootstrap test (Felsenstein, 1985). Bootstrap values for ML were calculated based on 1000 replications whereas on 2000 replications for MP. Generated trees were modified using TreeGraph2 software (Stöver and Müller, 2010).

2.2.3. Amino acid sequence analysis

For carrying out alignment analysis, some accessions belonging to previously annotated members of the CAB/ELIP/HLIP superfamily were used along with other ELIPs three of which belong to *L. brevidens* (not annotated) and three to *L. subracemosa* (not annotated) in addition to three Uniprot annotated accessions belonging to *Glycine max* CABS: CAB (P12471), CAB2 (P09755), and CAB3 (P09756). Amino acid sequence alignments were performed with the multiple alignment program Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (Sievers et al., 2011). Hydropathy plots were performed according to Kyte and Doolittle (1982) using ExPASy ProtScale (<https://web.expasy.org/protscale/>) (Gasteiger et al., 2005). FoldIndex online tool was used to predict the foldability (i.e., folded, partially folded or intrinsically unfolded) of the ELIPs, CABS, CBRS, and PSBSs (<https://fold.weizmann.ac.il/fldbin/findex>) (Prilusky et al., 2005). FoldIndex is based on mean residue hydrophobicity and mean net charge of the sequence and uses the parameters outlined by Uversky et al. (2000). The predicted charge at pH 7.00 for all proteins was evaluated using PROTEIN CALCULATOR v3.4 (<http://protcalc.sourceforge.net/>).

The length of amino acid sequences as well as their predicted molecular mass, predicted pI, and amino acid content were evaluated using the online tool ProtParam (<https://web.expasy.org/protparam/>) (Gasteiger et al., 2005). The identification PEST amino acid residues involved in the rapid destruction of the peptide in vivo were identified using the online tool Epestfind (<https://emboss.bioinformatics.nl/cgi-bin/emboss/epestfind>) (Rechsteiner M, Rogers, 1996). The online Protein Subcellular Localization Prediction tool, WoLF PSORT (<https://wolfsort.hgc.jp/>) was used to predict whether the signal peptides of ELIPs are imported into chloroplasts.

Primary amino acid data analysis was performed on Phyre2. The online tool was used to determine the length and position of signal peptides and transmembrane helices as well as construct the initial tridimensional protein models as PDB files. Then, the PDB files were downloaded from Phyre2 servers and used as input in PyMol (Student-licensed Version 2.3) to edit the tridimensional protein models and identify individual amino acid residues. Phyre2 Investigator was run to determine active sites the

tridimensional protein models. All ELIPs were modeled based on c6ijo2, the lhca2 of photosystem i of *chlamydomonas reinhardtii*. fPocket (<http://fpocket.sourceforge.net/>) (Le Guilloux et al., 2009), an algorithm built-in Phyre2 was used to preliminarily detect large pockets, a precursor which has been found to be associated with the location of active sites in protein domains. The obtained results for the positions of signal peptides and transmembrane helices were cross-checked by using Signal peptides were predicted using SignalP 5.0 Server (<http://www.cbs.dtu.dk/services/SignalP/index.php>) (Almagro Armenteros et al., 2019) and PRED-TMR (<http://athina.biol.uoa.gr/PRED-TMR/>) (Pasquier et al., 1999) respectively.

2.2.4. Plant treatment

L. brevidens Skan and *L. subracemosa* De Wild were grown by my colleagues as described by Bartels et al. (1990) and Dinakar and Bartels (2012). For the dehydration treatment, fully grown 6- to 8-week-old plants were gradually dehydrated in pots as described by Giarola (2008). Relative water content (RWC) measurements were made according to Bernacchia et al. (1996). For both species, dehydration treatment included: untreated (100% RWC), early dehydrated (60% RWC), late dehydrated (30% RWC), and fully desiccated (2% RWC). However, for *L. brevidens* the desiccated plant was rehydrated by submerging it in water for 24 h. RWC measurements were recorded by Dr. Valentino Giarola according to Giarola (2008) and Pieczynski et al. (2012):

$$\text{RWC [\%]} = \frac{\text{fresh weight} - \text{dry weight}}{\text{turgid weight} - \text{dry weight}} \times 100$$

Leaves from each dehydration treatment were detached, ground using liquid nitrogen, and stored in -80°C for future use.

2.2.5. Preparation of DEPC treated water

Diethylpyrocarbonate (DEPC) treated water was prepared by adding 1 ml of DEPC to 1 L of distilled water. The bottle was shaken rigorously for a couple of minutes and subsequently autoclaved. All solutions used for RNA analysis and total RNA extraction will be prepared with DEPC treated water to ensure that they are nuclease-free.

2.2.6. Preparation of colony-selection plates

LB agar was prepared by adding 35 g agar in 1 L of water. The solution was autoclaved and then let to cool up to 40-50°C. Next, 100 µl ampicillin (final concentration of 100 mg/l) was added and LB agar media was sequentially poured into plastic sterile plates (150 mm x 15 mm). Finally, plates were left to cool under the laminar flow hood and stored at 4°C for future use.

2.2.7. Total RNA isolation (Valenzuela-Avendano et al., 2005)

With the help of a pistol, mortar, and liquid nitrogen, a 200 mg of plant material was finely grinded into powder and subsequently transferred into pre-cooled 50 ml falcon tubes. Upon the successful grinding, 50 to 100 mg of the ground plant material was transferred to 2 ml Eppendorf tubes while the remaining plant material was stored at -80°C for future use. Next, 18.6 ml of the extraction buffer was mixed with 11.4 ml Tris-saturated phenol. This was followed by adding 1.5 ml of the obtained mixture into the ground tissue. The samples were homogenized and incubated at room temperature for 10 min then centrifuged at (10,000 g, 10 min, 4°C). The supernatants were transferred to fresh 1.5 ml Eppendorf tube and 300 µl of Chloroform/isoamylalcohol (24:1) was added. The final solutions were rigorously mixed by vortexing for 10 sec and centrifuged (10,000 g, 10 min, 4°C). The upper aqueous phase was carefully removed and transferred to new 1.5 ml Eppendorf tubes, 375 µl of isopropanol and 375 µl of 0.8 M sodium citrate were added to each Eppendorf tube, mixed thoroughly by vortexing for 30 sec, incubated at room temperature for 10 min, and subsequently centrifuged (12,000 g, 10 min, 4°C). The supernatants were discarded and the pellets were washed with ice-cold (-20°C) 70% Ethanol. After brief centrifugation (12,000 g, 10 min, 4°C), the supernatant was removed, the pellets were dried at room temperature and later dissolved in 100 µl DEPC treated water. Then 167 µl of 4 M LiCl was added followed by incubation on ice for 2 hours. After a long centrifugation (14,000 g, 20 min, 4°C), the supernatant was discarded and the pellet was briefly washed with 1 ml ice-cold (-20 °C) 70% Ethanol. The washed pellet was centrifuged for one last time (12,000 g, 10 min, 4°C), supernatant was discarded, and pellet was air dried at room temperature under the laminar flow hood. The air-dried pellets were dissolved in 25 µl DEPC-treated water. RNA concentration and purity were determined under by measuring the OD at 230, 260, and 280 nm.

2.2.8. cDNA synthesis using ReverseAid Reverse Transcriptase

One microgram of the isolated total RNA was used to synthesize cDNA.

2.2.8.1. DNase treatment

To ensure the removal of any contaminating DNA, the isolated total RNA was first treated with DNase. The procedure was carried out by mixing 1 µl RNase-free DNase I enzyme, 1 µl 10X DNase I reaction buffer with MgCl₂, 5 µl of PCR grade sterile water, and 4 µl of total RNA in PCR tubes. Subsequently, the mixture was incubated at 37°C for 10 minutes. Afterwards, the reaction was stopped by adding 1 µl of 25 mM EDTA and incubated at 65°C for 10 min. The resulting reaction mixture was briefly spun down, separated into two tubes of equal volumes (5 µl of reaction mixture); one designated for synthesizing cDNA (RT+) while the second will act as a negative control (RT-).

2.2.8.2. cDNA Synthesis

The synthesis of first-strand cDNA was performed by adding 1 µl oligo-dT primer (100pmol/ µl) and 0.5 µl MiliQ water were added to RT+ and Rt- reaction mixtures, gently mixing the PCR, and incubating at 65°C for 5 min. PCR tubes were cooled down on ice and briefly spun down. On one hand, for the treatment of RT+: 2 µl of 5x first strand buffer, 1 µl of 10 mM dNTP Mix, and 0.5 µl of Reverse Transcriptase enzyme were mixed, making a final total volume of 10 µl. On the other hand, for the treatment of RT-: 2 µl of 5x first strand buffer, 1 µl of 10 mM dNTP Mix, and 0.5 µl of MiliQ Water were mixed, making a total volume of 10 µl. The PCR tubes were gently mixed and subsequently incubated at 42°C for 60 min. Termination was achieved by incubating the samples at 72°C for 5 min. The cDNA product was diluted in 40 µl of MiliQ water to make a total volume of 50 µl.

2.2.9. Determination of nucleic acid concentration

The concentrations of RNA were determined using a Nanodrop machine. The machine automatically provides the concentration of RNA (DEPC water was used as blank) and measurements taken at OD260/230 which determine the quantity and the quality of extracted nucleic acid. If the OD260/230 ratio is significantly lower than 2.0-2.2, it may indicate the presence of contaminants such as proteins which absorb at 230 nm. The equation below were used to convert the OD values into concentration values:

$$1 \text{ OD260} = 40 \text{ ug RNA/ml}$$

2.2.10. RT-PCR and Realtime RT-qPCR analysis

The cDNA was used for amplification of Lbr and Lsu *ELIPs* using the forward and reverse primers described in the materials section. The amplified products were run on 1% (w/v) agarose gels.

Table 5. Reaction mixture (20 µl total volume) used for the amplification of Lbr and Lsu *ELIPs* as well as *EF5α*.

Ingredient	Amount
Template	1 µl
5 x PCR-Phusion buffer	4 µl
dNTPs (10 mM)	0.4 µl
Forward primer (10 mM)	1 µl
Reverse primer (10 mM)	1 µl
Phusion DNA polymerase	0.1 µl
Sterile H ₂ O	Add to 20 µl

Table 6. Touch down PCR with hot start was used for Lsu *ELIPs* 008078, 024731, 030118, and *EF5α*.

PCR Stage	Temperature	Time	Cycle
Initial denaturation	95°C	3 min	X1
Denaturation I	97°C	1 min	X1
Denaturation II	95°C	30 sec	
Primer Annealing I	70°C (Temp sequentially Decreases by 2°C to 54°C)	30 sec	X15
Elongation I	72°C	1 min	
Denaturation II	97°C	30 sec	
Primer Annealing II	55°C	30 sec	X25
Elongation II	72°C	1 min	
Final Elongation	72°C	5 min	
Storage	4°C	Indefinitely	

Table 7. PCR program used for Lbr *ELIPs* 016551, 025593, and 016696.

PCR Stage	Temperature	Time	Cycle
Initial denaturation	96°C	3 min	
Denaturation	95°C	30 sec	X38
Primer annealing	61°C	50 sec	
Elongation	72°C	1 min	
Final elongation	72°C	5 min	
Storage	4°C	Indefinitely	

Note: For *ELIP 016551*, primer annealing time was changed to 30 sec.

2.2.11. Agarose gel electrophoresis

To perform gel electrophoresis 1% (w/v) agarose gels were melted in 1X TAE buffer. The gels were used for checking the quality of extracted RNA, the size separation of nucleic acids, and the determination of quantity and quality of DNA for transformed clones. Immediately after pouring the hot solution, ethidium bromide (0.5 µg/ml) was added, and the agarose gel was let sit for 15-20 min to polymerize. To run RNA, 1 µl RNA was mixed with 1 µl loading buffer (final concentration of 1X) and 8 µl sterile water followed by loading into the gel and running the gel electrophoresis machine at 70 volts. Whereas to run DNA the amplified DNA product was mixed with 2 µl loading buffer (final concentration of 1X) followed by loading into the gel and running the gel electrophoresis machine at 120 volts. Gel bands were visualized using a UV illuminator. Photo were taken with a camera mounted on the UV illuminator and the built-in software was used to modify the images.

10 X loading buffer: 50% (v/v) glycerol; 0.25% (w/v) bromophenol blue

2.2.12. DNA extraction from agarose gels purification (**MACHEREY-NAGEL GmbH & Co. KG**)

NucleoSpin® Gel and PCR Clean-up protocol was carried out to extract amplicons from excised agar gels. To carry out the protocol, excised gel slices were transferred into 2 ml Eppendorf tubes and the weight of each excised gel slice was measured. Next, for each 100 mg of agarose gel 200 µl of Buffer NTI was added. Samples were incubated for 5-10 min at 50°C on a heating block until the gel slices were completely dissolved. In order to bind the amplicons to the silica membrane, the PCR clean-up column was placed in a collection tube (2 ml), loaded up to 650 µl, and centrifuged (11,000 g, 30 sec, RT). The

flow-through was discarded and the remaining samples were loaded and centrifuged as previously described. After discarding the flow-through, the silica membrane was washed with 650 µl of Buffer NT3 then centrifuged (11,000 g, 30 sec, RT). The previous step was repeated for a second time to minimize chaotropic salt-carry over and low A260/A230 values. Next, to ensure the total removal of Buffer NT3, the samples were centrifuged (11,000 g, 1 min, RT) and the flow-through was discarded. The silica membranes were completely dried by transferring the collection tube into a new 1.5 ml Eppendorf tube and incubating at 70°C for 2-5 min. Finally, 15 µl of NE buffer was added, let sit at room temperature for 1 min, centrifuged (11,000 g, 1 min, RT), and stored at -20°C for future use.

2.2.13. Ligation

To generate plasmid DNA constructs, the insert DNA was ligated to pJET1.2/blunt cloning vector. Ligation was carried out by mixing 4 µl insert DNA, 1 µl 10X ligase buffer, 0.5 µl PJet vector, and 0.5 µl T4 ligase making a total volume of 10 µl. The mixture was incubated at 22°C in a PCR Thermocycler for 8-10 min. The obtained constructs were stored at -20°C for future use.

2.2.14. Preparation of electrocompetent *Escherichia coli* cells (RbCl method)

Cells from *E. coli* DH10B strain were used to make electrocompetent cells. This was carried out by inoculating a single colony of *E. coli* into 5 ml LB liquid medium. Cells were incubated at 37°C and grown overnight with moderate shaking (220 rpm). Then, 1 ml of the over-night grown culture was used to inoculate 50 ml of LB liquid media in a sterile 500 ml flask. The bacterial culture was incubated at 37°C with moderate shaking (220 rpm) until the OD600 reached a value ranging between 0.35 and 0.45. After centrifugation (4000 rpm, 15 min, 4°C) the supernatant was discarded and the pellet was resuspended in 15 ml ice cold TBFI with a pipette and subsequently incubated on ice for 10 min followed by another round of centrifugation (4000 rpm, 10 min, 4°C). The supernatant was eliminated and the pellet re-suspended once more in 15 ml ice cold TBFI followed by another round of centrifugation (4000 rpm, 10 min, 4°C). The suspension was incubated on ice for 5 min and centrifuged at the same conditions as previously described. The pellet was resuspended in 2 ml TBFII. Finally, the suspension was aliquoted in 50 µl volumes in separate tubes, instantly frozen in liquid nitrogen, and stored at -20°C for future use.

TFB I: 30mM KAc; 100 mM RbCl; 10 mM CaCl₂ 2H₂O; 50 mM MnCl₂ 4H₂O; 15% (v/v) Glycerol. Adjust pH to 5.8 using 0.2 M acetic acid and filter sterilize.

TFB II: 10mM MOPS; 75 mM CaCl₂ 2H₂O; 10 mM RbCl; 15% Glycerol (v/v). Adjust pH to 6.5 using KOH and filter sterilize.

2.2.15. Transformation of electrocompetent *E. coli* cells

Before carrying out the procedure, it is imperative to prepare a fresh bucket of ice before fetching the DH10B cells from -80°C. Once the bucket has been prepared, the desired number of vials containing DH10B cells were removed from -80°C and immediately placed on ice to thaw. Similarly, all ligation mixtures were placed on ice. A timer is set so that after 5 min 3 µl of ligation mixture is added to one vial and immediately incubated at 42°C for exactly 50 sec to carry out the heat shock treatment. When the time is up, the vials were transferred to ice and let sit for 5 min. Next, 450 µl of SOC media is added and the final mixture is transferred into 15 ml falcon tubes followed by placing them in a shaking incubator (220 rpm) at 37°C for 1 hour. Then, each transformant was plated on 3 different selection plates under the laminar flow hood: the first plate contained 50 µl of transformant, the second plate contained 100 µl of transformants, and the third plate contained 250 µl of transformant. The plates were incubated at 37°C overnight then stored in a 4°C fridge for up to 3 weeks.

2.2.16. Colony PCR

This method was used to quickly screen for positive clones which are plasmids containing the desired insert directly from the bacterial colonies. To carry out colony PCR multiple random clones of the transformed ligation mixture were used. Autoclaved toothpicks were used to pick up individual colonies to inoculate the PCR reaction mixture. The latter was prepared as follows:

Table 8. Colony PCR reaction mixture.	
Ingredient	Volume
Sterile H ₂ O	11.9 µl
Template	Inoculation
5 x PCR-Phusion buffer	4 µl
dNTPs (10 mM)	1 µl
PJet forward primer (10 mM)	1 µl
PJet reverse primer (10 mM)	1 µl
Phusion DNA polymerase	0.1 µl

Table 9. Colony PCR program.

PCR Stage	Temperature	Duration	Number of cycles
Initial denaturation	98°C	30 sec	
Denaturation I	98°C	10 sec	
Primer Annealing I	70°C (Temp Dec.: 0.5°C)	15 sec	X6
Elongation I	72°C	20 sec	
Denaturation II	98°C	10 sec	
Primer Annealing II	67°C	15 sec	
Elongation II	72°C	20 sec	X35
Final Elongation	72°C	60 sec	
Storage	10°C	Indefinitely	

2.2.17. Plasmid Isolation

Two different methods were used for the isolation of plasmids from *E. coli* cultures.

2.2.17.1. Plasmid DNA mini-prep from *E. coli* (Brinboim and Doly, 1979)

Plasmid DNA was extracted from *E. coli* according to Brinboim and Doly (1979). The method was carried out by inoculating positive clones into 20 ml LB medium in a 50 ml Eppendorf falcon tube and subsequently adding 20 µl ampicillin (1:1000 dilution). Bacterial culture was grown overnight at 37°C with continuous shaking (220rpm). The next day bacterial cultures were centrifuged (8000 g, 3 min, RT) and the supernatant was discarded by pouring. To completely eliminate the supernatant a second round of centrifugation was performed at the same previous conditions followed by removing the remaining supernatant with a pipette. Next, 250 µl of Buffer 1 was added and the mixture was rigorously vortexed then transferred into 2 ml Eppendorf tubes. Into each 2 ml Eppendorf tube 250 µl of freshly prepared Buffer 2 was added and inverted gently multiple times until a thick viscous mixture appeared. After incubating the resulting mixture on ice for 3 min, 350 µl of Buffer was added, mixed gently, and centrifuged (15,000 g, 10 min, 4°C). The supernatant was transferred into new 1.5 ml Eppendorf tubes. Next, under the laminar flow hood, 300 µl of Chloroform and 300 µl of Tris-saturated phenol were added into each 2 ml Eppendorf tube, vortexed rigorously for 30 sec, and centrifuged (15,000 g, 3 min, RT). The upper phase was transferred into new 1.5 ml Eppendorf tubes, 0.7 volume of isopropanol was added,

rigorously vortexed and centrifuged (15,000 g, 15 min, 4°C). The supernatant was eliminated and 1 ml of ethanol was added. Samples were centrifuged again (15,000 g; 5 min; 4°C). The supernatant was discarded and to eliminate residual ethanol another centrifugation round was performed (15,000 g; 5 min; RT). The supernatant was once again discarded and the samples were air dried under the laminar flow hood. Finally, the pellets were resuspended in 40 µl of TE/RNase A buffer and samples were stored at -20°C.

Buffer 1: 50 mM Tris-HCL (pH 7.5); 10 mM EDTA (pH 8); 100 µg/ml RNase A.

Buffer 2: 0.2 M NaOH; 1% SDS.

Buffer 3: 0.9 M potassium acetate. Adjust pH acid to 6.8 using acetic acid.

2.2.17.1. Alkaline lysis mini-prep (Engebrecht and Brent, 1998)

Plasmid DNA was extracted from *E. coli* according to Engebrecht and Brent (1998) with minor modifications. A single positive bacterial colony was used to inoculate 25 ml sterile LB medium in a 50 ml Eppendorf tube.

The solution was incubated at 37°C for approximately 18 hours. Cells were centrifuged at maximum speed for 30 sec and collected in the same Eppendorf tube. After discarding the supernatant, the cells were resuspended in 200 µl GTE solution, vigorously shaken, and incubated at room temperature for 5 min. The obtained solution was equally separated into two 1.5 ml Eppendorf tubes. This was followed by adding 200 µl NaOH/SDS solution, vigorous shaking, and incubation on ice for 5 min. Next, 150 µl of potassium acetate solution was added and the final mixture was incubated on ice for another 5 min. The mixture was then centrifuged (13,000 g, 6 min, RT) and the supernatant was transferred into new 1.5 ml Eppendorf tube. One volume (450 µl) of phenol/chloroform/isoamyl alcohol (25/24/1) was added and vortexed rigorously for 10 sec. After a brief centrifugation (13,000 g, 1 min, RT) the upper phase was carefully transferred into new 1.5 ml Eppendorf tubes, mixed with two volumes of 95% ethanol, let sit 2 min at room temperature, and centrifuged (13,000 g, 10 min, RT) to precipitate the plasmids. The DNA pellet was subsequently washed with 1 ml of 70% ethanol, air dried under the laminar flow hood, and dissolved in 30 µl RNase-free TE buffer. The resuspended DNA was incubated in an incubator at 37°C for 1 hour to completely remove the RNAs and subsequently stored at -20°C.

2.2.18. Purification of Isolated Plasmids (MACHEREY-NAGEL GmbH & Co. KG)

NucleoSpin® Gel and PCR Clean-up protocol was carried out to purify the isolated plasmids before sending them for sequencing. To carry out the protocol, two volumes (60 µl) of Buffer NTI was added to the 30 µl resuspended plasmids. The mixture was briefly homogenized by gently tapping the Eppendorf tubes several times. The mixture was then transferred to the collection tube and centrifuged (11,000 g, 30 sec, RT) to bind the plasmids to the silica membrane. The flow-through was eliminated and 650 µl of Buffer NT3 was added then centrifuged (11,000 g, 30 sec, RT). The flow-through was eliminated and the previous step was repeated for a second time to minimize chaotropic salt-carry over and low A260/A230 values. The silica membrane was then dried by centrifugation (11,000 g, 1 min, RT) and the flow-through was eliminated. The spin column was transferred into new 1.5 ml Eppendorf tubes and incubated at 70°C on an heating block for 2 to 5 min prior to elution to completely remove residual ethanol from the spin column. To optimize plasmid elution the following steps were carried out: 20 µl of NE buffer was added to the spin column and subsequently incubated at 70°C for 5 min. The Eppendorf tubes were slowly centrifuged (100 g, 1 min, RT) then an additional 10 µl of NE buffer was added and again slowly centrifuged (100 g, 1 min, RT). In the final centrifugation the speed was increased to collect the purified plasmids (11,000 g, 2 min, RT). The samples were stored at -20°C for future use or sent for sequencing.

2.2.19. Sending purified plasmids for sequencing

The eluted and purified plasmids were sent for sequencing in two different tubes with each tube containing 12.5 µl purified plasmids and 2.5 sterile water to make a total volume of 15 µl. One tube will be amplified using the forward primer pJET1.2for while the second will be amplified using pJET1.2rev. No primers were added in-house and the sequencing relied on primers added by the company itself. To get the best results, plasmid samples destined for sequencing should have a concentration ranging between 50-100 ng/µl. Plasmids were transferred into new 1.5 ml Eppendorf tubes and sequencing labels (prepaid) were attached to them.

4. Results

4.1. Characterization of ELIPs

The ELIPs of many species such as the eudicot *Arabidopsis thaliana*, the monocot *Crocus sativus*, the chlorophyte *Dunaliella salina*, the lycophyte *Selaginella moellendorfii*, and the bryophyte *Physcomitrella patens* have been characterized, however, the ELIPs of the newly sequenced *L. brevidens* and *L. subracemosa* have not been previously described. Therefore, to gain more insight into the structure and role of these light-induced proteins in desiccation tolerance, the RNA-seq data published by VanBuren et al. (2018a). Out of the 26 ELIPs identified in *L. brevidens* (VanBuren et al., 2018a) 22 isoforms were identified using sequence homologies with CpELIP set as the query. On the other hand, all four ELIP isoforms identified by VanBuren et al. (2018a) were successfully identified. Afterwards, a cDNA expression library was used to amplify three *ELIPs* in *L. brevidens* (Lbr *ELIP* 025593, 016551, and 016696) and 3 ELIPs in *L. subracemosa* (Lsu *ELIPs* 008078, 030118, and 024731). The untranslated regions (1000bp) for each *ELIP* accession were identified upstream and downstream the transcription initiation and termination sequences of the open reading frame. Transcripts were obtained by using the primers and PCR programs outlined in Tables 6 and 7.

Based on nucleotide sequence homologies, Lbr and Lsu *ELIPs* showed significant similarity to members of the CAB/ELIP/HLIP family including *CpELIP* from *C. plantagineum* (Bartels et al., 1992), *ELIP1* and *ELIP2* from *A. thaliana* as well as to other *ELIPs* belonging to other monocots and eudicots such as: *Camellia sinensis*, *Capsicum annum*, *Citrus sinensis*.

The lengths of Lbr *ELIPs* were 843 bp for *ELIP* 016551, 803 bp for *ELIP* 025593, and 805 bp for *ELIP* 016696. Their corresponding sequences contained a single continuous open reading frame running from 71 to 653, from 243 to 803, and from 56 to 653.

In addition, the lengths of Lsu ELIPs were 681 bp for *ELIP* 008078, 807 bp for *ELIP* 024731, and 597 bp for *ELIP* 030118. Similar to Lbr *ELIPs*, their corresponding sequences contained a single continuous open reading frame running from 100 to 681, from 243 to 803, and from 1 to 597.

The properties of the open reading frames (ORFs) of the surveyed amino acid sequences of Lbr and Lsu ELIPs in addition to the properties of their corresponding mature proteins are provided below.

The ORFs of Lbr ELIPs:

1. The ORF of ELIP 016551 encoded a deduced polypeptide of 193 amino acids with a predicted molecular mass of 20.95 kDa and predicted pI of 9.00. The amino acid composition shows high content of (Alanine – Glycine – Leucine - Serine).
2. The ORF of ELIP 025593 encoded a deduced polypeptide of 186 amino acids with a predicted molecular mass of 19.99 kDa and predicted pI of 8.66. The amino acid composition shows high content of (Alanine – Glycine – Leucine - Serine).
3. The ORF of ELIP 016696 encoded a deduced polypeptide of 198 amino acids with a predicted molecular mass of 21.283 kDa and predicted pI of 8.85. The amino acid composition shows high content of (Alanine – Glycine – Leucine - Serine).

The ORFs of Lsu ELIPs:

1. The ORF of ELIP 008078 encoded a deduced polypeptide of 193 amino acids with a predicted molecular mass of 20.356 kDa and predicted pI of 9.3. The amino acid composition shows high content of (Alanine – Glycine – Leucine - Serine).
2. The ORF of ELIP 024731 encoded a deduced polypeptide of 189 amino acids with a predicted molecular mass of 20.128 kDa and predicted pI of 7.73. The amino acid composition shows high content of (Alanine – Glycine – Leucine).
3. The ORF of ELIP 030118 encoded a deduced polypeptide of 198 amino acids with a predicted molecular mass of 21.365 kDa and predicted pI of 9.05. The amino acid composition shows high content of (Alanine – Glycine – Leucine).

Table 10. The mature lengths, predicted molecular masses and pI values for mature length Lbr and Lsu ELIPs.

Protein	Mature protein length	Predicted molecular mass	Predicted pI
Lbr ELIP 016551	154	16.801 KDa	5.49
Lbr ELIP 025593	149	16.032 KDa	5.01
Lbr ELIP 016696	160	17.349 KDa	6.56
Lsu ELIP 008078	156	16.546 KDa	8.61
Lsu ELIP 024731	151	16.252 KDa	5.50
Lsu ELIP 030118	158	17.097 KDa	8.56

The aforementioned six ELIPs contain ORFs that were of comparable length to known ELIPs and shared the same pI range (7 to 9) (supplementary table 1). Their encoded putative primary amino acid sequences showed high identity to ELIPs from other species but low identity to PSBS, CABs, and CBRs (supplementary table 2).

Mature lengths of Lbr and Lsu ELIPs are of comparable sizes and their predicted molecular mass ranges between 16 and 17 KDa. However, their predicted pI showed significant variability with Lbr ELIP 016551, Lbr ELIP 025593, and Lsu ELIP 024731 exhibiting a relatively lower pI compared to other ELIPs (table 10).

Out of all identified ELIPs in *L. brevidens* and *L. subracemosa*, ELIP 016551 was the only protein which contains a target PEST sequence was predicted in between residues 56 and 74 (REDDETQQQLDSSELTPNR).

Sequence homologies revealed that Lbr and Lsu ELIPs contained a signal peptide and chlorophyll a/b binding domain. Their signal peptides are highly similar and judging from ELIP sequence homologies, the amino acid residues identified within the LHCs as core active sites are conserved in ELIPs and they are predicted to bind four Chl a molecules. The putative chlorophyll a-binding residues are distributed within helices I and III (a1: Glutamic acid in position 100; a2: Asparagine in position 103; a3: Glutamic acid in position 167; a4: Asparagine in position 170) (Figure 1).

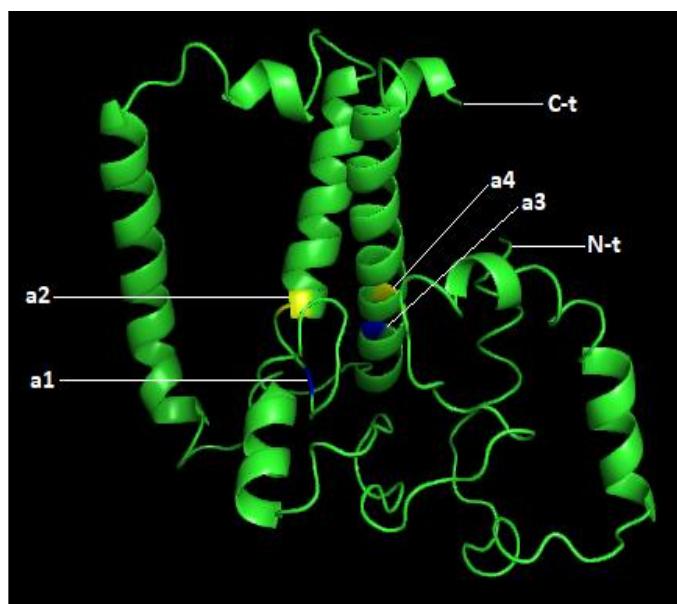


Figure 1. Tridimensional model of CpELIP showing the N-terminal (N-t), C-terminal (C-t) and the four Chl a binding sites (a1, a2, a3, and a4).

Moreover, sequence homology with the LHCII type I CAB-8 (also called LHCb1; Uniprot accession number: P27490) of *Pisum sativum* (Garden pea) revealed one lutein binding site (Tryptophan) which conform with the wide-spread theory that ELIPs are pigment-binding proteins (figure 2).

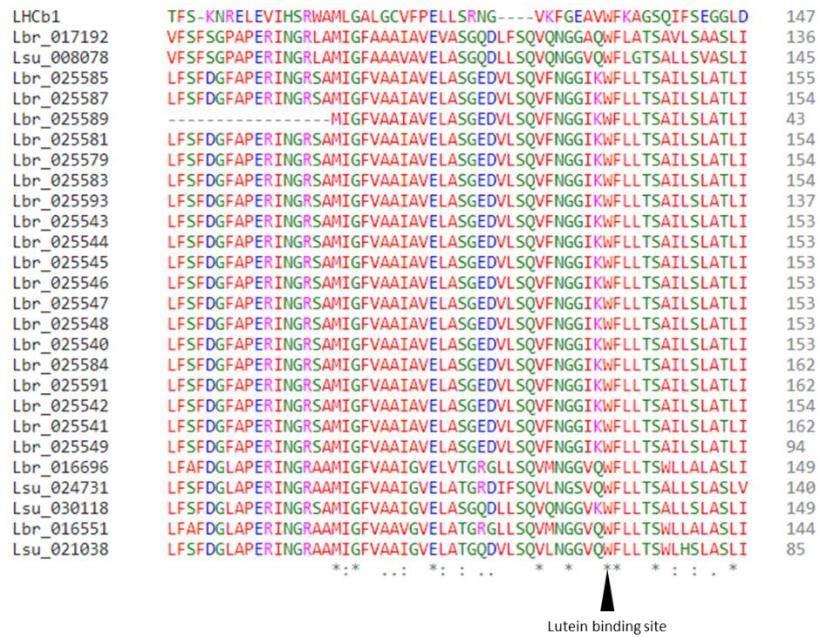


Figure 2. Amino acid sequence alignments of *L. subracemosa* and *L. brevidens* ELIPs in addition to LHCII type I CAB-8 (LHCb1). The black arrow indicates the lutein binding site.

ELIPs are known to target the thylakoids membranes. Due to the relatively high identity with most members of the ELIP family, it could be inferred that ELIPs of *L. brevidens* and *L. subracemosa* are primarily localized in the thylakoid membranes. However, no localization studies have been conducted to support the last statement. Due to the lack of experimental data, WoLF PSORT, a software designed to predict protein subcellular localization was used to whether the signal peptides of ELIPs are imported into chloroplasts. By running the program against *L. brevidens* and *L. subracemosa* ELIPs and comparing with results demonstrating the highest identity, it has been predicted that all surveyed ELIPs are targeted to chloroplasts.

The amino acid sequence of all identified *L. brevidens* and *L. subracemosa* ELIPs derived from the ORFs were compared with sequences in the NCBI GenBank. The comparison included a comprehensive amino acid alignment detailing the different motifs, domains, and conserved amino acid residue across different species. The alignment was composed of many proteins belonging to the superfamily. It included CBRs, CABs, PSBSs, and ELIPs belonging to a variety of clades including monocots, eudicots, and chlorophytes.

Hydropathy plots of the 6 ELIPs revealed the presence of three transmembrane domains designated helix I, helix II, and helix III (Figure 3). FoldIndex predicted that the majority of Lbr and Lsu ELIPs are partially unfolded (Figure 4; Supplementary table 1). Disordered residues were present in the N-terminal and unfoldability was the lowest in this region except for a small folded region at the start of the sequence. This is in accord with the fact that ELIPs contain an alpha-helix in its N-terminal followed by a short-disordered region preceding the three transmembrane helices. Since ELIPs are produced in the cytoplasm, it is logical to assume that foldability is ought to be high in the transmembrane helices because their amino acid composition is highly hydrophobic.

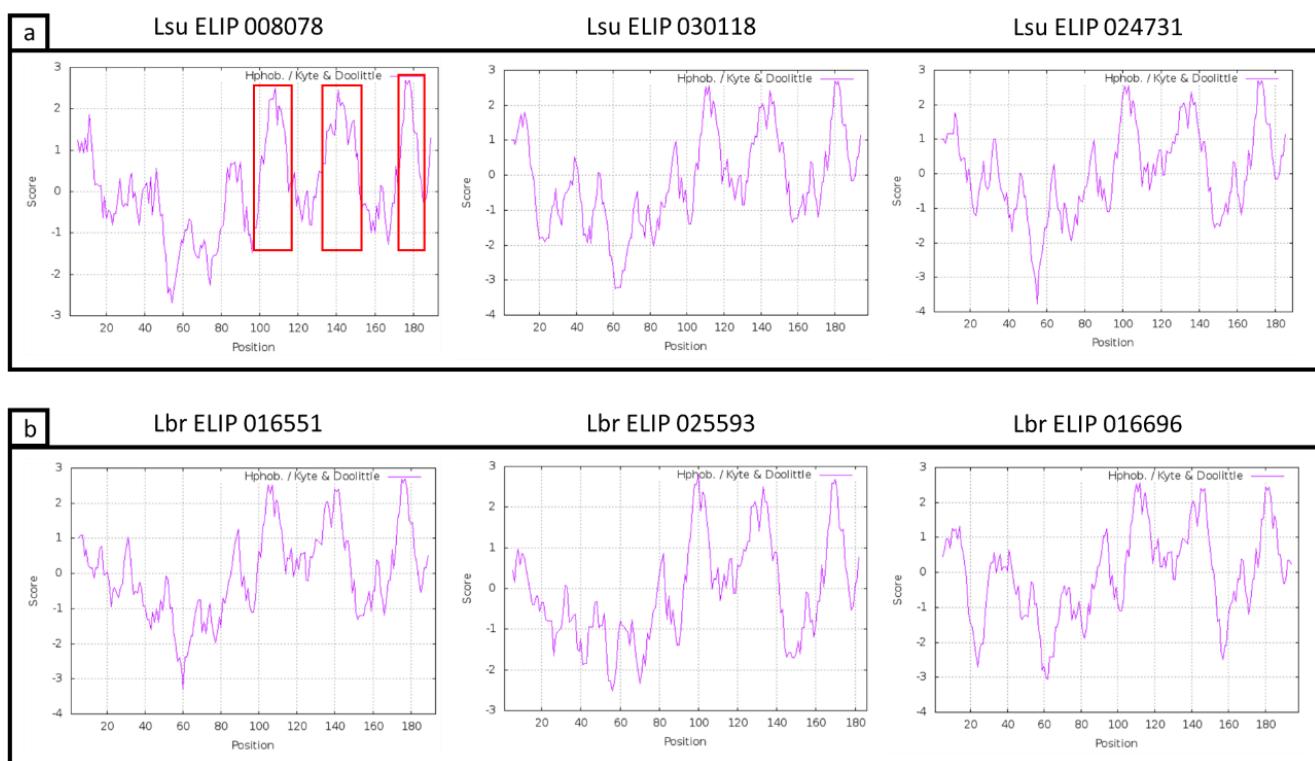


Figure 3. Hydropathy plots for the six investigated *L. subracemosa* (a) and *L. brevidens* (b) ELIPs. Red boxes indicate the hydrophobic transmembrane regions.

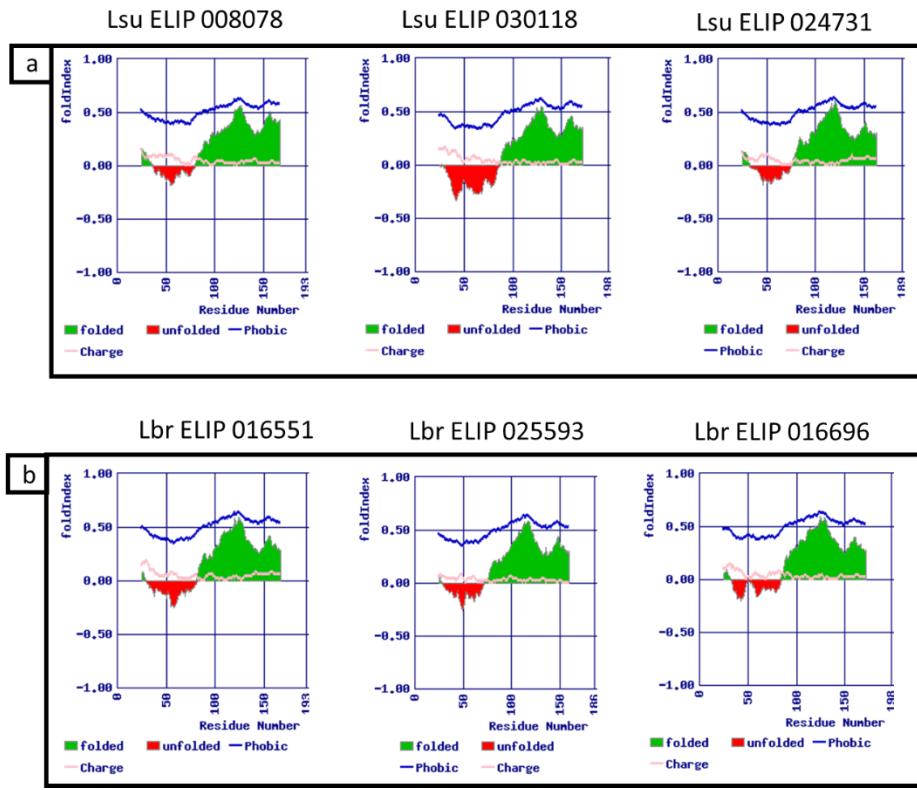


Figure 4. Foldability plots for the six investigated *L. subracemosa* (a) and *L. brevidens* (b) ELIPs.

Sequence homologies and alignments showed that a conserved chlorophyll a/b binding motif is found in Helices I and III and their generic motif is composed of 19 amino acid residues designated ERINGRLAMIGFVAALAVE and ELWNNGRFAM-LGLVALAFTE respectively. Whereas the generic motif of helix II is composed 22 amino acids (FLLTSAILSLATLIPYRGLSP).

Table 11. Important domains of the cloned *L. brevidens* and *L. subracemosa* unmatured ELIPs.

Protein	Signal peptide	Helix I	Helix II	Helix III	Chlorophyll a/b binding domain
Lbr ELIP 016551	1 - 39	98-113	126 - 146	168 - 184	76 - 153
Lbr ELIP 025593	1 – 37	91 – 106	119 – 139	161 - 177	135 - 177
Lbr ELIP 016696	1 – 38	103 – 118	131 – 148	173 - 189	58 - 164
Lsu ELIP 008078	1 – 37	99 – 114	128 - 145	169 - 185	74 - 152
Lsu ELIP 024731	1 – 38	94 – 109	121 – 142	164 - 180	82 - 147
Lsu ELIP 030118	1 – 40	103 – 118	131 – 149	173 - 189	91 - 189



Figure 5. Amino acid sequence alignments of all investigated members of the CAB/ELIP/HLIP superfamily. The figure shows the N-terminal region.

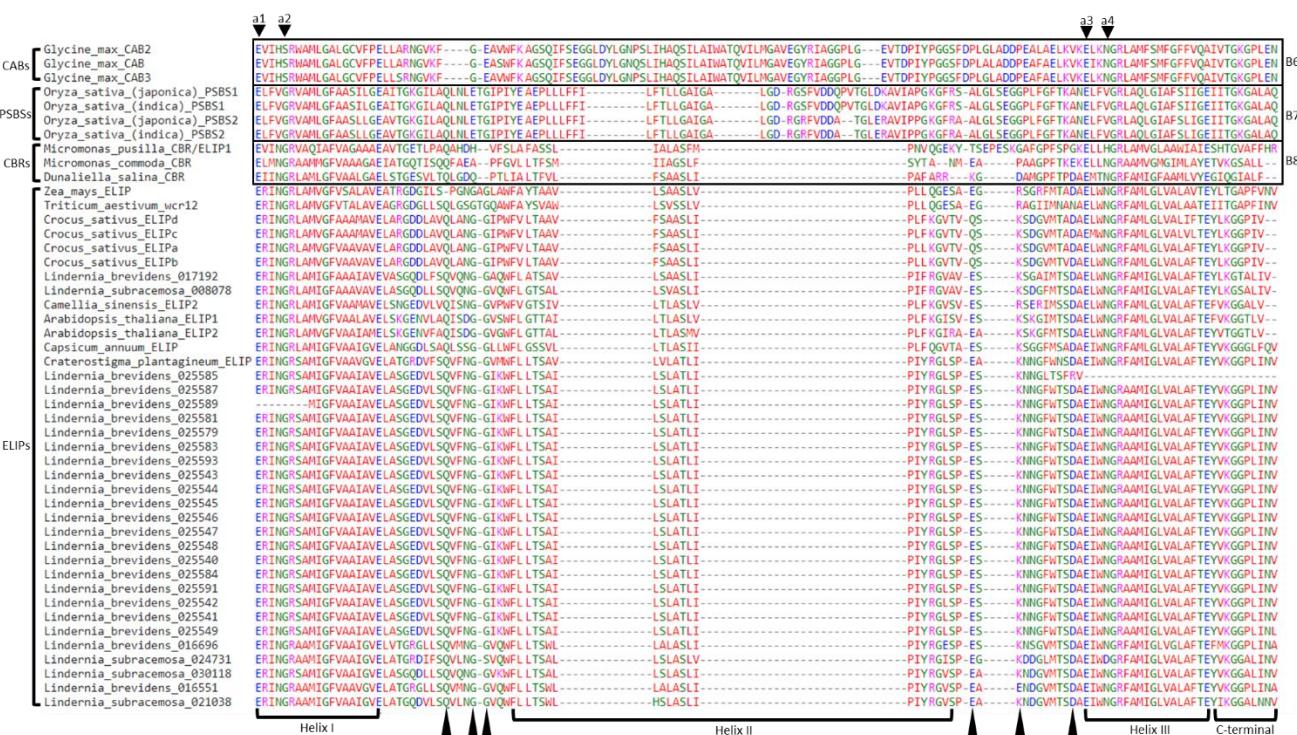


Figure 6. Amino acid sequence alignments of all investigated members of the CAB/ELIP/HLIP superfamily. The figure shows the full transmembrane helices and inter-helical regions and part of the C-terminal.

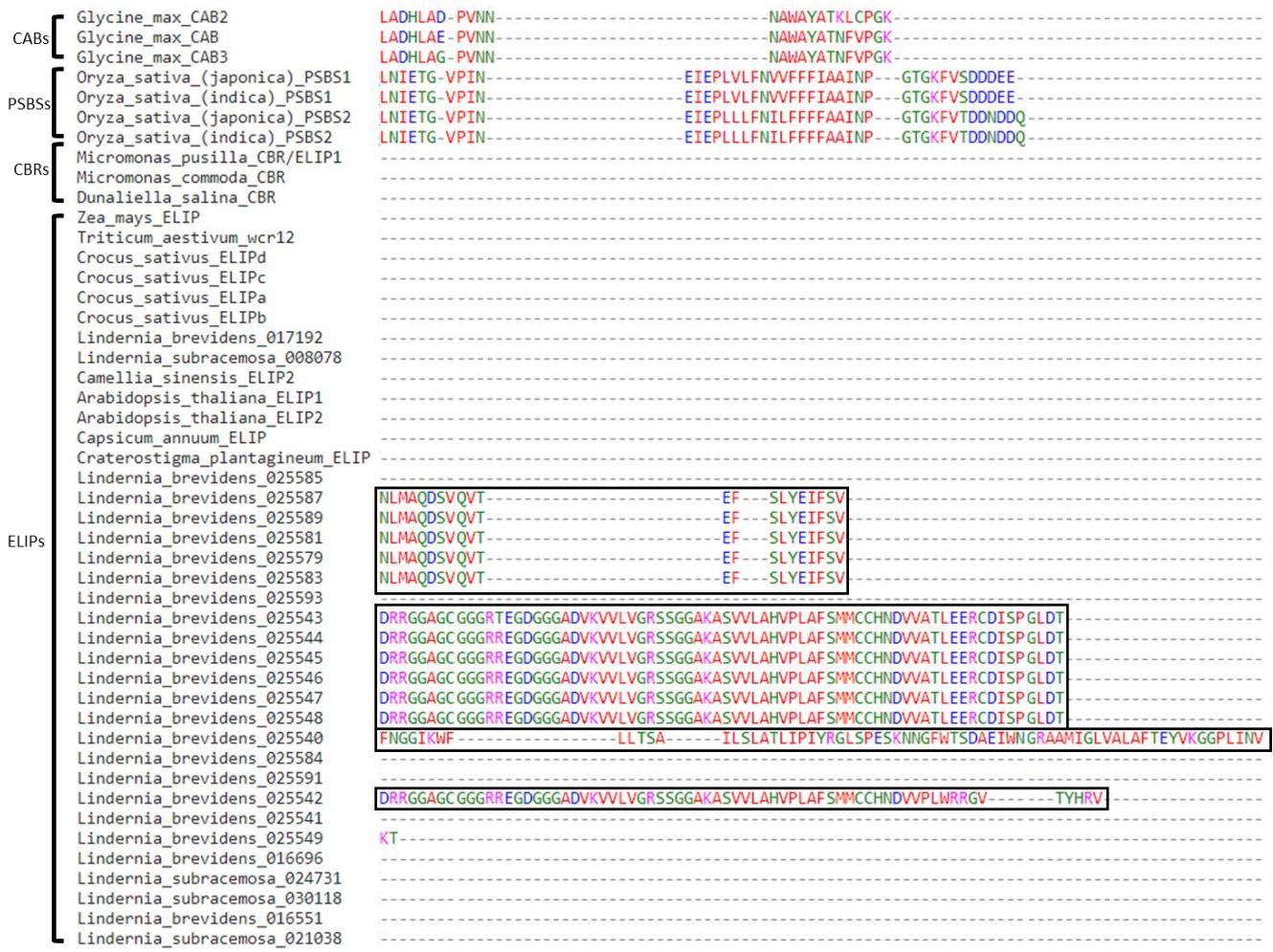


Figure 7. Amino acid sequence alignments of all investigated members of the CAB/ELIP/HLIP superfamily. The figure shows the C-terminal region. Each black box represents a unique C-terminal sequence.

Helix I and III contain several universally conserved amino acids for most known ELIPs:

1. The conserved amino acid residues of Helix III include: three Alanines (A), two glutamic acids (E), two glycines (G), one arginine (R), one methionine (M), one Isoleucine (I), one Asparagine (N), and one Phenylalanine (F). In helix I, unlike other ELIPs, CpELIP and Lbr ELIP 025593, Lsu ELIP 030118 share the Serine (S) residue.
2. The conserved amino acid residues of Helix III include: three Alanines (A), two glutamic acids (E), two glycines (G), one arginine (R), one methionine (M), one Leucine (L), one Tryptophan (W), one Asparagine (N) (except Lsu ELIP 024731), and one Phenylalanine (F) (except for Lbr ELIP 025593).

3. The conserved amino acid residues of Helix II include: one Phenylalanine (F), one Proline (P), and one Glycine (G) while the remaining residues vary between different species and within isoforms of the same species.

The inter-helical regions (region between the transmembrane helices) of ELIPs are similar and contain a number of universally conserved amino acids (figure 6; black arrows and helix II). The region with highest variation within ELIPs can be found in the N- and C-terminals. Monocot and eudicot ELIPs appear to have a proline rich gene with an unknown function (figure 5; black box). This characteristic is also shared by *L. brevidens* and *L. subracemosa*, namely Lsu ELIP 008078 and Lbr ELIP 017192. A region of highly conserved amino acids can be found upstream the generic motif AFSGPAP (Figure 5; green box). The region is composed of one extended motif which contains several conserved amino acids (Figure 5; red lines). Moreover, all eudicot and monocot ELIPs in addition to Lbr and Lsu ELIPs possess the generic motif AFSGPAP at its N-terminal (N-t) with the exception of Lsu ELIP 008078 and Lbr ELIP 017192 (Figure; Motif 1). The conservation of A and D of this motif appear to be a distinct property belonging to *L. brevidens* and *L. subracemosa*. This could imply that the substitution of an A to S and S to D is a characteristic that is unique to the *Lindernia* genus (family: Scrophulariaceae; order: Lamiales). However, the uniqueness of the S residue to the *Lindernia* genus could be debated because the *Capsicum annuum* ELIP, which belongs to the order Solanales, also possesses A to S substitution. Both the extended N-terminal domain and the generic AFSGPAP motif are missing in CBRs, CABs, and PSBSs.

Unlike Lsu ELIPs, the C-terminals (C-t) of most Lbr ELIPs contain an extended motif. Four distinct groups could be immediately recognized with each group possessing its unique amino acid signature (Figure 7; black boxes). The high variations in the amino acid content of the C-t within *L. brevidens* could imply that each group has a different function, different binding partners (e.g., proteins or ligands), or simply represents different protein degradation pathways. More research should be carried out to investigate these assumptions. The N- and C-terminal domains are longer and show the greatest divergence relative to the other members of the CAB/ELIP/HЛИP superfamily. Furthermore, the N- and C-terminals as well as the transmembrane spanning regions show less conservation when ELIPs are compared to other members of the family (CABs, PSBSs, and CBRs). The variation is most noticeable within CABs and PSBSs in which each protein subfamily shares a conserved domain (Figure 5; blue boxes). However, unlike CABs, it appears that first segment of the N-terminal of PSBS1 and PSBS2 is conserved across species (Figure 5; red boxes).

Additionally, the three transmembrane helices of CABs along with the inter-helical regions (amino acid residues between the transmembrane helices) are noticeably different than their counterparts in ELIPs. In particular, their inter-helical regions were much longer than their ELIP counterpart. CBRs did not contain a consistent and conserved number of transmembrane helices. Namely, each of *Micromonas commoda* CBR and *Micromonas pusilla* CBR/ELIP1 possessed one predicted transmembrane helix: FGVLLTFSMIIAGSLFSYT and VFSLAFASSLIALASF respectively. Whereas *Dunaliella salina* CBR contained two predicted transmembrane helices (PTLIALTFVLFSAASLIPAF and FAMIGFAAMLVYEGIQGIALF). PSBSs exhibited a unique 4 transmembrane helices (Table 12) and three inter-helical regions of which two are short and one is long (Table 13). Lastly, the C-terminal of CABs and PSBSs were conserved yet different in their composition from their homologues in *L. brevidens* and *L. subracemosa*.

Table 12. Generic sequences of transmembrane helices of CAB/ELIP/HLIP superfamily.

	Protein subfamily			
	ELIPs	CABs	CBRs	PSBSs
Transmembrane helix I	ERINGRLAMI GFVAALAVE	VIHSRWAMLG ALGCVFPELLA	Variable composition	VAMLGFAASILGEA ITGKGIL
Transmembrane helix II	FLLTSAILS LA TLIPIYRGLSP	LIHAQSILAIWA TQVILMGAV	Variable composition	AEPLLLFFILFTLLG AIGALG
Transmembrane helix III	ELWNGRFAM- LGLVALAFTE	LAMFSMFGFFF QAIVTGKGPL	Variable composition	LFVGRLAQLGIAFSI IGEIIIT
Transmembrane helix IV	X	X	X	PINEIEPLVLFNVVF FFIAAI

Table 13. Generic sequences of inter-helical regions of CAB/ELIP/HLIP superfamily.

	Protein subfamily			
	ELIPs	CABs	CBRs	PSBSs
Inter-helical region I	LATGQDV LSQVLNG GVQW	RNGVKFGEAV WFKAGSQIFSE GGLDYLGNPS	Variable composition	AQNLNETGIPIYE

Inter-helical Region II	PEAKNDG VMTSDA	EGYRIAGGPLG EVTDPIYPGGS FDPLGLADDPE ALAEELKVKEL KNGR	Variable composition	RGSFVDDQPVTG LDKAVIAPGKGF RSALGLSEGGPL FGFTKANE
Inter-helical Region III	X	X	X	GKGALAQLNIET GV

A prominent feature which distinguishes *C. plantagineum* ELIP (CpELIP) from its close relatives is the presence of a highly hydrophobic motif in the N-terminal region (EKEEQQQQQKQQQ) composed entirely of glutamine or glutamic acid residues interspersed with two lysine residues between amino acid positions 57 and 68. This domain does not resemble any known ELIP, CAB, CBR, or PSBS protein domain in the protein data bases and till now it has no assigned function. Whether or not this motif is conserved in resurrection plants or a particular genus within this lineage of angiosperms is still an unresolved topic. Lbr and Lsu ELIPs do not contain this hydrophobic motif, however Lsu ELIP 030118 contains four glutamic acids (Q) residues in its extended N-terminal domain which are reminiscent of the residues found in *C. plantagineum*. The hydrophobic segment is followed by a proline rich region spanning the amino acid positions 74 and 84.

4.2. Phylogenetic analysis

The deduced polypeptides of the six ELIPs under investigation were similar to homologues from *A. thaliana*, *C. plantagineum*, and *Crocus sativus*. High similarity with homologues of the ELIP family indicate an intertwined phylogenetic history and supports a relationship with a common ancestor. Therefore, to gain more insight into the structural relationship between ELIP and ELIP-like members of the CAB/ELIP/HLIP superfamily, the deduced amino acid sequences were used to construct two phylogenetic trees: (1) Maximum Likelihood tree using the JTT matrix-based model and the (2) maximum parsimony tree using Tree-Bisection-Rerooting (TBR) algorithm. The rooted ML tree with the highest log likelihood depicted in figure 8 and the most parsimonious rooted MP tree depicted in figure 9 reflect standard groupings of ELIPs, CABs, CBRs, and PSBSs.

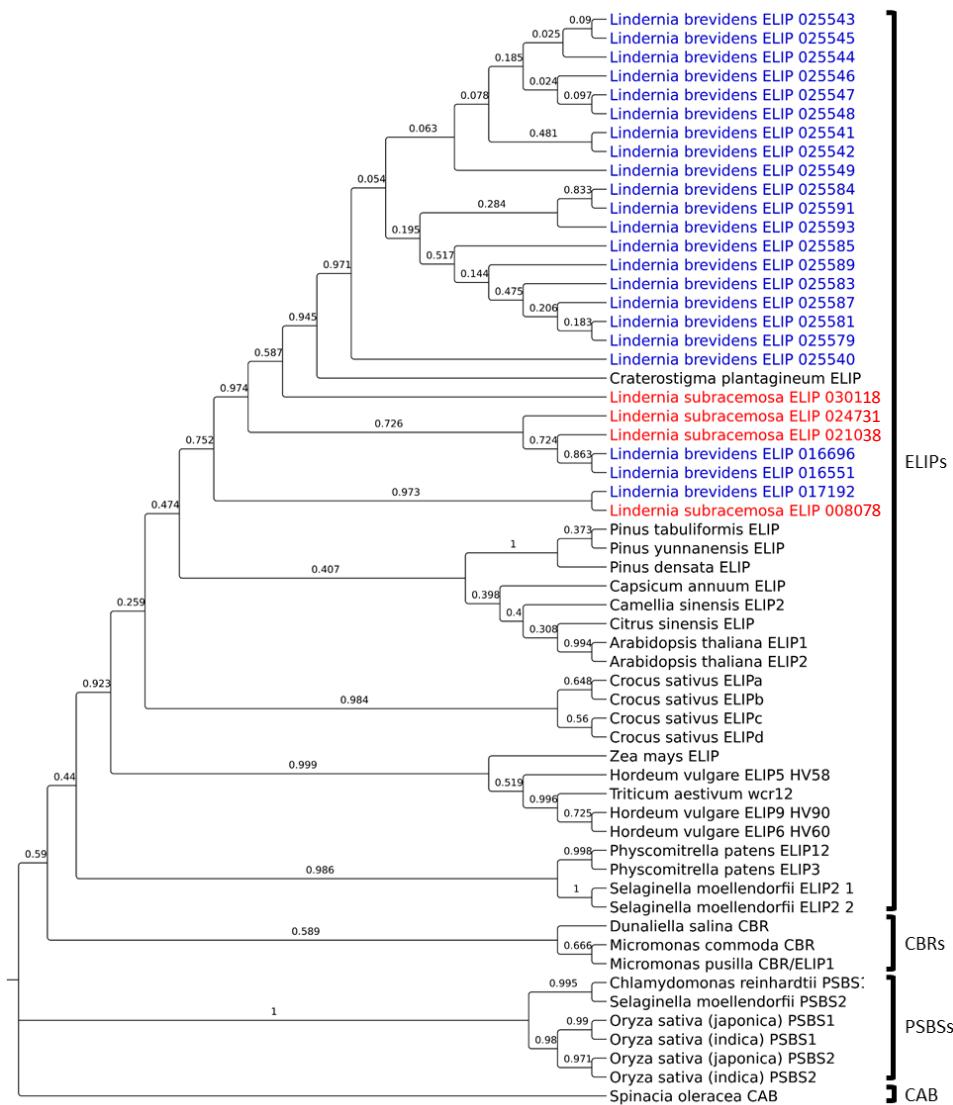


Figure 8. Maximum likelihood rooted tree. *L. brevidens* ELIPs are labelled in blue whereas *L. subracemosa* ELIPs in red.

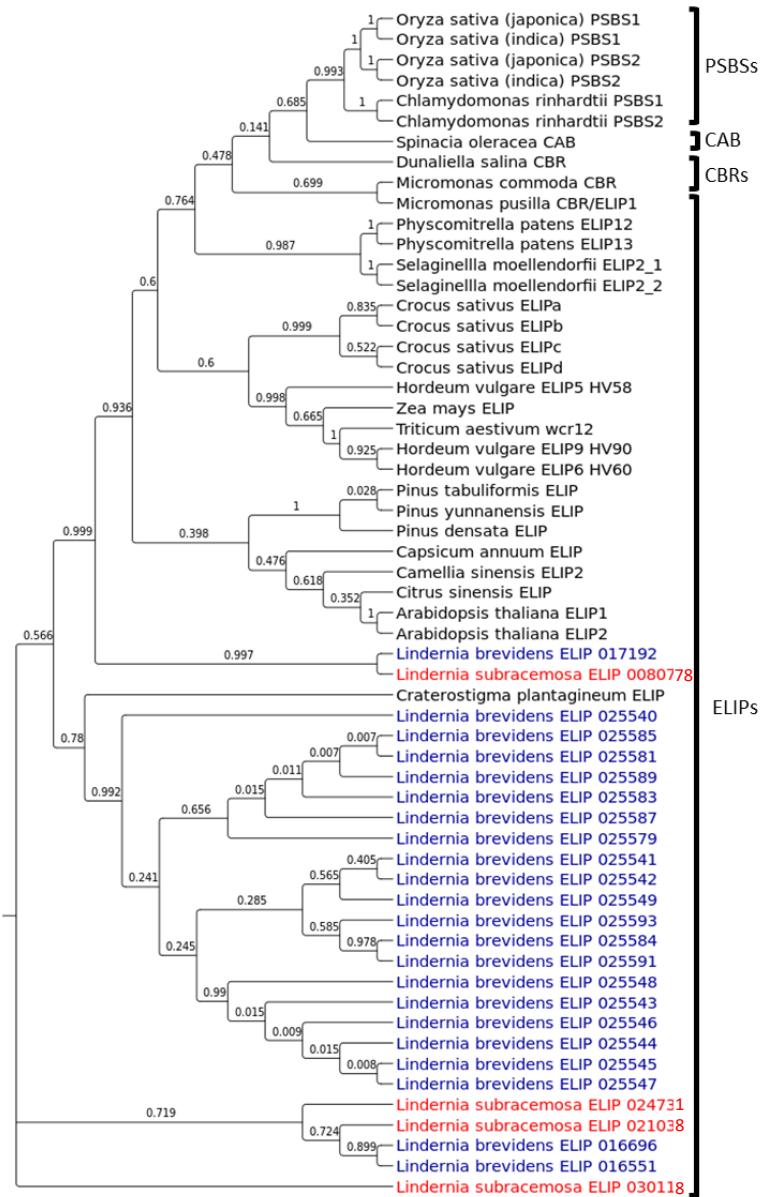


Figure 9. Maximum parsimony rooted tree. *L. brevidens* ELIPs are labelled in blue whereas *L. subracemosa* ELIPs in red.

In the ML and MP trees, the ELIPs labelled in blue grouped with *C. plantagineum*, thereby creating an association of ELIP sequences derived from desiccation-tolerance tissues. However, in the MP tree, the groupings were slightly different. The following ELIPs: Lbr ELIP 01696, Lbr ELIP 016551, Lsu ELIP 024731, Lsu ELIP 021038, and Lsu ELIP 0301, had an independent cluster. Whereas ELIPs labeled in blue clustered with ELIPs of CAB/ELIP/HLIP superfamily rather than with their ELIP isoforms. In the two trees, Lbr ELIP 017192 and Lsu ELIP 008078 clustered together. The phylogenetically nearest protein homologs belonged to spermatophytes (*Pinus densata*, *Pinus yunnanensis*, and *Pinus tabuliformis*) and eudicots (*Capsicum annuum*, *Camellia sinensis*, *Citrus sinensis*, and *Arabidopsis thaliana*).

4.3. RNA stability

Two distinct bands, one belonging to 28S and the other to 18S, can be clearly seen in figure 10. The yield of RNA in *L. brevidens* is much higher and more stable than in *L. subracemosa*. It appears that RNA is completely degraded in fully desiccated leaves of *L. subracemosa* whereas RNA was retained intact in fully desiccated and rehydrated leaf tissues of *L. brevidens*.

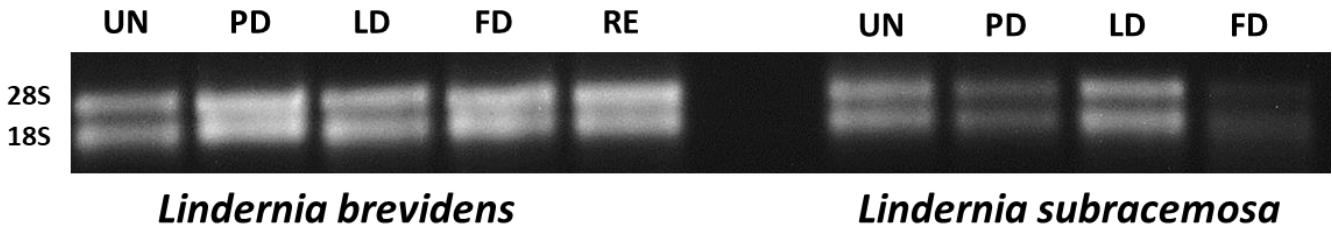


Figure 10. Ethidium bromide-stained agarose gels to monitor the quality of RNA extracted from leaves of *L. brevidens* and *L. subracemosa*. 1 µg total RNA was loaded on to the gel. UN: untreated, PD: partially dehydrated, LD: late dehydrated, FD: fully desiccated, and RE: rehydrated.

4.4. Transcript expression of Lbr and Lsu *ELIPs* in response to dehydration treatments

The gene expression profiles of *ELIPs* in untreated, early dehydrated, late dehydrated, and fully desiccated leaf tissues from *L. brevidens* and *L. subracemosa* were investigated. Expression levels in rehydrated leaf tissues were investigated only in *L. brevidens*. Transcripts of Lbr *ELIPs* 016551 and 025593.27 were constitutively expressed in the untreated leaf tissues. Their expression levels elevated rapidly upon further dehydration and peaked in the desiccated tissues yet started to decline gradually upon rehydration. The amplification of the partial gene of Lbr *ELIP* 016551 showed similar results to the full amplification of the same gene. However, only the partial amplification yielded a positive sequence result matching the coding sequence of the protein (Figure 14). Two primer combinations were used to amplify Lbr *ELIP* 025593, one was a partial gene amplification (PG – Lbr *ELIP* 025593.16) while the other was a full gene amplification (FG – Lbr *ELIP* 025593.27). However only the full gene amplification yielded a positive result (Figure 15). Lbr *ELIP* 016696 was neither expressed in the untreated nor the rehydrated tissues, however its expression was induced in an advanced state of dehydration (partial dehydration) and continued to increase upon further dehydration. SANGER sequencing of Lbr *ELIPs* 025593.16 and 016696 failed to match the desired coding sequences. A long deletion could be observed in the plasmids of transformed constructs of Lbr *ELIP* 016696 (Figure 16). Despite the failure of properly amplifying the partial Lbr *ELIP* 025593.27, the full gene was accurately amplified (Figure 15). On the other hand, Lsu *ELIPs* 008078 (partial gene amplification) and 030118 (full gene and partial gene amplifications) were

expressed at a steady-state level in the untreated tissues and throughout the dehydration process. Lsu *ELIP* 024731 was in untreated leaf tissues, but its expression was induced in partially dehydrated leaf tissues, peaked in the late dehydrated tissues, and completely disappeared in desiccated tissues. Nevertheless, since the partial amplification of Lsu *ELIP* 024731 did not yield consistent results when more replications were carried out, its amplification was abandoned. Both *L. brevidens* and *L. subracemosa* showed relatively stable expression levels of the positive control *EF5α*.

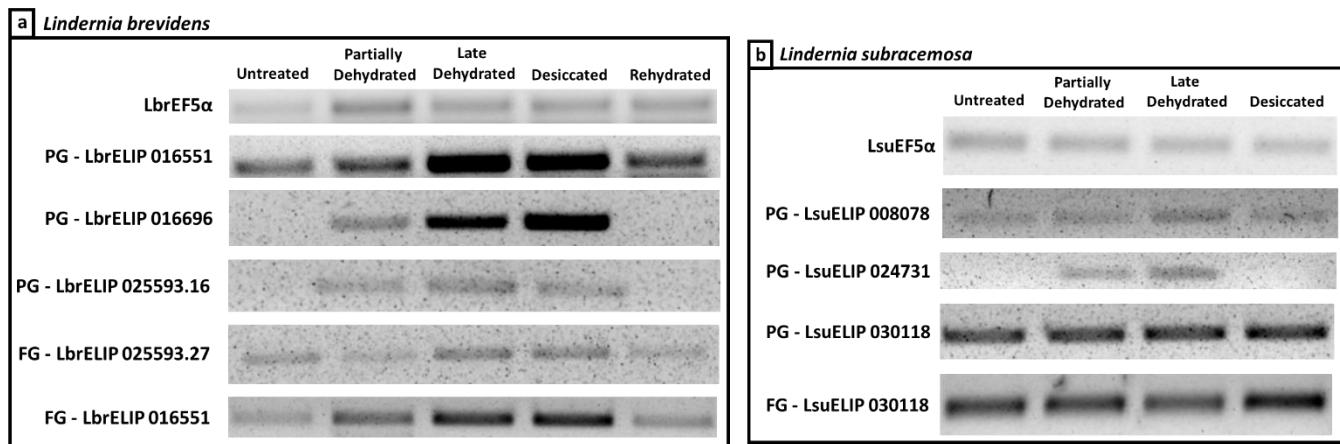


Figure 11. Ethidium bromide-stained agarose gels to monitor *ELIP* transcript expression. Expression profiles were determined for leaf tissues taken from untreated, partially dehydrated, late dehydrated, fully desiccated, and rehydrated plants. (a) *ELIP* transcript expression in *L. brevidens*. (b) *ELIP* transcript expression in *L. subracemosa*. PG: partial gene amplification. FG: full gene amplification. EF5 α was used as positive control.

4.5. SANGER sequencing results

Plasmid constructs were sent for sequencing at Eurofins MWG Operon. Below are the results.

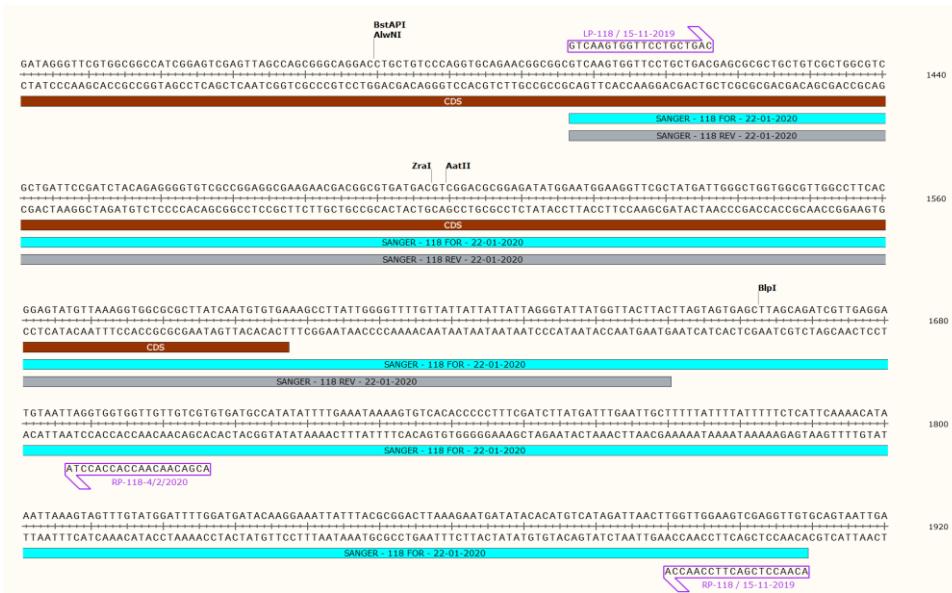


Figure 12. SANGER sequencing of Lsu *ELIP* 030118 (partial gene amplification). Coding sequence is colored dark red and the result of SANGER sequencing is colored light blue.

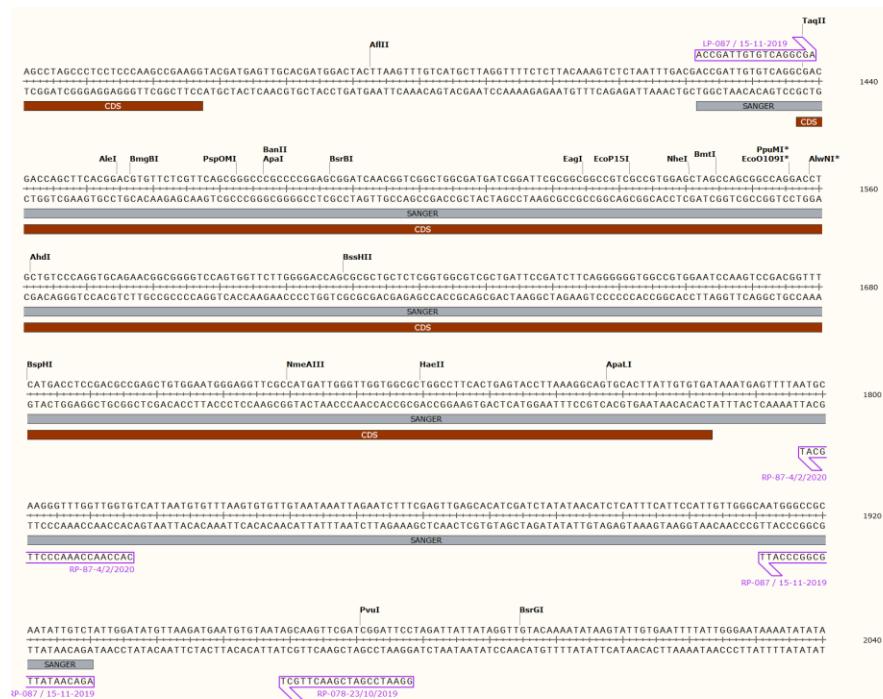


Figure 13. SANGER sequencing of Lsu *ELIP* 008078 (partial gene amplification). Coding sequence is colored dark red and the result of SANGER sequencing is colored light grey.



Figure 14. SANGER sequencing of Lbr *ELIP* 016551 (partial gene amplification). Coding sequence is colored dark red and the result of SANGER sequencing is colored light grey.

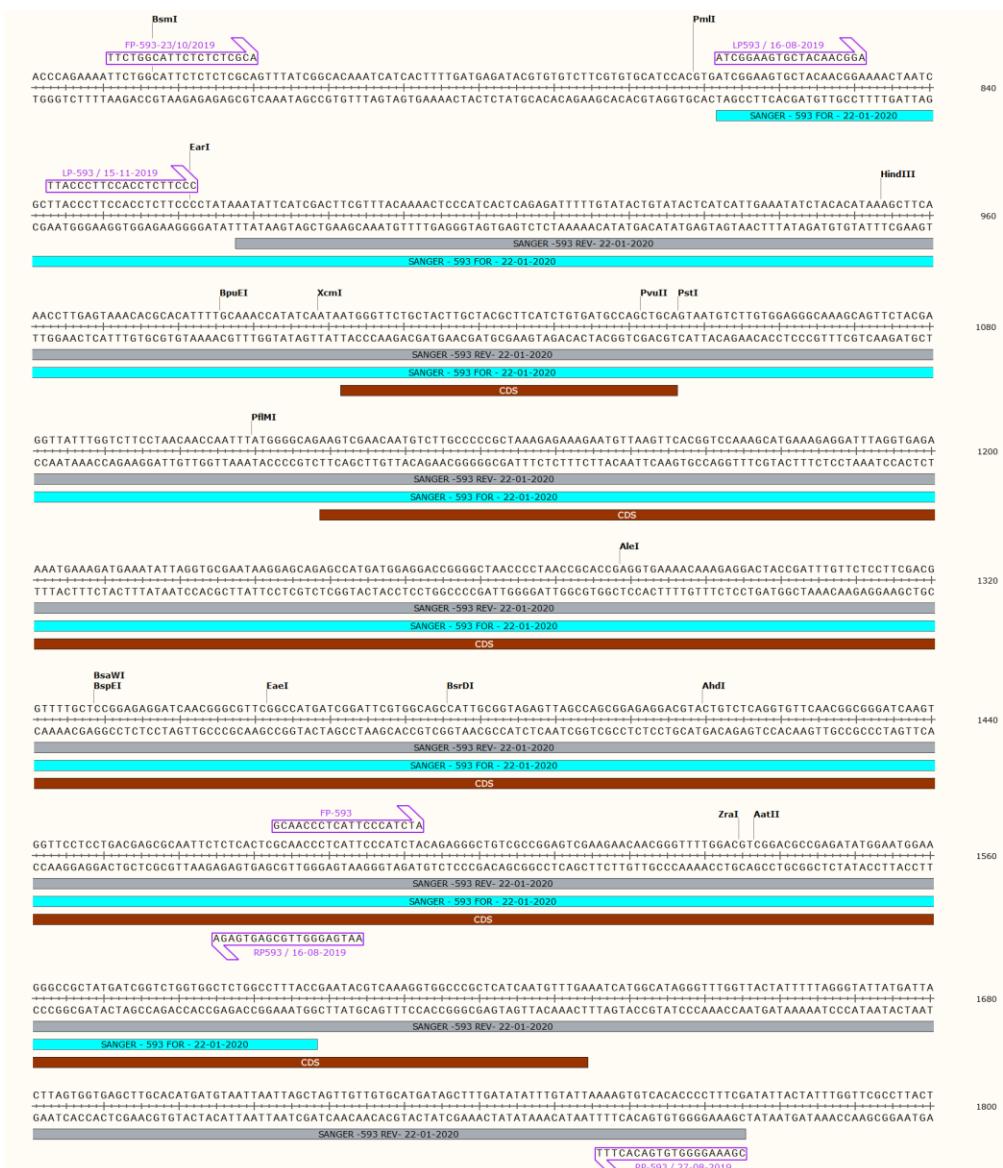


Figure 15. SANGER sequencing of Lbr *ELIP* 025593.27 (full gene amplification). Coding sequence is colored dark red and the result of SANGER sequencing is colored light blue.

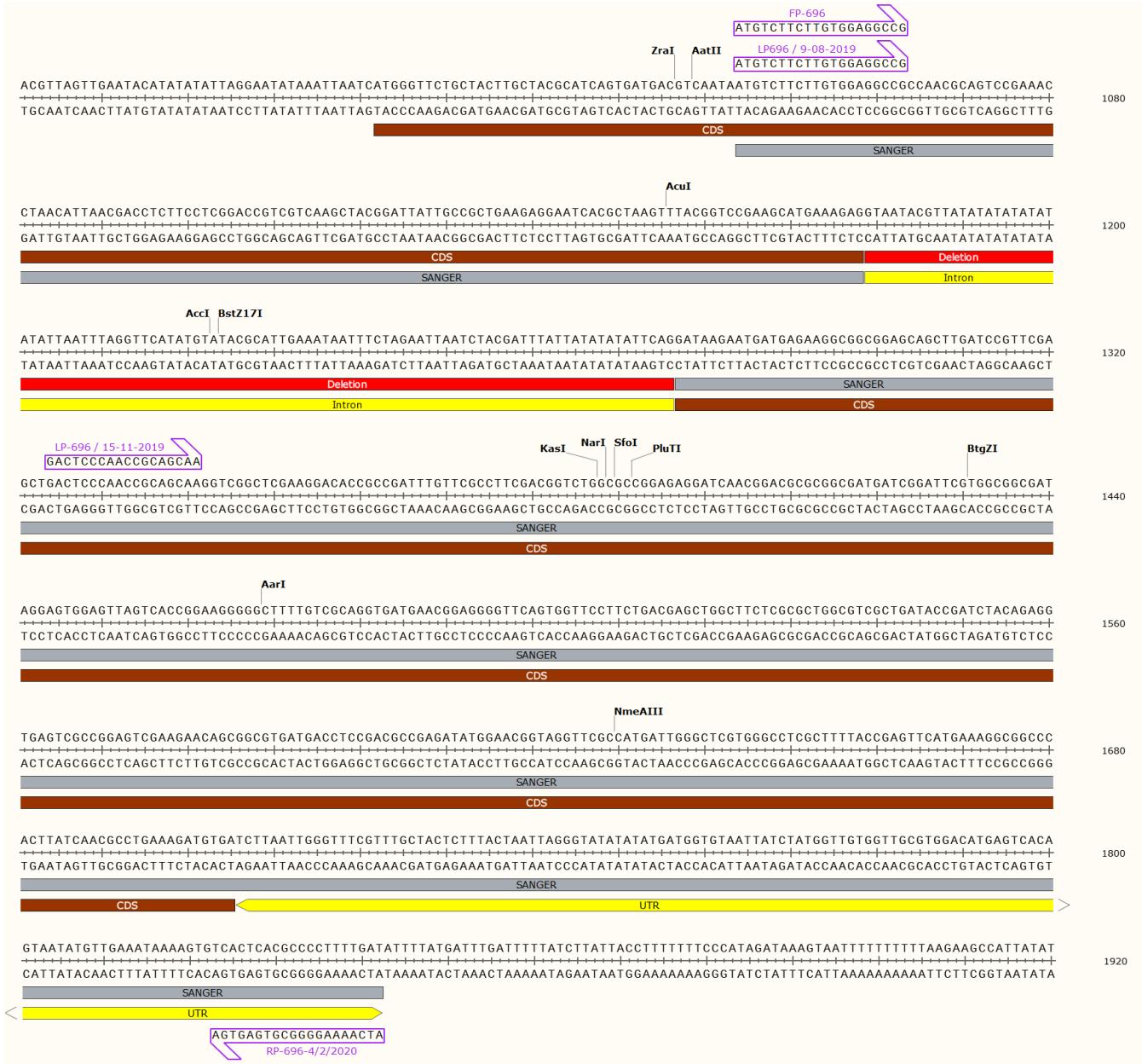


Figure 16. SANGER sequencing of Lbr ELIP 01696 (partial gene amplification). Coding sequence is colored dark red, the result of SANGER sequencing is colored light blue, the UTR is colored yellow, and the deletion segment is colored light red.

5. Discussion

Plants are sessile organisms and severe environmental conditions limit both the productivity and range of habitats available to plants. One of the major stresses straining agricultural production is drought stress (Bartels and Salamini, 2001). Moreover, the increasingly fluctuating weather phenomena influenced by climate change and global warming along with Earth's unconstrained population growth and unfettered overconsumption are adding additional strains on agricultural production further limiting the availability of fresh and uncontaminated water. Unusually dry and warm weather hugely favor infestations with pathogens and negatively affect plant productivity and fertility. Besides, locally adapted crop genotypes need to develop new adaptations to survive new biotic stress factors such as the climate-driven migration of pathogens and pests. Furthermore, a combination of abiotic and biotic stresses may result in an unusual trade off in responses which makes the plant tolerant to some stresses but susceptible to others (Dresselhaus and Hückelhoven, 2018). Therefore, plants have developed a number of adaptations to cope with changing environmental conditions. Nonetheless, these adaptations adversely affect yield parameters which in turn requires plant to sacrifice biomass, growth, and productivity in favor of survival (Bartels and Salamini, 2001). Consequently, and due to the advent of more severe and extreme climate and weather patterns, understanding the physiological and molecular background of such tolerance has become a task of utmost importance for plant breeders, insofar that developing this knowledge will help them breed cultivars which are able to withstand the continuously changing environmental conditions and produce higher crop yields (Dresselhaus and Hückelhoven, 2018; Janiak et al., 2018). In particular, it has become imperative to unravel the molecular pathways governing drought tolerance and the mechanisms by which drought-tolerant plants are capable of adapting to water deficit while retaining their productive and reproductive capacities.

Survival during drought stress relies on a multiplicity of factors which include moisture maintenance, control of water uptake, water use efficiency, osmotic adjustment, production of secondary metabolites, up-regulation of protective enzymes, and regulation of chlorophyll content (Fang and Xiong, 2015). There are many forms acclimation to water deficit in the plant kingdom amongst them are (1) drought avoidance (DA), (2) drought tolerance (DT), (3) drought escape (DE), and (4) drought recovery (DR) (Lawlor, 2103). However, the two major mechanisms that govern the processes conferring plant drought tolerance are DA and DT (Yue et al., 2006). Notwithstanding, desiccation is the most acute form of water deficit (Bartels and Salamini, 2001). Drought and desiccation tolerance are two terms that should not be interchangeably

used (Alpert, 2005). On one hand, drought tolerance is defined as the ability to endure and survive low environmental tissue water availability while maintaining high internal water content through the induction of adaptive traits including the maintenance of cell turgor and increasing protoplasmic resistance (Morgan, 1984; Berjak, 2006). On the other hand, desiccation tolerance is defined as the ability to survive drying to equilibrium with low environmental tissue water availability while maintaining low intracellular water content (Berjak, 2006). Desiccation tolerance occurs when most of the protoplasmic water is lost and only a very small amount of tightly bound water remains in the cell (Bartels and Salamini, 2001).

In land plants, desiccation tolerance can be induced in two types of tissues: (1) the whole plant which is called vegetative desiccation and (2) some parts of the plant (e.g., spores, pollen, and seeds) and that is called reproductive desiccation (Alpert and Oliver, 2002; Giarola et al., 2018). Despite being prevalent in some plant lineages, vegetative desiccation tolerance is not a common feature in the plant kingdom and the vegetative tissues of most flowering angiosperm plants are desiccation sensitive whereas their reproductive pollen, spores, and seeds are desiccation tolerant. The great majority of vegetative desiccation-tolerant plant species are found in the less complex clades (e.g., algae, lichens and bryophytes) (Oliver et al., 2000). However, vegetative desiccation is found within the more complex groups of vascular land plants in fern/fern-allies (60-70 species) (Bewley and Krochko, 1982) and angiosperms (approximately 350 species) (Proctor and Tuba, 2002).

A class of desiccation tolerant plants within angiosperms, resurrection plants, has evolved to adapt to seasonal rainfall and long dry periods. Nevertheless, one resurrection plants belonging to the genus *Lindernia*, specifically *L. brevidens*, is endemic to the montane rainforests of Tanzania and Kenya, where it never experiences seasonal dry periods. Resurrection plants could either exhibit homiochlorophyllly (retain chlorophyll during dehydration) or poikilohydrophyllly (lose chlorophyll during dehydration) (Bartels and Salamini, 2001). *L. brevidens* belongs to the former category and is recognized as a dicotyledonous homoiochlorophyllous desiccation tolerant plant.

The photosystems of homoiochlorophyllous plants survive desiccation essentially intact, even though chloroplasts are very susceptible to oxidative damage caused by water stress (Halliwell, 1987). It is postulated that mechanisms involved in the protection and/or repair of chloroplast localized structures play a fundamental role in vegetative desiccation-tolerance (Oliver et al., 2000). This led us to investigate

the genes responsible for conferring desiccation tolerance and protecting the photosystems from oxidative damage.

The expression of genes in response to dehydration mainly depends on the transcriptional regulation of stress-related genes. These dehydration-induced genes encode proteins that play a vital role in protecting resurrection plants from the subsequent dehydration-induced damage and facilitates their recovery after rehydration (Giarola, 2014). It has been postulated that several protective pathways and regulatory networks are involved in protecting resurrection plants under stress (Alamillo and Bartels, 2001). Nevertheless, identifying individual components of these complex regulatory networks is a daunting task and the identification of some components remains elusive to this day.

In green plants, an array of genes coding for LHCs that collect solar energy and are capable of binding to chlorophyll and carotenoid such as lutein and zeaxanthin. It appears that their primary function is light harvesting via the excitation of chlorophyll and the transfer of absorbed energy to PSII reaction centers (Montane and Kloppstech, 2000). Many newly discovered families of relatives to LHCs have been identified in angiosperms, ferns, algae, and cyanobacteria such as three-helix ELIPs, two-helix SEPs, and the one-helix proteins OHPs (Adamska, 1997; Adamska, 2001). ELIPs of the first nuclear-encoded light-inducible proteins are detectable within the thylakoid membrane system (Meyer and Kloppstech, 1984). Despite the similarity between LHCs and ELIPs at the amino acid sequence level (Green and Kühlbrandt, 1995; Jansson, 1999) ELIPs differ from the LHCs by their transient expression under high-light stress (Potter and Kloppstech, 1993). The gene expression of LHCs is tightly regulated by light induction (Escoubas et al., 1995; Niyogi, 1999). Likewise, it has been shown that high light intensities inhibit transcription of LHCs and activate the synthesis of ELIPs (Potter and Kloppstech, 1993). Unlike LHCs, the gene expression of ELIPs is responsive to a variety of stress signals which depends on the physiological process induced by the signal (Bruno and Wetzel, 2004).

It has been found recently that ELIPs, which were previously thought to be expressed solely under light stress are also upregulated upon dehydration and desiccation (Ahrazem et al., 2016; VanBuren et al., 2018a; VanBuren et al., 2019). ELIPs are not only limited to homiochlorophyllous desiccation tolerant plants such as the mosses *Tortula ruralis* (Zeng et al., 2002) and *Syntrichia caninervis* (Liu et al., 2019), the eudicot *Craterostigma plantagineum* (Bartels et al., 1992; Alamillo and Bartels, 2001; Bartels and Salamini, 2001) but have been also identified in poikilohydrous desiccation tolerant plants such as

Xerophyta viscosa (Costa et al., 2017). Yet, the structures (promoters, active sites, etc.) and functions of ELIPs (ELIP1 and ELIP2) which belong to a non-desiccation tolerant plant, *Arabidopsis thaliana*, have become the most extensively studied as compared to desiccation tolerant plants (Jansson, 1999; Harari-Steinberg et al., 2001; Hulin et al., 2003; Casazza et al., 2005; Rossini et al., 2006; Heddad et al., 2006; Tzvetkova-Chevolleau et al., 2007; Rizza et al., 2011; Rus Alvarez-Canterbury et al., 2014). The characterization of the structures and functions of ELIPs in desiccation tolerant and sensitive plants has been hampered by the lack of genomic data and the differences in genome sizes manifested by their polyploidy. For instance, only two research papers detailing the structure and function of ELIPs in *C. plantagineum* have been published (Bartels et al., 1992; Alamillo and Bartels, 2001). This lack of research published on *C. plantagineum* could be attributed to their massive octoploid genomes (880 Mbp) (Giarola et al., 2017) and lack of genomic sequencing data which discourages researchers from conducting research to investigate the roles of stress-induced genes. For instance, it wasn't until two decades later that someone attempted to conduct extensive research on ELIPs in *C. plantagineum* (Raj, 2020). Therefore, it has become imperative to find alternative desiccation tolerant model organisms whose genomes have been sequenced and are relatively smaller than *C. plantagineum*'s.

A large collection of ELIPs have been recently identified in the newly sequenced genomes of desiccation tolerant including *Boea hygrometrica* (Xiao et al., 2015), *Oropetium thomaeum* (VanBuren et al., 2015), *Xerophyta viscosa* (Costa et al., 2017), *Lindernia brevidens* (VanBuren et al., 2018a), *Selaginella lepidophylla* (VanBuren et al., 2018b), and *Selaginella tamariscina* (Xu et al., 2018), *Physcomitrella patens* (Rensing et al., 2008) and *Marchantia polymorpha* (Bowman et al., 2017) as well as some of desiccation sensitive plant species such as *Lindernia subracemosa* (VanBuren et al., 2018a).

For a long time, the homoiochlorophyllous desiccation tolerant plant *C. plantagineum* has been labelled as a well-suited system for molecular analysis because gene expression of desiccation tolerant genes can be studied and compared in two genetically identical systems (undifferentiated callus cultures and differentiated plants) (Bartels et al., 1990). Nonetheless, the physiological basis of desiccation tolerance in resurrection plants is governed by extremely intricate networks (Bartels and Salamini, 2001) and the size of *C. plantagineum*'s genome makes it difficult to perform transformation and in vitro expression analysis.

Comparative analysis is a powerful tool for dissecting the complex molecular basis of physiological traits. They utilize closely related desiccation tolerant and sensitive species making which facilitates distinguishing between drought and desiccation responses as well as reveal the evolutionary history of desiccation tolerance (VanBuren et al., 2018a). Yet, not all desiccation tolerant species are suitable for comparative analysis. High-quality reference genomes are available for some plants but their huge estimated divergence times deem them incompatible for comparative analysis (Baniaga et al., 2016). Despite the abundance of genomic resources for resurrection plants, comparative analysis is also limited by the lack of sequenced genomes of desiccation sensitive plants.

To solve this issue, two species belonging to the genus *Lindernia*: *L. brevidens* and *L. subracemosa*, stand out as appropriate model plant candidates for this task. The two homiochlorophyllous desiccation tolerant species with their smaller tetraploid genomes, 270 and 250 Mbp respectively, have become of particular interest to the desiccation tolerance scientific community. Carrying out experiments on them is advantageous namely due to the architecture of their small genomes, small gene number, and the retaining of genes as singletons or duplicates after their shared WGD event, make it easy to perform large-scale comparative analysis. Additionally, no significant differences in rRNAs, composition of repetitive elements, or clustering of desiccation-related genes have been observed in the two species (VanBuren et al., 2018). It has been thought that their genomic parameters were responsible for conferring desiccation tolerance (Xiao et al., 2015; Costa et al., 2017) however the similarity in genomic architecture between *L. brevidens* and *L. subracemosa* has led to the hypothesis which states that other elements of the genome architecture are responsible that trait (VanBuren et al., 2018). In our project, *ELIPs* belonging to two members of the Linderniaceae family, *L. brevidens* and *L. subracemosa*, were analyzed and reviewed.

5.1. Analysis of ELIP amino acid sequences

The ORFs of Lbr ELIPs encode deduced polypeptides averaging 222 amino acids in length. Their molecular mass, pI, and estimated charge at pH 7.00 averaged 23.709 KDa, 8.32, and 2.97 respectively. The ORFs of Lsu ELIPs encode deduced polypeptides averaging 178 amino acids in length. Their molecular mass, pI, and estimated charge at pH 7.00 averaged 19.111 KDa, 8.78 and 2.7 respectively (Supplementary table 1). The amino acid composition of Lbr and Lsu ELIPs contain a high content of glycine, leucine, alanine, and serine residues, a pattern exhibited by many other ELIPs.

The N-t of CpELIP have a predicted structure resembling a typical chloroplast-targeting transit peptide, a

random coil followed by a helix. Moreover, all ELIPs possess an extended N terminal (N-t) domain but the main differences among them were present in their signal peptides. Unlike the N-t domain of mitochondrial signal peptides, Arginine and Leucine are significantly underrepresented in the N-t portion of signal peptides of chloroplast-targeted proteins (Zhang and Glaser, 2002). Nevertheless, the overall amino acid distribution in mitochondrial pre-sequences resembles the distribution of these amino acids in the residual part of the transit peptides (Bhushan et al., 2006). Mitochondrial and chloroplast signal peptides are rich with hydroxylated, hydrophobic and positively charged amino acids and very few acidic amino acid residues are frequently interspersed throughout the sequence (Zhang and Glaser, 2002). This led to hypothesize whether or not deleting the N-t chloroplast precursors in the signal peptides would redirect the truncated proteins to the mitochondria. Three proteins were examined: ELIP, PetC and Lhcb2.1. In vivo and in vitro analysis revealed that the truncated ELIP was neither imported into chloroplasts nor miss-targeted to mitochondria, implying that the entire transit peptide is not only required for correct targeting to plastids but also miss-sorting to mitochondria (Bhushan et al., 2006).

The first helix is preceded by an extended N-t domain and a transit peptide whose amino acid sequence is yet to be determined experimentally. Previous studies have pointed out that ELIPs are type II membrane proteins with the N-t and C terminal domains oriented to the stromal and luminal sides of the thylakoid membrane respectively (Adamska, 2001; Grimm et al., 1989). The rate of insertion of ELIPs in thylakoid membranes is independent of temperature and appears to be regulated through the rate of processing of these proteins (Adamska, 1997). In most ELIPs of the higher plants the C-terminal contains a conserved L followed by a V. The length of the signal peptide was previously predicted to be 50 amino acids long (Zeng et al., 2002) however our analysis using Phyre2 revealed that the total length ranges from 37 to 40 amino acids. Since CpELIP is targeted to the thylakoid membranes, we would expect it to contain two signal peptides, the first cleaved after its importation into the plastids and the second cleaved after its incorporation into the thylakoid membranes. However, only cleavage site for the signal has been found. It could be possible that the first signal peptide is required to insert the protein through the chloroplast membrane and the insertion into thylakoid membranes could be mediated by Chl binding. It has been previously demonstrated that the stable insertion of ELIPs into etioplast membranes requires binding to Chl a (Adamska et al., 2000). Some Lbr ELIPs contained an extended C-terminal domain whose function or role in photoprotection has not been elucidated.

Hydropathy plots revealed that all surveyed ELIPs contained three transmembrane helices (Figure 3).

Moreover, charge vs hydropathy plots (Figure 4) demonstrated that all surveyed Lsu ELIPs were partially unfolded except for Lsu ELIP 021038 which is folded. Similarly, Lbr ELIPs were partially unfolded except for two folded proteins Lbr ELIP 025549 and 025589 which are completely folded. However, Lbr ELIP 025589 was missing the entire N-terminal domain while Lbr ELIP 025549 was missing the signal peptide. Lbr ELIP 025585 was missing the entire third helix region. The folded nature of the protein could be an artifact caused by these missing domains. It's noteworthy to mention that the foldability of the protein could affect its migration on the gel (Dries et al., 2011).

A characteristic property of ELIPs is the presence of two internally duplicated transmembrane helices (I and III) (Adamska, 2001) located at an angle of 56° to each other and a second helix (II) which is almost perpendicular to the thylakoid membrane plane (Green and Kühlbrandt, 1995; Adamska, 1997) and appears to contain several polymorphic regions but overall is conserved. Despite the diversity of its residue content, the second helix shows significant conservation within monocot or dicot plant species (Adamska, 2001). A conserved chlorophyll-binding motif (LHC motif) is found in Helices I and III. The generic LHC motif is composed of 19 amino acid residues (Helix I: ERINGRLAMIGFVAALAVE and Helix III: ELWNNGRFAMLGLVALAFTE) (Jansson, 1999; Adamska et al., 2001; Heddad et al., 2012) and is homologous to Helix I and III of all CAB proteins (Green and Kuhlbrandt, 1995) that form the peripheral light-harvesting antenna system of PSI and PSII (Heddad and Adamska, 2000; Tzvetkova-Chevolleau et al., 2007). These two helices form the core of the complex and are bound together by Arg–Glu salt bridges between them (Jansson, 1999). Several universally conserved amino acids for all known ELIPs have been identified including: three alanines (A), two glutamic acids (E), two glycines (G), one arginine (R), one methionine (M) in the first and third helices. The conserved residues: G, A, and M which are essential for close packing of helices I and III in Lhcbl (Kühlbrandt et al., 1994; Green and Kühlbrandt, 1995) might have similar function in ELIPs (Adamska, 2001). Similar to the CAB proteins, the two E residues flanking ELIP consensus sequences in the first and the third helices might act as a stop-transfer signal (Green and Pichersky, 1994). Outside the conserved motifs and helices, the amino acid residues: 3Gs, 1W and 1Q downstream of Helix 1, the E downstream helix II, and the first G downstream helix III are conserved throughout ELIPs.

Many close relatives of CpELIP share a highly hydrophobic motif that contains several glutamic acids (Q) arranged consecutively. This feature (figure 5, Extended N-terminal domain) appears to be exhibited by two relatives belonging to the genus Lindernia (Lbr ELIP 017192 and Lsu ELIP 008078). The repeated

presence of proline-rich peptides is characteristic of extensins which represent a family of hydroxyproline-rich glycoproteins, one of the primary components of plant cell walls (Varner and Lin, 1989). However, there is no evidence pointing out to the presence of glycoproteins in chloroplast membranes or to the role of ELIPs in cell wall-mediated interactions. Keegstra and Cline (1982) showed that thylakoid membranes from pea chloroplasts lack glycosylated polypeptides. All ELIPs have analogous functions all of which will be discussed later, yet further evidence is needed to indicate if the proline-rich regions in ELIP-like, CBR, and CAB proteins impart other functions that are dissimilar or complementary to the original function of ELIPs. However, unlike other close relatives of ELIPs which are mostly light-induced, a wide variety of stress conditions appear to induce the transient expression of ELIPs proteins, a characteristic unique to them (Adamska, 1997). No published research has ascribed this unique expression pattern to the presence of proline-rich motifs. The amino acid sequence of *L. brevidens* and *L. subracemosa* shows high homology to ELIPs belonging to other monocots and eudicots but not CBRs, CABs, and PSBSs.

5.2. ELIPs role as pigment binding proteins

Several pigment molecules including Chl a, chlorophyll b (Chl b), and some carotenoids can bind to the homologous regions across the LHC family (Jansson, 1999) but ELIPs can only bind to carotenoids, Chl a, and lutein (Harari-Steinberg et al., 2001). Under conditions of high light stress, ELIPs bind to free Chl a thus keeping them at low levels (Hutin et al., 2003). It has been visually demonstrated that ELIPs form chlorophyll-pigments complexes but the binding was not experimentally confirmed (Raj, 2020). This observation could be attributed to the weak association of pigments with ELIPs and their low excitonic coupling (Adamska et al., 2020). Nonetheless, the stable insertion of ELIPs into etioplast membranes requires binding to Chl a (Adamska et al., 2000). Therefore, a new experimental setup, that takes into consideration the weak binding nature of ELIPs, could shed more light on its binding capacities.

Eight Chl a, six Chl b, four carotenoids and two lipids can bind to LHCII monomers (Kühlbrandt et al., 1994). ELIPs contain conserved residues which could also be involved in Chl a binding. Therefore, based on the homology between LHC and ELIP family members it is expected that ELIPs will be able to bind at least four Chl a molecules (Green and Kuhlbrandt, 1995). Judging from the CpELIP sequence, the putative chlorophyll-binding residues in CpELIP are distributed within helices I and III (a1: Glutamic acid in position 100; a2: Asparagine in position 103; a3: Glutamic acid in position 167; a4: Asparagine in position 170) (Figure 1). In contrast to LHCI/II proteins, Helix II and the C-t domains were lacking the chlorophyll-binding domains (Green and Kuhlbrandt, 1995).

It was assumed that the localization of carotenoids in close proximity to the chlorophylls aids in preventing photodamage (Green and Kuhlbrandt, 1995) and protect triplet chlorophyll formation. Earlier reports suggested that binding zeaxanthin might allow ELIPs partake a critical role in the non-photochemical quenching of light energy (Krol et al., 1999). Due to the increase in abundance of zeaxanthin pigments and their binding affinity to members of the LHC family, it has been suggested that the accumulation of ELIPs in lhc mutants is concomitant with ELIPs binding zeaxanthin. Furthermore, binding to zeaxanthin might allow ELIPs partake a critical role in the non-photochemical quenching of light energy (Krol et al., 1999). Similarly, it has been shown that CpELIP co-localizes with the xanthophyll cycle pigment zeaxanthin and the inhibition of the latter causes a sharp decline in the accumulation of CpELIP (Alamillo and Bartels, 2001). Interestingly, in the green alga *Dunaliella bardawil* an ELIP homologous protein CBR is up-regulated in response to stress at the same time that there is accumulation of the carotenoid beta-carotene and activation of the photoprotective xanthophyll cycle (Lers et al., 1991, 1993). There is evidence that CBR binds zeaxanthin and that CBR is degraded as violaxanthin is de-epoxidated to zeaxanthin (Levy et al., 1993), suggesting a function for CBR in the xanthophyll cycle.

However, no evidence has been provided to support the existence of ELIP-zeaxanthin complexes under any light stress conditions. The mere correlation manifested in the up-regulation of the two could be an indicator of a yet unknown cross-regulatory pathway which requires the up-regulation of ELIPs. As a matter of fact, no zeaxanthin binding site have been found in ELIPs. Quantitative pigment-binding experiments support a role for ELIPs in pigment-binding, particularly two molecules of the xanthophyll cycle pigment lutein (Adamska et al., 2008). Lutein appears to be the primary carotenoid involved in ELIP-pigment complexes (Adamska et al., 2008). However, in our project we were able to identify only one lutein-binding site (Tryptophan) via sequence homology analysis with LHCBl (Figure 2).

The presence of the xanthophyll lutein in ELIPs could be an important requirement for the binding of free chlorophylls (Adamska, 1997). The aforementioned observation is in accordance with Green and Kuhlbrandt (1995), who postulated that ELIPs are capable of binding only four molecules of Chl a, no chlorophyll b, two luteins, and a number of other non-localized carotenoids. Despite the lack of evidence for ELIP-zeaxanthin binding, the suppression of both ELIP1 and ELIP2 in elip1/elip2 mutant of *A. thaliana* caused perturbations in the levels of chlorophyll and zeaxanthin, implying that ELIPs could possibly disturb the stability or synthesis of those pigments (Tzvetkova-Chevolleau et al., 2007). This is in line with what has been reported by Alamillo and Bartels (2001) in regards to the role of CpELIP in *C.*

plantagineum. The protein could either maintain a hydrophobic environment that promotes the stabilization of zeaxanthin or protects PSII by stabilizing zeaxanthin during rehydration. Nonetheless, there is no experimental evidence to support the previous claims.

These results demonstrate that the structure of Lbr and Lsu ELIPs resemble a typical ELIP protein. They are nuclear-encoded chloroplast-targeted type II membrane proteins which contain a chlorophyll a binding domain capable of binding to four Chl a molecules and capable of co-localizing with the xanthophyll cycle pigment lutein.

5.3. Expansion and expression of *ELIPs* in *L. brevidens* and *L. subracemosa*

The number of *ELIPs* in desiccation tolerant land plants ranges from 10 to 26 with an average of 20.7 whereas the number ranges from 1 to 8 with an average of 3.1 per genome for desiccation sensitive land plants (VanBuren et al., 2019). The number of *ELIPs* in *L. subracemosa* is similar to that in other desiccation sensitive angiosperms (Hayami et al., 2015).

Expansion of *ELIPs* via gene enrichment have been reported in many desiccation tolerant species such as *B. hygrometrica* (Xiao et al., 2015), *S. lepidophylla* (VanBuren et al., 2018b), and *L. brevidens* (VanBuren et al., 2018), yet not all resurrection plants have experienced this enrichment. For instance, no *ELIP* enrichment has been reported in *O. thomaeum* (VanBuren et al., 2015), *X. viscosa* (Costa et al., 2017), or *S. tamariscina* (Xu et al., 2018). After closer examination of the resurrection plants involved in the aforementioned study, we could clearly visualize why this differential enrichment behavior exists. Preferential *ELIP* expansion appears to favor enrichment in homochlorophyllous resurrection plants such as *B. hygrometrica*, *S. lepidophylla*, and *L. brevidens*. On the contrary to homochlorophyllous plants, no enrichment of *ELIPs* has been reported in poikilohydrous resurrection plants such as *O. thomaeum* (VanBuren et al., 2015), *X. viscosa* (Costa et al., 2017), or *S. tamariscina* (Xu et al., 2018). This suggests that the role *ELIPs* play in desiccation tolerance is not completely universal (VanBuren et al., 2019) and that *ELIPs* in each category exhibit roles which may not be involved in protecting the photosynthetic apparatus. Despite having a high copy number, it appears that *ELIPs* in poikilohydrous plants are not capable of protecting the photosynthetic apparatus and retain their chlorophyll content. This is noticeable in *O. thomaeum* which has 22 newly annotated *ELIPs*, in *S. tamariscina* which could possibly have 74 unannotated *ELIPs* (Xu et al., 2018), and in *X. viscosa* which has 10 *ELIPs* (VanBuren et al., 2019). This imply that *ELIPs* are expressed but there might be a deficiency in interacting agents or

promoters that either act down- or up-stream the *ELIP* signalling pathways. Another possibility could be that ELIPs in poikilohydrous plants serve a function different from that found in homoiochlorophyllous plants. More research needs to be conducted to discern the differences of ELIP transcript and protein expression profiles of between homoiochlorophyllous and poikilohydrous resurrection plants.

The classification of *ELIPs* genes as singleton or tandemly duplicated based on their physical proximity in the genome appears to correlate with level of enrichment in singleton *ELIPs* in homoiochlorophyllous and poikilohydrous plants. For example, a large portion of *B. hygrometrica*, *S. lepidophylla*, and *L. brevidens* *ELIPs* are tandemly duplicated but small portion is classified as singletons. Whereas the *ELIPs* of the three poikilohydrous plants *O. thomaeum*, *X. viscosa*, and *S. tamariscina* are all tandemly duplicated. Nevertheless, the level of expression of *ELIPs* in homoiochlorophyllous and poikilohydrous plants does not correlate with their classification as singleton or tandemly duplicated. Despite exceeding the expression threshold (the minimum number of *ELIPs* needed to confer desiccation tolerance properties) none of the poikilohydrous plants retain their chlorophyll. Similarly, even though *ELIP* transcript expression was elevated in poikilohydrous plants (Figure 17), chlorophyll is not retained.

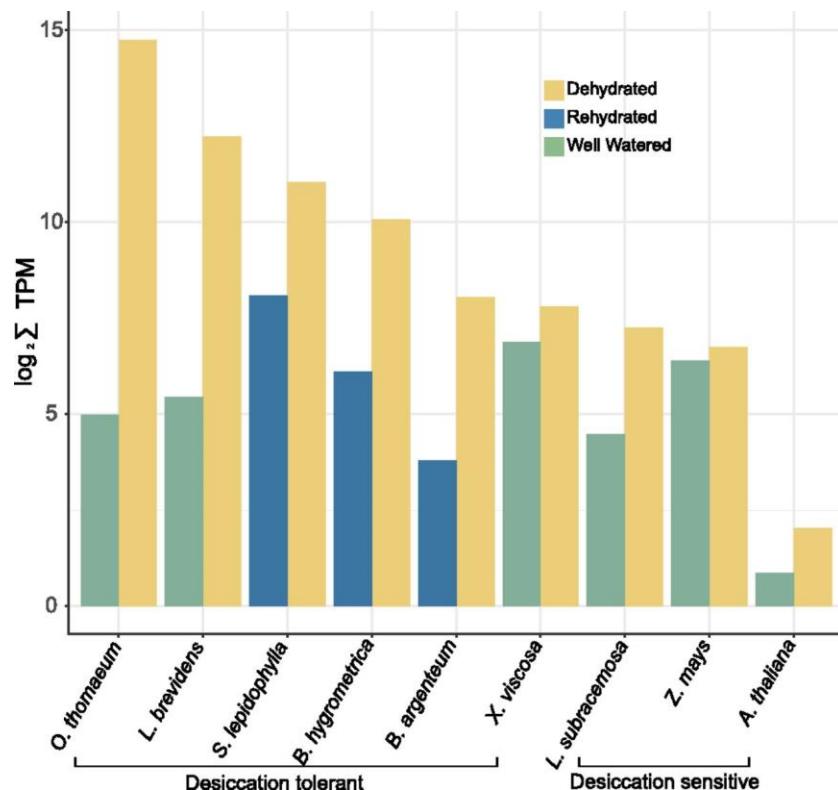


Figure 17. *ELIPs* transcript expression levels in desiccation tolerant and sensitive species. Expression levels under drought/desiccation (yellow), well-watered (green), and rehydrated plant tissue (blue) [Adopted from VanBuren et al. (2019)].

It is likely that ELIPs in resurrection plants have maintained their ancestral photoprotective role, but tandem gene duplication may have increased the absolute abundance of *ELIP* transcripts and improved their photoprotective capacity (VanBuren et al., 2019). The previous evidence supports a photoprotective role for ELIPs that does not involve binding to chlorophyll in poikilohydrous resurrection plants.

The presence of 24 ELIP proteins in *L. brevidens* and only 4 in its desiccation sensitive counterpart, *L. subracemosa*, raised the question about their redundant physiological function. Therefore, to elucidate whether a correlation exists between dehydration/rehydration and *ELIP* transcript abundance, *ELIP* expression levels have been investigated. Gene expression of leaf tissues experiencing dehydration and subsequent rehydration was assessed using RT-PCR and real-time PCR. Expression levels of *ELIPs* whose sequences have been verified shall be discussed. To improve the specificity of the primers, it is recommended that amplifications are re-run after conducting melting curve analysis.

In *L. brevidens*, gene expression is relatively stable from severe dehydration to desiccation (VanBuren et al., 2018a). Agarose gel electrophoresis of total RNA of *L. subracemosa* RNA further consolidates this claim (Figure 10). Lbr *ELIP 016551*, Lbr *ELIP 025593.27*, Lsu *ELIP 008078*, and Lsu *ELIP 030118* were constitutively expressed at low levels in untreated well-watered plants. This is consistent with the observations of VanBuren et al. (2019) which showed that in all surveyed tolerant and sensitive species, *ELIPs* either have low or undetectable expression under well-watered conditions. Transcript abundance increased rapidly when the *L. brevidens* were dehydrated and peaked in the leaves of desiccated plant whereas *L. subracemosa* *ELIPs* retained a steady-state level of expression in dehydrated and desiccated plants. *ELIP* expression returned to basal levels in the rehydrated tissues of *L. brevidens* (Figure 11). The previous observation in *L. brevidens* is similar to what has been previously reported. *ELIP* expression is elevated upon dehydration treatment and is the highest in desiccated tissues (Figures 17 and 18) (VanBuren et al., 2018a; VanBuren et al., 2019). In the ML and MP trees (Figures 8 and 9) it looks like all tandemly duplicated Lbr *ELIPs* clustered as one group (Figure 18) implying that they might share the same ancestral history.

Agarose gel electrophoresis of total RNA reveals the degradation of RNA in the desiccated leaf tissues of the desiccation sensitive plant *L. subracemosa*. This could be caused by decreased synthesis or degradation of RNAs. On average *ELIPs* have a 622-fold higher expression abundance in desiccation tolerant plants than desiccation sensitive species under water deficit (VanBuren et al., 2019). In *L.*

subracemosa, it has been shown that a high proportion of genes were differentially expressed during the transition from well-watered to desiccation (VanBuren et al., 2018a). However, *ELIP* transcript abundance did not decline in response to further dehydration and remained uniform even in desiccated tissues (Figure 10).

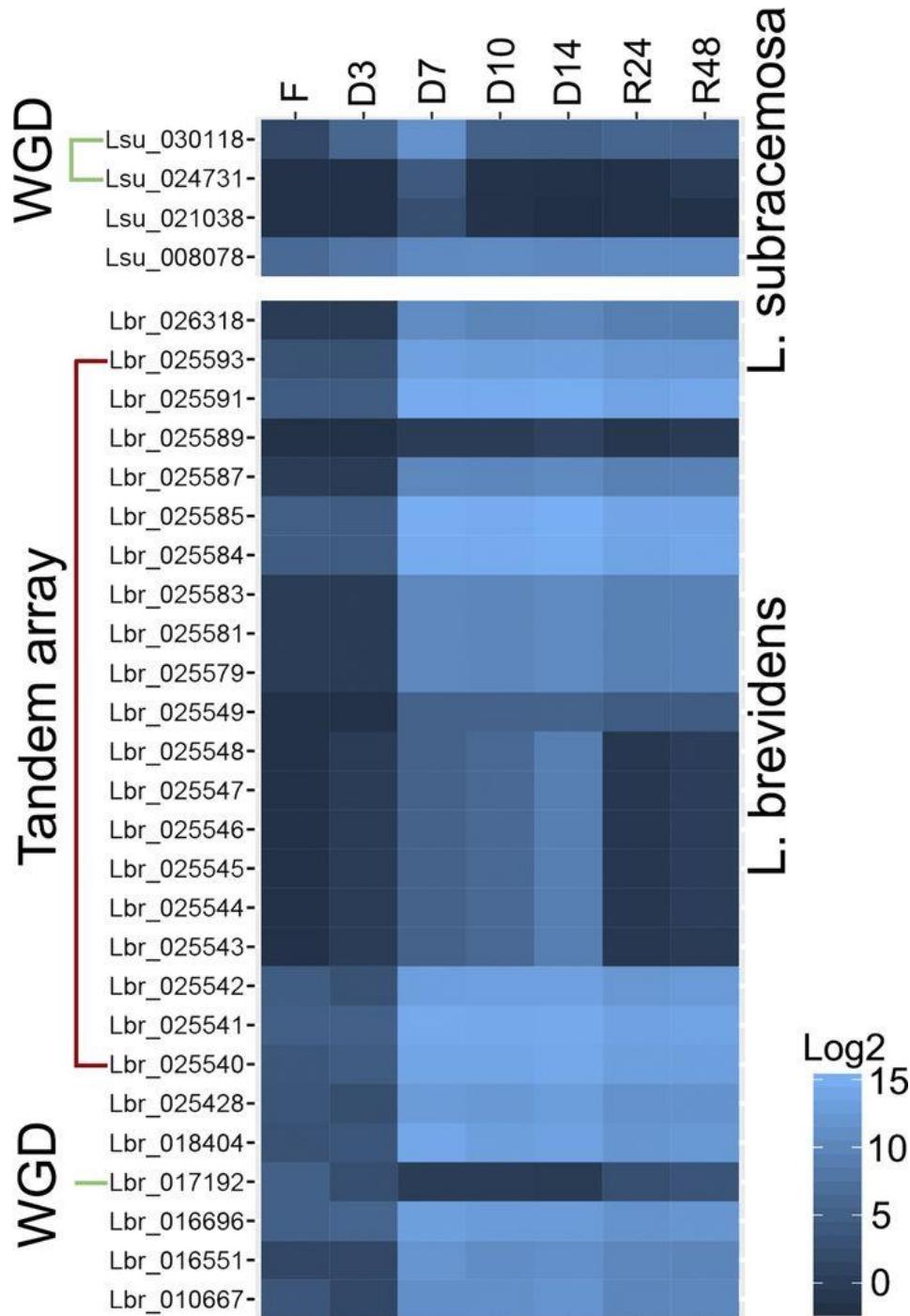


Figure 18. Log2 transformed transcript expression profiles of Lbr and Lsu *ELIPs*. The tandem array and syntenic whole-genome duplicates are shown [Adopted from VanBuren et al. (2018a)].

Even though the decline in transcript expression of constitutively expressed genes suggests that mRNAs are unstable in *L. subracemosa* (Dinakar and Bartels, 2012) *ELIP* transcript abundance did not vary. This observation could be interpreted by figure 17 adopted from VanBuren et al. (2019). On the contrary to the down-regulation of some transcripts such as those encoding antioxidants in desiccated *L. subracemosa* leaf tissues (Dinakar and Bartels, 2012) *ELIP* transcript levels in desiccation sensitive plants (e.g., *L. subracemosa*, *Z. mays*, and *A. thaliana*) does not differ between well-watered and dehydrated tissues. The slight increase in *ELIP* expression in desiccation sensitive plants could not be observed on the agarose gels and future experiments need to quantify the expression level using fluorescence measurements in real-time quantitative PCR. The aforementioned observations suggest that in *L. subracemosa*, the uniformity of *ELIP* expression could be mediated by an unknown ELIP-stabilizing mechanism or could be a characteristic of a gene expression profile that reflects *L. subracemosa* imminent death. The expression of *ELIPs* in *L. subracemosa* appears to be regulated by normal light illumination and genes could be constitutively expressed upon daily light illumination. However, this could be an artifact resulting from using a wide-spectrum wavelength during illumination. That is, it is still unclear which light wavelength affect level of *ELIP* expression in *L. brevidens* or *L. subracemosa* and further experiments to elucidate the role of photoreceptors in ELIP induction need to be conducted. This is important because differential expression of *ELIPs* in response to different wavelengths of light stimuli has been previously reported in eudicots and mosses (Harari-Steinberg et al., 2001; Rizza et al., 2011; Liu et al., 2020).

Environmental factors such as developmental stage and rate of drying in addition to dehydration priming influence the recovery rate *L. brevidens*. Post-rehydration, *L. brevidens* plants were viable whereas *L. subracemosa* plants were largely dead (VanBuren et al., 2018a). Transcript abundance of Lbr *ELIPs* 016551 and 025593.27 declined upon dehydration but they were not completely degraded. In this report, no data has been provided regarding rehydrated leaf tissues in *L. subracemosa*. Nevertheless, our results are consistent with the findings of VanBuren et al. (2018a) which showed that nearly the transcript abundance of nearly all Lbr *ELIP* genes increased during severe dehydration, desiccation, and rehydration, but was hardly noticeable in well-watered and mildly dehydrated tissue. This suggests that Lbr *ELIPs* might play a significant role in the recovery of the function of the photosynthetic apparatus upon rehydration before they are completely degraded. It has been previously proposed that ELIPs could help reassemble chlorophyll in the photosynthetic complexes (Adamska, 2001). Hulin et al. (2003) suggested a similar role for ELIPs which involves binding newly formed chlorophyll molecules then integrating them into the photosynthetic complexes (Hulin et al., 2003). However, a counter argument has posed the

possibility that ELIPs are not involved directly in the synthesis and assembly of photosynthetic complexes, but instead influence the biogenesis of chlorophyll-binding complexes. Interaction between ELIPs and the chlorophyll biosynthesis pathway has been identified in *A. thaliana*. Two possible candidates which might interact with ELIPs are Glu tRNA reductase (GluTR) and 5-aminolevulinate (ALA) biosynthetic pathways. In ELIP2-overexpressing *A. thaliana* plants, the protein level of GluTR and the synthesis rate of ALA were drastically reduced, suggesting that ELIP2 interferes with the initial rate-limiting steps of chlorophyll synthesis (Tzvetkova-Cheolleau et al., 2007).

6. Conclusions

L. brevidens is a dicotyledonous homochlorophyllous resurrection plant that belongs to the family Linderniaceae and exhibits modified desiccation tolerance in its vegetative tissues. Many stress-induced genes are induced when the plant experiences water deficit. At the center of these stress-induced genes lies *ELIPs*. A wide variety of environmental stresses result in the accumulation of *ELIP* transcripts such as heat, cold, and high light stress. Similarly, *ELIPs* accumulate rapidly upon dehydration but they are subsequently and gradually de-regulated upon rehydration. Due to the presence of ELIPs in rehydrated tissues it is postulated that ELIPs of *L. brevidens* might play a vital role in the recovery of the function of the photosynthetic apparatus, the reassembly of chlorophyll in their photosynthetic complexes, or the biogenesis of chlorophyll-binding complexes. Despite the degradation of RNA in leaf desiccated tissues, the *ELIPs* of *L. subracemosa* appear to be constitutively expressed upon normal light illumination and neither dehydration nor desiccation affect their transcript expression profiles. This could be caused by an ELIP-stabilizing mechanism or could simply reflect the imminent death of *L. subracemosa*. ELIPs belong to the CAB/ELIP/HLIP superfamily of distantly related polypeptides. Our results revealed that Lbr and Lsu ELIPs are nuclear-encoded chloroplast-targeted type II membrane proteins which are composed of three transmembrane helices. They contain a chlorophyll a binding domain capable of binding to four chlorophyll a molecules and capable of co-localizing with the xanthophyll cycle pigment lutein. All of the investigated amino acid sequences of *L. brevidens* and *L. subracemosa* ELIPs presented noteworthy similarity to monocot and eudicot nuclear-encoded chloroplast-targeted ELIPs which are actively expressed upon dehydration, desiccation, and after the beginning of illumination. ELIPs might have a photoprotective role that involves binding to chlorophyll and lutein but thus far neither the mechanism nor the binding have been experimentally proven.

7. References

- Adamska, I. (1997). ELIPs - Light-induced stress proteins. *Physiologia Plantarum*, 100: 794-805.
- Adamska, I. (2000). The ELIP family of stress proteins in the thylakoid membranes of pro and eukaryotes, Regulation of Photosynthesis, Aro, E.M., Andersson, B., Eds., Dordrecht: Kluwer, 2001, pp. 487–505.
- Adamska, I. and Kloppstech, K. (1991). Evidence for an association of the early light-inducible protein (ELIP) of pea with photosystem II. *Plant Molecular Biology*, 16: 209–223.
- Adamska, I., et al. (2008). Isolation of Pigment-Binding Early Light-Inducible Proteins from Pea.” *European Journal of Biochemistry*, 260(2): 453–460.
- Adamska, I., Kruse, E., and Kloppstech, K. (2000). Stable Insertion of the Early Light-induced Proteins into Etioplast Membranes Requires Chlorophylla. *Journal of Biological Chemistry*, 276(11): 8582-8587.
- Adamska, I., Lindahl, M., Roobol-Bóza, M., and Andersson, B. (1996). Degradation of the light-stress protein is mediated by an ATP-independent, serine-type protease under low-light conditions. *European Journal of Biochemistry*, 236: 591–599.
- Adamska, I., Ohad, I., and Kloppstech, K. (1992). Synthesis of the early light-inducible protein is controlled by blue light and related to light stress. *Proceedings of the National Academy of Sciences*, 89: 2610–2613.
- Adamska, I., Roobol-Bóza, M., Lindahl, M., and Andersson, B. (1999). Isolation of pigment-binding early light-inducible proteins from pea. *European Journal of Biochemistry*, 260(2): 453-460.
- Ahmadizadeh, M., et al. (2012). Behavior of durum wheat genotypes under normal irrigation and drought stress conditions in the greenhouse. *African Journal of Biotechnology*. 11(8): 1912-1923.
- Ahrazem, O., Argandoña, J., Castillo, R., Rubio-Moraga, Á., and Gómez-Gómez, L. (2016) Identification and Cloning of Differentially Expressed SOUL and ELIP Genes in Saffron Stigmas Using a Subtractive Hybridization Approach. *PLOS ONE*, 11(12): e0168736.
- Alamillo, J.M. and Bartels, D. (2001). Effects of Desiccation on Photosynthesis Pigments and the ELIP-like Dsp 22 Protein Complexes in the Resurrection Plant *Craterostigma Plantagineum*.” *Plant Science*, 160(6): 1161–1170.

- Almagro Armenteros, J.J., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology*, 37: 420–423.
- Alpert P. (2000) The discovery, scope, and puzzle of desiccation-tolerance in plants. *Plant Ecology*, 151: 5–17.
- Alpert P. (2006) Constraints of tolerance: why are desiccation tolerant organisms so small or rare? *Journal of Experimental Biology*, 209: 1575– 1584.
- Alpert, P. and Oliver, M.J. (2002). Drying without dying. In Desiccation and Survival in Plants: Drying Without Dying (ed. M. Black and H. W. Prichard), pp.3 -43. Wallingford, UK: CAB International.
- Al-Sadi, A.M., Al-Masoudi, R.S., Al-Habsi, N., Al-Said, F.A., Al-Rawahy, S.A., Ahmed, M., and Deadman, M.L. (2010). Effect of salinity on pythium damping-off of cucumber and on the tolerance of Pythium aphanidermatum. *Plant Pathology*, 59: 112–120.
- Al-Whaibi, M.H. (2011). Plant heat-shock proteins: a mini review. *Journal of King Saud University Science*, 23: 139–150.
- Andersson, U., Heddad, M., and Adamska, I. (2003). Light stress-induced one-helix protein of the chlorophyll a/b-binding family associated with photosystem I. *Plant Physiology*, 132: 811-820.
- Araus, J.L., Slafer, G.A., Reynolds, M.P., and Royo, C. (2002). Plant Breeding and Drought in C3 Cereals: What Should We Breed For? *Annals of Botany*, 89:925-940.
- Asada, K. (1994). Mechanism for scavenging reactive molecules generated in chloroplasts under light stress. In Photoinhibition of Photosynthesis, from Molecular Mechanisms to the Field (Baker, N.R. and Bowyer, J.R., eds), pp. 129–142. Bios Scientific Publishers, Oxford.
- Atkinson, N.J. and Urwin, P.E. (2012). The interaction of plant biotic and abiotic stresses: from genes to the field. *Journal of Experimental Botany*, 63: 3523–3543.
- Atkinson, N.J., Lilley, C.J., and Urwin, P.E. (2013). Identification of Genes Involved in the Response of Arabidopsis to Simultaneous Biotic and Abiotic Stresses. *Plant Physiology*, 162(4): 2028–2041.
- Baniaga, A.E., Arrigo, N., and Barker, M.S. (2016). The small nuclear genomes of Selaginella are associated with a low rate of genome size evolution. *Genome Biology and Evolution*, 8: 1516–1525.

Banks, J.A., et al. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science*, 332: 960–963.

Barczak-Brzyzek, A.K., et al. (2017). Cross-Talk between High Light Stress and Plant Defence to the Two-Spotted Spider Mite in *Arabidopsis Thaliana*. *Experimental and Applied Acarology*, 73(2): 177–189.

Barry G. H. (2013). Building Phylogenetic Trees from Molecular Data with MEGA, *Molecular Biology and Evolution*, 30(5): 1229–1235.

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

Bartels, D. and Hussain, S.S. (2011). Chapter 16 Resurrection Plants: Physiology and Molecular Biology. In U. Luttge et al. (Eds.), Title of work: Plant Desiccation Tolerance, Ecological Studies (pp. 339-364). Springer-Verlag Berlin Heidelberg 2011. DOI 10.1007/978-3-642-19106-0_16.

Bartels, D. and 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: 1346–1353.

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

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

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

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

Bewley, J.D. and Krochko, J.E. (1982). In Encyclopedia of Plant Physiology, Vol 12B, Physiological Ecology II; Lange, O. L.; Nobel, P. S.; Osmond, C. B.; Ziegler, H., Eds.; Springer-Verlag: Berlin, pp 325-378.

- Bhushan, S., Kuhn, C., Berglund, A.-K., Roth, C., and Glaser, E. (2006). The role of the N-terminal domain of chloroplast targeting peptides in organellar protein import and miss-sorting. *FEBS Letters*, 580(16): 3966–3972.
- Bilgin, D.D., et al. (2010). Biotic stress globally downregulates photosynthesis genes. *Plant, Cell and Environment*, 33: 1597–1613.
- Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic acids research*, 7(6): 1513–1523.
- Bohnert, H.J., Nelson, D.J., and Jensen, R.G. (1995). Adaptations to environmental stresses, *Plant Cell*, 7: 1099–1111.
- Bostock, R.M., Pye, M.F., and Roubtsova, T.V. (2014). Predisposition in plant disease: Exploiting the nexus in abiotic and biotic stress perception and response. *Annual Review of Phytopathology*, 52: 517–549.
- Bowman, J.L., et al. (2017). Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell*, 171: 287–304. e215.
- Boyer, J.S. (1970). Leaf enlargement and metabolic rates in corn, soybean, and sunflower at various leaf waterpotentials. *Plant Physiology*, 46: 233–235.
- Braun, P., Banet, G., Tal, T., Malkin, S., and Zamir, A. (1996). Possible Role of Cbr, an Algal Early-Light-Induced Protein, in Nonphotochemical Quenching of Chlorophyll Fluorescence. *Plant Physiology*, 110: 1405–1411.
- Bruno, A.K. and Wetzel C.M. (2004). The Early Light-Inducible Protein (ELIP) Gene Is Expressed during the Chloroplast-to-Chromoplast Transition in Ripening Tomato Fruit. *Journal of Experimental Botany*, 55(408): 2541–2548.
- Casazza, A.P., et al. (2005). Mutational and Expression Analysis of ELIP1 and ELIP2 in *Arabidopsis thaliana*. *Plant Molecular Biology*, 58(1): 41–51.
- Chaves, M. M., Flexas, J., and Pinheiro, C. (2008). Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Annals of Botany*, 103(4): 551–560.

- Cohen, S.P., and Leach, J.E. (2019). Abiotic and biotic stresses induce a core transcriptome response in rice. *Scientific Reports*, 9(1).
- Costa, M., et al. (2017). A footprint of desiccation tolerance in the genome of *Xerophyta viscosa*. *Nature Plants*, 3: 17038.
- Cramer, G.R., Urano, K., Delrot, S., Pezzotti, M., and Shinozaki, K. (2011). Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biology*, 11(1).
- Czolpinska, M. and Rurek, M. (2018). Plant Glycine-Rich Proteins in Stress Response: An Emerging, Still Prospective Story. *Frontiers in Plant Science*, 9.
- Dai, A. (2012). Increasing drought under global warming in observations and models. *Nature Climate Change*, 3: 52–8.
- Demmig-Adams, B. and Adams, W.W. (2002). Food and photosynthesis: antioxidants in photosynthesis and human nutrition. *Science*, 298: 2149–2153.
- Deyholos, M.K. (2010). Making the most of drought and salinity transcriptomics, *Plant, Cell and Environment*, 33: 648-654.
- Dinakar, C. and Bartels, D. (2012). Light response, oxidative stress management and nucleic acid stability in closely related Linderniaceae species differing in desiccation tolerance. *Planta*, 236: 541–555.
- Dinakar, C. and Bartels, D. (2013). Desiccation tolerance in resurrection plants: New insights from transcriptome, proteome, and metabolome analysis. *Frontiers in Plant Science*, 4: 1–14.
- Dresselhaus, T. and Hückelhoven, R. (2018). Biotic and Abiotic Stress Responses in Crop Plants. *Agronomy*. 8, 267.
- Dries, Niels Van Den, et al. (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.
- Durnford, D.G., Deane, J.A., Tan, S., McFadden, G.I., Gantt, E., and Green, B.R. (1999). A phylogenetic assessment of the eukaryotic light-harvesting antenna proteins, with implications for plastid evolution. *Journal of Molecular Evolution*, 48: 59–68.

Engebrecht, J. and Brent, R. (1998). Preparation of Plasmid DNA. *Current Protocols in Protein Science*, A.4C.1–A.4C.3. doi:10.1002/0471140864.psa04cs13.

Escoubas, J.M., Lomas, M., LaRoche, J., and Falkowski, P.G. (1995). Light intensity regulation of cab gene transcription is signaled by the redox state of the plastoquinone pool. *Proceedings of the National Academy of Sciences*, 92: 10237–10241.

Fang, Y. and Xiong, L. (2015). General mechanisms of drought response and their application in drought resistance improvement in plants. *Cellular and Molecular Life Science*, 72: 673-689.

Farooq, M., Wahid, A., Fujita, N.K.D., and Basra, S.M.A. (2009). Plant drought stress: effects, mechanisms and management. *Agronomy for Sustainable Development*, 29: 185-212.

Felsenstein, J. (1985). Phylogenies and the comparative method. *American Naturalist*, 125: 1–15.

Fischer, E. (2004). Scrophulariaceae. in: Kadereit J. W. (ed.), The families and genera of vascular plants 7. Flowering plants. Dicotyledons. Lami-ales (except Acanthaceae including Avicenniaceae). – Berlin, Heidelberg, New York: Springer. pp. 333 – 432.

Flexas, J., Bota, J., Loreto, F., Cornic, G., and Sharkey, T.D. (2006). Diffusive and metabolic limitations to photosynthesis under drought and salinity in C3 plants, *Plant Biology*, 6: 269-279.

Foyer, C.H., Rasool, B., Davey, J., and Hancock, R.D. (2016). Cross tolerance to biotic and abiotic stresses in plants: a focus on resistance to aphid infestation. *Journal of Experimental Botany*, 67: 2025–2037.

Funk C. (2001). The PsbS protein: A Cab protein with a function of its own. In Advances in Photosynthesis and Respiration — Regulation of Photosynthesis, vol 11, Aro E-M, Andersson B (eds). Kluwer Academic: Dordrecht, Boston, London; 453–467.

Funk, C., Adamska, I., Green, B.R., Andersson, B., and Renger, G. (1995). The nuclear-encoded chlorophyll-binding photosystem II-S protein is stable in the absence of pigments. *Journal of Biological Chemistry*, 270: 30141–30147.

Gaff, D. F. and Oliver, M. (2013). The evolution of desiccation tolerance in angiosperm plants: a rare yet common phenomenon. *Functional Plant Biology*, 40(4): 315 – 328.

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

- Gassmann, W., Appel, H. M., and Oliver, M. J. (2016). The interface between abiotic and biotic stress responses. *Journal of Experimental Botany*, 67(7): 2023–2024.
- Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., and Bairoch A. (2005). Protein Identification and Analysis Tools on the ExPASy Server; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press. pp. 571-607.
- Ghosh, D. and Xu, J. (2014). Abiotic stress responses in plant roots: a proteomics perspective. *Frontiers in Plant Science*, 5.
- Giarola, V. (2008). Expression study of Ls11-24 a dehydration induced gene in Lindernia subracemosa (Linderniaceae). (Master thesis dissertation). University of Bonn.
- Giarola, V., Hou, Q., and Bartels, D. (2017). Angiosperm Plant Desiccation Tolerance: Hints from Transcriptomics and Genome Sequencing. *Trends in Plant Science*, 022(8): 705–717.
- Giarola, V., Jung, N., Singh, A., Satpathy, P., and Bartels, D. (2018). Analysis of pcC13-62 promoters predicts a link between cis-element variations and desiccation tolerance in Linderniaceae, *Journal of Experimental Botany*, 69(15): 3773–3784.
- Giarola, V. (2014). Identification and molecular characterization of unknown genes involved in desiccation tolerance mechanisms in the resurrection plant *Craterostigma plantagineum* (Doctoral dissertation). University of Bonn.
- Green, B.R. and Kühlbrandt, W.O. (1995). Sequence conservation of light harvesting and stress response proteins in relation to the three-dimensional molecular structure of LHCII, *Photosynthesis Research*, 4: 139–148.
- Green, B.R. and Pichersky, E. (1994). Hypothesis for the evolution of three-helix chlorophyll a/b and chlorophyll a/c light harvesting antenna proteins from two-helix and four-helix ancestors. *Photosynthesis Research*, 39: 149–162.
- Grimm, B. and Kloppstech, K. (1987). The early light-inducible proteins of barley. Characterization of two families of 2-h-specific nuclear-coded chloroplast proteins. *European Journal of Biochemistry*, 167: 493–499.

Grimm, B., Kruse, E., and Kloppstech, K. (1989). Transiently expressed early light-inducible thylakoid proteins share transmembrane domains with light-harvesting chlorophyll binding proteins. *Plant Molecular Biology*, 13: 583–593.

Gruszecki W. I. and Strzalka K. (2005). Carotenoids as modulators of lipid membrane physical properties. *Biochimica et Biophysica Acta*, 1740: 108–115.

Halliwell, B. (1987). Oxidative damage, lipid peroxidation and antioxidant protection in chloroplasts. *Chemistry and Physics of Lipids*, 44: 327–340.

Harari-Steinberg, O., Ohad, I., and Chamovitz, D.A. (2001). Dissection of the light signal transduction pathways regulating the two early light induced protein genes in Arabidopsis, *Plant Physiology*, 127: 986–997.

Hayami N., et al. (2015). The responses of Arabidopsis early light-induced protein2 to ultraviolet B, high light, and cold stress are regulated by a transcriptional regulatory unit composed of two elements. *Plant Physiology*, 169: 840–855.

Heddad, M. and Adamska, I. (2000). Light stress regulated two helix proteins in Arabidopsis thaliana related to the chlorophyll a/b binding gene family, *Proceedings of the National Academy of Sciences*, 97: 3741–3746.

Heddad, M., Engelken, J., and Adamska, I. (2012). Light stress proteins in viruses, cyanobacteria and photosynthetic eukaryota. In J Eaton-Rye, BC Tripathy, TD Sharkey, eds, *Photosynthesis: Plastid Biology, Energy Conversion and Carbon Assimilation*, Vol 34. Springer, Dordrecht, The Netherlands, pp 299–317.

Heddad, M., et al. (2006). Differential Expression and Localization of Early Light-Induced Proteins in Arabidopsis. *Plant Physiology*, 142(1): 75–87.

Hsiao, T. (1973). Plant responses to water stress. *Annual Review of Plant Biology*, 24: 519–570.

Hutin, C., et al. (2002). Double Mutation cpSRP43 –/cpSRP54– Is Necessary to Abolish the CpSRP Pathway Required for Thylakoid Targeting of the Light-Harvesting Chlorophyll Proteins. *The Plant Journal*, 29(5): 531–543.

- Hutin, C., et al. (2003). Early Light-Induced Proteins Protect Arabidopsis from Photooxidative Stress. *Proceedings of the National Academy of Sciences*, 100(8): 4921–4926.
- Illing, N., Denby, K.J., Collett, H., Shen, A., and Farrant, J.M. (2005). The signature of seeds in resurrection plants: a molecular and physiological comparison of desiccation tolerance in seeds and vegetative tissues. *Integrative and Comparative Biology*, 45: 771–787.
- Jansson, S. (1999). A guide to the Lhc genes and their relatives in Arabidopsis. *Trends in Plant Science*, 4(6): 236–240.
- Jansson, S., Andersson, J., Kim, S.J., and Jackowski, G. (2000). An Arabidopsis thaliana protein homologous to cyanobacterial high-light-inducible protein. *Plant Molecular Biology*, 42: 345–351.
- Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences* 8: 275-282.
- Jordan, W.R. Clark, R.B., and Seetharama, N. (1984). The Role of Edaphic Factors in Disease Development. In: Proceedings of the Consultative Group Discussion on Research Needs and Strategies for Control of Sorghum Root and Stalk Rot Diseases, 27 Nov - 2 Dec 1983, Bellagio, Italy.
- Kappen, L. and Valladares, F.J. (2007). 2 Opportunistic Growth and Desiccation Tolerance: The Ecological Success of Poikilohydrous Autotrophs.
- Keegstra, K. and Cline, K. (1982). Evidence that Envelope and Thylakoid Membranes from Pea Chloroplasts Lack Glycoproteins. *Plant physiology*, 70(1): 232–237.
- Kissoudis, C., et al. (2015). Combined biotic and abiotic stress resistance in tomato. *Euphytica*, 202(2): 317–332.
- Kissoudis, C., van de Wiel, C., Visser, R.G.F., and Van Der Linden, G. (2014). Enhancing crop resilience to combined abiotic and biotic stress through the dissection of physiological and molecular crosstalk. *Frontiers in Plant Science*, 5: 207.
- Kneebone, W., Mancino, C., and Kopec, D. (1992). Water requirements and irrigation. Turfgrass. *Alliance of Crop, soil, and Environmental Science Societies*, 1992.

- Kolanus, W., Scharnhorst, C., Kühne, U., and Herzfeld, F. (1987). The structure and light-dependent transient expression of a nuclear-encoded chloroplast protein gene from pea (*Pisum sativum* L.). *Molecular and General Genetics*, 209: 234–239.
- Krol, M., et al. (1995). Chlorophyll a/b-Binding Proteins, Pigment Conversions, and Early Light-Induced Proteins in a Chlorophyll b-Less Barley Mutant. *Plant Physiology*, 107(3): 873–883.
- Krol, M., et al. Exposure of *Dunaliella salina* to Low Temperature Mimics the High Light-Induced Accumulation of Carotenoids and the Carotenoid Binding Protein (Cbr). *Plant and Cell Physiology*, 8(2): 213–216.
- Król, M., Ivanov, A.G., Jansson, S., Kloppstech, K., and Huner, N.P.A. (1999). Greening under high light or cold temperature affects the level of xanthophyll-cycle pigments, early light-inducible proteins, and light-harvesting polypeptides in wild-type barley and the Chlorina f2 mutant. *Plant Physiology*, 120: 193–203.
- Kuhlbrandt, W., Wang, D.N., and Fujiyoshi, Y. (1994). Atomic model of plant light-harvesting complex by electron crystallography. *Nature*, 367: 614–621.
- Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35: 1547-1549.
- Kyte, J. and Doolittle, R. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157: 105–132.
- Lawlor, D.W. (2013). Genetic engineering to improve plant performance under drought: physiological evaluation of achievements, limitations, and possibilities. *Journal of Experimental Botany*, 64: 83–108.
- Le Guilloux, V., Schmidtke, P., and Tuffery, P. (2009). Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics*, 2009, 10: 168.
- Lee, J., Kim, T.J., Lim, Y.P., Bang, J.W., and Hur, Y. (2006). Molecular cloning and characterization of an early light-inducible gene, BrELIP, from *Brassica rapa*, and its overexpression protects *Arabidopsis*. *Korean Journal of Genetics*, 28: 207–220.

Lers, A., Levy, H., and Zamir, A. (1991). Co-regulation of a gene homologous to early light-induced genes in higher plants and β -carotene biosynthesis in the alga *Dunaliella bardawil*. *Journal of Biological Chemistry*, 266: 13698–13705.

Levitt, J. (1980). Responses of Plants to Environmental Stresses. Water, Radiation, Salt and Other Stresses, vol. 2, Academic Press, New York.

Levy, H., Tal, T., Shaish, A., and Zamir, A. (1993). Cbr, an algal homolog of plant early light-induced proteins, is a putative zeaxanthin binding protein. *Journal of Biological Chemistry*, 268: 20892–20896.

Li, S., Yu, X., Cheng, Z., Yu, X., Ruan, M., Li, W., and Peng, M. (2017). Global Gene Expression Analysis Reveals Crosstalk between Response Mechanisms to Cold and Drought Stresses in Cassava Seedlings. *Frontiers in Plant Science*, 8.

Liu, X., et al. (2020). Functional Aspects of Early Light-Induced Protein (ELIP) Genes from the Desiccation-Tolerant Moss *Syntrichia caninervis*. *International Journal of Molecular Sciences*, 21(4).

Mafakheri, A., Siosemardeh, A., Bahramnejad, B., Struik, P. and Sohrabi, Y. (2010). Effect of drought stress on yield, proline and chlorophyll contents in three chickpea cultivars. *Australian Journal of Crop Science*, 4: 580.

Maiti, R. and Satya, P. (2014). Research advances in major cereal crops for adaptation to abiotic stresses. *GM Crops Food*, 5: 259–279.

Martins, M.Q., et al. (2016). Protective Response Mechanisms to Heat Stress in Interaction with High [CO₂] Conditions in Coffea Spp. *Frontiers in Plant Science*, 7.

Melotto, M., Underwood, W., Koczan, J., Nomura, K. and He, S.Y. (2006). Plant stomata function in innate immunity against bacterial invasion. *Cell*, 126, 969–980.

Meyer, G. and Kloppstech, K. (1984). A rapidly light-induced chloroplast protein with a high turnover coded for by pea nuclear DNA. *European Journal of Biochemistry*, 138: 201–207.

Mishler, B.D. and Churchill, S. P. (1985). Transition to a land flora: phylogenetic relationships of the green algae and bryophytes. *Cladistics*, 1: 305–328.

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

Miyashita, K., Tanakamaru, S., Maitani, T., and Kimura, K. (2005). Recovery responses of photosynthesis, transpiration, and stomatal conductance in kidney bean following drought stress. *Environmental and Experimental Botany*, 53: 205-214.

Montane M.H., Dreyer, S., and Kloppstech, K. (1996). Post-Translational Stabilization of ELIPs and Regulation of other Light Stress Genes under Prolonged Light and Cold Stress in Barley. In: Grillo S., Leone A. (eds) Physical Stresses in Plants. Springer, Berlin, Heidelberg.

Montane, M.H. and Kloppstech, K. (2000). The family of light-harvesting-related proteins (LHCs, ELIPs, HLIPs): was the harvesting of light their primary function? *Gene*, 258: 1.

Montane, M.H., Dreyer, S., Triantaphyllides, C., and Kloppstech, K. (1997). Early light-inducible proteins during long-term acclimation of barley to photooxidative stress caused by light and cold: high level of accumulation by posttranscriptional regulation. *Planta*, 202: 293–302.

Morgan, J.M. (1984). Osmoregulation and Water Stress in Higher Plants. *Annual Review of Plant Physiology*, 35: 299–319.

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

Neale, A.D., Blomstedt, C.K., Bronson, P., Le, T.N., Guthridge, K., Evans, J., Gaff, D.F., and Hamill, J.D. (2000). The isolation of genes from the resurrection grass *Sporobolus stapfianus* which are induced during severe drought stress. *Plant, Cell and Environment*, 23: 265–277.

Nei, M. and Kumar, S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Niyogi, K.K. (1999). Photoprotection revisited: genetic and molecular approaches. *Annual Review of Plant Physiology and Plant Molecular Biology*, 50: 333–359.

Oliver, M. J. and Bewley, J. D. (2010). Desiccation-Tolerance of Plant Tissues: A Mechanistic Overview. *Horticultural Reviews*, 171–213.

Oliver, M.J. (2007). Lessons on dehydration tolerance from desiccation tolerant plants. In: Jenks, M., Wood, A., editors. Plant Desiccation Tolerance. Ames, IA: Wiley-Blackwell Publishing. p. 11-50.

- Oliver, M.J., Tuba, Z. and Mishler, B.D. (2000). The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology*, 151: 85–100.
- Oliver, M.J., Velten, J., and Mishler, B.D. (2005). Desiccation tolerance in bryophytes: A reflection of the primitive strategy for plant survival in dehydrating habitats? *Integrative and Comparative Biology*, 45: 788–799.
- Pan Y., Li J., Jiao L., Li C., Zhu D., and Yu J. (2016). A non-specific *Setaria italica* lipid transfer protein gene plays a critical role under abiotic stress. *Frontiers in Plant Science*, 7: 1752.
- Pandey, P., Irulappan, V., Bagavathiannan, M.V., and Senthil-Kumar, M. (2017). Impact of Combined Abiotic and Biotic Stresses on Plant Growth and Avenues for Crop Improvement by Exploiting Physiomorphological Traits. *Frontiers in Plant Science*, 8.
- Panigrahi, K.C., Mishra, P., Kumari, K., and Kumar, A. (2012). Environmental stress influencing plant development and flowering. *Frontiers in Bioscience*, S4(4): 1315–1324.
- Pasquier, C., Promponas, V.J., Palaios, G.A, Hamodrakas, J.S., and Hamodrakas, S.J. (1999). A novel method for predicting transmembrane segments in proteins based on a statistical analysis of the SwissProt database: the PRED-TMR algorithm. *Protein Engineering*, 12(5): 381-5
- Peters, K., Breitsameter, L., and Gerowitt, B. (2014). Impact of climate change on weeds in agriculture: a review. *Agric. Sustain. Dev.* 34, 707–721. doi: 10.1007/s13593-014-0245-2
- Phillips, J.R., Fischer, E., Baron, M., van den Dries, N., Facchinelli, F., Kutzer, M., Rahmazadeh, R., Remus, D., and Bartels, D. (2008). *Lindernia brevidens*: a novel desiccation-tolerant vascular plant, endemic to ancient tropical rainforests. *The Plant Journal*. 54: 938–948
- Pieczynski, M., et al. (2012) Down-regulation of CBP80 gene expression as a strategy to engineer a drought-tolerant potato. *Plant Biotechnology Journal*, 11(4): 459-69.
- Pinheiro, C. and Chaves, M.M. (2011). Photosynthesis and drought: can we make metabolic connections from available data? *Journal of Experimental Botany*, 62 (3): 869-882.
- Porembski, S. (2011) Evolution, diversity, and habitats of poikilohydrous vascular plants. In Plant Desiccation Tolerance. Springer, pp 139–156.

Porembski, S. and Barthlott, W. (2000), (eds) *Inselbergs — Biotic Diversity of Isolated Rock Outcrops in Tropical and Temperate Regions*. Springer-Verlag, Berlin.

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

Potter, E. and K. Kloppstech. (1994). Effects of light stress on the expression of early light-inducible proteins in barley. *European Journal of Biochemistry*, 214: 779–786.

Prilusky, J., Clifford E. Felder, Tzviya Zeev-Ben-Mordehai, Edwin H. Rydberg, Orna Man, Jacques S. Beckmann, Israel Silman, Joel L. Sussman, FoldIndex©: a simple tool to predict whether a given protein sequence is intrinsically unfolded, *Bioinformatics*, Volume 21, Issue 16, Pages 3435–3438, <https://doi.org/10.1093/bioinformatics/bti537>.

Proctor, M.C.F. and Tuba, Z. (2002). Poikilohydry and homoihydry: antithesis or spectrum of possibilities. *New Phytology*, 156, 327–349

Rahmanzadeh, R., et al. (2005). The Linderniaceae and Gratiolaceae Are Further Lineages Distinct from the Scrophulariaceae (Lamiales). *Plant Biology*, 7(1): 67–78.

Raj, Srijana. (2020). Expression of recombinant Early Light-Induced Protein (ELIP) from Craterostigma plantagineum and investigation of putative target molecules (Master thesis). University of Bonn.

Rajkumar, M., Bruno, L.B., and Banu, J.R. (2017). Alleviation of environmental stress in plants: The role of beneficial *Pseudomonas* spp. *Critical Reviews in Environmental Science and Technology*, 47(6): 372–407.

Ramegowda, V., and Senthil-Kumar, M. (2015). The interactive effects of simultaneous biotic and abiotic stresses on plants: mechanistic understanding from drought and pathogen combination. *Journal of Plant Physiology*, 176: 47–54.

Rechsteiner, M. and Rogers, S.M. (1996). PEST sequences and regulation by proteolysis. *Trends in Biochemical Sciences*, 21, 267–271.

Rizhsky, L., Liang, H.J., Shuman, J., Shulaev, V., Davletova, S., and Mittler, R. (2004). When defense pathways collide. The response of *Arabidopsis* to a combination of drought and heat stress. *Plant Physiology*, 134: 1683–1696.

Rizza, A., et al. (2011). Inactivation of the ELIP1 and ELIP2 Genes Affects Arabidopsis Seed Germination. *New Phytologist*, 190(4): 896–905.

Rossini, S., et al. (2006). Suppression of Both ELIP1 and ELIP2 in Arabidopsis Does Not Affect Tolerance to Photoinhibition and Photooxidative Stress. *Plant Physiology*, 141(4): 1264–1273.

Rus Alvarez-Canterbury, A.M., Flores, D.J., Keymanesh, K., To, K., and Brusslan, J.A. (2014). A double SORLIP1 element is required for high light induction of ELIP genes in *Arabidopsis thaliana*. *Plant Molecular Biology*, 84(3): 259–267.

Sah, S.K., Reddy, K.R., and Li, J. (2016). Abscisic Acid and Abiotic Stress Tolerance in Crop Plants. *Frontiers in Plant Science*, 7.

Saijo, Y., and Loo, E. P. I. (2019). Plant immunity in signal integration between biotic and abiotic stress responses. *New Phytologist*, 225(1): 87–104.

Saúl Fraire-Velázquez, Raúl Rodríguez-Guerra and Lenin Sánchez-Calderón (August 29th 2011). Abiotic and Biotic Stress Response Crosstalk in Plants, Abiotic Stress Response in Plants - Physiological, Biochemical and Genetic Perspectives, Arun Shanker and B. Venkateswarlu, IntechOpen, DOI: 10.5772/23217. Available from: <https://www.intechopen.com/books/abiotic-stress-response-in-plants-physiological-biochemical-and-genetic-perspectives/abiotic-and-biotic-stress-response-crosstalk-in-plants>

Scott, M., Madden, T.L. BLAST: at the core of a powerful and diverse set of sequence analysis tools, Nucleic Acids Research, Volume 32, Issue suppl_2, 1 July 2004, Pages W20–W25, <https://doi.org/10.1093/nar/gkh435>

Sievers, F., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7: 539.

Skirycz, A., and Inze, D. (2010). More from less: plant growth under limited water. *Current Opinion in Biotechnology*, 21(2): 197-203.

Smékalová, V., Doskočilová, A., Komis, G., and Samaj, J. (2014). Crosstalk between secondary messengers, hormones and MAPK modules during abiotic stress signalling in plants. *Biotechnology Advances*, 32: 2–11.

Smirnoff, N. (1995) Antioxidant Systems and Plant Response to the Environment. In: Smirnoff, V., Ed., Environment and Plant Metabolism: Flexibility and Acclimation, BIOS Scientific Publishers, Oxford, 217-243.

Soliman, M.F. and Kostandi, S.F. (1998). Effect of saline environment on yield and smut disease severity of different corn genotypes (*Zea mays* L.). *Journal of Phytopathology*, 146: 185–189.

Stephen, F. et al. (2005). Protein database searches using compositionally adjusted substitution matrices. *The FEBS Journal*, 272: 5101-5109.

Stöver, B.C. and Müller, K.F. (2010). TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics*, 11:7.

Suzuki, N., Rivero, R. M., Shulaev, V., Blumwald, E., and Mittler, R. (2014). Abiotic and biotic stress combinations. *New Phytologist*, 203: 32–43.

Takhtajan, A.L. (1997). Diversity and classification of flowering plants. New York: Columbia University Press.

Tester, M. and Bacic, A. (2005). Abiotic stress tolerance in grasses. From model plants to crop plants. *Plant Physiology*, 137(3): 791-793.

The PyMOL Molecular Graphics System, Version 2.3 Schrödinger, LLC.

The UniProt Consortium, UniProt: a worldwide hub of protein knowledge, Nucleic Acids Research, Volume 47, Issue D1, 08 January 2019, Pages D506–D515, <https://doi.org/10.1093/nar/gky1049>.

Tiago, B. and Kühlbrandt. W. (2009). Crystallisation, Structure and Function of Plant Light-Harvesting Complex II.” *Biochimica Et Biophysica Acta – Bioenergetics*, 1787(6): 753–772.

Tubiello, F.N. et al. (2000). U.S. National Assessment Technical Report Effects of Climate Change on U.S. Crop Production Part I: Wheat, Potato, Corn, and Citrus. <http://www.usgcrp.gov/usgcrp/nacc/agriculture/TubielloEtal2000.pdf>.

TzvetkovaChevolleau, T., et al. (2007). The light stress induced protein ELIP2 is a regulator of chlorophyll synthesis in *Arabidopsis thaliana*. *The Plant Journal*., 50: 795– 809.

Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2000). Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins: Structure and Function Genetics*, 41: 415–427.

- VanBuren, R., et al. (2015). Single-molecule sequencing of the desiccation-tolerant grass *Oropetium thomaeum*. *Nature*, 527: 508–511.
- Vanburen, R., et al. (2019). Massive Tandem Proliferation of ELIPs Supports Convergent Evolution of Desiccation Tolerance across Land Plants. *Plant Physiology*, 179(3): 1040–1049.,
- VanBuren, R., Wai, C.M., Ou, S., Pardo, J., Bryant, D., Jiang, N., Mockler, T.C., Edger, P., and Michael, T.P. (2018b). Extreme haplotype variation in the desiccation-tolerant clubmoss *Selaginella lepidophylla*. *Nature Communications*, 9: 13.
- VanBuren, R., Wai, C.M., Pardo, J., Giarola, V., Ambrosini, S., Song, X., and Bartels, D. (2018a). Desiccation tolerance evolved through gene duplication and network rewiring in Lindernia. *The Plant Cell*, 30: 2943–2958.
- Varner J.E. and Lin L.S. (1989). Plant cell wall architecture. *Cell*, 56: 231–239.
- Vurukonda, S.S.K.P., Vardharajula, S., Shrivastava, M., and Skz, A. (2016). Enhancement of drought stress tolerance in crops by plant growth promoting rhizobacteria. *Microbiological Research*, 184: 13–24.
- Wallace, J.S., Acreman, M.C., and Sullivan, C.A. (2003). The sharing of water between society and ecosystems: from conflict to catchment-based co-management. *Philosophical Transactions of The Royal Society B Biological Sciences*, 358: 2011–2026.
- Walters, C., Hill, L.M., and Wheeler, L.J. (2005). Dying while dry: Kinetics and mechanisms of deterioration in desiccated organisms. *Integrative and Comparative Biology*, 45: 751-758.
- Wheeler, T. and Von Braun, J. (2013). Climate change impacts on global food security. *Science*, 341: 508–513.
- Wolfe, G.R., Cunningham, F.X., Durnford, D.G., Green, B.R., and Gantt, E. (1994). Evidence for a common origin of chloroplasts with light-harvesting complexes of different pigmentation. *Nature*, 367: 566–568.
- Xiao, L., et al. (2015). The resurrection genome of Boea hygrometrica: A blueprint for survival of dehydration. *Proceedings of the National Academy of Science*, 112: 5833–5837.
- Xu, Z., et al. (2018). Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of desiccation tolerance. *Molecular Plant*, 11: 983–994.

- Yue, B., Xue, W., Xiong, L., Yu, X., Luo, L., Cui, K., Jin, D., Xing, Y., and Zhang, Q. (2006). Genetic basis of drought resistance at reproductive stage in rice: separation of drought tolerance from drought avoidance. *Genetics*, 172: 1213–1228.
- Zeng, Q. (2002). Two early light-inducible protein (ELIP) cDNAs from the resurrection plant *Tortula ruralis* are differentially expressed in response to desiccation, rehydration, salinity, and high light. *Journal of Experimental Botany*, 53(371), 1197-1205.
- Zhang, X.P. and Glaser, E. (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends in Plant Science*. 7: 14–21.
- Zhu, J.K. (2002). Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology*, 53: 247–273.
- Zhu, J.K. (2016). Abiotic Stress Signaling and Responses in Plants. *Cell*, 167(2): 313–324.

8. Acknowledgments

I am immensely grateful to my supervisor Prof. Dr. Dorothea Bartels for being incredibly patient, encouraging, and motivating throughout the completion of this project. Your assistance, guidance, and insightful comments have improved my research skills.

I would like to acknowledge with gratitude, the active support and motivation of Dr. Valentino Giarola, whom without his help in the beginning, I wouldn't have been able to start working on this project.

It behooves me to include a special note of thanks to Srijana, Aishwarya, and Patricija for their inspiring motivation and unceasing support throughout the past year.

Finally, I would like to extend my appreciation and gratitude to all current and former members of Prof. Dr. Bartels group whose invaluable assistance was necessary for the fruitful and timely completion of this project.

9. Supplementary material

Supplementary table 1. Properties of some members of the CAB/ELIP/HLIP superfamily. Proteins were classified based on their respective phylum (E: eudicots; M: monocots; C: chlorophytes; S: spermatophytes; B: bryophytes).

Phylum	Protein	Length	Predicted molecular mass (kDa)	Predicted pI	Predicted charge at pH 7.00	Foldability
E	Lbr ELIP 025540	272	29.279	9.16	5.1	PU
	Lbr ELIP 025541	211	22.740	9.41	5.1	PU
	Lbr ELIP 025542	272	28.884	9.39	9.8	PU
	Lbr ELIP 025543	274	28.837	8.01	2.5	PU
	Lbr ELIP 025544	274	28.892	8.24	3.5	PU
	Lbr ELIP 025545	274	28.892	8.24	3.5	PU
	Lbr ELIP 025546	274	28.892	8.24	3.5	PU
	Lbr ELIP 025547	274	28.892	8.24	3.5	PU
	Lbr ELIP 025548	274	28.892	8.24	3.5	PU
	Lbr ELIP 025549	145	15.603	6.06	-0.8	F
	Lbr ELIP 025584	211	22.671	8.97	3.1	PU
	Lbr ELIP 025585	175	18.790	9.51	6.1	PU
	Lbr ELIP 025587	224	24.151	7.94	1.1	PU
	Lbr ELIP 025589	113	12.190	4.41	-5.1	F
	Lbr ELIP 025591	211	22.671	8.97	3.1	PU
	Lbr ELIP 025593	186	19.999	8.62	2.1	PU
	Lbr ELIP 017192	184	19.442	9.41	4.9	PU
	Lbr ELIP 016696	198	21.283	8.45	2.1	PU
	Lbr ELIP 016551	193	20.950	8.85	2.9	PU
	Lbr ELIP 025581	224	24.151	7.94	1.1	PU
	Lbr ELIP 025579	224	24.151	7.94	1.1	PU
	Lbr ELIP 025583	198	21.365	8.88	3.8	PU

	<i>Average of LbrELIPs</i>	222	23.709	8.32	2.97	X
	Lsu ELIP 021038	134	14.595	9.10	2.2	F
	Lsu ELIP 008078	193	20.356	9.20	3.9	PU
	Lsu ELIP 024731	189	20.128	7.94	0.9	PU
	Lsu ELIP 030118	198	21.365	8.88	3.8	PU
	<i>Average of LsuELIPs</i>	178	19.111	8.78	2.7	X
	<i>C. plantagineum</i> ELIP	199	21.991	8.45	2.1	PU
	<i>Ci. sinensis</i> ELIP	208	22.392	9.32	5.1	PU
	<i>Ca. sinensis</i> ELIP2	175	18.791	9.41	4.9	F
	<i>C. annuum</i> ELIP	160	17.010	8.88	2.1	F
	<i>A. thaliana</i> ELIP1	195	20.324	9.66	5.9	F
	<i>A. thaliana</i> ELIP2	193	20.344	9.65	4.9	F
	<i>S. oleracea</i> CAB	261	27.839	6.58	-0.6	F
	<i>G. max</i> CAB	245	26.208	5.24	-5.4	F
	<i>G. max</i> CAB2	256	27.122	6.20	-2.2	F
	<i>G. max</i> CAB3	263	27.861	5.64	-3.4	F
M	<i>C. sativus</i> ELIPa	185	19.448	9.30	4.1	F
	<i>C. sativus</i> ELIPb	184	19.374	8.92	3.1	F
	<i>C. sativus</i> ELIPc	179	18.931	9.30	4.4	F
	<i>C. sativus</i> ELIPd	172	17.973	8.97	2.9	F
	<i>Z. mays</i> ELIP	182	18.527	11.20	7.9	F
	<i>T. aestivum</i> wcr12	174	17.699	10.23	3.9	F
	<i>H. vulgare</i> ELIP9	172	17.553	9.67	1.9	F
	<i>H. vulgare</i> ELIP6	167	17.132	10.19	2.9	F
	<i>H. vulgare</i> ELIP5	231	24.053	9.62	4.1	F
	<i>O. sativa</i> (J) PSBS1	268	27.903	7.00	0.0	F
	<i>O. sativa</i> (J) PSBS2	254	26.822	6.25	-0.8	F
	<i>O. sativa</i> (In) PSBS1	268	27.903	7.00	0.0	F
	<i>O. sativa</i> (In) PSBS2	254	26.892	7.28	0.2	F
C	<i>C. reinhardtii</i> PSBS1	245	26.133	7.00	0.0	F

	<i>D. salina</i> CBR	172	17.886	8.97	2.4	F
	<i>M. commode</i> CBR	208	21.392	5.27	-2.1	F
	<i>M. pusilla</i>	174	17.742	8.50	2.1	F
	CBR/ELIP1					
S	<i>P. densata</i> ELIP	120	12.616	5.05	-2.1	F
	<i>P. tabuliformis</i> ELIP	120	12.615	6.62	-0.1	F
	<i>P. yunnanensis</i> ELIP	120	12.615	6.62	-0.1	F
B	<i>P. patens</i> ELIP12	225	24.723	8.64	2.2	PU
	<i>P. patens</i> ELIP13	226	24.767	9.13	4.1	PU

Supplementary table 2. Percentage identity of surveyed proteins to known ELIPs, PSBS, CABs, and CBRs.

	<i>A. thaliana</i> ELIP2	<i>C. plantagineum</i> ELIP	<i>O. sativa</i> (J) PSBS1	<i>S. oleracea</i> CAB	<i>D. salina</i> CBR
Lbr ELIP 025593	48.77	35.00	33.33	30.00	37.98
Lbr ELIP 017192	76.00	46.52	41.94	41.18	46.90
Lbr ELIP 016696	45.14	62.07	38.60	32.26	38.46
Lbr ELIP 016551	56.49	68.39	36.36	29.03	39.24
Lsu ELIP 021038	60.17	67.92	33.33	35.29	49.12
Lsu ELIP 008078	47.83	41.18	41.94	41.18	47.83
Lsu ELIP 024731	39.19	71.00	46.67	32.14	39.19
Lsu ELIP 030118	58.14	73.89	43.33	35.29	40.44