

**Deciphering the Role(s) of CBS Domain Containing Proteins of
Unknown Function in Modulating Stress Tolerance in Plants**

Thesis submitted

to

Jawaharlal Nehru University
New Delhi, India

For the award of the degree of

Doctor of Philosophy

by

Ashish Subba



Plant Stress Biology Group
International Centre for Genetic Engineering and Biotechnology
New Delhi, India
2021



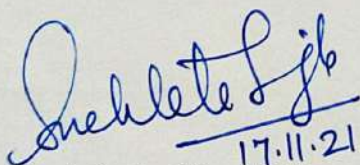
INTERNATIONAL CENTRE FOR GENETIC ENGINEERING AND BIOTECHNOLOGY

ICGEB Campus, Aruna Asaf Ali Marg
New Delhi - 110 067, India
<http://www.icgeb.org>

Tel. : 91-11-26741358/61
91-11-26742357/60
91-11-26741007
Fax : 91-11-26742316
UIN No.: 0717UNO00161UNZ
E-mail : icgeb@icgeb.res.in

Certificate

This is to certify that the research work embodied in this thesis titled **“Deciphering the role(s) of CBS domain containing proteins of unknown function in modulating stress tolerance in plants”** has been carried out by Ashish Subba under my supervision at the Plant Stress Biology Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, India, for the award of the degree of Doctor of Philosophy. The presented work is original and has not been submitted in any part or full, to this or any other University, for the award of any other degree or diploma.



17.11.21

Dr. Sneh Lata Singla-Pareek

Supervisor
Plant Stress Biology Group
ICGEB
New Delhi-110067
India



Dr. Dinakar M. Salunke

Director
ICGEB
New Delhi-110067
India

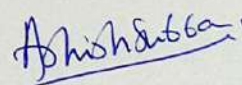


Declaration

I hereby declare that the research work embodied in this thesis titled **“Deciphering the role(s) of CBS domain containing proteins of unknown function in modulating stress tolerance in plants”** has been carried out by me under the supervision of Dr. Sneh Lata Singla-Pareek at the Plant Stress Biology Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, India.

Date: 17/11/2021

Place: ICGEB, New Delhi



Ashish Subba

Pre-doctoral fellow
Plant Stress Biology Group
ICGEB
New Delhi- 110067
India

Abstract

Cystathionine β -Synthase (CBS) domain, named for its presence in human Cystathionine β -Synthase enzyme, exists in all three domains of life in highly diverse protein families. This domain is known to function as a regulatory unit, mainly through the binding of adenosine-derived nucleotides. Mutations in this domain or broadly the loss of function of the CBS domain containing proteins (CDCPs) have been identified to be associated with severe disorders in different organisms, which signifies its physiological importance. However, on plant CDCPs, very little knowledge from only a few studies is available. The present study is, therefore, implemented to understand the preliminary functions of two important CDCP members from rice: (a) *OsCBSX3* (contains only a pair of CBS domains) and (b) *OsCBSCLC6* {a chloride channel (CLC) with transmembrane domain and a pair of cytosolic CBS domains}. We found *OsCBSX3* as a shoot expressive and diurnally regulated gene, that appeared to negatively regulate the expression of *OsCYP450*, a gene that overlaps with *OsCBSX3* on its opposite strand. Its transcript level increased in some stresses, mainly desiccation. *OsCBSX3* was observed to form a homo-oligomer, probably a homodimer, and showed chloroplast localization. Accordingly, it interacted specifically with the putative chloroplastic Thioredoxin y protein (*OsTrx y*). Importantly, its N-terminal residues preceding the first CBS domain were observed to be essential for both homo-oligomerization and interaction with *OsTrx y*. The transgenic plants (both *Arabidopsis* and rice) overexpressing *OsCBSX3* exhibited a normal growth phenotype and performed better under salinity and drought stresses. In the case of *OsCBSCLC6*, the expression was observed to be root-specific, while highly induced in the shoots under desiccation stress, and the protein was identified to localize in the tonoplast. The sequence-based analysis indicated *OsCBSCLC6* to possibly function as Cl^- channel and its cytosolic CBS domains to have a potential binding site for one adenosine nucleotide. The transgenic rice plants overexpressing *OSCBSCLC6* exhibited a reduction in plant height under control growth conditions; however, under salinity and drought stresses, the transgenic plants showed healthier physiological performances and higher yield than the WT. These findings present the fundamental information and shed light on the functional aspects of these two CDCPs, which would be highly contributing to further work on these proteins as well as in the characterization of other plant CDCPs.

Acknowledgment

And finally, writing this page after years of learning, training, experimenting, and researching on my valuable journey of Ph.D., feels very pleasant. And remembering now all the helping and supporting hands behind this successful pursuit, feels so much grateful and blessed. Through this column, although so small, I wholeheartedly extend my sincere gratitude to everyone and everything for their abundant assistance and constant encouragement, which made this part of my time very blissful and beautiful.

ICGEB, I adore the most, I am fortunate to get admission here, and I thank this institute for providing the best environment and facilities for learning, researching, and well-being. I acknowledge the Department of Biotechnology, Govt. of India, for fulfilling my eligibility to apply for ICGEB, and I am grateful for financially aiding me with the fellowship for five years. In ICGEB, I am very fortunate to work under the supervision of Dr. Sneh Lata Singla-Pareek, my Ph.D. supervisor. Ma'am, I am thankful to you for trusting in me, accepting me, nurturing me, motivating me, and endowing me with all your best efforts and blessings to groom me to the best of me.

I extend my sincere gratitude to the ICGEB Director, Dr. Dinakar M. Salunke for delivering all the necessities and the facilities required for a productive working atmosphere. I also offer my deep respect and gratitude to Prof. Sudhir K. Sopory, former Director, ICGEB, for his presence has been the essence of positive energy and inspiration to me.

I am grateful to have Dr. M. K. Reddy and Dr. Suresh Nair as my Doctoral Committee members, for they bestowed on me their valuable guidance, support, and lab facilities. I also acknowledge Dr. Sadhu Leelavathi, Dr. Neeti Sanan-Mishra, and Dr. Tanushri Kaul, for I acquired every needed support from their labs.

I also want to extend my profound gratitude to Prof. Ashwani Pareek from the School of Life Sciences, JNU, for extending his experienced knowledge on my research work, for his encouraging remarks, and for offering me to utilize his entire lab facilities.

For the course work I attended at ICGEB, I thank all the Scientists for sharing their knowledge, especially Dr. Dinkar Sahal and Dr. Sunil Mukherjee, both of whom I highly admire. And for different facilities in ICGEB that I was fortunate to avail, I extend my profound gratitude to Dr. Purnima for scanning numerous slides of mine under confocal microscopy, Sir Bhupinder for transforming my countless samples with Particle bombardment, Dr. Ranjan

Nanda, Sandeep, and Faruquee, for training me and allowing me to run my samples independently in GC-MS and ICP-MS, Dr. Inderjeet for LC-MS/MS, Garima for GPC, and Sir Ashok, Sir Rakesh, Patra Ji, Munna Ji from Animal House Facility, for assisting me in antibody production.

For my plants so precious and dear to me, I thank Rajeev Ji, Ramesh Ji, Harivansh Ji, and all other Green House and Pest Control staff, for delivering the best care to plants and for maintenance of the Green House functions. I also like to honour every single ICGEB personals from Administration, Procurement, IT service, Accounts, and Maintenance, especially Sir Mayank, Sir Sachin, Ma'am Pratibha, Ma'am Nidhi, Sir SP Rawat, Sir Sumit, Sir Omkar, Sir Patwal, and Rashmi, for assisting me in my academic activities in various ways. Also, I am grateful to all the ICGEB Guest House staff for making my stay healthy and to all the Security staff for making my stay safe.

Coming to my lab where I spent most of the time, its happy and healthy ambience, filled with talented, joyful, and cheering people, has been a vital source of energy and refreshment. I deeply thank my loving lab and all my lab members, Dr. Brijesh Gupta, Dr. Rohit, Dr. Charanpreet, Dr. Kamlesh, Dr. Sampurna, Dr. Manjari, Dr. Amit, Dr. Ramsong, Dr. Khalid, Brijesh Kumar, Surabhi, Ray S Rathore, Yashika, Yajnaseni, Himani, Bidisha, Nancy, Pooja, Ramesh Ji, Baldev Ji, Shafik, and Ranjeet, for sharing with me all the laughter, emotions, knowledge, cooperation, and everything that exists in a happy family. And sharing the same floor and bonded with the friendship, I also thank all the members from Plant Biology Group, including Dr. Mohan, Dr. Vara, Dr. Dhiren, Dr. Donald, Dr. Aakriti (my dear sister), Dr. Babu, Dr. Pankaj, Dr. Jitender, Dr. Isha, Ravi, Himangini, Aafreen, Sangeeta, Sahil, Sambhu, Meeraj, Himani, Rashi, Monica, Brother Henry, Shruti, Mamta, Soniya, Khalid, Aayushi, Jyotsna, Arul, Sudhir, Anita, Ambreen, and Yunieth for good their cordial and helping nature.

And now remembering of my friends Pradeep, Bhabesh, Kashif, Sir Rokeb, Manish Bhai, Jaspreet Bhabhi, Veena, Ekta, Niharika, Garima, Sneha, and Sandeep, I feel overjoyed, for our every meet and every gathering was so much of fun and memorable. I cheer for our friendship, and I thank you all for your best support, affection, and those happy days.

In the course of time, I am also fortunate to have met my beloved friend, Pragati, whom I owe a lot for her unconditional help, advice, and care. I am immensely thankful to you for

your strong support, surprises, and most importantly for your concern to complete this thesis as well.

I express my gratitude to my seniors, Sir Dileep and Sir Sachin for their affection, constant support, and invaluable suggestions. And I also thank the Living Science Group (Dr. Sarita and the team) for every show I attended was so motivating to me.

I am also very lucky to have met my wonderful friend, Wungrampha, from JNU, who exceedingly helped to learn to operate a Handy PEA machine that I thoroughly used in my experiments. I also want to acknowledge the contribution of the 'Arabidopsis Biological Research Centre' (ABRC) from where I received *Arabidopsis* seeds and 'Sci-hub' from where I obtained many research papers.

I thank my old school friends: Adang, Bilin, Dawa, Donald, Dorjee, Ekum, Gaurav, Henuka, Karma P, Karma T, Kenzong, Pempa, Phurba, Prakash, Tadi, Tashi, and Visal, and brothers: Bhaichung and Aakash, for their moral support and encouragement. I express my gratitude to my beloved teachers from my old schools and colleges, for their teaching got me to this place. I also thank all my relatives, uncles and aunts, brothers and sisters, for relentlessly cheering me.

Successful completion of this thesis work would mean a lot more to my family members. I deeply thank my family members, mom, dad, brother, and sister for their eternal support, motivation, patience, blessings, and prayers.

I also express my gratitude to my late Grandparents for their love and blessings that remain in me.

And lastly, I extend my humble gratitude to Almighty God for bringing me this far, for blessing me with abundant goodness, and with your blessings, I shall keep on moving ahead again...

Abbreviations and Symbols

| | |
|----------|---|
| % | Percentage |
| °C | Degree Celsius |
| 3-AT | 3-Amino-1,2,4-triazole |
| ADP | Adenosine diphosphate |
| AMP | Adenosine monophosphate |
| ATP | Adenosine triphosphate |
| BCIP-NBT | 5-bromo-4-chloro-3-indolyl phosphate-nitro blue tetrazolium |
| bp | Base pair |
| BSA | Bovine serum albumin |
| CBS | Cystathionine β -Synthase |
| CDCPs | Cystathionine β -Synthase domain containing proteins |
| cDNA | Complementary deoxyribonucleotide |
| CDS | Coding sequence |
| Ci | Curie |
| cm | Centimeter |
| CT | Cycle threshold |
| CTAB | Cetyltrimethylammonium bromide |
| DEPC | Diethylpyrocarbonate |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| dNTP | Deoxy ribonucleotide triphosphate |
| DO | Drop-out |
| DTT | Dithiothreitol |
| EtBr | Ethidium bromide |
| g | Gram |
| GFP | Green fluorescent protein |
| hr | Hour(s) |
| IPTG | Isopropyl β -D-1-thiogalactopyranoside |
| Kb | Kilo base pair |
| kDa | Kilo Dalton |
| kPa | Kilopascal |
| LB | Luria-Bertani |

| | |
|----------|--|
| M | Molar |
| m | Meter |
| mg | Milligram |
| min | Minute(s) |
| ml | Milliliter |
| mM | Millimolar |
| mm | Millimeter |
| mRNA | Messenger RNA |
| MS | Murashige and Skoog |
| ng | Nanogram |
| O.D. | Optical density |
| PCR | Polymerase chain reaction |
| qRT-PCR | Quantitative Real time polymerase chain reaction |
| RNA | Ribonucleic acid |
| RNase A | Ribonuclease A |
| rpm | Revolutions per minute |
| RT | Reverse transcriptase |
| SDS-PAGE | Sodium dodecyl sulphate–polyacrylamide gel electrophoresis |
| Sec | Second(s) |
| Taq | Thermus aquaticus |
| U | Enzyme unit |
| UDG | Uracil-DNA glycosylase |
| V | Volts |
| WT | Wild type |
| x g | times gravity (relative centrifugal force) |
| μCi | Microcurie |
| μg | Microgram |
| μl | Microliter |
| μm | Micrometer |
| μM | Micromolar |

Table of Contents

| | |
|--|--------------|
| Chapter 1: Introduction | 1-4 |
| Chapter 2: Review of Literature | 5-42 |
| 2.1 CBS domain- discovery and basic structural aspects | |
| 2.2 Functional aspects of CBS domain in CDCPs | |
| 2.2.1 Proteins containing only a single pair of CBS domains | |
| 2.2.2 ‘ γ -subunit’ of AMPK / SNFK / SnRK | |
| 2.2.3 CBS domains in Cystathionine β Synthase enzyme | |
| 2.2.4 CBS domains in Inosine Monophosphate Dehydrogenase (IMPDH) | |
| 2.2.5 CBS domains in CNNM, CorC and MgtE: Proteins involved in Mg^{2+} transport | |
| 2.2.6 CBS domains in Inorganic Pyrophosphatase (PPase) | |
| 2.2.7 CBS domains in other CDCP members | |
| 2.3 Chloride channels (CLCs) | |
| 2.3.1 Structural features and conserved residues governing the functions of CLCs | |
| 2.3.2 Functional features of mammalian CLCs | |
| 2.3.3 CLCs from plants | |
| 2.3.3.1 <i>Plant CLCs involved in nitrate transport and storage</i> | |
| 2.3.3.2 <i>Plant CLCs with canonical chloride transport function</i> | |
| 2.3.3.3 <i>Phylogenetically distinct members of plant CLCs</i> | |
| 2.3.3.4 <i>Plant CLCs in pathogen interaction</i> | |
| Chapter 3: Materials and Methods | 43-80 |
| 3.1 Gene expression analysis | |
| 3.1.1 Plant growth and treatments for expression analysis | |
| 3.1.2 RNA isolation, UV/Vis spectrophotometry, and denaturing gel electrophoresis | |
| 3.1.3 First-strand cDNA synthesis | |
| 3.1.4 Gene expression analysis by qRT-PCR | |
| 3.1.5 Gene expression profiling using publicly available expression datasets | |
| 3.2 Agarose gel preparation and gel electrophoresis | |
| 3.3 Glycerol stock preparation | |

3.4 Gene amplification, cloning, and transformation in *E. coli*

- 3.4.1. PCR amplification of the gene/insert
- 3.4.2. Extraction and purification of DNA fragments from agarose gel
- 3.4.3. Restriction digestion of insert and vector
- 3.4.4. Purification of DNA inserts
- 3.4.5. Alkaline phosphatase treatment
- 3.4.6. Ligation of DNA insert and the vector
- 3.4.7. *E. coli* competent cells preparation and transformation
 - 3.4.7.1 *Preparation of E. coli competent cells*
 - 3.4.7.2 *Transformation of E. coli competent cells*
- 3.4.8. Colony PCR of the *E. coli* clones
- 3.4.9. Plasmid isolation from *E. coli*
- 3.4.10. *Agrobacterium* competent cells preparation and transformation
 - 3.4.10.1 *Preparation of Agrobacterium competent cells*
 - 3.4.10.2 *Transformation of Agrobacterium competent cells*

3.5 In planta protein subcellular localization analysis

- 3.5.1 Protein subcellular localization in Tobacco epidermal cells
- 3.5.2 Protein subcellular localization in rice protoplast
 - 3.5.2.1 *Isolation of protoplast from rice*
 - 3.5.2.2 *Transfection of rice protoplast*

3.6 Protein expression in *E. coli*

- 3.6.1. Protein SDS-PAGE
- 3.6.2 CBB staining of SDS-PAGE gel
- 3.6.3 Western blot analysis
- 3.6.4 Bradford assay
- 3.6.5 Cloning and transformation of gene constructs for protein expression in *E. coli*
- 3.6.6 Optimization of protein expression in *E. coli*
- 3.6.7 Purification of 6x-His tagged OsCBX3(-sp) from soluble fraction by Ni²⁺-NTA affinity chromatography
- 3.6.8 Solubilization of 6x-His tagged proteins from insoluble fraction followed by Ni²⁺-NTA affinity chromatography purification
- 3.6.9 Gel permeation chromatography (GPC)

3.7 ATP binding analysis

3.8 Yeast Two-Hybrid (Y2H) assay

- 3.8.1. Protein homo-oligomerization analysis using Y2H
- 3.8.2. Screening of putative interacting partners from rice cDNA library by Y2H
- 3.8.3. Interaction analysis of OsCBSX3 with full-length of putative interacting partners by Y2H
- 3.8.4. Transformation of *Saccharomyces cerevisiae*
- 3.8.5. Isolation of plasmid from yeast cells

3.9 Interaction analysis of OsCBSX3 and OsTrx y by Bimolecular Fluorescence Complementation (BiFC) in planta in onion epidermal cells

- 3.9.1. Gene and vector constructs used in BiFC assay
- 3.9.2. Onion epidermal strips preparation
- 3.9.3. Microcarrier-DNA preparation, and bombardment on onion epidermal cells

3.10 Generation of transgenic rice overexpressing *OsCBSX3* and *OsCBSCLC6* and their subsequent screening

- 3.10.1 Gene and vector constructs used for transformation of rice
- 3.10.2 Rice calli production, its transformation and regeneration into plants
- 3.10.3 PCR screening of putative transgenic plants
- 3.10.4 Analysis of transgene integration by Southern blot
 - 3.10.4.1 *Genomic DNA (gDNA) isolation*
 - 3.10.4.2 *Restriction digestion and gel electrophoresis of gDNA*
 - 3.10.4.3 *Preparation of gel for transfer*
 - 3.10.4.4 *Preparation of membrane and blotting papers for transfer*
 - 3.10.4.5 *Transfer of DNA from gel to membrane*
 - 3.10.4.6 *PCR amplification of Digoxigenin (DIG) labelled probe*
 - 3.10.4.7 *Pre-hybridization, hybridization and immunodetection*
- 3.10.5 Screening of T₁ and T₂ transgenic lines
- 3.10.6 Expression analysis of transgene by qRT-PCR

3.11 Plant growth analysis under control and stress conditions

- 3.11.1 Plant growth and stress treatments
- 3.11.2 Plant physiological and yield parameter analyses

3.12 Generation and growth analysis of transgenic *Arabidopsis* overexpressing *OsCBSX3*

3.13 Primer designing

- 3.14 Heat map preparation**
- 3.15 Multiple alignment and phylogenetic tree**
- 3.16 Statistical analysis**

Chapter 4: Results

81-135

4.1 Functional characterization of OsCBSX3

- 4.1.1 Expression analysis of *OsCBSX3*
- 4.1.2 Subcellular localization of *OsCBSX3*
- 4.1.3 Biochemical characterization of *OsCBSX3* and identification of its interacting proteins
 - 4.1.3.1 Expression and purification of OsCBSX3 in E. coli*
 - 4.1.3.2 Oligomeric status of OsCBSX3*
 - 4.1.3.3 Sequence analysis for ligand binding residues and ATP binding assay for OsCBSX3*
 - 4.1.3.4 Identification of OsCBSX3 interacting partners by yeast two-hybrid assay*
 - 4.1.3.5 Interaction analysis OsCBSX3 and OsTrx y by BiFC assay*
- 4.1.4 Generation and growth assessment of transgenic rice plants overexpressing *OsCBSX3*
- 4.1.4 Growth assessment of transgenic *Arabidopsis* overexpressing *OsCBSX3*

4.2 Functional characterization of OsCBSCLC6

- 4.2.1 Gene expression analysis of *OsCBSCLC6*
- 4.2.2 Subcellular localization of *OsCBSCLC6*
- 4.2.3 CLCs functional motif sequence analysis in *OsCBSCLC6*
- 4.2.4 Expression of *OsCBSCLC6* in *E. coli*
- 4.2.5 Homo-oligomeric status of *OsCBSCLC6*
- 4.2.6 Generation and growth assessment of transgenic rice plants overexpressing *OsCBSCLC6*

Chapter 5: Discussion

136-142

Chapter 6: Summary and Conclusions

143-145

Chapter 7: Bibliography

146-165

Chapter 8: Appendices

166-178

Chapter 9: Publications

179

1. Introduction

Time and evolution have endowed us with a rich diversity in plant genetic resources that every plant genotype within a species possesses some superior traits over the others. The benefit of such diversity had been exploited since the adoption of crop cultivation during human civilization and is continuing even now in the science of crop improvements. Among such resources in the gene pool of rice, one of the most important cereal crops, is “Pokkali” which has originated and is being cultivated in the saline coastal tracts of Kerala, South India. This is a highly salt-tolerant landrace of rice, a trait that has made it extremely useful in research and breeding programs (Walia et al, 2005; Mishra et al, 2020). Seeking to gain insight into the plant responses and tolerance mechanisms to salinity stress, our lab has carried out the transcriptome profiling of Pokkali against the contrasting salt-sensitive IR64 variety of rice, employing cDNA subtractive hybridization and northern hybridization approaches (Kumari et al, 2009). The study identified novel genes being differentially regulated in these two genotypes, among which included the genes encoding Cystathionine β -Synthase (CBS) domain containing proteins (CDCPs). Members of this gene family showed early responsiveness to salinity in Pokkali in contrast to late induction in IR64 upon salinity stress treatment at the seedling stage (Kumari et al, 2009). Expression analysis of genes encoding CDCPs from rice and *Arabidopsis* using existing Microarray and Massively Parallel Signature Sequencing (MPSS) databases, further exhibited genes encoding some members of CDCPs to be highly responsive to plant stress (Kushwaha et al, 2009). These findings stimulated our lab to pursue deeper into the functions of CDCPs that have remained almost unexplored, particularly in the plant kingdom.

The CBS domain containing proteins (CDCPs) comprise functionally diverse protein members that share conserved Cystathionine β - Synthase (CBS) domain. This protein domain exists in all kingdoms of life. It spans a length of approximately 60 amino acid residues and was first observed by Bateman (1997) while characterizing the genome sequences of *Methanocaldococcus jannaschii*, an archaeobacterium. It is named as CBS domain because Bateman (1997) also observed this domain to occur in human Cystathionine β - Synthase, an enzyme involved in cysteine biosynthesis. Generally, CBS domains are present in two or four tandem copies, mostly co-existing with other non-homologous domain(s) in the polypeptide (Bateman, 1997; Baykov et al, 2011). Each pair of CBS domains associates tightly, forming a structure called a ‘Bateman domain’ or ‘Bateman module’. Two Bateman modules further interact to form a structure consisting of four CBS domains known as a ‘CBS module’, which in the case of CDCPs with a single pair of CBS domains, results in the formation of a

homodimer (Baykov et al, 2011; Ereño-Orbea et al, 2013a). CBS domains possess an affinity for various adenosine-derived molecules, and the ligand-induced conformational changes in CBS domains are known to regulate functions of the associated domain(s) in the same or another interacting polypeptide (Baykov et al, 2011; Ereño-Orbea et al, 2013a). Mutations in the CBS domains of enzymes and membrane transport proteins in humans have been found to cause several hereditary disorders (Ignoul and Eggermont, 2005; Giménez-Mascarell et al, 2019). Plants also exhibit various physiological defects associated with the loss of function of CDCPs (Yoo et al, 2011; Shin et al, 2020; Zafar et al, 2020; Subba et al, 2021; Zafar et al, 2021). These signify an indispensable role of the CBS domain (or broadly CDCPs) in various physiological processes associated with growth and development in all living organisms.

In plants, the first genome-wide analysis of CDCPs was carried out from our lab which identified 34 CDCPs (encoded by 33 genes) and 59 CDCPs (encoded by 37 genes) in *Arabidopsis* and rice, respectively (Kushwaha et al, 2009). Like in other systems, CDCPs from these two plants possess either (a) a single pair of CBS domains or (b) two pairs of CBS domains. Further, these two groups of CDCPs either retain or lack additional non-homologous domain(s). The other non-homologous domains identified in plant CDCPs with single CBS domain pair are transmembrane Chloride Channel (CLC), Cyclin M transmembrane N-terminal domain / Domain of Unknown Function 21 (CNNM/DUF21), transporter associated CorC_HlyC, Inosine Monophosphate Dehydrogenase (IMPDH), Sugar Isomerase (SIS) and Pentatricopeptide Repeats (PPR). In the case of plant CDCPs with two CBS domain pairs, Phox and Bem1p (PB1) is the only additional domain identified to be present (Kushwaha et al, 2009).

Our lab then initiated the preliminary study on OsCBSX4, a protein containing only a single pair of CBS domains (we have referred to such plant CDCPs containing only a single pair of CBS domains as CBSXs). OsCBSX4 is one of the CDCPs, the transcript of which showed early responsiveness to salinity stress in Pokkali (Kumari et al, 2009). The overexpression of *OsCBSX4* in transgenic tobacco conferred enhanced tolerance towards salinity, oxidative, and heavy metal stresses (Singh et al, 2012). In *Arabidopsis*, *AtCBSX1* has been observed to be induced by various stresses (Yoo et al, 2011). The overexpression of *AtCBSX1* imparted better growth to the plants; however, these plants exhibited severe sterility associated with anther indehiscence, caused by reduced lignin deposition in the secondary wall of anther endothecium. This lignin deposition in the anther wall was suggested to be maintained by *AtCBSX1* in association with Thioredoxin (Trx) proteins in the chloroplast in a redox-dependent manner (Yoo et al, 2011). The same group also reported another CBSX homolog

from *Arabidopsis*, AtCBSX3, to regulate the redox homeostasis in mitochondria by interacting with mitochondria specific Thioredoxin (Trx-o2) (Shin et al, 2020). Realizing the involvement of these CDCP members in plant development as well as stress tolerance, we opted to characterize another distinct single CBS domain pair containing protein from rice, “OsCBSX3”, as one of the CDCPs in the present study. OsCBSX3 is phylogenetically distant from the other CBSX homologs including OsCBSX4, AtCBSX1 and AtCBSX3. Notably, its gene (*OsCBSX3*; LOC_Os02g57280) completely overlaps in an antiparallel arrangement with *OsCYTOCHROME P450* (*OsCYP450*; LOC_Os02g57290) in the genome, which implies possible interactions among these two at expression and functional level.

Another important family members of the CDCP superfamily are Chloride Channel proteins (CLCs). They possess a transmembrane CLC domain at the N-terminal region, followed by the cytosolic C-terminal region containing a single pair of CBS domains. However, CLCs from some prokaryotes lack CBS domains. Interestingly, not to be deceived by its nomenclature, the CLC family includes both channel and exchanger proteins. Besides, diverged from their canonical chloride (Cl^-) transport function, some CLC members from plants and bacteria execute the transport of nitrate (NO_3^-) and fluoride (F^-), respectively (Subba et al, 2021). The structural and functional features of the CLCs have been obtained mostly from mammalian CLCs, where the loss of function of each CLC member results in severe pathological disorders (Jentsch, 2015). In plants, CLCs have been identified to function mainly in nitrogen assimilation, stomatal regulation, photosynthetic efficiency, and stress tolerance (Subba et al, 2021). Despite such importance of CLCs, their studies are still scarce in plants and mostly limited to *Arabidopsis*. In this scenario, the present study also focused to understand the basic functions of “OsCBSCLC6”, a member of the CLC family protein from rice. OsCBSCLC6 shares maximum sequence identity with AtCLC-g from *Arabidopsis*, which has been suggested to function in Cl^- sequestration and its tolerance (Nguyen et al, 2016).

Keeping all these recent developments in mind, the present work targets the studies on the two above-mentioned CDCPs from rice, namely, OsCBSX3 (having only a single pair of CBS domains) and OsCBSCLC6 (having transmembrane CLC domain followed by a single pair of CBS domains). The studies were drafted with the following objectives:

1. Expression analysis of *OsCBSX3* and *OsCBSCLC6*, and subcellular localization of their respective proteins.

2. Biochemical characterization of OsCBSX3 and OsCBSCLC6, and identification of their interacting partners.
3. Functional validation of *OsCBSX3* and *OsCBSCLC6* in transgenic rice plants under stress conditions.

This study has gained important preliminary knowledge on both OsCBSX3 and OsCBSCLC6. It has opened a new direction and thoughts on the possible functions of these two proteins. The experiments and outcomes of the study would be useful in advancing further towards understanding the physiological functions of OsCBSX3 and OsCBSCLC6 in rice.

2. Review of Literature

The landmark reached by humankind in the field of science and development has been remarkable through all this time. This comprises immense efforts and the achievements gained in agricultural sciences that have managed to balance the equilibrium between plant harvest and the rising human population. Importantly, the introduction of high-yielding varieties, agrochemicals, farm machinery, and irrigation practices, during the so-called “Green Revolution”, reformed the concept of agriculture (Davies, 2003). The later breakthroughs in molecular biology and biotechnology have further transformed the conventional method of plant breeding to advanced marker-assisted breeding and genetic engineering towards crop improvement. At present, the genome of hundreds of plant species has been sequenced and several genes associated with agriculturally important traits have been characterized, facilitating better understanding and precise engineering of traits and/or pathways in plants.

Despite this significant progress, agricultural science foresees the threat imposed by a global human population which is expected to rise by 25%, reaching 10 billion in the next 30 years (Hickey et al, 2019). Besides, the indiscriminate exploitation of the environment by several human activities is leading to climate change and degradation of soil, which is further challenging the domain of agriculture (Pareek et al, 2020). Hence, sooner the world may need “Green Revolution-II” to introduce new crop varieties with higher yield potential and enhanced tolerance to different stresses, but most importantly, not compromising the sustainability. Participating to resolve this issue among many research labs around the globe, our lab is working on understanding the plant physiology and tolerance mechanisms towards abiotic stresses, particularly in rice (*Oryza sativa*).

The plant responds to abiotic stress, which is caused by environmental factors, such as drought, salinity, heat, cold, and strong light intensity, by altering gene expression and metabolism as an adaptive or survival strategy. The genes induced in response to stress are generally classified into two major groups. The first one comprises the genes encoding functional proteins that directly provide tolerance through cellular detoxification or by synthesis of the osmoprotectants. While another group encodes regulatory proteins such as kinases and transcription factors, which execute downstream signaling and induction of the former group of genes directly involved in stress tolerance (Hirayama and Shinozaki, 2010; Akpinar et al, 2012; Nakashima et al, 2012). Exploiting the genetic diversity of rice, our lab carried out a comparative study between salt-tolerant Pokkali and salt-sensitive IR64 at

transcriptome level and identified many genes differentially induced in Pokkali under salt stress treatment. Among them, are the genes encoding members of the protein family containing the conserved Cystathionine β -Synthase (CBS) domain (Kumari et al, 2009). Since then, our lab has been working on understanding the role of these uncharacterized CBS domain containing proteins (CDCPs) from rice (Kushwaha et al, 2009; Singh et al, 2012; Kumar et al, 2018). In this study, we have attempted to explore the functions of two important members of rice CDCPs, namely, OsCBSX3 (polypeptide with only a single pair of CBS domains) and OsCBSCLC6 {polypeptide with N-terminal Chloride Channel (CLC) transmembrane domain followed by a single pair of CBS domains at C-terminal}.

This chapter, therefore, presents the literature available on structural and functional features of CBS domains from different CDCPs which have remained highly conserved during evolution and would be useful in the preliminary characterization of novel CDCPs. The literature available on structural and functional aspects of CLCs (including their cytosolic CBS domains) from diverse organisms is presented separately in section 2.3 of this chapter.

2.1 CBS domain- discovery and basic structural aspects

Cystathionine β -Synthase (CBS) domain (Pfam ID: PF00571) is a small intracellular protein domain of about 60 amino acid residues, present in several functionally diverse proteins across all three domains of life. This domain was identified for the first time by Bateman (1997) from the sequences of archaebacterium *Methanocaldococcus jannaschii*, which was also observed to occur in some known mammalian and bacterial proteins. Among these mammalian proteins was a human Cystathionine β -Synthase (CBS) enzyme, which is involved in cysteine biosynthesis. Thus, Bateman (1997) termed this domain as the “CBS domain”. It consists of a typical β 1- α 1- β 2- β 3- α 2 core (numbered according to their arrangement from N-terminal towards C-terminal), in which β 1 is always preceded by a small helix (α 0) and often a short β -strand (β 0) at its N-terminus, while few proteins lack β 1 strand in the domain (Fig. 2.1a; Bateman, 1997; Baykov et al, 2011; Ereño-Orbea et al, 2013a).

CBS domain exists primarily in two or four tandem copies, while mostly co-existing with a wide variety of other non-homologous domain(s) in the polypeptide (Bateman, 1997; Baykov et al, 2011). Altogether, the CBS domain containing proteins (CDCPs) constitute a large protein superfamily of which some families are conserved across all domains of life, while others are specific to a kingdom or lower groups within the kingdom (Table 2.1). Two tandem CBS domains in a polypeptide associate through their β -sheets to a rigid structure of

approximately 2-fold symmetry, called as “Bateman module” (Fig. 2.1b). Notably, the CBS domains from various CDCPs have maintained high conservation in the tertiary structure of their Bateman module, despite the existence of less conservation among them at the sequence level (Baykov et al, 2011; Ereño-Orbea et al, 2013a).

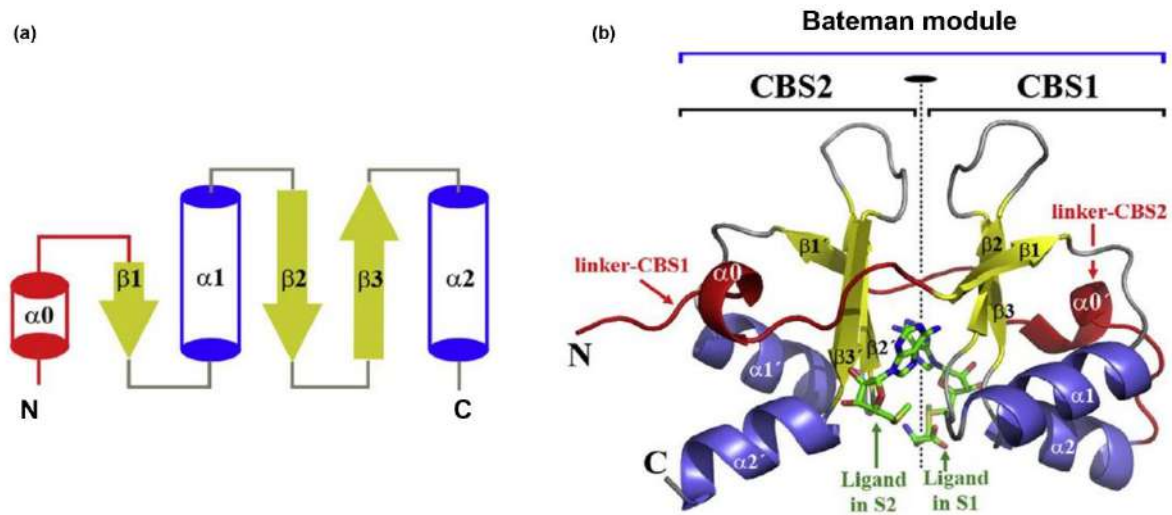


Fig. 2.1: Structure of CBS domains and Bateman module. (a) General topology of a CBS domain that consists of a three β - strands and two α -helices arranged as $\beta1$ - $\alpha1$ - $\beta2$ - $\beta3$ - $\alpha2$. CBS domains are always preceded by a linker of variable length that includes small helix ($\alpha0$, in red). (b) Structure of a Bateman module formed by association of two CBS domains (referred as CBS1 and CBS2) through interactions between their β -sheets. Two CBS domains in the Bateman module are related by pseudo-2-fold symmetry and they form two ligand binding cavities {shown as S1 (for site-1) and S2 (for site-2), respectively, each with the bound ligand shown as a stick model} which are usually occupied by adenosine nucleotides. The flexible linker preceding each CBS domain is shown in red. The represented structure corresponds to Bateman module of MJ0100 from *M. jannaschii* (PDB 3KPC). Figure is reproduced from Ereño-Orbea et al (2013a).

Table 2.1: Domain organization of CBS domains in some common CDCP families and their conservation across different kingdoms.

| Kingdom | CBS | CLC-CBS | IMPDH-CBS | CNNM(DUF21)-CBS | CNNM(DUF21)-CBS-CorC_HlyC | SIS-CBS | PPR-CBS | CNNM(DUF21)-CBS-cNMP-BD | PLP-CBS | MgtE_Nter-CBS-MgtE | DHH-CBS-DHHA2 | CBS-CBS | GBD-CBS-CBS | CBS-CBS-PB1 |
|----------|-----|---------|-----------|-----------------|--|---|---------|--------------------------------|---------|--------------------|---------------|---------|-------------|-------------|
| Plant | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ |
| Animal | ✓ | ✓ | ✓ | ✓ | ✗ (identified in <i>Caenorhabditis japonica</i>) | ✗ (identified in <i>Apis mellifera</i>) | ✗ | ✓ | ✓ | ✗ | ✗ | ✓ | ✗ | ✗ |
| Fungi | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ | ✗ | ✓ | ✗ | ✓ |
| Protista | ✓ | ✓ | ✓ | ✓ | ✓# | ✗ | ✗ | ✓# | ✓# | ✗ | ✗ | ✓ | ✗ | ✗ |
| Archaea | ✓ | ✓# | ✓# | ✓# | ✓# | ✗ | ✗ | ✓# (present as cNMP_BD-CBS) | ✗ | ✓# | ✓# | ✓ | ✗ | ✗ |
| Bacteria | ✓ | ✓# | ✓ | ✓# | ✓ | ✓# | ✗ | ✓# (present as cNMP_BD-CBS) | ✓# | ✓ | ✓# | ✓# | ✗ | ✗ |

The data is based on CDCPs present in representative organisms from each kingdom as retrieved from SUPERFAMILY 1.75 database (Wilson et al, 2009), which is further analyzed using Scanprosite (de Castro et al, 2006). ‘CBS’: a single pair of CBS domains; ‘CLC’: Chloride Channel transmembrane domain; ‘IMPDH’: Inosine Monophosphate Dehydrogenase domain; ‘CNNM/DUF21’: Cyclin M transmembrane N-terminal domain; ‘CorC_HlyC’: CorC_HlyC transporter associated domain; ‘SIS’: Sugar Isomerase domain; ‘PPR’: Pentatricopeptide Repeats; ‘cNMP BD’: Cyclic nucleotide-binding domain; ‘PLP’: Pyridoxal-5'-Phosphate binding domain; ‘MgtE_Nter’: Mg²⁺ transporter N-ter domain; ‘MgtE’: Mg²⁺ transporter transmembrane domain; ‘DHH’: DHH domain ; ‘DHHA2’: DHHA2 domain; ‘CBS-CBS’: two pairs of CBS domains; ‘GBD’: Glycogen binding domain; ‘PB1’: Phox and Bem1p domain. ‘✓’ and ‘✗’ denote the presence or absence of a given CDCP member, respectively; ‘#’ denotes less conservation of a given CDCP member among organisms within the kingdom. Note that several other domain organizations also exist, specific to certain members within the kingdom, and particularly in archaea and bacteria, which are not represented in this table.

From the crystal structure of most of the CDCPS, it is observed that two Bateman modules in a polypeptide (containing two pairs of CBS domains) or from two polypeptides (each containing a single pair of CBS domains) associate with each other through their corresponding α -helices. This structure consisting of four CBS domains is referred to as the “CBS module” (Fig. 2.2; Mahmood et al, 2009). In the case of a polypeptide containing a single pair of CBS domains, this association leads to the formation of a homodimer. Generally, three types of structural arrangements have been identified for the CBS modules from various CDCPs. The most frequently observed arrangement is “head-to-head” or “parallel” orientation in which CBS1 (first CBS domain) of one Bateman module interacts with CBS1’ of other associated Bateman module, and accordingly, CBS2 interacts with CBS2’ (Fig. 2.2 a).

Antagonistic to this, some CBS modules exhibit “head-to-tail” or “antiparallel” assembly in which CBS1 and CBS2 of one Bateman module interact with CBS2’ and CBS1’ of the other, respectively (Fig. 2.2 b). Nevertheless, both these arrangements produce an elliptical disk-shaped CBS module. The third structural arrangement of the CBS module exhibited by a few members of CDCPs involves the association between either CBS1 and CBS1’ or CBS2 and CBS2’ among their Bateman modules, resulting in a V-shaped structure (Fig. 2.2 c; Baykov et al, 2011; Ereño-Orbea et al, 2013a).

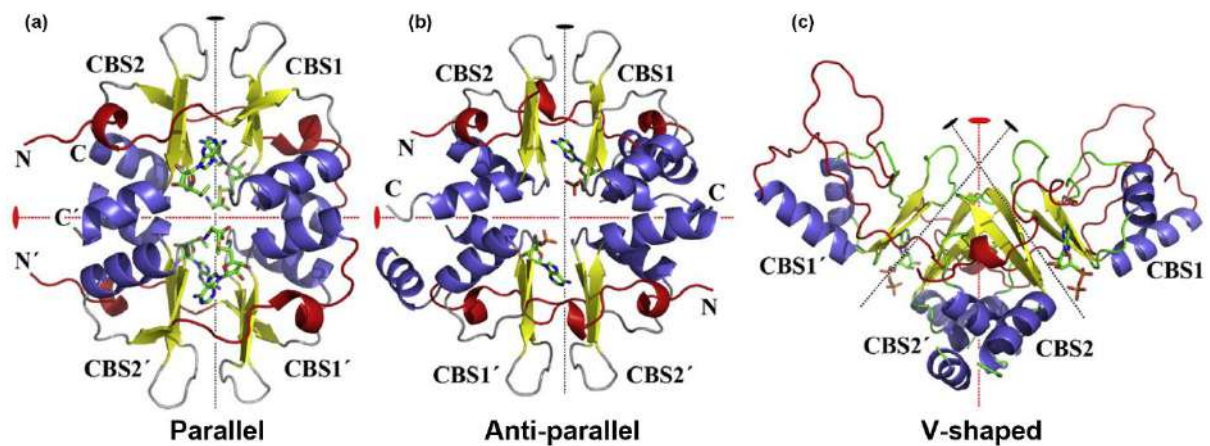


Fig. 2.2: Structure of CBS modules formed by association of two Bateman modules. (a) CBS module with “head-to-head” or “parallel” arrangement in which CBS1 of one Bateman module interacts with CBS1’ of other and likewise, CBS2 interacts with CBS2’ among their Bateman modules. (b) CBS module with “head-to-tail” or “antiparallel” arrangement in which CBS1 of one Bateman module interacts with CBS2’ of other and CBS2 interacts CBS1’ of other. (c) “V-shaped” CBS module in which two Bateman modules associates only by contacts between either CBS1 and CBS1’ or CBS2 and CBS2’. As two CBS domains forming a Bateman module are related by pseudo-2-fold symmetry (represented by black dotted line), in all three conformations of the CBS module, their two constituting Bateman modules are also related by a pseudo-2-fold symmetry axis (represented by a red dotted line). The represented structures correspond to (a) MJ0100 from *M. jannaschii* (PDB 3KPC), (b) AtCBSX2 from *A. thaliana* (PDB 4GQY) and (c) human CLC-5 (PDB 2JA3). The bound ligands are shown as a stick model. Figure is reproduced from Ereño-Orbea et al (2013a).

CBS domains are known to bind various nucleotides, with AMP being most commonly observed. However, CBS domains from a few CDCPs also show an affinity for metal ions and nucleic acids. Two canonical adenosine nucleotide-binding sites are present at the interface region between two CBS domains in a Bateman module. Thus, four potential binding sites are present in a CBS module. At any single binding site, the residues from both the CBS domains in a Bateman module and also those from other Bateman module associated in a CBS module coordinates with the binding of a nucleotide. While all the binding sites are occupied by ligands in some CDCPs, they remain unbound or only partly bound in others. Notably, non-canonical

ligand binding sites also exist in the Bateman module of some CDCPs (Ereño-Orbea et al, 2013a).

Structural and sequence comparison CBS domain-ligand complexes have identified each canonical binding site in a Bateman module to be mainly composed of the residues from three structural blocks: (i) residues from flexible loop preceding the first strand of β -sheet ($\beta 1/\beta 1'$), (ii) residues from the second strand of β -sheet ($\beta 2/\beta 2'$), and (iii) residues from the third strand of β -sheet ($\beta 3/\beta 3'$) and first two turns of the following $\alpha 2/\alpha 2'$ helix (Fig. 2.3). The first block presents the highest diversity in residues and often possesses a conserved Thr/Ser residue, while substituted by Arg/Lys in some CDCPs, which interacts with the hydroxyls of a ribose ring. This block also possesses residue that interacts with the adenine ring of the adenosyl ligands. The second block, consisting of the conserved motif “h-y-y'-h'-P” (‘h’: hydrophobic, ‘y’: any residue), accommodates the adenine ring with hydrophobic residues, and the residue (y') interacts with the adenine ring while impeding the potential binding of guanosine derivatives. The third block exhibits the highest amino acid conservation and contains ribose phosphate-binding motif “G-h-h'-T/S-y-y'-D/N” (‘h’: hydrophobic residue, ‘y’: any residue). The conserved ‘D’ (or ‘N’ in a few CDCPs) in this motif H-bonds with the ribose of a nucleotide and its substitution with other residues impairs the nucleotide-binding. Besides, the residue (y') preceding this conserved ‘D’ has also been considered to determine the binding of phosphate- or non-phosphate-adenosyl ligands (Lucas et al, 2010; Ereño-Orbea et al, 2013a).

The binding of ligand results in subtle changes in the orientation of amino acid side chains or the secondary structural elements residing mainly at dimer interface region between two Bateman modules in a CBS module. This minute change significantly affects the overall structure of the CBS module through the formation of ‘open’, ‘closed’, ‘bent’, or ‘flat’ conformation at the interface region between two Bateman modules. Subsequently, the ligand-induced structural change either activates or represses the activity of the associated functional domain in the same polypeptide or other hetero-polypeptide in a protein complex (Baykov et al, 2011; Ereño-Orbea et al, 2013a). The CBS domain residues involved in ligand binding at both canonical and non-canonical binding sites from some structurally characterized CDCPs are concisely presented in Table 2.2.

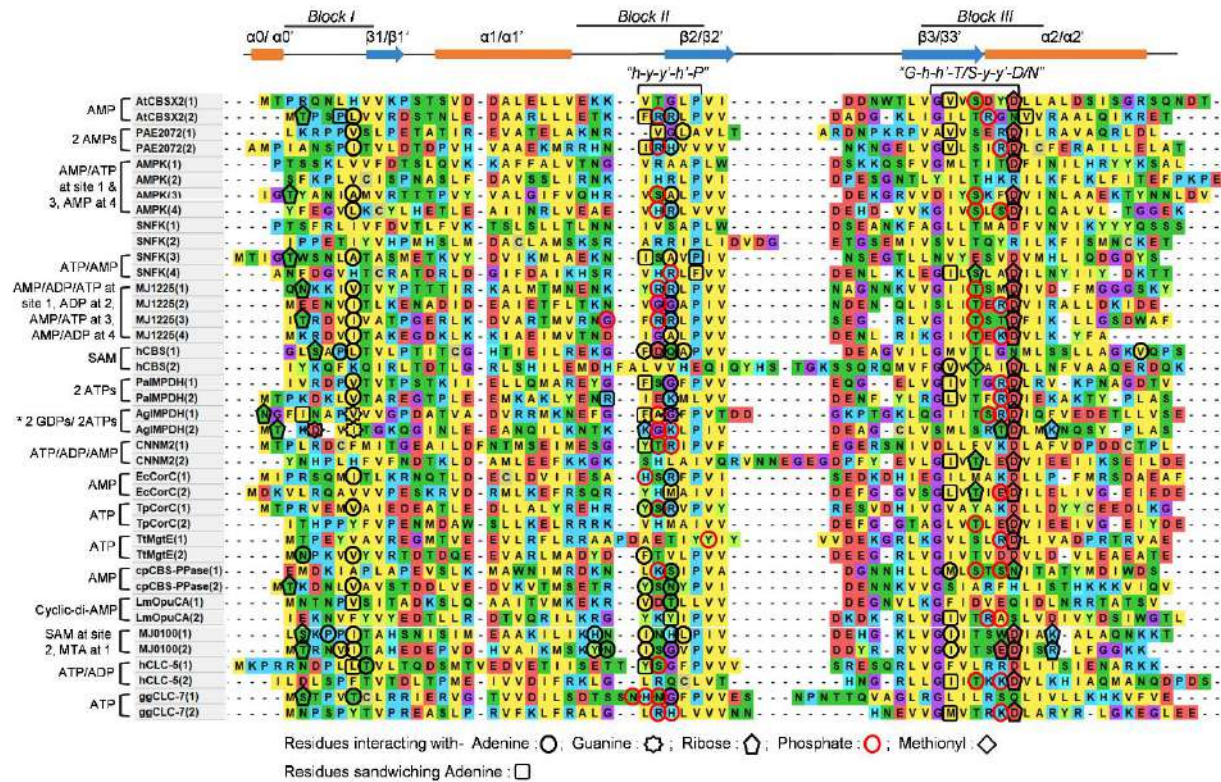


Fig. 2.3: CBS domain residues involved in ligand binding. Sequence alignment of the CBS domains from different CDCPs wherein the residues interacting with ligands have been determined structurally (see Table. 2.2 for more details). CBS domains from each CDCP have been numbered as (1), (2), (3) and (4), according to their arrangement from N-terminus to C-terminus in a polypeptide. Three blocks in a CBS domain that form the ligand binding sites and the two conserved motifs, “h-y-y’-h’-P” and “G-h-h’-T/S-y-y’-D/N” (‘h’: hydrophobic, ‘y’: any residue) present in Block II and Block III, respectively are highlighted on the top. The Block I exhibit low sequence conservation. Different ligands known to bind CBS domains in respective CDCPs are shown on the left, while the residues that interact with the ligands are highlighted by enclosing in different shapes. Note that the residues from both CBS domains in a Bateman module interacts with a ligand in each binding site. For e.g., only one ATP binds to AtCBSX2 (protein with a single pair of CBS domains), however residues from both CBS1 and CBS2 domains interacts with the ligand (as shown in the Fig. above). Besides, the CBS domain residues from another subunit also interacts with the ligand bound to associated subunit in a CBS module. The sequence alignment is performed by ClustalW using Mega X program.

Table 2.2: CBS domain-ligand complex interaction.

| Protein and PDB ID | Ligand | CBS domain residues interacting with the ligand |
|--|-------------------------------|--|
| AtCBSX2, <i>Arabidopsis thaliana</i> ; 4GQY (Jeong et al, 2013a) | AMP (1 per Bateman module) | Adenine (sandwiched between F202, V125 and P181): R204, L182. Ribose: D130, T178. Phosphates: S127, R203, R220’, R204, R203’. |
| PAE2072, | AMP | Adenine {sandwiched between V33, V113 (AMP1) and V50, I98 (AMP2)}; V13, L35 (AMP1), |

| | | |
|--|--|---|
| <i>Pyrobaculum aerophilum</i> ; 2RIF (King et al, 2008) | (2 per Bateman module) | I78, H100 (AMP2). Ribose: D118 (AMP1), D55 (AMP2). Phosphate: R117 (AMP1), R99, R117' (AMP2), H100, R99' both AMPs. |
| γ -AMPK, <i>Rattus norvegicus</i> ; 2V8Q, 2V92, 2V9J (Xiao et al, 2007) 2V9J | AMP/ATP at site 1 and 3, and AMP at site 4. At site 2, R170 substitutes the position of conserved D. | Site 1: D89 (Ribose). Site 3: R298, L276 (Adenine); D244, R268 (Ribose); S241, R298, H297, R151, K169 (Phosphate). Site 4: A226, A204 (Adenine); D316, T199 (Ribose); S313, S315, S225 (Phosphate). |
| SNFK γ -subunit, <i>Schizosaccharomyces pombe</i> ; 2OOX, 2OOY (Townley and Shapiro, 2007) | 1 ATP/AMP at CBS3 and CBS4 interface in the CBS module | ATP/AMP- Adenine (sandwiched between I216, P220, I303, F292): A196 and A218. Ribose: T191, D308, and S305. Phosphate: R290, R139, R141. |
| MJ1225, <i>Methanocaldococcus jannaschii</i> ; 3KH5, 3LFZ (Gómez García et al, 2010) | AMP/ADP/ATP (site 1), ADP (site 2), AMP/ATP (site 3), AMP/ADP (site 4). Site 2 showed least affinity to nucleotides and mostly remained unoccupied. | Adenine: I95, G117 (site 1), V15, R37 (site 2), I234, A256 (site 3), and I159, R181 (site 4). Ribose: D56, D135, D198, and D274 (at site 1, 2, 3, and 4, respectively) as well as N10 in site 2 and T155 in site 4. α -phosphate: T53, T132, T195 and T271 (at site 1 to 4, respectively) as well as with R134 and R36 at site 2 and 3, respectively. β -phosphates: R37, G116, R180, K273 (site 1) and G178, R180 and R181 (site 4). AMP at non-canonical binding site- Adenine (accommodated by F60, L69, I89, K66, K73): E88; Ribose: E72; Phosphate: K73. The corresponding binding site (6) in adjacent Bateman module lacks Lys in equivalent position of K73 in site (5), which is anticipated to preclude nucleotide binding. |
| CBS, <i>Homo sapiens</i> ; 4PCU, 4UUU | SAM per Bateman module | Adenine: 'FDQ' motif (443-445), V533, P422, L423, A446. Ribose: D538 from 'GVVTAID' motif, S420. Methionyl carboxyl: T535. |

| | | |
|--|---------------------------|--|
| (Ereño-Orbea et al, 2013b; McCorvie et al, 2014) | | Methionyl nitrogen: D444'. |
| PaIMPDPH, <i>Pseudomonas aeruginosa</i> ; 4DQW (Labesse et al, 2013) | 2 ATPs per Bateman module | Adenine {sandwiched between F118, L194 (ATP 1), I132, I179 (ATP 2)}: V98, G120 (ATP 1), V159, K181 (ATP 2). Ribose: D137, D199. Phosphates: R136, K181, R198 |

Continuation of Table 2.

| Protein and PDB ID | Ligand | CBS domain residues interacting with the ligand |
|---|---|--|
| AgIMPDPH, <i>Ashbya gossypii</i> ; 4Z87 (Buey et al, 2015) | 3 GDPs per Bateman module (GDP1 and 2 at canonical binding site, GDP3 at non-canonical binding site). | GDP1- Guanine (sandwiched between K208, T184): D186, I188. Ribose: D168, T184, K208. Phosphates: S166, G209, R167'. GDP2- Guanine (sandwiched between F145, I121): V125, G147. Ribose: N118, T227, D228, K231. Phosphate: G147, A146, K210. GDP3 at non-canonical binding site- Guanine (accommodated by G119, L196, L224; sandwiched by L229, K245): N118, N200, K240. |
| AgIMPDPH, <i>Ashbya gossypii</i> ; 5mcp (Buey et al, 2017) | 2ATPs per Bateman module | ATP1- Adenine (sandwiched between K208, I163): I188. Ribose: D168. Phosphates: S166, R167, G209, K210. ATP2- Adenine (sandwiched between F145, I121): V125; Ribose: D228; Phosphates: A146, G147, K210. |
| CNNM2, <i>Mus musculus</i> and <i>Homo sapiens</i> ; 4P1G, 4IY0, 4P10 | 1ATP/ADP/AMP per Bateman module. | Adenine: sandwiched between Y478, I566 in a hydrophobic pocket comprising of P482, I481, C456, F457, M458. Ribose: T568, D571. |

| | | |
|--|--|--|
| (Corral-Rodríguez et al, 2014) | Steric hindrance by F496, K498 in S1 (in positions equivalent to T568 and E570 in S2) is anticipated to impair nucleotide binding at S1. | α -phosphate: R480; β/γ -phosphate: T479. Mg ²⁺ is coordinated by E570. |
| EcCorC, <i>Escherichia coli</i> ; 5YZ2 (Feng et al, 2018) | 1 AMP per Bateman module | Adenine (packed against I74, M165, L178, H100): I80, R102. Ribose: T180 and D183. Phosphate: H100, E182. |
| TpCorC, <i>Thermus parvatiensis</i> ; 7CFI (Huang et al, 2021) | 1 ATP per Bateman module | Adenine: V235, R257, Y255. Ribose: D339. Phosphates: S256, R257, T336 |
| TtMgtE, <i>Thermus thermophilus</i> ; 5X9G (Tomita et al, 2017) | 1 ATP per Bateman module | Adenine: N203, V207, F227. Ribose: D188. Phosphates: Y170 and R187 |
| cpCBS-PPase, <i>Clostridium perfringens</i> ; 3L31 (Tuominen et al, 2010) | 1 AMP per Bateman module. Also bind AP4A, one molecule per dimer, occupying the AMP binding pocket, but with altered contacts at CBS1 domain (Ribose: S116, S118; Phosphate: S279, K100') | Adenine: M114, Y278, N280, V258, S101. Ribose: N119, T253. Phosphate: S116, S118, K100', S101, N280. |

Continuation of Table 2.

| Protein and PDB ID | Ligand | CBS domain residues interacting with the ligand |
|---|---|---|
| LmOpuCA, <i>Listeria monocytogenes</i> ; 5KS7 (Huynh et al, 2016) | 1 cyclic-di-AMP per dimer | Adenine (packed between Y342, I355, V260, V280): V280, T282. Phosphate: R358, A359. |
| MJ0100, <i>Methanocaldococcus jannaschii</i> ; 3KPC (Lucas et al, 2010) | SAM at site 2; MTA at site 1 in a Bateman module. P397 (at site 1) and N458 (at site 2) are suggested to sterically preclude the accommodation of guanine derivatives. | Adenine: I399, I419, L422, I495, H421, P397, H417, S395 (at site 1); I460, I480, I434, I460, G482, Y478, N479 (at site 2). Ribose: D500, S395, R503, E499 (at site 1), D439, T456, K442 (at site 2). |
| hCLC-5, <i>Homo sapiens</i> ; 2J9L, 2JA3 (Meyer et al, 2007) | 1 ATP/ADP per Bateman module | Adenine (packed between Y617, I722, L595): T596. Ribose: D727. Phosphates: S618, T724, K726. |
| ggCLC-7, <i>Gallus gallus</i> ; 7JM6 (Schrecker et al, 2020) | 1 ATP per Bateman module | Adenine (packed between H654, M778): T632, G656. Ribose: S628, D783. Phosphates: H654, N653, K782, N655, H765, R764. |

Note that (') denotes the residue from the associated Bateman module in a CBS module structure.

2.2 Functional aspects of CBS domain in CDCPs

The current knowledge on the functions of CBS domains has mostly been acquired from the studies on CDCPs from mammalian and bacterial systems, and especially from the CDCPs containing additional non-homologous domain(s). The studies and understandings on the CBS domain or CDCPs from plants are still scarce. Nevertheless, the structure and function of CBS domains appear to be conserved in these functionally diverse CDCPs throughout all three domains of life. This domain is suggested to act as a regulatory unit that perceives various cellular signals mainly through interaction with the ligands and accordingly modulates the

activity of other functional domain(s) present in the same or associated polypeptide. Hence, we have referred, analyzed, and presented the know-how on CBS domains from various characterized CDCPs from different organisms in the following sections. Note that the information of CBS domains from chloride channels (CLCs) has been presented in the latter section (section 2.3) describing CLCs.

2.2.1 Proteins containing only a single pair of CBS domains

This family of CDCPs contains only the basic unit of CDCPs, i.e., a single pair of CBS domains in a polypeptide. For this reason, Kushwaha et al (2009) termed these protein members as “CBSX”, so we have used ‘CBSX’ herein to refer to the protein family containing only a single pair of CBS domains. CBSX is present in all organisms, from prokaryotes to eukaryotes; however, only a few members from Archaea and plants have been studied thus far. In archaebacterium (*Pyrobaculum aerophilum*), PAE2072, a CBSX protein, has been structurally characterized to form a homodimer with ‘parallel’ arrangements and bind AMP at all four binding sites in a dimer or CBS module (King et al, 2008). MJ0729, a CBSX protein from another archaeon, *Methanocaldococcus jannaschii*, has been reported to exhibit a pH-dependent oligomerization equilibrium, with dominant species of dimer at pH 7, while tetramer at pH 4.5–4.8 (Martínez-Cruz et al, 2009). Notably, this protein is identified to bind dsDNA (calf-thymus DNA and E-box DNA sequences recognized by basic HLH transcription factors) (Aguado-Llera et al, 2010).

In the *Arabidopsis* genome, 6 out of a total of 33 genes encoding CDCPs have been identified to be composed of only a single pair of CBS domains, which are referred to as *AtCBSX1-AtCBSX6* (Kushwaha et al, 2009). The polypeptides encoded by these six genes showed different subcellular localization: *AtCBSX1* and *AtCBSX2* are localized in the chloroplast, *AtCBSX3* in mitochondria, *AtCBSX4* in the cytosol, and *AtCBSX5* and *AtCBSX6* in the endoplasmic reticulum (Yoo et al, 2011; Ok et al, 2012; Shin et al, 2020). Consistent with the localization, *AtCBSX1* and *AtCBSX2* activated the redox proteins ‘Thioredoxins’ (Trxs) present in the chloroplast, namely, Trx f, Trx m, Trx x, and Trx y. Activation of these Trxs by *AtCBSX1* and *AtCBSX2* is augmented in the presence of AMP, but not by ATP/ADP. Similarly, *AtCBSX3* activated the mitochondrial Trx o2; however, AMP/ADP/ATP inhibited *AtCBSX3* induced Trx o2 activation (Yoo et al, 2011; Shin et al, 2020). *AtCBSX1* is observed to be expressed mainly in the anther tissues, while in leaves, its expression is detected only in trichomes. However, *AtCBSX2* is found to be highly expressed in the leaves (Yoo et al, 2011).

Notably, the expression of *AtCBSX1* is found to be induced in response to NaCl, H₂O₂, and mannitol treatments in *Arabidopsis* seedlings (Yoo et al, 2011).

Both *AtCBSX1* and *AtCBSX2* form a homodimer and their crystal structures exhibited two monomers to be arranged in ‘antiparallel’ orientation in the dimer. This dimeric structure in apo-state, presented a characteristic 155° and 120° bent along the side in *AtCBSX1* and *AtCBSX2*, respectively, contrasting to 180° flat structure in other known structures of CBS modules (Yoo et al, 2011; Jeong et al, 2013a; Jeong et al, 2013b). However, the binding of AMP triggered subtle rearrangement of residues in the dimer interface which resulted in a conformational change from bent to flat structure (Jeong et al, 2013a). Two AMP molecules are found to bind in each CBS module or a dimer of *AtCBSX2* (Jeong et al, 2013a).

When *AtCBSX1* was overexpressed in *Arabidopsis*, the transgenic plants exhibited severe sterility due to anther indehiscence, while its T-DNA insertional mutant line (*atcbx1*) showed normal fertility (Yoo et al, 2011). It was found that the anther from *AtCBSX1* overexpressing lines have reduced H₂O₂ level which caused defective lignin deposition and failure of secondary wall thickening in the anther endothecium, thus resulting in anther indehiscence linked sterility. Whereas the anthers from *atcbx1* lines showed a high level of H₂O₂ content, more lignin deposition, and relatively thicker endothecium as compared to the wild-type (WT) plants. Besides, the *AtCBSX1* overexpressing lines also exhibited an increased number of flower buds and delayed apical senescence (a typical fertilization defective characteristic). They also showed faster growth under normal long-day conditions and less growth retardation of their seedlings when germinated on a sucrose-free medium as compared to *atcbx1* and the WT plants. In contrast, the *cbx1* mutants showed severe growth retardation than the WT plants on a sucrose-free medium. Based on these effects of *AtCBSX1* in plant growth and its ability to activate Trxs, it is suggested to have a role in Trx-mediated redox regulation in the cellular processes, including the photosynthesis process (Yoo et al, 2011).

Antagonistic to *AtCBSX1*, the plants overexpressing mitochondrial *AtCBSX3* showed normal fertility; however, its knockdown (*atcbx3*) lines exhibited nearly sterile and shorter siliques (Shin et al, 2020). This sterility, as associated with *AtCBSX1* but in overexpressing lines instead (Yoo et al, 2011), is also identified to be due to anther indehiscence resulting from insufficient ROS accumulation which led to decreased lignin deposition in the secondary cell wall of the anther endothecium (Shin et al, 2020). The *AtCBSX3* overexpressing lines with higher ROS accumulation showed retarded plant growth, while the knockdown *atcbx3* lines

with reduced ROS accumulation exhibited accelerated plant growth than the WT plants (Shin et al, 2020). Based on these plant phenotypes, and the identified interaction of AtCBSX3 with Trx-o2 and Trx-o2 with SDH1 (flavoprotein subunit of succinate dehydrogenase 'SDH') which is a major source of ROS, Shin et al (2020) suggested CBSX3-Trx-o2 as a ROS regulator in mitochondria.

In rice (*Oryza sativa* ssp. *Japonica*), its genome contains 37 genes encoding CDCPs, and 12 of them (referred to as *OsCBSX1- OsCBSX12*) have been identified to encode polypeptides with only a single pair of CBS domains (Kushwaha et al, 2009). *OsCBSX3* and *OsCBSX4* have been suggested to have important functions in various stresses (Wang et al, 2004; Singh et al, 2012; Mou et al, 2015). Transcripts of *OsCBSX3* (previously annotated as *OsBi1*) are observed to be upregulated in response to brown planthopper (BPH; *Nilaparvata lugens*) feeding in the shoots of B5 rice (BPH resistant variety), significantly after 36 h of BPH infestation. The gene is identified to be induced mainly in tissues around the vascular bundle in stems. In addition to BPH, the induction of *OsCBSX3* is observed also by ethephon (ethylene releasing compound) treatment and by water-deficit stress in plants, but not by mechanical wounding or by *Magnaporthe grisea* (rice blast fungus) inoculation (Wang et al, 2004). However, the study by Mou et al (2015) reported significant induction of *OsCBSX3* in the leaves of Nipponbare rice by *M. oryzae* at 1-5 days post-inoculation (dpi). Besides, the studies by Wang et al (2004) and Mou et al (2015) observed contrasting expression patterns of *OsCBSX3* as no induction and induction, respectively to jasmonic acid and salicylic acid treatment. However, the difference in plant genotypes, the concentration of molecules and the time points of treatments, etc., may account for these contrasting results. Mou et al (2015) reported the localization of *OsCBSX3* into the plasma membrane. Besides, they also observed higher expression of plant defense-associated marker genes and enhanced resistance to *M. oryzae* inoculation in transgenic rice overexpressing *OsCBSX3*.

OsCBSX4 is observed to be highly upregulated, particularly in response to NaCl and Methyl Viologen (MV) (>5-fold), and also to ZnCl₂ (up to 3-fold) stress treatments in Pokkali rice (salinity-tolerant). However, in IR64 rice (salinity-sensitive), it was upregulated only slightly by NaCl (2-fold) stress, while downregulated and remained unaffected by ZnCl₂ and MV stress treatments, respectively (Singh et al, 2012). The leaf of transgenic tobacco overexpressing *OsCBSX4* when treated with NaCl, ZnCl₂, and MV solutions showed reduced chlorophyll degradation. Accordingly, the *OsCBSX4* overexpressing transgenic tobacco seedlings exhibited better growth and lower level of H₂O₂ accumulation as compared to the

WT seedlings under NaCl, ZnCl₂, and MV stress. Further, when grown under NaCl stress, the transgenic plants exhibited healthy vigor, higher seed setting, and maintained lower Na⁺/K⁺ ratio and higher photosynthetic capacity as compared to WT counterparts (Singh et al, 2012).

2.2.2 ‘ γ -subunit’ of AMPK / SNFK / SnRK

Adenosine monophosphate-activated protein kinase (AMPK) and its orthologs ‘sucrose, non-fermenting kinase’ (SNFK) in yeast, and ‘SNF related protein kinase’ (SnRK) in plants, are heterotrimeric serine/threonine kinase complex present in all eukaryotes, except for a few intracellular parasites. They are composed of a catalytic ‘ α ’ subunit and two regulatory subunits: ‘ β ’ and ‘ γ ’, of which ‘ γ ’ subunit possesses only two pairs of CBS domains (Lin and Hardie, 2018). Conventionally, AMPK orthologs in yeasts and plants are called SNF1 and SnRK1 complex; however, because of the existence of multiple isoforms of their subunits and their multiple combinations in complexes, we have found it appropriate to refer to them collectively as SNFK and SnRK, respectively in this thesis.

In mammals, these enzyme complexes function in regulating cellular energy homeostasis. The enzyme is activated in the state of low cellular energy level or fall in ATP:AMP ratio, which subsequently triggers the ATP generating catabolic pathways while turning off the ATP consuming anabolic pathways, to maintain the energy homeostasis. The activation of this complex requires phosphorylation of the conserved Threonine (T172 in rat AMPK) at α -subunit by the upstream AMPK kinase (AMPKK). AMP assists the activation of AMPK by (i) promoting its phosphorylation by AMPKK, (ii) inhibiting its dephosphorylation by phosphatases, and (iii) inducing allosteric activation through binding at the ‘ γ ’ subunit (Crozet et al, 2014; Lin and Hardie, 2018). Although the former two mechanisms also apply in SNFK and SnRK complex, however, the allosteric activation by AMP binding to the γ subunit appears to be absent in them (Mackintosh et al, 1992; Wilson et al, 1996; Crozet et al, 2014).

Yeast SNFKs are required for yeast growth on sucrose or non-fermentable (SNF) carbon sources, hence called SNF kinases. The enzyme is activated upon glucose starvation and thus it senses glucose availability. It then derepresses the glucose-repressed genes necessary for the metabolism of alternative carbon sources (Wilson et al, 1996). The two Bateman modules of AMPK/SNFK γ -subunit, associate in a ‘parallel’ arrangement, forming a CBS module (Amodeo et al, 2007; Townley and Shapiro, 2007; Xiao et al, 2007). The mammalian γ -AMPK have been identified to bind three adenine nucleotides per CBS module (Xiao et al, 2007), whereas the yeast γ -SNFK exhibited the binding of only one AMP/ATP in

it (Amodeo et al, 2007; Townley and Shapiro, 2007). Further, contrasting to the mammalian AMPK complex, which is a monomer of heterotrimer (Xiao et al, 2007), that in yeast is observed to form a dimer of heterotrimers (Townley and Shapiro, 2007).

Plant SnRKs are associated with several processes, ranging from metabolism, development, and stress responses (Polge and Thomas, 2007). Exceptional to mammalian and yeast counterparts, plants have, in addition to a canonical β -subunit and a γ -subunit, a truncated β -subunit without Glycogen Binding Domain (GBD), and a hybrid $\beta\gamma$ -subunit-containing GBD at N-terminal of the four CBS tandem repeats of γ -subunit (Lumbreras et al, 2001). Although both canonical and truncated β -subunits complemented the function of β -subunit of yeast (Polge et al, 2008), only hybrid $\beta\gamma$ -, but not canonical γ -subunit, restored phenotype of γ -subunit deleted yeast mutant (Ramon et al, 2013). Consistently, in phylogenetic analysis, hybrid $\beta\gamma$ clustered within a monophyletic clade comprising of yeast and animal γ -SNFKs/AMPKs, while plant genes encoding only two pairs of CBS domains (CBSCBS) formed a separate clade, implying that hybrid $\beta\gamma$ functions as the canonical γ subunits in plant SnRK (Ramon et al, 2013). Surprisingly, GBD in hybrid $\beta\gamma$ -subunit is found to be non-essential for complex formation and function (Lumbreras et al, 2001; Ramon et al, 2013). Nevertheless, GBD in hybrid $\beta\gamma$ -subunit is identified to bind starch (as it does in β -subunit) that resulted in inhibition of the enzyme activity of the SnRK complex formed by α - and hybrid $\beta\gamma$ -subunits as a heterodimer, but not requiring β -subunit (Avila-Castañeda et al, 2014). Alike in yeast, GBD of hybrid $\beta\gamma$ -subunit has also been suggested to sense glucose abundance and regulate the enzyme activity (Lumbreras et al, 2001). Notably, Arabidopsis hybrid $\beta\gamma$ -subunit has been reported to be essential for the biogenesis of mitochondria and peroxisomes in pollen to maintain the ROS level critical for pollen germination (Gao et al, 2016). Besides, its homologs in maize, ZmAKIN $\beta\gamma$ -1 and ZmAKIN $\beta\gamma$ -2, have been identified to interact with the P8 protein of Rice Black-Streaked Dwarf Virus (RBSDV), the knockdown of which promoted the accumulation of RBSDV in maize (Li et al, 2020).

Although AMPK complex has not been observed in the archaeal system, they also possess CDCPs composed of only two pairs of CBS domains, such as MJ1225 in *M. jannaschii*, which has been proposed to function as eukaryotic γ -AMPKs, regulating the kinases or other metabolic enzymes (Gómez García et al, 2010). Its two Bateman modules to associate in a head-to-head arrangement, forming a CBS module, and bind three adenosine nucleotides, while the fourth site remained mostly unoccupied. Besides, it also exhibited an additional AMP binding at a non-canonical site (Gómez García et al, 2010).

2.2.3 CBS domains in Cystathionine β Synthase enzyme

Cystathionine β Synthase, from where the CBS domain derived its name, is an enzyme that catalyzes the production of cystathionine from serine and homocysteine substrates in the cysteine biosynthesis pathway. This enzyme also catalyzes alternative reactions leading to H₂S production, an important signaling molecule in cardiovascular and the nervous systems. It is composed of the N-terminal heme-binding domain (absent in lower eukaryotes), the central catalytic pyridoxal-5'-phosphate (PLP) binding domain, and a pair of CBS domains at the C-terminal. Its two CBS domains assemble to form a Bateman module and bind S-adenosyl L-methionine (SAM / AdoMet), which results in allosteric activation of the enzyme. In an unbound state, the Bateman module acts as an autoinhibitory unit by blocking the catalytic site of an enzyme (Anashkin et al, 2017; Zuhra et al, 2020). Consistently, the deletion of this regulatory module yields a highly active catalytic core that is insensitive to SAM (Shan and Kruger, 1998). In addition, its Bateman module assists in the formation of a homotetrameric enzyme, while the truncation of the Bateman module leads to homodimerization (Kery et al, 1998; Jhee et al, 2000). Exceptionally, the Bateman module in the CBS enzyme of *Drosophila melanogaster* does not bind to SAM and the enzyme is a homodimer (Koutmos et al, 2010). Whereas the homotetrameric yeast CBS enzyme, although binds SAM in its Bateman module, is insensitive to SAM binding (Jhee et al, 2000; Anashkin et al, 2017).

Crystal structures of the modified full-length human CBS enzyme (hCBS) {deletion of 516-525 amino acid residues in CBS domain, that facilitated protein expression and crystal growth, but as a homodimer (Oyenarte et al, 2012)} revealed that in SAM unbound state, the Bateman module of one subunit blocks the catalytic site of the other subunit, thus limiting the flux of substrates into the cavity in SAM. However, the binding of SAM induced displacement of the Bateman module away from an entrance to the catalytic cavity and activates an enzyme. Besides, the SAM-induced structural arrangement resulted in the association of two Bateman modules to form a CBS module in an 'antiparallel' orientation (Ereño-Orbea et al, 2014; McCorvie et al, 2014). This SAM bound structure of hCBS resembled that of *D. melanogaster* which remains constitutively active (Ereño-Orbea et al, 2014).

The structural studies reported the binding of one SAM molecule per Bateman module of hCBS, i.e., four SAMs per homotetramer (Ereño-Orbea et al, 2013b; Ereño-Orbea et al, 2013a; McCorvie et al, 2014). In contrast, the Isothermal Titration Calorimetry (ITC) study suggested the binding of six SAMs per homotetramer of full-length CBS: two with high

binding affinity, apparently for kinetic stabilization of the regulatory domains, while the remaining four with low affinity for enzyme activation (Pey et al, 2013). This difference has been inferred to be due to the deletion of 516–525 amino acid residues from the CBS enzyme in crystal structural studies which caused the elimination of two high-affinity binding sites responsible for kinetic stabilization of the Bateman module, thus resulting in the loss of binding of two SAMs and dimer formation (Pey et al, 2016).

The SAM binding pocket in the human CBS enzyme is formed by ‘FDQ’ motif (443–445 residues in CBS1), GVVTAID motif (532–538 residues in CBS2), and 414–423 amino acid residues present in the linker region between CBS1 and CBS2 (Ereño-Orbea et al, 2013a; McCorvie et al, 2014). Mutagenesis study has confirmed the importance of ‘FDQ’ motif in SAM binding, of which F443 and D538 are found to be involved in SAM-driven inter-domain communication as well (McCorvie et al, 2014).

2.2.4 CBS domains in Inosine Monophosphate Dehydrogenase (IMPDH)

IMPDH is a rate-limiting enzyme in the *de novo* biosynthesis of guanine nucleotides that is conserved from prokaryotes to eukaryotes. It catalyzes the NAD⁺-dependent oxidation of inosine monophosphate (IMP, also a precursor of adenine nucleotides) to xanthosine monophosphate. It consists of two domains: a catalytic domain composed of a (β/α)₈ barrel, and a pair of CBS domains as a regulatory unit; the latter is inserted within the loop of the catalytic domain (Anashkin et al, 2017). IMPDH exists as an octamer through the association of two tetrameric units. The octameric arrangement is mainly facilitated and stabilized by the interactions between the Bateman modules (arranged antiparallely) from two tetramers in the ligand-bound state (Labesse et al, 2013; Buey et al, 2015).

IMPDH Bateman module from *Pseudomonas aeruginosa* (PaIMPDH) is identified to bind two ATP molecules in the presence of Mn²⁺ or Mg²⁺, which allosterically increased its activity. Accordingly, deletion of the Bateman module raised the basal activity of the enzyme to ATP activated level and rendered the enzyme ATP insensitive (Labesse et al, 2013). Such deletion of the Bateman module in EcIMPDH (from *E. coli*) dramatically impaired the cellular ATP/GTP pool (Pimkin and Markham, 2008).

In eukaryotes, the IMPDH Bateman module from a fungus, *Ashbya gossypii* (AgIMPDH), is identified to bind GTP/GDP (Buey et al, 2015) as well as ATP/ADP/AMP (Buey et al, 2017). The Bateman module of IMPDH from parasitic protozoan (*Leishmania*

donovani, LdIMPDH) is also shown to bind GMP/GTP as well as ATP (Smith et al, 2016). The binding of GTP/GDP allosterically inhibited the enzyme activity of AgIMPDH as well as that of two human IMPDH isoforms, HsIMPDH1 and HsIMPDH2, while the binding of AMP/ADP/ATP activated these enzymes (Buey et al, 2015; Buey et al, 2017). In bovine IMPDH1, the phosphorylation at T159/S160 residues in its Bateman domain has been reported to desensitize it from allosteric inhibition by GDP/GTP (Plana-Bonamaisó et al, 2020). On the contrary, the activity of prokaryotic IMPDHs, EcIMPDH, and BsIMPDH (from *Bacillus subtilis*), remained unaffected to GTP/GDP (Buey et al, 2015).

Like prokaryotic homolog, AgIMPDH showed binding of two nucleotides per Bateman module (Buey et al, 2015; Buey et al, 2017). Notably, Buey et al (2015) also identified an additional non-canonical binding site specific to a GDP molecule in each Bateman module. They found this non-canonical GDP binding site to be structurally different from the one identified in MJ1225 from *M. jannaschii* (mentioned in section: 2.2.2); however, conserved in eukaryotic IMPDHs, but not in prokaryotic ones (Buey et al, 2015). Importantly, the function of AgIMPDH is suggested to be mainly determined by nucleotide binding in this non-canonical site and the second canonical site in its Bateman module (Buey et al, 2017).

Furthermore, the Bateman modules from IMPDHs have also been reported to bind dinucleoside polyphosphate (Fernández-Justel et al, 2019) and single-stranded nucleic acids (both DNA and RNA) (Cornuel et al, 2002; McLean et al, 2004; Mortimer et al, 2008; Kozhevnikova et al, 2012). The former molecule has also been reported to bind in the Bateman module of the bacterial CBS-PPase enzyme (described in the later section 2.2.6). In the case of plants, not much is known about IMPDH at the molecular level. The gene has been cloned from *Arabidopsis thaliana* (Collart et al, 1996) and *Glycine max* (Cao and Schubert, 2001), and that from the latter has been observed to complement the growth of *E. coli* KLC381 (a strain lacking *IMPDH/AguaB*) on purine deprived medium (Cao and Schubert, 2001).

2.2.5 CBS domains in CNNM, CorC and MgtE: Proteins involved in Mg²⁺ transport

CNNM/cyclin M proteins, formerly referred to as “ancient conserved domain proteins”, function as the divalent metal cation transport mediators. They are composed of four components: N-terminal extracellular region, CNNM or DUF21 membrane-spanning domain, followed by a pair of CBS domains and a cyclic nucleotide monophosphate (cNMP) binding-like domain at the C-terminal. Four CNNM members (CNNM1–4) are present in humans which are suggested to be involved in maintaining the Mg²⁺ homeostasis, either directly as

transporters or as intracellular Mg^{2+} sensors by indirectly regulating other yet unidentified transporters. Importantly, CBS domains in CNNMs are indispensable for the Mg^{2+} transport function (Hirata et al, 2014; Giménez-Mascarell et al, 2019).

The Bateman modules from two CNNM polypeptides have been identified to co-associate to form the CBS module in ‘parallel’ orientation (Corral-Rodríguez et al, 2014; Gulerez et al, 2016). Each CNNM Bateman module is identified to bind one adenosine nucleotide (ATP/ADP/AMP) (Corral-Rodríguez et al, 2014), except for CNNM3 unto which no ligand binding is detected (Hirata et al, 2014). The binding of ATP is favored by Mg^{2+} , which neutralized the negative charge repulsion between its phosphates and the acidic residues in the binding pocket (Corral-Rodríguez et al, 2014; Hirata et al, 2014). In nucleotide unbound state, the CBS module adopted ‘V or Y-shaped’ conformation, mediated by contacts between only the CBS2 domains of two Bateman modules. However, the binding of nucleotides changed the conformation of the CBS module to a ‘disc-like/flat’ structure with extensive interactions among both CBS1 and CBS2 from two subunits (Corral-Rodríguez et al, 2014; Chen et al, 2020). This conformational change in the Bateman module has been presumed to be transmitted to the transmembrane region and thus regulates its activity (Corral-Rodríguez et al, 2014).

Through their CBS domains, CNNMs have been found to interact with PRLs (phosphatases of regenerating liver), which are considered as the oncogenic protein-tyrosine phosphatases (PTPs) (Hardy et al, 2015; Giménez-Mascarell et al, 2017). Specifically, a region has been identified, unique to CBS2 of CNNM family proteins, which is involved in their interaction PRLs (Hardy et al, 2015; Gulerez et al, 2016; Kostantin et al, 2016; Giménez-Mascarell et al, 2017; Zhang et al, 2017). The interaction between these two proteins resulted in the flat disc conformation of the CNNM Bateman module, corresponding to the one adopted in Mg^{2+} ATP bound state (Giménez-Mascarell et al, 2017; Zhang et al, 2017). Besides, this complex formation has been observed to be important for Mg^{2+} transport *in vivo* (Hardy et al, 2015; Gulerez et al, 2016; Kostantin et al, 2016).

The CorC and the MgtE families of CDCPs are also known to be involved in Mg^{2+} transport. In the bacterium *Staphylococcus aureus*, MpfA (CorC homolog) and MgtE exhibited Mg^{2+} export and import function, respectively (Trachsel et al, 2019). The CorC family proteins, alike CNNMs, possess the N-terminal DUF21 transmembrane domain and cytoplasmic CBS domain pair; however, one more domain, CorC/HlyC, exists at its C-terminus (Armitano et al,

2016; Huang et al, 2021). As in CNNMs, the Bateman module from CorC has been identified to bind one molecule of ATP/AMP (Feng et al, 2018; Huang et al, 2021), and these two protein families share highly conserved residues in adenosine binding sites (Giménez-Mascarell et al, 2019; Huang et al, 2021). The MgtE Mg^{2+} transport proteins are composed of an N-terminal N domain, a pair of CBS domains, and a transmembrane domain at the C-terminal (Hattori et al, 2007). These channel proteins, although exist in both prokaryotes and eukaryotes, the homologs from the latter (represented as SLC41 family) lack both N domain and CBS domains (Moomaw and Maguire, 2008). Bacterial MgtE forms a homodimer and functions as a highly selective Mg^{2+} channel. The cytosolic domains of MgtE contain several Mg^{2+} binding sites as well as an ATP/ADP binding site in the Bateman module (Hattori et al, 2009; Tomita et al, 2017). It exhibited a higher affinity for Mg^{2+} in ATP bound state, and likewise, ATP binding affinity is increased in the presence of Mg^{2+} (Tomita et al, 2017). The binding of ATP resulted in inhibition of the channel gating, while the dissociation of ATP facilitated Mg^{2+} influx. Based on this observation and the positive correlation between ATP and Mg^{2+} at the cellular level, the Bateman module in MgtE is suggested to act as an ATP-dependent intracellular Mg^{2+} sensor (Tomita et al, 2017).

2.2.6 CBS domains in Inorganic Pyrophosphatase (PPase)

PPases catalyze the hydrolysis of inorganic pyrophosphate (PPi) to orthophosphate (Pi). These enzymes exist in two forms, soluble and integral membrane-bound, and the soluble PPases are further sub-divided into two non-homologous families. The family I PPases are present in all living organisms, whereas the family II PPases are found only in some bacterial lineages. It is nearly within a quarter of the presently known type II PPases, a pair of CBS domain is present, thus called as CBS-PPases (Jämsen et al, 2007; Tuominen et al, 2010). These family II PPases are Co^{2+} or Mn^{2+} metalloenzymes and they require Mg^{2+} for catalysis. They consist of the N-terminal DHH domain and C-terminal DHHA2 domain, while in the case of CBS-PPase, two CBS domains are inserted in the DHH domain. Interestingly, the majority of the CBS-PPases also possess a DRTGG domain of about 120 residues, inserted between the two CBS domains (Tuominen et al, 2010). Like other type II PPases, CBS-PPases were previously suggested to form a homodimer (Jämsen et al, 2007; Tuominen et al, 2010). However, a recent study revealed that they form a homotetramer instead, mediated by CBS domains (Dadinova et al, 2020).

Unlike the family II PPase members lacking CBS domains, the activity of CBS-PPases is strongly inhibited by AMP and ADP, while activated by ATP and linear diadenosine polyphosphates, through their binding at CBS domains of these proteins (Jämsen et al, 2007; Tuominen et al, 2010; Jämsen et al, 2012). Two molecules of AMP or one molecule of diadenosine polyphosphate (AP4A) have been identified to bind in each CBS module. The binding of AP4A induced opening of the Bateman module dimer interface as compared to AMP bound state (Tuominen et al, 2010). Notably, the DRTGG domain has been observed to be essential for the binding of diadenosine polyphosphates in the Bateman module of CBS-PPase (Anashkin et al, 2015). The crystal structure of the regulatory region consisting of CBS1-DRTGG-CBS2 from *Clostridium perfringens* cpCBS-PPase exhibited its Bateman module to form a homodimer in ‘parallel’ arrangement, with DRTGG domain positioned adjacent to CBS1 of each Bateman module (Tuominen et al, 2010).

2.2.7 CBS domains in other CDCP members

In rice, *OsCBSCBSPB4* encoding a nucleocytoplasmic protein consisting of two pairs of CBS domains followed by a Phox/Bep1 (PB1) domain at C-terminus, have been observed to be induced in response to various abiotic stress conditions, particularly in heat, cold, oxidative, and salinity stresses (Kumar et al, 2018). The PB1 domain is a protein interaction module conserved in eukaryotes. This domain facilitates interactions among diverse proteins containing it, either through oligomerization between their PB1 domains (Sumimoto et al, 2007; Mutte and Weijers, 2020). Heterologous expression of *OsCBSCBSPB4* in *E. coli* and tobacco conferred unto them the tolerance to various abiotic stresses (Kumar et al, 2018). Another CDCP from rice, *OsCBSDUF1*, consisting of a pair of CBS domains and a DUF21 domain has been reported to interact specifically with two mitochondrial thioredoxins (*OsTrx1* and *OsTrx20*) and are suggested to regulate the redox-dependent cuticle development in both leaf and anther (Zafar et al, 2020; Zafar et al, 2021). The loss of function mutation in this gene caused a reduction in panicle fertility, hence the gene was termed as *Degenerated Panicle and Partial Sterility (DPS1)* by Zafar et al (2020).

In bacteria, the cytoplasmic ATPase subunit (*OpuCA*) of osmolyte transporter ‘*OpuC*’ is composed of N-terminal ATP-binding cassette (ABC) and two CBS domains at its C-terminus. Its CBS domains have been identified to bind cyclic-di-AMP (c-di-AMP; Huynh et al, 2016; Schuster et al, 2016), which is suggested to inhibit *OpuC* osmolyte uptake function (Sikkema et al, 2020). The crystal structure exhibited dimerization of the Bateman modules

from *Listeria monocytogenes* LmOpuCA, in ‘anti-parallel’ arrangement, and each dimer or CBS module is observed to bind a molecule of c-di-AMP (Huynh et al, 2016). The crystal structure of the CBS domains of MJ0100, a protein consisting of DUF39 domain and a pair of CBS domains from *Methanocaldococcus jannaschii*, exhibited dimerization of the Bateman modules with ‘parallel’ arrangement (Lucas et al, 2010). Each Bateman module is identified to bind S-adenosyl-L-methionine (SAM) and S-methyl-5'-thioadenosine (MTA) at two different sites. The serine/threonine residue preceding the conserved aspartate in the recognition motif GhxS/TxS/TD (x-any residue and h-hydrophobic residue) is found to be substituted by negatively charged (Glutamate, E) and a bulky hydrophobic (Tryptophan, W) residue in each of its two CBS domains, respectively. Accordingly, two novel recognition motifs: “GhxS/TxhD/E” and “GhxS/TxED”, are suggested for binding of SAM and MTA, respectively in CBS domains (Lucas et al, 2010).

2.3 Chloride channels (CLCs)

CLC comprises an anion conducting protein family members, present ubiquitously in all organisms, from prokaryotes to eukaryotes. These proteins are composed of the N-terminal transmembrane domain and C-terminal cytosolic region containing two CBS domains. Notably, most of the CLCs from prokaryotes lack cytosolic CBS domains. CLC was first identified by White and Miller (1979) as a voltage-dependent Cl^- selective anion channel during the electrophysiological experiment in membrane vesicles from the electric organ of *Torpedo californica* (electric ray fish). Later, Jentsch et al. (1990) isolated the gene encoding this anion channel from *T. marmorata* and called it ‘CLC-0’, the founder of all CLCs. Initially, all CLCs were considered to function as channels as has been named (substrate diffuses passively down their electrochemical gradients), until the identification of exchanger function (Cl^-/H^+) in CLC-ec1 from *E. coli*, through which the substrate moves against their concentration gradients by using energy coupled to the movement proton (Accardi & Miller, 2004). Thereafter, many other CLCs, from both prokaryotes and eukaryotes have been identified to function as an exchanger. Hence not to be confused with its name, CLCs include channel as well as exchanger proteins.

Nine CLCs are present in mammals, of which four are channel proteins (CLC-1, 2, Ka, and Kb) while the remaining five (CLC-3 to CLC-7) function as transporters. Interestingly, the former group localizes in the plasma membrane, whereas the exchangers localize into the intracellular vesicles (Jentsch, 2015). In plants, studies on CLCs are mainly conducted on

Arabidopsis, which possesses seven CLCs (AtCLC-a to AtCLC-g). Contrasting to mammalian homologs, all seven AtCLCs exhibited organellar membrane localization (Lv et al, 2009). Among them, AtCLC-a and AtCLC-b have been identified to function as exchangers; however, selective to nitrate (NO_3^-/H^+ exchanger) (De Angeli et al, 2006; von der Fecht-Bartenbach et al, 2010). All these CLC exchangers from mammals and plant (*Arabidopsis*) as well as from red alga (*Cyanidioschyzon merolae*) and *E. coli*, have been identified to exhibit an exchange stoichiometry of 2 anions:1 proton (Accardi and Miller, 2004; Scheel et al, 2005; De Angeli et al, 2006; Graves et al, 2008; Feng et al, 2010; Neagoe et al, 2010). Furthermore, the phylogenetically distinct member of bacterial CLCs, “CLC-Fs” exhibit fluoride (F^-) selectivity and are identified to function as an exchanger. However, unlike other CLC exchangers, CLC-Fs exhibited 1:1 exchange stoichiometry (Stockbridge et al, 2012). CLC-Fs are associated with fluoride riboswitches, that transport F^- into liposomes and impart fluoride tolerance in bacteria (Baker et al, 2012).

In *E. coli*, CLC homologs are known to confer acid tolerance through the removal of excess intracellular protons while neutralizing the intracellular positive charged products with the Cl^- influx (Foster, 2004). Yeast has only a single gene encoding CLC (*ScCLC* or *GEF1*). It is involved in Cl^- influx into the Golgi vesicles, which facilitates Cu^{2+} transport by *Ccc2* (Cu^{2+} transporting P-type ATPase) and subsequent Cu^{2+} loading onto the iron transport multicopper oxidase (*FET3*), a ferrous transporter protein involved in high-affinity iron uptake (Davis-Kaplan et al, 1998). Accordingly, *ScCLC* mutant (*Agef1*) yeast exhibits an iron-limited growth defect and is sensitive to different salts and higher pH (Davis-Kaplan et al, 1998; Gaxiola et al, 1998).

2.3.1 Structural features and conserved residues governing the functions of CLCs

In CLCs, the transmembrane domain is composed of 16–18 α -helices (herein referred to in alphabetical order), followed by the cytosolic CBS domains which are absent in most of the prokaryotic CLCs. These anion transport proteins exist as a homodimer and each monomer possesses its ion transporting pore, thus two pores exist in a dimer. Although CLCs share less similarity at the sequence level, the structure of CLCs, both channels and exchangers, are highly conserved across prokaryotic to eukaryotic organisms. (Dutzler et al, 2002; Feng et al, 2010; Park et al, 2017; Park and MacKinnon, 2018; Wang et al, 2019; Schrecker et al, 2020). The structure of one of the CLC-Fs, “CLC-eca” from *Enterococcus casseliflavus*, also

complemented well with the structure of other CLCs. However, its helices appeared to be relatively shorter than those in other CLC members (Last et al, 2018).

The first structures of CLCs are reported by Dutzler et al (2002) from *E. coli* and *Salmonella enterica* (Fig. 4). Each subunit/monomer consisted of 18 α -helices of variable lengths, that aligned tilted in the membrane. The N-terminal half of the polypeptide is structurally related to the C-terminal half. These two halves are oriented in an antiparallel architecture, bringing closer the amino acid residues from four different regions to form an ion conduction pathway. These pore-forming residues are highly conserved (called selectivity filter), comprising of the sequences ‘GSGIP’ (106–110 in helix D), ‘G(K/R) EGP’ (146–150 in helix F), and ‘GXFXP’ (355–359 in helix N), as well as ‘Y445’ in helix R (residues are numbered based on *E. coli* CLC-ec1). The Cl^- ion is coordinated by S107, I356, F357, and Y445 (Dutzler et al, 2002).

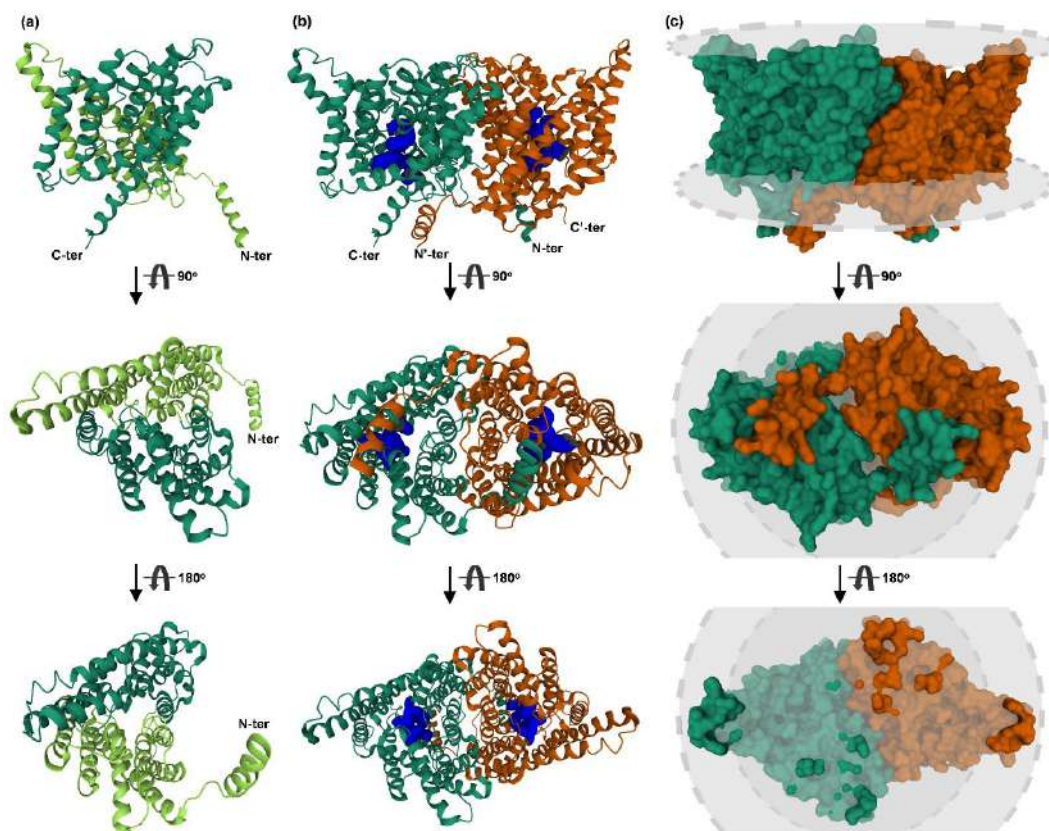


Fig. 2.4: Structure of CLC-ec1 from *E. coli*. (a) Structure of a single subunit or monomer of CLC-ec1 in three different views. The N-terminal and C-terminal halves that are structurally related in the transmembrane domain of CLCs are coloured as light green and dark green, respectively to depict their orientation. (b) and (c) illustrates the structure of CLC-ec1 in homodimeric conformation as cartoon and molecular surface model, respectively. One subunit is represented in green colour and the other in orange with its terminal ends marked by ('). The conserved residues that form ion conduction pathway are highlighted in blue in the cartoon model. The membrane surfaces are highlighted in grey in surface model. The structure is obtained from and analyzed in Protein Data Bank (PDB ID: 1KPK).

In CLC exchangers, two conserved glutamates have been identified to be essential for their exchanger function. One on the extracellular side (E148 in CLC-ec1), which when protonated is displaced away to permit Cl^- transport and is called the gating glutamate (Dutzler et al, 2003). The other resides on the intracellular surface (E203 in CLC-ec1) that exchanges the proton with gating glutamate, called proton glutamate (Accardi et al, 2005). Substitution mutation of the gating glutamate abolished H^+ but not Cl^- transport in *E. coli*, mammalian, as well as in plant CLC transporters (Accardi and Miller, 2004; Scheel et al, 2005; Bergsdorf et al, 2009; Leisle et al, 2011). Similarly, substitution mutation of proton glutamate eliminated H^+ transport with a reduced Cl^- conduction in CLC-ec1 (Accardi et al, 2005). Whereas in mammalian and plant CLC transporters, mutation of proton glutamate abolished both Cl^- and H^+ transport, however an additional mutation of the gating glutamate restored their Cl^- transport function (Zdebik et al, 2008; Bergsdorf et al, 2009; Leisle et al, 2011). Nevertheless, exceptions may occur at the conserved proton glutamate, which is replaced by threonine in CmCLC from *C. merolae* (Feng et al, 2010), isoleucine in CLC-ck2 from *Citrobacter koseri* (Phillips et al, 2012), and valine (or rarely, isoleucine) in bacterial CLC-Fs (Stockbridge et al, 2012), yet all three functions as CLC exchangers.

In CLC channels, valine occupies the residue corresponding to proton glutamate (Accardi et al, 2005), however, its substitution with glutamate does not convert CLC channels to exchangers (Zdebik et al, 2008). Whereas the gating glutamate is present even in some CLC channel proteins, and its substitution results in a constitutively open type of channel (Dutzler et al, 2003; Niemeyer et al, 2003). As in CLC transporters, the gating glutamate present in CLC channels is also suggested to permit Cl^- transport on protonation of its side chain. Thus, CLC channels have also been proposed to act as ‘broken’ Cl^-/H^+ exchangers that have become leaky for Cl^- . Besides, both CLC channels and exchangers are suggested to possess evolutionary conserved physically distinct routes for Cl^- and H^+ transport (Lísal and Maduke, 2008; Leisle et al, 2020). Interestingly, the structure of mammalian CLC channels, CLC-K (Park et al, 2017) and CLC-1 (Park and MacKinnon, 2018; Wang et al, 2019) exhibited high similarity to that of CLC transporters, with slight differences in the residues forming an ion-conducting pore. In CLC-K (from bovine), the loop connecting α -helices C and D is displaced, directing the serine residue (S121) of the ion-conducting pore towards the cytosol instead, which is suggested to remove the kinetic barrier in Cl^- conduction pathway and widen the pore (Park et al, 2017). Similarly, in human CLC-1, the presence of amino acids with smaller hydrophilic side chains towards the Cl^- conducting pore, and shift in the position of ion binding residue Y578, is

proposed to form a wider pore size for channel function (Park and MacKinnon, 2018; Wang et al, 2019). Accordingly, altering the residues in CLC-ec1 transporter, corresponding to the ones observed to widen the pore in CLC-K and CLC-1, led to a channel like activity of CLC-ec1 with remarkably increased Cl^- transport (Park and MacKinnon, 2018), supporting the notion that CLC channels indeed are a ‘broken’ form of Cl^-/H^+ antiporters.

As *Arabidopsis* AtCLC-a is found to function as NO_3^-/H^+ exchanger (De Angeli et al, 2006), its sequence comparison with the Cl^-/H^+ exchangers (human hCLC5 and CLC-ec1) as well as Cl^- channel CLC-0 led to the identification of ‘Proline’ (P160) in AtCLC-a instead of conserved ‘Serine’ residue in Cl^- selective CLCs (S168 in hCLC-5) in the conserved ‘GSGIP’ ion binding motif (Zifarelli and Pusch, 2009). Substitution of this proline to serine (P160S) in AtCLC-a resulted in its higher conductance to Cl^- over NO_3^- , with both ions efficiently coupled to proton transport (Bergsdorf et al, 2009; Wege et al, 2010). Likewise, the substitution of conserved serine to proline resulted in the transformation of CLC-ec1, CLC-5, and CLC-7 from Cl^-/H^+ to NO_3^-/H^+ exchanger, and increased the NO_3^- conductance in CLC-0 channel (Bergsdorf et al, 2009; Picollo et al, 2009; Zifarelli and Pusch, 2009; Leisle et al, 2011), signifying the role of these two residues in anion selectivity in both CLC transporters and channels. The sequence analysis of plant CLCs identified at least one CLC in each seed-producing plant examined to possess proline in the selective filter, but not in the CLCs from green algae, bryophyte, and lycophyte. Thus, NO_3^- -specific CLCs are suggested to have evolved in higher plants (Wege et al, 2010). Interestingly, in CLC-Fs, the equivalent position of ‘Serine/Proline’ is occupied by methionine or asparagine, which have been proposed to account for F^- selectivity in these proteins (Brammer et al, 2014).

The C-terminal cytosolic domains of CLCs containing two CBS domains also share high structural conservation among both CLC channels and exchangers. Like in CBS domains from various CDCPs, two CBS domains of CLC associate into a Bateman module. However, homo-dimerization of Bateman modules from two CLCs involves major interactions between their respective CBS2 domain, resulting in a ‘V-shaped’ structure instead of mostly observed ‘disk-shaped’ structure (Meyer and Dutzler, 2006; Markovic and Dutzler, 2007; Meyer et al, 2007; Feng et al, 2010). Besides, the CBS2 domain also makes extensive contact with the transmembrane helices that extend into the ion conduction pathway (Feng et al, 2010; Wang et al, 2019). Consistently, many disease-causing mutations have been observed to be linked to the CBS-transmembrane interface region of CLCs. Therefore, CBS domains are proposed to influence ion conduction as well as the interactions between two transmembrane subunits of

CLCs (Feng et al, 2010). Each Bateman module of CLC has been identified to bind one ATP molecule (Meyer et al, 2007; Schrecker et al, 2020; Fig. 2.5). Schrecker et al (2020) also identified binding of phosphoinositol-3-phosphate (PI3P, a phosphatidylinositol lipid species predominant in endolysosomal membranes) at the interface between the transmembrane domain and the cytosolic domain in ggCLC-7 (Fig. 2.5). The binding of such lipid molecules has been anticipated to modulate the ion transport function of CLCs (Schrecker et al, 2020).

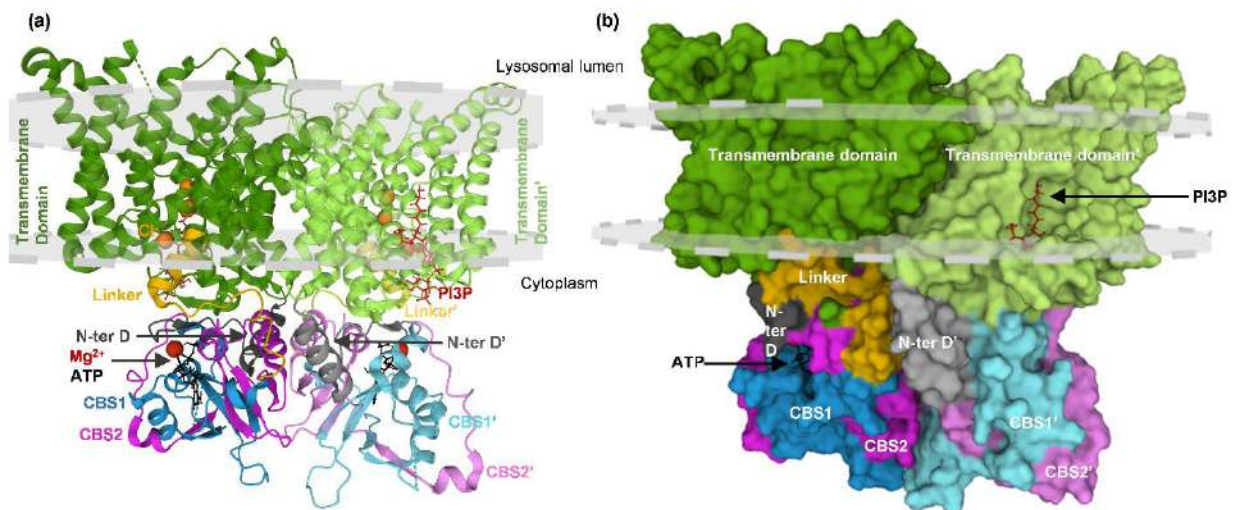


Fig. 2.5: Structure of chicken CLC-7 (ggCLC-7) bound to phosphoinositol-3-phosphate (PI3P), ATP and Mg^{2+} . (a) Cartoon illustration of the structure of ggCLC-7 homodimer (a view from within the membrane). Different regions are highlighted with different colours, whereas same colour but in contrasting shades are used for depicting corresponding domains in two subunits. The colour for one subunit (positioned on left side) is kept deeper while in other subunit (on the right side), the same colour is adjusted to lighter scale to distinguish the domains from two subunits. The N-terminal domain (N-ter D) is represented in grey, transmembrane domain in green, linker in yellow, first CBS domain (CBS1) in blue, second CBS domain (CBS2) in violet. In each subunit, three chloride ions are shown as orange coloured sphere, magnesium ion as red sphere, ATP and PI3P are shown as ball and stick model in black and red colours, respectively. The regions or domains in subunit on the right side are marked with ('). (b) Surface representation of the structure of ggCLC-7 homodimer. Different regions in two subunits are depicted using the colour combinations corresponding to its illustration in cartoon model (a). ATP bound to one subunit (left) and PI3P bound to other subunit (right) are shown by arrows. The membrane surfaces are highlighted in grey facing lysosomal lumen by the upper surface and cytoplasmic region by the lower surface. The structure is obtained from and analyzed in Protein Data Bank (PDB ID: 7JM6).

2.3.2 Functional features of mammalian CLCs

Based on sequence identity, nine CLCs in mammals have been divided into three groups. The first group includes CLC-1, 2, Ka/K1, and Kb/K2, which functions as Cl^- channels and are plasma membrane localized. CLC-3, 4, and 5 constitute the second group, whereas CLC-6 and 7 make the third group. The CLCs from the latter two groups (CLC-3 to 7) function as Cl^-/H^+ -exchangers and are localized on intracellular compartments, mainly in the endosomal/lysosomal pathway (Jentsch, 2015). Altogether, these mammalian CLCs have been identified to be involved in a diverse range of functions, and the loss of function of each one

of them is associated with severe disorders or growth defects (Jentsch, 2015; Jentsch and Pusch, 2018).

CLC-1 is expressed predominantly in the skeletal muscles, where it stabilizes the membrane voltage and repolarizes the membrane during action potentials (Steinmeyer et al, 1991). CLC-1 is negatively regulated by binding of ATP/NAD⁺ to its CBS domain in a redox-dependent manner and by phosphorylation with protein kinase C at T891, S892, and T893 residues (Pedersen et al, 2016; Altamura et al, 2020). CLC-2 has a wide tissue expression, and its activity is also inhibited by ATP (Stölting et al, 2013). Notably, for its membrane localization and transport function, its association with MLC1 (an integral membrane protein) and GLIALCAM (an Ig-like cell-adhesion molecule) has been identified to be essential (López-Hernández et al, 2011a; López-Hernández et al, 2011b; Jeworutzki et al, 2012; Hoegg-Beiler et al, 2014). CLC-Ka and CLC-Kb (CLC-K1 and CLC-K2, respective orthologs in rodents) are very closely related CLCs. CLC-K1 is specifically expressed in the thin limb of Henle's loop and are associated with nephrogenic diabetes insipidus (Uchida et al, 1995; Matsumura et al, 1999), while CLC-K2 is expressed in the basolateral membranes of the thick ascending limb of Henle's loop and distal tubules and are linked to type III-Bartter's syndrome (Kobayashi et al, 2001; Simon et al, 1997). Both CLC-Ka and CLC-Kb form complex with barttin, and this association is found to be requisite for their membrane insertion and transport function (Waldegger et al, 2002; Scholl et al, 2006). CLC-K/barttin performs Cl⁻ reabsorption in the kidney and K⁺ secretion by recycling Cl⁻ in the inner ear (stria) (Estévez et al, 2001). Interestingly, association with barttin is also identified to be required for intracellular processing and targeting of CLC-5 (Wojciechowski et al, 2018).

CLC-3 and CLC-4 exhibit wide expression (Borsani et al 1995; Kawasaki et al, 1994), (van Slegtenhorst et al, 1994), whereas CLC-5 is highly expressed in kidney and epithelial intestinal cells (Vandewalle et al, 2001; Steinmeyer et al, 1995). These three CLCs showed high co-localization in transfected cells, and the co-immunoprecipitation analysis indicated that they form hetero-oligomers (Suzuki et al, 2006). Moreover, CLC-3 and CLC-4 are found to interact *in vivo* whereby CLC-3 facilitated the export of CLC-4 from ER and its stability (Weinert et al, 2020). Based on this interaction and their expression patterns, CLC-3/CLC-4 are suggested to function as heterodimers in the brain, while CLC-3 functions as homodimers in other tissues (Bose et al, 2021). Exchanger activity of CLC-3, CLC-4, and CLC-5 are shown to be regulated by adenine nucleotide in the presence of Mg²⁺. Cytosolic ATP and ADP increased their ion transport activity, whereas AMP decreased their function (Grieschat et al,

2020). In CLC-5, disruption of the ATP binding by substitution mutation (Y617A) impaired its processing and trafficking (Wellhauser et al, 2011).

CLC-6 and CLC-7 also have wide expression patterns (Brandt and Jentsch, 1995), but CLC-6 is most predominantly expressed in the neurons and is localized in late endosomes (Poët et al, 2006). Recently CLC-6 has also been reported to localize in Golgi and its loss of function exhibited a reduction of Golgi calcium stores (Klemens et al, 2021). CLC-7 forms a complex with Ostm1 (a highly glycosylated type I single-pass membrane protein) and co-localizes to late endosome/lysosome (Lange et al, 2006). This complex formation is found to be important for the stability and Cl^-/H^+ exchange activity of CLC-7 (Lange et al, 2006; Leisle et al, 2011).

2.3.3 CLCs from plants

In plants, the pioneering work on CLCs was carried out independently by Lurin et al (1996) in tobacco and Hechenberger et al (1996) in *Arabidopsis*. Lurin et al (1996) isolated *NtCLC-1* that elicited anion-selective conductance when expressed in *Xenopus* oocytes. It exhibited the highest expression in the stem (Lurin et al, 1996) and is localized in the ER (Sun et al, 2018). Hechenberger et al (1996) found AtCLC-d among four *Arabidopsis* CLCs (AtCLC-a – d) to be able to rescue the iron-limited growth phenotype of yeast *Agef1* mutant. This functional complementation by AtCLC-d was impeded by substitution mutation corresponding to disease-linked mutation in human CLCs, suggesting its function in Cl^- transport (Hechenberger et al, 1996).

The seven CLCs present in *Arabidopsis* (AtCLC-a to -g) form two major phylogenetic groups: AtCLC-a to -d and -g, form group I, while AtCLC-e and AtCLC-f, sharing higher identity with the prokaryotic CLCs, form group II (Lv et al, 2009). Based on the sequence homology, 10 genes encoding CLCs have been identified in *Oryza sativa* (Kushwaha et al, 2009), 17 in *Nicotiana tabacum* (Zhang et al, 2018), 8 in *Camellia sinensis* (Xing et al, 2020), 11 CLCs in *Brassica rapa*, 10 in *B. oleracea*, and 22 in *B. napus* (Liao et al, 2018), 7 in *Suaeda fruticosa* (Nedelyaeva et al, 2018) and 8 in *Glycine max* (Wei et al, 2019). Like in AtCLCs, most of the CLCs from these plants formed a major phylogenetic group (group I), while few CLCs formed a separate group (group II) that are more related to bacterial CLCs. These two groups of plant CLCs are suggested to have originated from two distinct ancestral genes (Lv et al, 2009). Altogether, the plant CLCs are involved in several functions ranging from nitrate uptake and assimilation to stomatal regulation, Cl^- sequestration, and plant-pathogen interactions, which are presented in the following sub-sections.

2.3.3.1 Plant CLCs involved in nitrate transport and storage

Nitrate (NO_3^-) is among the primary forms of mineral nitrogen available to plants. It is either stored in the vacuole or reduced to nitrite (NO_2^-) and then to ammonium (NH_4^+) in the cytosol for the subsequent assimilation of nitrogen into organic compounds. Besides, NO_3^- also acts as a signaling molecule in regulating plant metabolism and development (Gojon et al, 2011). In *Arabidopsis*, AtCLC-a and AtCLC-b share high sequence identity, both are tonoplast localized, and function as NO_3^-/H^+ exchangers (De Angeli et al, 2006; von der Fecht-Bartenbach et al, 2010). AtCLC-a is critical for nitrate transport into the vacuole. Its transcript level is upregulated in response to NO_3^- application in both roots and shoots. Accordingly, its mutant, *clc-a*, showed a significant reduction in NO_3^- content in both roots and shoots under NO_3^- treatment and exhibited more sensitivity to the herbicide sodium chlorate, an analog of NO_3^- (Geelen et al, 2000). Consistent with this plant phenotype, De Angeli et al (2006) subsequently confirmed AtCLC-a as a NO_3^-/H^+ (2:1) exchanger by a patch-clamp method in the vacuole isolated from mesophyll cell. They observed a significant reduction in the activity of AtCLC-a by ATP, while AMP suppressed the ATP-mediated inhibition, thus suggesting the regulation of AtCLC-a mediated NO_3^- transport by cellular ATP:AMP ratio. The substitution mutation in the CBS domains of AtCLC-a (D753A), corresponding to the ATP binding residue of hCLC-5 (Meyer et al, 2007), eliminated the inhibitory effect of ATP on its activity, confirming ATP regulation via binding to the CBS domains of AtCLC-a (De Angeli et al, 2009).

AtCLC-a showed strong expression in the guard cells, and its loss of function mutant lines (*clc-a*) exhibited a reduced stomatal aperture in response to light illumination, and a defect in ABA-responsive stomatal closing (Wege et al, 2014). Based on the former phenotype, they proposed the role of AtCLC-a in NO_3^- influx into the vacuole in response to light for stomatal opening. While the latter phenotype indicated a probable role of AtCLC-a in NO_3^- efflux from the vacuole as well in response to ABA for stomatal closing. Consistently, they identified AtOST1 (OPEN STOMATA 1 / SnRK2.6), an SnRK2 family kinase that phosphorylates and subsequently activates different plasma membrane anion channels (Geiger et al, 2009; Imes et al, 2013), to phosphorylate AtCLC-a at T38 residue which induced NO_3^- efflux from the vacuole (Wege et al, 2014). Accordingly, the C5 catalytic subunit of protein phosphatase 2A (PP2A-C5) has been found to interact with all tonoplast localized AtCLCs. It is proposed that PP2A-C5 dephosphorylates these AtCLCs, thereby facilitating anions influx into the vacuole (Hu et al, 2017). Furthermore, Carpaneto et al (2017) observed inhibition of AtCLC-a activity

upon treatment with phosphatidylinositol-3,5-bisphosphate (PI(3,5)P2), a signaling lipid molecule that induces vacuolar convulsion and acidification of the lumen, important during rapid ABA-induced stomatal closure (Bak et al, 2013). The PI(3,5)P2 treatment decreased the vacuolar pH in WT but not in *clc-a* mutants. Hence, Carpaneto et al (2017) proposed that inactivation of CLC-a by PI(3,5)P2 leads to acidification of the vacuolar lumen and favors stomatal closing.

AtCLC-b, with proline residue in the selectivity filter, is another electro-physiologically characterized NO_3^-/H^+ exchanger. However, unlike *clc-a*, no reduction in NO_3^- content is observed in *clc-b* mutants (von der Fecht-Bartenbach et al, 2010). As NO_3^- is assimilated by nitrate reductases in the cytosol during the day, while stored in the vacuole at night (Cookson et al, 2005), von der FechtBartenbach et al. (2010) observed diurnal variations in the expression of *AtCLC-a* and *AtCLC-b* from the gene expression database. *AtCLC-b* showed peak expression at the beginning of the light period, while *AtCLC-a* reached maximum expression during the dark. Based on gene expression data and electrophysiological results, they suggested that AtCLC-b mediates the NO_3^- release from vacuoles during the day, which is then assimilated in the cytoplasm, while AtCLC-a stores NO_3^- into the vacuole at night.

In *Brassica napus*, the genes encoding BnaCLC-a members (BnaCLCa1-4, that share high sequence similarity with AtCLC-a) also showed strong induction in both shoots and roots upon NO_3^- replenishment following N starvation. BnaCLC-b was induced in the shoots, while BnaCLC-c to -g members were either downregulated or showed no changes to NO_3^- treatment (Liao et al, 2018). Even in *Suaeda altissima*, a euhalophyte, the expression of *SaCLC-a1* (shares high sequence similarity with AtCLC-a and AtCLC-b) increased in the roots under NO_3^- deficiency (Nedelyaeva et al, 2019). *SaCLC-a1* possesses both conserved gating glutamate and proton glutamates, and proline in the selectivity filter, thus are proposed to function as NO_3^-/H^+ exchanger. Besides, as observed for AtCLC-a (Wege et al, 2010), *SaCLC-a1* failed to complement *Agef1* mutant, however substituting proline with serine (P188S) in the selectivity filter conferred partial growth restoration of *Agef1* mutant (Nedelyaeva et al, 2019), implying a conserved NO_3^-/H^+ exchanger function in *SaCLC-a1*.

2.3.3.2 Plant CLCs with canonical chloride transport function

In *Arabidopsis*, AtCLC-c, AtCLC-d, and AtCLC-g possess serine in selectivity filter. AtCLC-c and AtCLC-d have both gating and proton glutamates, indicating their function as Cl^-/H^+ transporter, while the gating glutamate in AtCLC-g is replaced by alanine, denoting it as a

probable Cl^- channel. AtCLC-c and AtCLC-g share high sequence identity, and both are localized in the tonoplast (Jossier et al, 2010; Nguyen et al, 2016), whereas AtCLC-d is localized in the trans-Golgi network (von der Fecht-Bartenbach et al, 2007). Upon NaCl treatment, expression of AtCLC-d and AtCLC-g are reported to be induced in the shoots, while AtCLC-c is induced in both shoots and roots (Jossier et al, 2010).

AtCLC-c is identified to locate in the major QTL for nitrate accumulation in *Arabidopsis*. It is observed to be down-regulated upon nitrate treatment, while its mutant, *clc-c*, exhibited reduced NO_3^- , Cl^- and citrate content (Harada et al, 2004). However, Jossier et al (2010) observed no difference in NO_3^- level between *clc-c* mutant and wild type (WT) plants, contrasting to the observation by Harada et al (2004), which they reasoned to be due to different accessions used in these two studies. AtCLC-c showed high expression in guard cells and pollens, and its expression in guard cells increased furthermore by ABA treatment; however, its expression is observed to be weak in roots (Jossier et al, 2010). The *clc-c* mutants, like *clc-a* (Wege et al, 2014), exhibited impaired stomatal opening to light and insensitivity to ABA-induced stomatal closing. This defect in stomatal regulation is rescued by the application of NO_3^- , but not Cl^- . Also, Cl^- content is found to be significantly less in the guard cells of *clc-c* mutants (Jossier et al, 2010). Thus, they proposed that AtCLC-c confers Cl^- influx into the vacuole and regulates the stomatal opening and that impaired Cl^- accumulation in the vacuole of *clc-c* mutants affected its efflux from the vacuole, subsequently affecting the stomatal closing as well. Besides, *clc-c* mutant is also reported to exhibit strong hypersensitivity to Cl^- , exhibiting a reduction in the shoot as well as root biomass (Jossier et al, 2010), while the overexpression of *AtCLC-c* conferred salt tolerance to the plant (Hu et al, 2017). AtCLC-c (cytosolic N- and C-terminal) is identified to interact with a C5 catalytic subunit of the protein phosphatase 2A (PP2A-C5), the overexpression of which also improved salt tolerance, while the mutant (*pp2a-c5*) showed a salt-sensitive phenotype. PP2A-C5 is suggested to function upstream of AtCLC-c since the overexpression of AtCLC-c in *pp2a-c5* mutant could not rescue its salt-sensitive phenotype. Besides, PP2A-C5 also showed interactions with the other three tonoplast-localized AtCLCs (AtCLC-a, -b, and -g) and are proposed to activate the vacuolar CLCs by dephosphorylation (Hu et al, 2017).

Like *AtCLC-c*, the *AtCLC-g* mutant (*clc-g*) also showed Cl^- sensitive growth retardation. The expression of AtCLC-g is observed to be higher in the mesophyll cells while AtCLC-c expression is prominent in the guard cells (Jossier et al, 2010, Nguyen et al, 2016). Both *clc-c* and *clc-g* mutants are reported to accumulate higher Cl^- in the shoots upon NaCl

treatment (Nguyen et al, 2016), contrasting to the observation by Jossier et al (2010) who reported no difference in Cl^- accumulation in the shoots from *clc-c* and WT plants under NaCl treatment. The double mutant plants (*clc-c:clc-g*) exhibited no difference from the single mutant of each gene to NaCl treatment (Nguyen et al, 2016). Contrary to the observation by Jossier et al (2010), *AtCLC-g* expression is observed to be repressed in the *clc-c* mutant background in both controlled and NaCl-treated conditions. *AtCLC-c* expression also decreased in *clc-g* mutants but only in NaCl-treated conditions. Thus, Nguyen et al (2016) suggested that *AtCLC-g* functions in Cl^- sequestration into the vacuole of mesophyll cells and its gene expression to coordinate with the expression of *AtCLC-c* in guard cells for stomatal regulation. In mature leaves, *AtCLC-g* expression is observed in phloem cells and hydathodes, and it is hypothesized to prevent Cl^- overaccumulation in photosynthetically active leaves by partitioning the Cl^- into different organs via the phloem and by excreting it through hydathodes, respectively (Nguyen et al, 2016).

SaCLC-c1, a homolog of *AtCLC-c* in *S. altissima*, is observed to be induced in leaves by higher Cl^- treatment. *SaCLC-c1* possesses both gating and proton glutamates and serine in the selectivity filter, indicating its role as Cl^-/H^+ exchanger (Nedelyaeva et al, 2019). When expressed in Δgef1 yeast mutant, *SaCLC-c1* complemented the iron deficient, pH-sensitive, and salt-sensitive growth phenotype of the mutant (Nedelyaeva et al, 2018, 2019); however, *AtCLC-c* failed to complement Δgef1 (Hechenberger et al, 1996). In rice, *OsCLC-1* and *OsCLC-2*, which share high sequence similarity with *AtCLC-c*, showed localization in the tonoplast. Both partially rescued the Δgef1 yeast mutant on various salts and higher pH, indicating their possible role in Cl^- transport (Nakamura et al, 2006). The expression of *OsCLC-1* is found to be induced by NaCl treatment (Diédhiou and Golldack, 2006; Nakamura et al, 2006). This induction is subsequently observed to reduce in the shoots, however, at a slow rate in Pokkali (salt-tolerant) than in IR29 (salt-sensitive). In roots, its expression under NaCl stress decreased in IR29, whereas it increased in Pokkali (Diédhiou and Golldack, 2006). Although no difference was observed in normal conditions, under NaCl treatment, Pokkali shoots accumulated significantly less Cl^- than in IR-29 (Diédhiou and Golldack, 2006). The Tos17 insertional mutant lines of both genes, *osclc-1* and *osclc-2* exhibited a reduction in height at all plant stages (Nakamura et al, 2006).

In *Nicotiana tabacum* 17 genes encoding CLCs have been identified of which *NtCLC-2*, *NtCLC-3*, and *NtCLC-12* showed induction in both root and shoot, while *NtCLC-3*, *NtCLC-13*, and *NtCLC-15* showed induction only in shoot upon NaCl treatment. Accordingly, *NtCLC-*

2- silenced plants showed reduced Cl^- content, but it was increased in *NtCLC-13*-silenced plants (Zhang et al, 2018). *NtCLC-1* (*AtCLC-c* homolog) is proposed to act as a Cl^-/H^+ exchanger based on conserved residues in the selectivity filter. Double knockout lines of two *NtCLC-1* homologs, *NbCLC-1a:NbCLC-1b*, in *N. benthamiana*, resulted in a slight reduction in plant height and smaller flowers, which was fully rescued by the expression of either *NtCLC-1*, *NbCLC-1a*, or *NbCLC-1b*; however, not by *NtCLC-1* possessing mutations at either gating or proton glutamates, or at both of these glutamates (Sun et al, 2018). *NtCLC-1* is identified to localize in the endoplasmic reticulum (ER), and in *nbcl-1a: nbcl-1b* knockout lines, the pH of ER lumen was slightly lower (7.05 ± 0.13) than that of WT (7.37 ± 0.14), which was restored by *NtCLC-1* (Sun et al, 2018). *CsCLC-c* from trifoliate orange (*Poncirus trifoliata*; shares high identity with *AtCLC-c*) also exhibited gene induction to NaCl as well as to ABA and cold treatment. Its overexpression in *Arabidopsis AtCLC-c* mutant (*clc-c*) improved seedling growth better than WT under salinity stress, with reduced Cl^- accumulation in both roots and shoots of the overexpressing lines (Wei et al, 2013).

In cultivated soybean (*Glycine max*) and its salt-tolerant wild relative (*G. soja*), *GmCLC-1* and *GsCLC-c2*, respectively, both localized in the tonoplast, have been reported to function in Cl^- transport (Li et al, 2006; Wei et al, 2019). *GmCLC-1* expression is observed to be induced in leaf by NaCl, polyethylene glycol (PEG), cold and ABA treatments (Li et al, 2006; Zhou and Qiu, 2010) and induced in roots by Cl^- treatment (Wong et al, 2013). *GmCLC-1* possesses both gating and proton glutamates and exhibited a pH-dependent Cl^- conductance (higher Cl^- current at pH 7.5 than at pH 5.5) in *Xenopus* oocytes (Wong et al, 2013). It enhanced the growth of *Agefl* yeast mutant on different Cl^- salts (Wei et al, 2016) and its overexpression in tobacco BY-2 cell lines conferred tolerance against NaCl treatment with higher Cl^- accumulation into the vacuole (Li et al, 2006). Its overexpression also conferred salt tolerance in soybean hairy roots as well as in *Arabidopsis* and poplar, by reducing Cl^- accumulation in the shoots through its compartmentation into the vacuole in roots (Zhou and Qiu, 2010; Sun et al, 2013; Wei et al, 2016). However, *GmCLC-1* possesses a proline residue in the anion selectivity filter and exhibited NO_3^- conductance equivalent to that of Cl^- on *Xenopus* oocytes (Wei et al, 2019), which implies its role as a NO_3^-/H^+ exchanger. In the case of *GsCLC-c2* from *G. soja* (homolog of *AtCLC-c* and *AtCLC-g*), the sequence comparison with its ortholog from *G. max* identified the occurrence of a few substitution mutations. The orthologs are also expressed differentially in two species. *GsCLC-c2* possesses a serine and both gating and proton glutamates in the selectivity filter, indicating it as a probable Cl^-/H^+ exchanger. It

showed higher Cl^- conductance over NO_3^- in *Xenopus* oocytes and rescued the growth of *Δgefl* yeast mutant under Cl^- stress conditions, further indicating its function as a Cl^-/H^+ exchanger. Its overexpression conferred salt tolerance to plants by sequestering Cl^- into the vacuole in roots, thus lowering the Cl^- level in the shoots (Wei et al, 2019). In maize, *ZmCLC-c* (*AtCLC-c* homolog) is observed to be induced by cold. Distinguished from all other CLCs, *ZmCLC-c* has been reported to localize in the mitochondria where anti-*ZmCLC-c* antibody treatment is observed to inhibit Cl^- and I^- fluxes. It is suggested to be a plant inner membrane anion channel (PIMAC; Tampieri et al, 2011).

AtCLC-d is localized in trans-Golgi and exhibits a wide expression pattern, but relatively higher in seedling roots (von der Fecht-Bartenbach et al, 2007). It fully complemented the iron-sensitive growth phenotype of *Δgefl* yeast mutant (Hechenberger et al, 1996). Its mutant, *clc-d*, showed a defect in root growth, but not in anion content in the tissue (von der Fecht-Bartenbach et al, 2007). As shorter root length and reduction in cell elongation is also caused by partial loss of another trans-Golgi localized proton transporting V-ATPase activity (Schumacher et al, 1999; Padmanaban et al, 2004), von der Fecht-Bartenbach et al (2007) treated the plants with concanamycin A (V-ATPase inhibitor). Interestingly, this treatment resulted in shorter hypocotyls in *clc-d* mutants as compared to WT. Thus, *AtCLC-d* is suggested to maintain V-ATPase-mediated acidification of the trans-Golgi network (von der Fecht-Bartenbach et al, 2007). Its homolog in maize, *ZmCLC-d*, is reported to be induced by cold, drought, heat, NaCl, ABA, and H_2O_2 treatments, and its overexpression in *Arabidopsis* enhanced the tolerance of plants towards cold, drought, and salt stresses (Yang et al, 2011; Wang et al, 2015).

2.3.3.3 Phylogenetically distinct members of plant CLCs

Members of plant CLCs that are homologs of *Arabidopsis* *AtCLC-e* and *AtCLC-f* form a separate phylogenetic group with a closer identity to prokaryotic CLCs. Both *AtCLC-e* and *AtCLC-f* possess conserved gating glutamate, but the proton glutamate in *AtCLC-e* and *AtCLC-f* is replaced by similar yet different residues, serine, and threonine, respectively. Notably, the proton glutamate in *CmCLC* (Cl^-/H^+ exchanger) from a thermophilic red alga (*C. merolae*) is also substituted by threonine (Feng et al, 2010), and the higher identity of these two *AtCLCs* with bacterial CLCs implies their exchanger function. *AtCLC-e* is localized in the thylakoid membrane in chloroplasts (Marmagne et al, 2007). Its mutant, *clc-e*, exhibited an altered chlorophyll fluorescence from the dark-adapted leaves when illuminated, which is

proposed to be due to altered ion homeostasis in the lumen of the thylakoid (Marmagne et al, 2007; Herdean et al, 2016). The dark-adapted mutants also exhibited a defect in the thylakoid network and a large stromal region in the chloroplast, while no such differences were observed in mutants and WT chloroplasts from light-adapted plants. Pre-treatment with KCl but not KNO₃ restored Chl-a fluorescence kinetics in mutants. Thus, Herdean et al (2016) proposed a probable function of AtCLC-e in Cl⁻ efflux from thylakoid to the stroma of the chloroplast after a transition from light to dark. Contrastingly, Monachello et al (2009) suggested the role of AtCLC-e in NO₃⁻ transport into thylakoid lumen with the observation of reduced NO₃⁻ accumulation in *clc-e* mutants. They observed a reduction in net NO₃⁻ influx in roots, whereas the NO₂⁻ level was elevated in *clc-e* as well as in *clc-a* (NO₃⁻/H⁺ exchanger) and their double mutant *clc-a:clc-e*. The expression of NO₃⁻ transporter genes, *AtNRT2.1* and *AtNRT1.1*, downregulated in the roots of both *clc-a* and *clc-e* mutants, which was suggested to be due to an increase in NO₂⁻ level (Monachello et al, 2009). However, Herdean et al (2016) inferred the study by Monachello et al (2009) as an indirect effect resulting from possible alteration in the expression of NO₃⁻/H⁺ exchangers and instead emphasized the role of AtCLC-e in Cl⁻ transport.

AtCLC-f is localized in Golgi vesicles, and it complemented the iron as well as pH-sensitive growth phenotype of *Agefl* yeast mutant (Marmagne et al, 2007). Its ortholog in spinach is reported to localize in the outer membrane of the chloroplast and its gene is observed to express only in shoots, but not in roots. The treatment with p-chloro-phenoxy-acetic acid (CLC inhibitor) impaired photosynthetic activity as visible by a reduction in oxygen evolution and altered chlorophyll-a fluorescence in the isolated chloroplast and intact plant, respectively (Teardo et al, 2005).

2.3.3.4 Plant CLCs in pathogen interaction

Plant CLCs have also been found to be involved in plant-pathogen interactions. Among six *AtCLC* mutants (except for *clc-f*), Guo et al. (2014) identified *clc-d* to have an enhanced response to flg22, a conserved N-terminal region of flagellin which act as an elicitor of defense response (Felix et al, 1999). Accordingly, *AtCLC-d* overexpressing lines were less responsive to flg22, indicating that AtCLC-d negatively regulates the pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) (Guo et al, 2014). When inoculated with *Pseudomonas syringae* pv. tomato DC3000, *clc-d* mutants showed resistance, while *AtCLC-d* overexpressing lines exhibited susceptibility. The expression of *AtCLC-d*, as well as all other

AtCLCs, was found to downregulate by flg22 treatment. While in *fls2* mutant plants [FLS2 is a plant cell surface pattern recognition receptor (PRR) for flg22 that triggers PTI], the expression of *AtCLC-d* was higher but remained unaffected by flg22. In contrast, the expression of *AtCLC-d* was reduced in *FLS2* overexpressing line. These indicated that inhibition of *AtCLC-d* expression requires pre-recognition of flg22 by FLS2 (Guo et al, 2014). Correlating with the role of *AtCLC-d* in the acidification of trans-Golgi network (von der Fecht-Bartenbach et al, 2007), Guo et al (2014) proposed that AtCLC-d probably affects the FLS2 endocytic transport and trafficking via the trans-Golgi network which is critical for its PTI function.

Tobacco NtCLC-1 has been reported to interact and co-localize with the POTATO VIRUS Y (PVY) 6 K2 protein in the ER. This interaction resulted in the alkalinization of ER, which is reported to be necessary for viral intracellular replication and systemic infection (Sun et al, 2018). In tomato, *CLC-b-like* mRNA is identified to be a target of Potato Spindle Tuber Viroid (PSTVd)-derived small RNA (vd-sRNA). The PSTVd infection led to the cleavage of *CLC-b-like* mRNA, significantly reducing its transcript level. Accordingly, the virus-induced gene silencing (VIGS)-mediated *CLC-b-like* mRNA silenced plants exhibited a stunting and leaf curling phenotype like viroid infection (Adkar-Purushothama et al, 2017). Sasvari et al (2013) observed impaired replication of Tomato Bushy Stunt Virus (TBSV) in Δ *gefl* yeast mutant, which was found to be due to higher Cu_2^+ accumulation in the cytosol, an indirect effect of Δ *gefl*, which mainly affected the assembly of viral replicase complex (VRC). Similarly, chemical blocking of CLCs in *Nicotiana benthamiana* using SITS (4-acetamido-4'-isothiocyano-2,2'-disulfonic stilbene) also reduced TBSV RNA accumulation in both protoplasts and leaves of the whole plant, which is suggested to be due to cellular ionic imbalance caused by an alteration in Cl^- homeostasis (Sasvari et al, 2013). The CLCs of plant microbes, both fungi, and bacteria, also appear to be essential for infecting plants, as disruption of their CLCs are reported to cause a significant decrease in their virulence (Rojas-Jiménez et al, 2005; Cañero and Roncero, 2008).

3. Materials and Methods

3.1 Gene expression analysis

3.1.1 Plant growth and treatments for expression analysis

For stress treatments, seeds of Pokkali and IR64 rice genotypes were treated with Bavistin (0.5%) for 20 min followed by thorough washing with water. Seeds were then soaked in water for 24 hr in the dark and germinated on germination paper by incubating in the dark at 28°C for the first 2 days. Once the coleoptile emerged out of the rolls, they were maintained at 28°C with 14 hr light /10 hr dark photoperiod. After the fifth day of germination, water in the container was replaced by Yoshida medium (Appendix 1, Table A1.1). The two-week-old seedlings were subjected to various stress treatments. For salinity (NaCl, 200 mM), oxidative (H₂O₂, 10 mM), biotic (Salicylic acid, 100 µM), and heavy metal (CuCl₂, 100 µM) stress responses, respective salts or reagents were supplemented into the Yoshida medium. For cold and heat stresses, the seedlings were incubated at 4°C and 42°C, respectively. The desiccation stress is imposed by placing the seedlings on a dry 3 mm Whatman paper. The seedlings were also maintained under stress untreated or control conditions as experimental controls. The seedlings were harvested at specific time points of stress treatments, snap-frozen in liquid nitrogen, and stored at -70°C. For gene expression study at tillering and reproductive stages, plants were grown on soil in earthen pots in the green house at 28±2°C temperature and 14 hr light /10 hr dark photoperiod. Different tissues at specific stages were harvested, weighed, snap-frozen in liquid nitrogen, and stored at -70°C.

3.1.2 RNA isolation, UV/Vis spectrophotometry, and denaturing gel electrophoresis

Total RNA was isolated using TRIzol RNA isolation reagent (Invitrogen, 15596018) following the manufacturer's procedure. The tissue sample (100 mg for shoot sample, 200 mg for root sample) was ground to a fine powder in liquid nitrogen using mortar and pestle. TRIzol reagent (1 ml) was added to the sample, mixed well and the homogenized sample was then incubated at room temperature for 5 min in a 1.5 ml microcentrifuge tube. Next, 200 µl of chloroform was added to the homogenate and mixed thoroughly by vortexing for 15 sec. Tubes were then incubated at room temperature for 5 min followed by centrifugation at 12,000 x g for 15 min at 4°C. The upper aqueous phase was carefully pipetted up (avoiding contamination from the interphase) and transferred into a fresh microcentrifuge tube. The aqueous phase was treated one more time with 200 µl of chloroform, followed by centrifugation and transfer to another

microcentrifuge tube. After that, 500 μ l of isopropanol was added onto the aqueous phase, mixed gently, and incubated at room temperature for 10 min to precipitate the RNA. Tubes were then centrifuged at 12,000 $\times g$ for 10 min at 4°C. The supernatant was decanted, and the RNA pellet was washed with 1 ml of 75% ethanol (prepared in DEPC treated nuclease-free water) by vortexing briefly followed by centrifuging at 7500 $\times g$ for 5 min at 4°C. The supernatant was decanted and the ethanol washing step was repeated one more time. After the second ethanol wash, the RNA pellet was air-dried for about 15 min inside the laminar hood and subsequently resuspended in 50 μ l nuclease-free water. The RNA was dissolved completely by incubating at 60°C in a water bath for 10 min. The isolated RNA was stored at 4°C until a fraction of it was used for cDNA preparation which was performed usually on the same day or the next day, after which the RNA was stored at -70 °C.

The concentration and purity of the isolated RNA were determined by measuring the absorbance at 260 nm (A_{260}), 280 nm (A_{280}), and 230 nm (A_{230}) using a UV/Vis spectrophotometer (Ultrospec 2100 pro, Biochrom). An absorbance of 1 at 260 nm corresponds to the concentration of approximately 40 ng/ μ l of RNA, whereas the ratio of A_{260}/A_{280} and A_{260}/A_{230} should be around 2.1 and 2.1-2.3, respectively for pure RNA (Koetsier and Cantor, 2019). The integrity of RNA was analyzed by resolving it on 1.2 % denaturing agarose gel (Appendix 6, A6.1) in 1X MOPS running buffer (Appendix 6, Table A6.1). Before gel electrophoresis, the sample was denatured in 2X RNA loading dye (Thermo Scientific, R0641) by heating at 70°C for 10 min followed by chilling immediately on ice. The presence of 25S and 18S rRNAs as two discrete bands in approximately 2:1 concentration indicates the intactness of isolated RNA.

3.1.3 First-strand cDNA synthesis

First-strand cDNA from the isolated RNA was prepared using RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, K1622) based on the manufacturer's procedure. RNA (1 μ g from each sample) was first treated with 1 μ l of 1 U/ μ l DNase I (Thermo Scientific, EN0521) in a reaction mixture containing 1 μ l of 10X DNase I Buffer and nuclease-free water up to 10 μ l of the reaction volume in 0.2 ml PCR tube for 30 min at 37°C. The DNase I was then inactivated by adding 1 μ l of 50 mM EDTA followed by incubation at 65°C for 10 min. Onto the DNase-treated RNA, 1 μ l of 100 μ M Oligo(dT)₁₈ primer was added and annealed at 65°C for 5 min followed by snap chilling on ice. Then after, 1 μ l of 200 U/ μ l RevertAid M-MuLV Reverse Transcriptase, 1 μ l of 20 U/ μ l RiboLock RNase Inhibitor, 2 μ l of 10 mM dNTP mix,

and 4 μ l of 5X Reaction Buffer was added, making up a final volume of the content to 20 μ l. The reaction mixture was then incubated for 1 h at 45°C for first-strand cDNA synthesis. The reaction was terminated by heating at 70°C for 5 min. The reverse-transcribed cDNA was used in the downstream gene expression analysis by quantitative real-time PCR (qRT-PCR). It was also used for the amplification of genes of interest. For long-term storage, cDNA samples were stored at -70°C.

3.1.4 Gene expression analysis by qRT-PCR

Prior to qRT-PCR, the prepared cDNA samples were first analyzed for their quality as well as the dilutions to be used in the qRT-PCR, by carrying out thermal cycler PCR at different cDNA dilutions using primers specific to housekeeping gene [*Elongation Factor-1 α* (*eEF-1 α*)]. The primers to be used in qRT-PCR were also analyzed for their gene specificity and PCR efficiency by carrying out both thermal cycler and qRT-PCR (for accessing melt curve) followed by gel electrophoresis in 2% agarose gel. These, along with various other critical steps, to be considered to minimize the errors in qRT-PCR are described in detail by Taylor et al (2019).

The qRT-PCR mixture of 10 μ l volume was optimized, which comprised of 4 μ l of the template (20X diluted cDNA, section 3.1.3), 0.5 μ l of each 10 μ M primer, and 5 μ l of 2X PowerUp SYBR Green Master Mix (Applied Biosystems, A25742). The reaction was performed in 7500 Realtime PCR System (Applied Biosystems, 4351104) set with the following conditions: initial 50°C (2 min) and 95°C (2 min) for activation of UDG and Dual-Lock DNA polymerase, respectively, and 40 cycles of 95°C for 15 sec (denaturation) and 60°C for 1 min (annealing/extension), followed by instrument's default dissociation step. The gene expression data was analyzed based on the comparative CT method (Schmittgen and Livak, 2008) using *eEF-1 α* as an internal control.

3.1.5 Gene expression profiling using publicly available expression datasets

The expression profiling of genes based on publicly available expression datasets was performed using the Genevestigator tool (Hruz et al, 2008). The expression data acquisition was applied only from *Oryza sativa* wild-type (WT) genetic background in the data selection filter. The expression level of genes in different tissues at different developmental stages was procured from RNA-seq datasets which were represented in linear scale of expression {arbitrary units (a.u.) in Genevestigator}. The gene expression data under various stress

treatments were procured from both Affymetrix GeneChip and mRNA-seq dataset platforms in the Genevestigator. The p-value of the expression data was adjusted to < 0.05 and the expression unit was set to Log₂ of fold change (Log₂FC). The expression data were arranged in MS excel and presented either as a bar graph or as a heatmap using MultiExperiment Viewer (MeV) 4.9.0 tool (Saeed et al, 2003).

3.2 Agarose gel preparation and gel electrophoresis

Agarose gel (1.0 %) prepared in 1X TAE buffer (Appendix 6, Table A6.2) was used to run DNA samples in the study unless specifically mentioned other alternatives or different agarose concentrations. For 100 ml of agarose gel preparation, 1 g of agarose was dissolved in 100 ml of 1X TAE buffer by heating in a microwave oven for about 2 min. After cooling down the gel to about 55°C, 5 µl of 10 mg/ml EtBr was added. The gel solution was mixed thoroughly by gentle swirling and poured onto the gel casting tray. The DNA agarose gel electrophoresis was carried out in a 1X TAE buffer.

3.3 Glycerol stock preparation

For all the transformed bacterial or yeast clones, glycerol stock was prepared by aliquoting 0.5 ml of their culture grown till stationary phase in 1.5 ml of microcentrifuge tube and adding 0.5 ml of sterile 50% glycerol onto it. The content was mixed by inversion and stored at -70°C as a glycerol stock.

3.4 Gene amplification, cloning, and transformation in *E. coli*

3.4.1. PCR amplification of the gene/insert

The coding sequence (CDS) of the genes or any other inserts to be sequenced or cloned into the vector were PCR amplified using Phusion high-fidelity DNA polymerase (Thermo Scientific, F530S). A reaction mixture comprised of 5 µl of 5X Phusion buffer, 0.5 µl of 10 mM dNTPs, 0.5 µl of each 10 µM primer, 0.25 µl of 2 U/µL Phusion DNA polymerase, and DNA template {4 µl of 20X diluted cDNA (section 3.1.3) or 1 ng of plasmid} and sterile water up to 25 µl reaction volume. The PCR was performed following the conditions specific to Phusion DNA polymerase: initial denaturation at 98°C for 30 sec, 35 cycles of denaturation (98°C for 10 sec), annealing (temperature specific to a given primer set for 15 sec), and extension (72°C, time-based on product length, 30 sec/kb), followed by a final extension at

72°C for 10 min. The PCR product was resolved in an agarose gel, followed by gel extraction and purification.

3.4.2. Extraction and purification of DNA fragments from agarose gel

GeneJET gel extraction kit (Thermo Scientific, K0692) was used for the extraction and purification of DNA fragments from agarose gel slices as per the manufacturer's protocol. Gel slice containing the DNA fragment of interest (PCR product or vector digested with restriction enzymes) was excised from the agarose gel using a clean scalpel, transferred into a pre-weighed microcentrifuge tube, and the weight of the gel slice was determined. Binding buffer at the ratio of 1:1 (volume: weight of the gel slice) was added onto it (for gel slice with an agarose content greater than 2%, 2:1 volume of binding buffer was used). The gel slice was dissolved completely by incubating the content at 60°C in a water bath for 10-15 min with intermittent mixing by vortexing. The solubilized gel mixture was then transferred into a GeneJET purification column and centrifuged at 12,000 rpm for 1 min. The flow-through in the collection tube was discarded and 700 µl of Wash buffer was added into the column. The column was centrifuged at 12,000 rpm for 1 min and the flow-through was discarded. The empty column was centrifuged again at 12,000 rpm for 1 min to completely remove the residual Wash buffer. The column with its cap left open was placed onto a fresh 1.5 ml microcentrifuge tube for about 5 min for the residual ethanol from Wash buffer to evaporate completely. Finally, 40 µl of Elution buffer was added to the center of the purification column membrane, allowed to stand for 2 min, and then centrifuged at 12,000 rpm for 1 min. The concentration of eluted DNA was determined by taking its absorbance at 260 nm using a NanoVue spectrophotometer (GE Healthcare) and was visualized again by resolving on an agarose gel. The purified DNA fragment was used in downstream steps or stored at -20°C.

3.4.3. Restriction digestion of insert and vector

For cloning, the restriction digestion of DNA inserts and vectors was performed using the Fast-digest restriction enzymes (Thermo Scientific). Double digestion (digestion with two restriction endonucleases simultaneously) was generally followed with the reaction mixture comprising of 2 µg of plasmid or insert DNA, 1.5 µl of each fast-digest restriction enzyme, 4 µl of 10X fast-digest buffer, and sterile water making up the volume to 40 µl. Notably, for DNA vector wherein two restriction sites were separated apart by less than 10 nucleotides, sequential digestion was carried out in which DNA was first digested with one restriction enzyme, then purified using GeneJET PCR purification kit, followed by digestion with the

second restriction enzyme. The digestion was achieved by incubating in 37°C water-bath for 30 min. The digested insert DNA was then purified using a GeneJET PCR purification kit, while the digested vector DNA was excised from agarose gel and purified using a GeneJET gel extraction kit.

3.4.4. Purification of DNA inserts

Purification of the restriction digested DNA insert was performed using a GeneJET PCR purification kit (Thermo Scientific, K0702) following the manufacturer's procedure. An equal volume of Binding buffer (1:1 volume) was added to the DNA insert mixture and mixed thoroughly. The mixture was transferred into a GeneJET purification column and centrifuged at 12,000 rpm for 1 min. The flow-through in the collection tube was discarded. Next, the column was washed, DNA was eluted and analyzed by NanoVue plus spectrophotometer and agarose gel electrophoresis, in essentially the same way as followed for the extraction of DNA fragments from agarose gel (section 3.4.2).

3.4.5. Alkaline phosphatase treatment

In the case of cloning of an insert into the vector at a single restriction site, the nucleotides at two ends of the digested vector DNA were dephosphorylated to prevent recircularization during ligation by treatment with FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific, EF0651). The reaction mixture comprised of 200 ng of vector DNA, 1 µl of 1U/µl FastAP alkaline phosphatase, 1 µl of 10X FastAP buffer, and sterile water making up the volume to 10 µl. The reaction was performed by incubating the mixture at 37°C for 15 min and terminated by heating at 65°C for 15 min. The dephosphorylated vector DNA was used directly in a ligation reaction without the need for purification.

3.4.6. Ligation of DNA insert and the vector

For ligation reaction, about 100 ng of the vector was used, while the mass of insert was taken based on its molar ratio of 3:1 over the vector, which was calculated using the online ligation calculator tool (<https://nebiocalculator.neb.com>). The reaction mixture comprised of 1 µl of 5 Weiss U/µl T4 DNA ligase (Thermo Scientific, EL0011), 2 µl of 10X ligation buffer, the appropriate amount of insert and vector, and sterile water making up the reaction volume to 20 µl. The reaction was performed by incubating the mixture either at 22°C or room temperature for 30 min. Subsequently, the entire volume of the reaction mixture was used for the transformation of 100 µl of *E. coli* competent cells.

3.4.7. *E. coli* competent cells preparation and transformation

Different *E. coli* strains used in this study were Top10, BL21 (DE3), ArcticExpress (DE3), C43 (DE3), and Rosetta (DE3). Top10 has a streptomycin resistance gene. ArcticExpress (DE3) has a tetracycline resistance gene along with the gentamycin resistance gene which is incorporated in plasmid expressing chaperonins (cpn10/cpn60). Rosetta possesses plasmid encoding tRNAs for six rare codons and has a chloramphenicol resistance gene for selection. Accordingly, for growing these strains carrying antibiotic resistance genes, respective antibiotics were always added in the growth medium.

3.4.7.1 Preparation of *E. coli* competent cells

Competent *E. coli* cells were prepared using the calcium chloride (CaCl₂) method. All the steps were carried out in sterile conditions. For the growth of different *E. coli* strains, their respective antibiotics for selection were supplemented in both LB agar and LB broth. The *E. coli* cells collected after their required growth were always maintained cold by keeping on ice, all the centrifugation steps were operated at 4°C and ice-cold reagents were used during this process. A single colony of *E. coli* was picked from *E. coli* streaked LB agar plate grown overnight at 37°C and was inoculated into 5 ml LB broth. The culture was grown overnight in a 37°C incubator with shaking at 180 rpm. From this primary culture, 1% or 1 ml was inoculated into 100 ml LB broth and was grown further at 37°C for about 2 hr until the Optical Density at 600 nm (OD₆₀₀) reaches 0.6 - 0.8. The culture was then transferred into 50 ml polypropylene tubes and cooled on ice for 10 min. The cells were pelleted by centrifuging at 3,000 rpm for 5 min. The medium was decanted while the cells were resuspended in 20 ml of ice-cold 100 mM CaCl₂ solution by gently pipetting in and out the suspension using a 1 ml pipette with the end of the pipette tip cut by a sterile scalpel. The cell suspension from different polypropylene tubes was pooled into a single tube. The suspension was incubated on ice for 2 hr and centrifuged at 3,000 rpm for 5 min. The supernatant was discarded, and cells were resuspended in 3 ml of ice-cold 100 mM CaCl₂ solution by gently pipetting the suspension. The cell suspension was then incubated overnight on ice. The next day, 3 ml of ice-cold 50% glycerol was added onto the cell suspension and mixed gently. The competent cells were dispensed in aliquots of 100 µl into ice-cold 1.5 ml microcentrifuge tubes, frozen in liquid nitrogen, and stored at -70°C.

3.4.7.2 Transformation of *E. coli* competent cells

A microcentrifuge tube containing 100 μ l of *E. coli* competent cells was taken out from -70°C storage and thawed on ice for 15 min. The ligated product (20 μ l, section 3.4.6) or the plasmid DNA (approximately 50 ng) was added onto the competent cells, mixed gently by tapping the bottom of the tube 2-3 times with a finger, and incubated on ice for 20 min. The cells were then subjected to heat shock treatment at 42°C for 90 sec in a water bath, and immediately placed back on the ice and incubated for 5-10 min. Then, 1 ml of LB broth (without any antibiotics) was added into the tube and the cells were grown for 1 hr in a 37°C incubator shaking at 180 rpm. The cells were pelleted by centrifugation at 4,000 rpm for 5 min at room temperature. Much of the supernatant was decanted leaving behind about 100 μ l at the bottom of the tube with which the cells were resuspended and plated on an LB agar plate containing appropriate antibiotic(s). The plates were inverted and incubated overnight at 37°C . Two or three colonies growing on a plate were confirmed for the presence of recombinant plasmid by colony PCR with the appropriate primers. Additionally, in the case of the *E. coli* Top10 strain, the PCR positive colony was further confirmed for transformation by restriction digestion of its plasmid.

3.4.8. Colony PCR of the *E. coli* clones

An individual bacterial colony was picked with a pipette tip and was slightly streaked on an LB agar plate containing the required antibiotic(s) (a replica plate) which was incubated at 37°C for 6-8 hr. The same tip was then dipped into 10 μ l sterile water in a PCR tube. The tip was discarded and 10 μ l of PCR mixture comprising of 0.25 μ l of 5U/ μ l Taq Polymerase (produced in Dr. M. K. Reddy's Lab, ICGEB, India), 2 μ l of 10X PCR buffer, 0.25 μ l of 10 mM dNTPs, 0.25 μ l of each 10 μ M primer, and 7 μ l of sterile water was added onto the tube, making up the total reaction volume of 20 μ l. The PCR tube was given a short spin and PCR was performed with the following conditions: initial denaturation at 95°C for 5 min, 35 cycles of denaturation at 95°C for 30 s, annealing at temperature specific to primer set for 30 s, and extension at 72°C (time dependent on product size, 1 kb/min), followed by final extension step at 72°C for 7 min. PCR product was analyzed by resolving on an agarose gel.

3.4.9. Plasmid isolation from *E. coli*

Plasmid isolation was carried out using GeneJET Plasmid Miniprep Kit (Thermo Scientific, K0503) according to the manufacturer's procedure. All the isolation steps and centrifugations were carried out at room temperature unless specified. Briefly, the bacterial clone from the

replica plate was inoculated into 5 ml of LB broth supplemented with the appropriate antibiotic(s) and grown overnight in a 37°C incubator shaking at 180 rpm. The cells were collected in a 15 ml polypropylene tube by centrifuging at 5,000 rpm for 5 min. The medium was decanted completely, while the cell pellet was resuspended in 250 µl of Resuspension solution by pipetting and vortexing. Resuspended cells were transferred to a 1.5 ml microcentrifuge tube and mixed with 250 µl of lysis solution by inverting the tube 4-6 times (vortexing or incubating for more than 5 min at this step was avoided). Thereafter, 350 µl of the Neutralization solution was added and mixed immediately and thoroughly by inverting the tube 10-12 times. The suspension was cooled on ice for 5 min followed by centrifugation at 12,000 rpm for 5 min. The supernatant was carefully transferred to the GeneJET spin column and passed through it by centrifuging at 12,000 rpm for 1 min. Flow-through was discarded and 500 µl of Wash Solution was passed through the column by centrifuging at 12,000 rpm for 1 min. Flow-through was discarded and the column was washed one more time with 500 µl of Wash Solution. After decanting the wash solution, the empty column was centrifuged for an additional 1 min to remove the residual wash solution. The column with its cap left open was then placed onto a fresh 1.5 ml microcentrifuge tube for about 5 min for the residual ethanol from the Wash buffer to evaporate completely. Finally, 40 µl of Elution buffer was added to the center of the purification column membrane, incubated for 2 min, and then centrifuged at 12,000 rpm for 2 min to elute the plasmid. The concentration of eluted plasmid DNA was determined by taking its absorbance at 260 nm using NanoVue plus spectrophotometer (GE Healthcare) and its intactness was analyzed by resolving on an agarose gel. The isolated plasmid was used in downstream steps or stored at -20 °C.

3.4.10. *Agrobacterium* competent cells preparation and transformation

Two different strains of *Agrobacterium tumefaciens* were used in the study. LBA4401 strain was used for the transformation of rice, whereas GV3101 strain was used for the transformation of tobacco and *Arabidopsis*. The former has Streptomycin and Rifampicin selection genes, while the latter has Gentamycin and Rifampicin selection genes.

3.4.10.1 Preparation of *Agrobacterium* competent cells

All the steps during competent cell preparation were carried out in sterile conditions. For the growth of two different *Agrobacterium* strains, their respective antibiotics for selection were supplemented in both YEM agar and YEM broth. The cells after harvesting from the culture medium were always maintained cold by keeping on ice, all centrifugation steps were operated

at 4°C, and ice-cold reagents were used during the process. A single colony of *Agrobacterium* was picked from the streaked YEM agar plate grown at 28°C for three days and was inoculated in 5 ml of YEM broth (primary culture). The culture was grown for two days (till the stationary phase) at 28°C in an incubator shaking at 180 rpm. From this primary culture, 0.5% or 0.5 ml was inoculated into 100 ml YEM broth {supplemented with appropriate antibiotic(s)} and was grown overnight in a 28°C incubator shaking at 180 rpm until the OD₆₀₀ reached 0.6 - 0.8. The culture was transferred into 50 ml polypropylene tubes and cooled on ice for 10 min. The cells were pelleted by centrifuging at 3,000 rpm for 5 min. The medium was decanted while the pelleted cells were resuspended in 20 ml of ice-cold 100 mM CaCl₂ solution by gently pipetting in and out the suspension using 1 ml pipette with the end of the pipette tip cut by a sterile scalpel. The cell suspension from different polypropylene tubes was pooled into a single tube. The suspension was incubated on ice for 2 hr and then centrifuged at 3,000 rpm for 5 min. The supernatant was discarded, and the cells were resuspended in 2 ml of ice-cold 10 mM CaCl₂ solution by gently pipetting the suspension. Thereafter, 2 ml of ice-cold 50% glycerol was added and mixed gently by pipetting. The competent cells were dispensed in aliquots of 100 µl into ice-cold microcentrifuge tubes, frozen in liquid nitrogen, and stored at -70°C.

3.4.10.2 Transformation of *Agrobacterium* competent cells

Agrobacterium competent cells (100 µl per tube) were taken out from -70°C and thawed on ice for 15 min. Plasmid DNA (approximately 500 ng) was added onto the competent cells, mixed gently by tapping the bottom of the tube 3-4 times with a finger, and incubated on ice for 30 min. The cells were then subjected to cold shock by freezing in liquid nitrogen for 1 min and immediately given heat shock at 37°C for 5 min in a water bath. Immediately, cells were placed back on the ice and incubated for 5-10 min. Thereafter, 1 ml of YEM broth (no antibiotics) was added and the cells were grown for 3 hr at 28°C in an incubator shaker maintained at 180 rpm. The cells were then pelleted by centrifugation at 4,000 rpm for 5 min at room temperature. Much of the supernatant was then decanted, leaving behind about 100 µl at the bottom of the tube with which the cells were resuspended and plated on a YEM agar plate containing appropriate antibiotic(s). The plates were inverted and incubated at 28°C for 2-3 days until healthy colonies appeared. Two or three colonies growing on a plate were confirmed for the presence of recombinant plasmid by colony PCR with the appropriate primer sets. *Agrobacterium* colony PCR procedure was essentially the same as that of *E. coli* (section 3.4.8).

3.5 *In planta* protein subcellular localization analysis

Subcellular localization of a protein was determined by overexpressing the protein of interest in fusion with GFP either transiently or stably (*OsCBSX3-GFP* overexpressing rice), and subsequently tracking the fluorescent fusion protein in tobacco leaf as well as in rice protoplast. To produce GFP fusion protein, full-length CDS of both *OsCBSX3* and *OsCBSCLC6* was cloned into the pCAMBIA1302 vector at *Bgl*III and *Spe*I restriction sites. Note that an additional nucleotide “G” was incorporated at the 5’ end of the CDS and the stop codon was excluded at its 3’ end, to maintain the frame with upstream translation initiation site and fusion with the downstream GFP, respectively. The construct was stabilized in *E. coli* Top10 strain and the recombinant plasmid was either used directly for the transfection of protoplasts or transformed into *Agrobacterium* GV3101 strain for Agroinfiltration of tobacco leaf epidermal cells. Empty pCAMBIA1302 plasmid or the *Agrobacterium* transformed with empty pCAMBIA1302 plasmid was used as a GFP control in protoplast transfection or tobacco Agroinfiltration experiment, respectively.

3.5.1 Protein subcellular localization in Tobacco epidermal cells

The tobacco leaf infiltration with the *Agrobacterium* transformed with recombinant plasmid was carried out according to the protocol described by Sparkes et al (2006) with minor modifications. The transformed *Agrobacterium* clone was inoculated into 5 ml of YEM broth supplemented with Gentamycin, Rifampicin, and Kanamycin and grown till stationary phase for 2 days at 28°C in an incubator shaking at 180 rpm. From this culture, 25 µl or 0.5% of it was inoculated into 5 ml of YEM broth supplemented with all three antibiotics and grown overnight until the OD₆₀₀ reached 0.6 - 0.8 in a 28°C incubator shaking at 180 rpm. After that, the culture was transferred into a 15 ml polypropylene tube and cells were pelleted by centrifuging at 4,000 rpm for 5 min at room temperature. The medium was discarded and the cells were washed by resuspending in 5 ml of infiltration medium (Appendix 1, Table A1.2) followed by centrifuging at 4,000 rpm for 5 min at room temperature. The washing step was repeated one more time to remove traces of antibiotics which would kill the leaf tissue after infiltration. Finally, the cells were resuspended in 5 ml of infiltration medium and incubated in the dark at room temperature with gentle shaking for about 2 hr prior to infiltration. Infiltration was performed in well expanded healthy leaves (third to fifth leaves from the apical meristem) of about 5 to 6-week-old *Nicotiana benthamiana* plants grown in a greenhouse at 25°C under 14 hr light /10 hr dark photoperiod. Infiltration was carried out in the evening time and the

plants were exposed to fluorescent light for about 15 min before infiltration. A small scratch was made on the lower or abaxial surface of the leaf using a syringe needle; however, care was taken as not to puncture the leaf. *Agrobacterium* suspension in infiltration medium was filled in 1 ml plastic syringe (without needle) and the tip was placed over the scratched region on the abaxial surface while supported by finger from the other side of the leaf, and the plunger was pressed down gently. Usually, an entire leaf can be infiltrated by 1 ml of *Agrobacterium* suspension with one or two infiltration point(s). On the third day of infiltration, the infiltrated leaf was excised, the lower epidermis of the leaf was peeled out using forceps. The peels were then mounted on glass slides using 30% glycerol. Slides were observed under a confocal microscope for the emission of GFP as well as for chlorophyll autofluorescence.

3.5.2 Protein subcellular localization in rice protoplast

For subcellular localization of the protein of interest in rice protoplast, the isolation and transfection of rice protoplast were performed following the procedure by Yoo et al (2007) and Zhang et al (2011). Both the steps were carried out in a sterile condition. The protoplasts were isolated from rice seedlings (PB1 var.). OsCBSX3-GFP was analyzed in the protoplast of transgenic rice stably overexpressing OsCBSX3-GFP fusion protein (section 3.10). However, *OsCBSCLC6-GFP_pCAMBIA1302* recombinant plasmids were transfected into rice protoplasts to analyze the subcellular localization of OsCBSCLC6-GFP. The protoplasts from non-transgenic or WT plants and the ones transfected with empty pCAMBIA1302 plasmid were used as control.

3.5.2.1 Isolation of protoplast from rice shoot

The protoplast was isolated from rice seedlings raised in a sterile environment. Briefly, the de-husked rice seeds were sterilized in the mercuric chloride sterilization solution (see section 3.10.2) and were germinated on Half-MS medium (0.5X MS medium, 1% sucrose, and 0.3% phytigel) in a Jam Bottle in the dark at 28°C. After germination, the seedlings were maintained under a cycle of 14 hr light /10 hr dark photoperiod. Roots and leaves from 12-15 two-week-old seedlings were trimmed, and their stems were piled up and cut together into thin strips of approximately 0.5 mm using a razor blade.

These strips were immediately transferred into a 100 ml conical flask containing 15 ml of enzyme solution for protoplast isolation (Appendix 2, Table A2.1) and mixed well to disperse the strips and immerse them completely into the solution. The flask was incubated for

6 hr in the dark on a horizontal shaker shaking at 90 rpm at room temperature. After enzymatic digestion, 15 ml of W5 solution (Appendix 2, Table A2.2) was added and the flask was shaken vigorously by hand for 10 sec. The protoplasts were then passed through a cell strainer of 70 μm pore size (Sigma-Aldrich; Z742103) and collected in a 50 ml polypropylene tube. The tissues that remained on the strainer or flask were rinsed again with 15 ml of W5 solution which is then passed through the same strainer, and the filtrate containing protoplasts was collected in the polypropylene tube. The protoplasts were pelleted by centrifugation at 1,500 rpm for 3 min in a swinging bucket rotor. The supernatant was pipetted out, and the protoplasts were washed again by resuspending in 15 ml of W5 solution followed by centrifugation. The supernatant was discarded and the protoplasts were resuspended in 0.5 ml of MMG solution (Appendix 2, Table A2.3). At this stage, the protoplasts stably overexpressing OsCBSX3-GFP were mounted on a glass slide and were observed under a confocal microscope. Whereas, to transiently overexpress OsCBSCLC6, the volume of protoplast suspension was adjusted further to bring about 2×10^5 protoplasts ml^{-1} in MMG solution after counting the protoplasts under the compound microscope using a hemacytometer. The protoplasts were then advanced for transfection.

3.5.2.2 Transfection of rice protoplast

For polyethylene glycol (PEG) mediated protoplast transfection, 10 μl of plasmid DNA (10 μg *OsCBSCLC6_pCAMBIA1302* or empty *pCAMBIA 1302*) and 100 μl of protoplasts suspension in MMG solution (2×10^4 protoplasts) were added into a microcentrifuge tube and mixed gently. Then, 110 μl of PEG solution (Appendix 2, Table A2.4) was added and mixed completely by gently tapping the tube. The mixture was incubated for transfection at room temperature for 15 min in the dark. Next, the transfection mixture was diluted with 440 μl of W5 solution and was mixed well by gently inverting the tube to stop the transfection. The protoplasts were then pelleted by centrifuging at $100 \times g$ for 2 min at room temperature. The supernatant was removed, and the protoplasts were resuspended gently in 1 ml of WI solution (Appendix 2, Table A2.5), followed by incubation for 16 hr in the dark at room temperature. The protoplasts were then harvested by centrifugation at $100 \times g$ for 2 min. Much of the supernatant was pipetted out, leaving behind about 100 μl with which the protoplasts were resuspended and visualized for GFP signal under the confocal microscope.

3.6 Protein expression in *E. coli*

3.6.1. Protein SDS-PAGE

SDS-PAGE and western blot analyses were carried out using Mini-PROTEAN Tetra Vertical Electrophoresis Cell and Mini Trans-Blot Module (Biorad). For SDS-PAGE gel preparation, the acrylamide solution for resolving gel (Appendix 5, Table A5.2) was poured into the assembled glass plates, leaving sufficient space for stacking gel (length of the teeth of the comb plus 1 cm). Isopropanol was poured on top of the acrylamide solution to prevent the diffusion of oxygen. The polymerization was completed in around 30 min, after which the isopropanol overlay was poured off and the top of the gel was washed well with deionized water. The water was drained out and the acrylamide solution for stacking gel (Appendix 5, Table A5.3) was poured directly onto the surface of polymerized resolving gel. Immediately the comb was inserted into the stacking solution, avoiding the formation of air bubbles. More stacking solution required to fill the spaces of the comb was further added. The stacking solution polymerized completely in about 30 min.

For running the gel, the comb was removed, the wells were washed well with deionized water and the gel was mounted on electrophoresis apparatus. Into the tank, 1X Tris-Glycine electrophoresis buffer (Appendix 5, Table A5.4) was added, while avoiding the trapping of air bubbles at the bottom of the gel. The protein samples and the protein marker denatured in the Laemmli sample buffer (Appendix 5, Table A5.1) were loaded into the wells. Initially, the gel was run at 60 V for about 20-30 min until the bromophenol blue dye migrated from stacking to resolving gel. Then, the voltage was increased to 100 V. The gel was run till the bromophenol blue reached the bottom of the resolving gel. The gel was then removed from the gel plate and was either visualized by staining with Coomassie Brilliant Blue or used for immunoblotting for the detection of the protein of interest.

3.6.2 CBB staining of SDS-PAGE gel

To visualize the gel by staining with Coomassie Brilliant Blue, the gel was first rinsed 2-3 times with deionized water in a staining box. After that, the gel was stained with Coomassie staining solution (Appendix 5, Table A5.5) for 30 min with gentle shaking on an orbital shaker. Thereafter, the gel was rinsed 2-3 times with deionized water followed by incubation in de-staining solution on a gently shaking shaker. The gel was de-stained for about an hour, with intermittent changing of the de-staining solution (Appendix 5, Table A5.5*) two to three times

until clear background in the gel was obtained. After de-staining, the gel was immersed in water for about 5-10 min, and then visualized by placing on a transilluminator.

3.6.3 Western blot analysis

For immunoblot or western blot, SDS-PAGE gel containing the resolved protein samples was immersed in 1X Western Transfer Buffer (Appendix 5, A5.6). A nitrocellulose membrane (Hybond-C Extra, 0.45 micron, Amersham, GE Healthcare, RPN303E) and four pieces of 3 mm Whatman paper, both cut equivalent to the size of the gel, and the western blot foam pads were also immersed in 1X Western Transfer Buffer. The gel and the membrane were then sandwiched between two sheets of Whatman papers and a foam pad on each side, ensuring no trapping of air bubbles between the layers by rolling a Blot Roller over the surface of each layer. The transfer cassette was placed in a transfer tank with the gel side facing the cathode and the membrane side facing the anode. An ice block was placed into the tank to avoid heating during transfer. An adequate 1X Western Transfer Buffer was added into the tank to cover the cassette. The electro-transfer was carried out at 80 V for 1 hr.

For the detection of His-tagged or GST-tagged protein expressed in bacteria, the membrane, after electro-transfer, was placed into a plastic box and rinsed with 1X PBS (Appendix 5, Table A5.7), and proceeded for immunodetection. All the incubation steps in this process were carried out at room temperature on a gently rotating shaker. The membrane was first blocked by incubating in a blocking solution (5% skimmed milk prepared in 1X PBS) for 1 hr. It was then incubated in Alkaline Phosphatase (AP)-conjugated anti-His or anti-GST antibodies (1:10,000 titration, prepared in blocking solution) for 1 hr. It was then washed three times in 0.3% PBST for 5 min each, followed by another three washing steps in 0.1% PBST for 5 min each. Finally, the membrane was rinsed well with 1X TBS (Appendix 5, Table A5.8) and then it was completely saturated with chromogenic BCIP-NBT (Merck, 203790), followed by incubation in dark for 5-10 min until a dark-colored protein band with little or no background color was developed. The reaction was stopped by washing the membrane thoroughly with deionized water. The membrane was air-dried and stored protected away from light.

For the detection of GFP-tagged OsCBSX3 protein from transgenic plants, the membrane after the blocking step mentioned above was incubated first in Anti-GFP antibodies produced in rabbit (Sigma-Aldrich, G1544; 1:10,000 titration, prepared in blocking solution) for 2 hr. It was then washed three times in 0.1% PBST for 5 min each, followed by incubation

in AP-conjugated anti-rabbit antibodies produced in goat (Invitrogen, G-21079; 1:10,000 titration, prepared in blocking solution) for 1 h. Next, the membrane was washed and proceeded for the detection step as mentioned above.

3.6.4 Bradford assay

The quantification of protein was carried out using Bradford Reagent (Sigma-Aldrich, B6916). The standard curve was prepared using different concentrations of BSA in 30 μ l volume that was mixed with 1 ml of Bradford Reagent in a microcentrifuge tube. For the protein of unknown concentration, 30 μ l of protein sample was mixed with 1 ml of Bradford Reagent. The reactions were incubated for 5 min, transferred into a glass cuvette and the absorbance at 595 nm (A_{595}) was noted. The protein concentration was determined by plotting A_{595} of sample into the standard curve made from A_{595} of BSA standards.

3.6.5 Cloning and transformation of gene constructs for protein expression in *E. coli*

The expression of full-length OsCBSX3 was carried out by cloning full-length CDS of *OsCBSX3* in pET28a and pGEX-4T-1 (at *Bam*HI and *Eco*RI restriction sites in both vectors) in frame with the upstream CDS for 6x His tag and GST tag respectively. To express OsCBSX3 with the deletion of '105 nucleotides' from 5' end of the CDS which corresponded to 35 amino acid residues at N-terminus of the protein that was predicted as a chloroplast localization signal peptide by TargetP 1.1 server (Emanuelsson et al, 2000), the 5' end deleted region of *OsCBSX3* CDS {referred as *OsCBSX3(-sp)*} was PCR amplified and cloned into pET28a vector at *Bam*HI and *Eco*RI sites. Similarly, the full-length CDS of *OsCBSCLC6* was cloned into the pET28a expression vector at *Eco*RI and *Xho*I restriction sites in frame with the upstream CDS for 6x His tag. To express the C-terminus cytosolic region of *OsCBSCLC6* containing CBS domain pair, the 5' end region of *OsCBSCLC6* CDS which corresponded to the transmembrane domain was deleted {(a) 1716 nucleotides deleted (*CLC* Δ 1716) and (b) 1752 nucleotides deleted (*CLC* Δ 1752)}. These deletion inserts were PCR amplified and cloned into pET28a at *Eco*RI and *Xho*I restriction sites in frame with the upstream CDS for 6X His tag.

All these constructs were first stabilized by transforming in *E. coli* Top10 strain and then the recombinant plasmids were transformed into *E. coli* BL21 (DE3) protein expression strain for protein expression. The expression of full-length *OsCBSX3* was also carried out in *E. coli* ArcticExpress (DE3) expression strain by transforming it with *OsCBSX3_pET28a*

recombinant plasmid. In the case of full-length OsCBSCLC6, besides BL21 (DE3), its expression was also attempted in C43 (DE3) and Rosetta (DE3) expression strains of *E. coli*.

3.6.6 Optimization of protein expression in *E. coli*

To initially optimize the protein expression condition, the primary culture of transformed *E. coli* expression strain was grown overnight in 5 ml LB broth supplemented with appropriate antibiotic(s) in a 37°C incubator shaking at 180 rpm. The next day, for secondary culture, 5 ml LB broth {supplemented with appropriate antibiotic(s)} was inoculated with 1% of the primary culture and grown in 37°C incubator shaker at 180 rpm for about 2 hr till the OD₆₀₀ reached approximately 0.6. Notably, for full-length OsCBSCLC6, the expression was attempted in Yeast extract Tryptone (2X YT) broth and Terrific Broth (TB) as well. Thereafter, 1 ml of culture was taken into a microcentrifuge tube and stored at 4°C until needed for SDS-PAGE analysis, which serve as an uninduced control sample. The remaining culture was then adjusted to different induction temperatures: 10°C {specific to ArcticExpress (DE3)}, 18°C, 28°C or 37°C, in incubation shaker for 10 min, followed by induction of protein expression with IPTG added to a concentration of 0.1 mM. The culture was then grown at respective induction temperatures for different time points. At different induction time points, the OD₆₀₀ of the induced culture was determined and the volume of induced culture equivalent to OD₆₀₀ of 0.6 in 1 ml of suspension was transferred into a microcentrifuge tube. Then, the tubes containing induced samples along with 1 ml of uninduced control sample stored at 4°C in microcentrifuge tube were centrifuged at 5000 rpm for 5 min. The supernatants were discarded, and the cells were resuspended in 0.1 ml of 1X Laemmli sample buffer. The samples were heated in a 95°C dry bath for 5 minutes followed by centrifugation at 12,000 rpm for 2 min. The protein supernatant from induced and uninduced samples (20 µl each) was loaded and resolved in SDS-PAGE gel.

To test the solubility of the expressed target protein, 1 ml of the induced culture was transferred into a microcentrifuge tube and the cells were pelleted by centrifugation at 5,000 rpm for 5 min. The supernatant was discarded, and the cells were resuspended in 100 µl of Bugbuster protein extraction reagent (Novagen, 70584) by pipetting and vortexing, followed by incubation at room temperature for 20 min with intermittent vortexing. The soluble fraction was then separated from the insoluble fractions and cellular debris by centrifuging at 12,000 rpm for 20 min at 4°C. The soluble fraction was transferred into a new microcentrifuge tube and mixed with 4X Laemmli sample buffer, whereas the insoluble fraction was solubilized in

100 µl of 1X Laemmli sample buffer. The two fractions were then heated at 95°C in a dry bath for 5 minutes followed by centrifugation at 12,000 rpm for 2 min. The proteins in soluble and insoluble fractions (15 µl each) were loaded and resolved in SDS-PAGE gel.

For large-scale expression of the target protein, secondary culture volume was increased to 1 L of LB broth (supplemented with appropriate antibiotic) which was inoculated with 10 ml of primary culture. The culture was then grown and induced for protein expression following the steps as mentioned above, at optimized incubation temperature and induction time period. Then, the cells were harvested by centrifuging at 4,000 rpm for 20 min at 4°C (Sorvall RC 6+, Thermo Scientific) and proceeded for protein purification.

3.6.7 Purification of 6x-His-tagged OsCBX3(-sp) from soluble fraction by Ni²⁺-NTA affinity chromatography

To purify *E. coli* expressed N-terminus 6x-His-tagged OsCBSX3(-sp) from the soluble fraction of the cell lysate, the pelleted cells from 1 L culture were first resuspended in 50 ml of binding buffer (50 mM Tris pH 8.0, 150 mM NaCl, 20% Glycerol, 1 mM PMSF, 0.1% Triton X-100, 1 mM DTT and 20 mM Imidazole). Onto it, 50 µl of Lysozyme (100 mg/ml stock) was added, followed by incubation for 30 min at 4°C on a rocking shaker for cell lysis. The cells were lysed further by sonication at 40% amplitude with a pulse cycle of ‘ON’ for 5 sec and ‘OFF’ for 10 sec while keeping the cell suspension on ice to maintain a low temperature. After complete lysis (sample becomes much clear and loses its viscosity), the lysate was centrifuged at 8,000 rpm for 30 min at 4°C. The supernatant was passed through a 0.45 µm syringe filter and collected in two 50 ml polypropylene tubes which were already kept on ice to proceed for purification. The Ni²⁺-NTA agarose resin slurry 2 ml (or 1 ml bed volume) for 1 L culture volume was pipetted into a 15 ml polypropylene tube and briefly centrifuged. The supernatant was decanted, and the resin was mixed with a 10 ml binding buffer. It was briefly centrifuged again, and the supernatant was decanted. The equilibrated resin was then resuspended in 4 ml of binding buffer which was then added equally into two tubes containing a soluble fraction of cell lysate. The suspension was mixed gently by shaking on a rotary shaker at 4°C for 30 min. The lysate–Ni²⁺-NTA mixture was then loaded into a purification column and the resin was allowed to settle down at the bottom of the column. Then, the outlet cap of the column was removed and the flow-through was collected and stored at 4°C for SDS-PAGE analysis. The outlet cap was plugged in again and the wash solution (Binding solution containing 30 mM Imidazole) was added carefully into the column (up to a maximum volume of the column) with

minimal interference to the resin. The outlet cap was removed, and the initial 5 ml of the washed fraction was collected and saved at 4°C for SDS-PAGE analysis. A total volume of 80 ml wash solution was used for washing. Finally, 6x-His _OsCBSX3(-sp) was eluted with elution buffer (binding buffer containing 50 mM, 100 mM, and 200 mM of Imidazole, respectively). The first two elutions, each of 1 ml fractions, were carried out in elution buffer containing 50 mM followed by 100 mM of Imidazole. Then the protein was eluted with 5 ml of an elution buffer containing 200 mM of Imidazole, and the eluates were collected in 5 fractions of 1 ml each in separate microcentrifuge tubes. The eluted fractions along with the flow-through and wash fractions (15 µl each) were mixed with 4X Laemmli sample buffer (5 µl each), heated at 95°C in a dry bath for 5 min, and analyzed by SDS-PAGE. Protein concentration in the eluted fractions was determined by the Bradford method. The purified proteins were stored at -70°C. Note that all purification steps, from binding to elution, were carried out in a cold room and the eluted fractions were immediately placed on ice. The binding buffer composition used in this mentioned method was optimized after screening the solubility of 6x-His_OsCBSX3(-sp) in HEPES and Phosphate buffers as well as at different concentrations of NaCl and Glycerol. The cells were harvested from 100 ml culture volume, lysed in 5 ml buffer, and purified with 0.5 ml of Ni²⁺-NTA resin slurry during the optimization of buffer composition.

3.6.8 Solubilization of 6x-His-tagged proteins from insoluble fraction followed by Ni²⁺-NTA affinity chromatography purification

The 6x-His tag_OsCBX3 (full length) and two 6x-His-tagged cytosolic variants of OsCBSCLC6: CLCA1716 and CLCA1752 (1716 and 1752 nucleotides deleted from 5' end of CDS, respectively), portioned in the insoluble fraction *E. coli*, despite trying various approaches to get them in the soluble fraction. Therefore, these proteins were solubilized from the insoluble fraction by chaotropic extraction, purified by Ni²⁺-NTA affinity chromatography, and attempted to refold by dialysis. The initial protein induction in 1 L culture and subsequent lysis by lysozyme and sonication, followed by separation of an insoluble fraction by centrifugation was performed in essentially the same way as mentioned in section 3.6.7. Thereafter, the insoluble pellet was resuspended in 30 ml of cold Wash Buffer (20 mM Tris pH 8.0, 0.5 M NaCl, 5 mM DTT, 2 M urea, and 2% Triton X-100) by sonication. The suspension was centrifuged at 8,000 rpm for 30 min at 4°C. The supernatant was discarded, and the pellet was washed two more times with wash buffer. The pellet was then washed once with cold wash buffer without Urea and Triton X-100. The insoluble protein from the washed

pellet was then solubilized in 30 ml of extraction buffer (20 mM Tris pH 8.0, 0.5 M NaCl, 6 M Guanidine hydrochloride, 5 mM Imidazole, and 2 mM β -mercaptoethanol) and incubated for 30 min on a gently shaking shaker at room temperature. The suspension was ultra-centrifuged at 30,000 rpm for 1 h at 4°C (Sorvall LYNX 6000, Thermo Scientific). The supernatant was passed through a 0.45 μ m syringe filter and collected in a 50 ml polypropylene tube.

The 6x-His-tagged proteins solubilized from insoluble fractions were then incubated with 1 ml of Ni²⁺-NTA agarose resin pre-equilibrated with extraction buffer for 30 min on a rotary shaker at room temperature. The suspension was then transferred into the purification column, and the resin was allowed to settle down, followed by the release of the flow-through which was collected and saved for SDS-PAGE analysis. Thereafter, the resin was washed with 40 ml of Wash Buffer II (20 mM Tris pH 8.0, 0.5 M NaCl, 4 M Urea, 2 mM β -mercaptoethanol) containing 10 mM of imidazole. Finally, the 6x-His-tagged target proteins were eluted with 10 ml of Wash Buffer II containing 200 mM of imidazole in an aliquot of 1 ml. The eluates were analyzed by SDS-PAGE. The eluted aliquots with higher homogeneity of the target protein (as per observation by SDS-PAGE) were pooled into Tube-O-DIALYZER, Medi (G-Biosciences, 786-616), and the proteins were subjected to refolding through gradual removal of Urea employing stepwise dialysis. The proteins in the dialysis tube were initially incubated in 2 L of Wash Buffer II containing 2 M of Urea after which the buffer was changed with the one with 1 M Urea, then to 0.5 M Urea and finally to Wash Buffer II lacking Urea. In each dialysis buffer, the proteins were incubated for 12 hr at 4°C.

3.6.9 Gel permeation chromatography (GPC)

Recombinant 6x-His_OsCBSX3(-sp) expressed in *E. coli* and purified using Ni²⁺-NTA affinity chromatography was run through GPC Superdex 75 column to determine its oligomeric status. Prior to GPC, the protein sample, and the GPC buffer (50 mM Tris pH 8.0 and 100 mM NaCl) were passed through a 0.22- μ m filter to clear the particles. The GPC buffer was degassed as well to remove the air bubbles. The Superdex 75 column was pre-calibrated with protein standards of well-characterized native molecular weights (Mw), viz., conalbumin (75.0 kDa), ovalbumin (44.0 kDa), carbonic anhydrase (29.0 kDa), ribonuclease A (13.7 kDa), and aprotinin (6.5 kDa). The void volume (V_o) and the bed volume (V_c) of the column were pre-determined to be 45.14 ml and 120 ml, respectively. Based on the elution volume (V_e) of the standards, and V_o and V_c of the column, the partition coefficient K_{av} {K_{av} = (V_e-V_o)/(V_c-

V_0) of these standards were pre-determined and plotted against their respective log molecular weights (K_{av} vs. Log Mw) to obtain a standard curve. Accordingly, the molecular weight of 6x-His_OsCBSX3(-sp) was determined by plotting its K_{av} which in turn was based on its V_e in the calibration curve.

3.7 ATP binding analysis

ATP-binding analysis of 6x-His_OsCBSX3(-sp) was carried out following the methods from Marín et al (2000) and Choudhury et al (2006) with few modifications. Briefly, 1 μ g of purified recombinant protein was added into ATP binding buffer consisting of 50 mM Tris pH 8.0, 10 mM $MgCl_2$, 20 mM NaCl, 1 mM DTT, 10% Glycerol, and 5 μ Ci of [α - ^{32}P] ATP (3000 Ci/mM, Perkin–Elmer Life Sciences), making up the reaction volume to 20 μ l. The reaction mixture was incubated on ice for 15 min followed by ultraviolet (UV) cross-linking (254 nm) for 15 min while still incubated on ice. The proteins were then denatured in the presence of Laemmli sample buffer and resolved on 12% SDS-PAGE gel. Then, the gel was covered with plastic wrap and exposed to a Phosphor imaging screen enclosed in the PhosphorImager cassette overnight. The screen was scanned on the PhosphorImager system to analyze the binding of ATP.

3.8 Yeast Two-Hybrid (Y2H) assay

3.8.1. Protein homo-oligomerization analysis using Y2H

To determine the homo-oligomerization of OsCBSX3, its full-length CDS as well as the CDS with deletion of 5' end '105 nucleotides' which corresponding to N-terminus 35 amino acid residues predicted as a chloroplast signal peptide {OsCBSX3(-sp)}, were cloned into pGADT7 and pGBKT7 at *EcoRI* and *BamHI* sites in both Y2H vectors. Likewise, to determine the homo-oligomerization among Bateman modules in OsCBSCLC6, its CDS with the deletion of 5' end 1716 nucleotides corresponding to the transmembrane domain (CLC Δ 1716) was cloned into pGADT7 (at *EcoRI* and *XhoI* sites) and pGBKT7 (at *EcoRI* and *PstI* sites). These constructs were stabilized in the *E. coli* Top10 strain. The autoactivation activity of these constructs was analyzed first by co-transforming each of these constructs cloned in pGBKT7 with empty pGADT7 vector into yeast cells (AH109 strain) followed by the screening of the transformed clones on 3 DO (SD/-Leu/-Trp/-His) and 4 DO (SD/-Trp/-Leu/-His/-Ade) minimal medium supplemented with 5 mM of 3-Amino-1,2,4-triazole (3-AT, a Histidine antimetabolite). Thereafter, the homo-oligomerization was analyzed by co-transforming AH109 cells with their

constructs in pGADT7 and pGBKT7 vectors, followed by screening the transformed clones on 3 DO and 4 DO minimal medium.

3.8.2. Screening of putative interacting partners from rice cDNA library by Y2H

To identify the interacting partners of OsCBSX3 by Y2H through screening of a collection rice cDNA library clones inserted in pAD-GAL4-2.1 vector, *OsCBSX3* full-length CDS was cloned into pBD-GAL4-Cam vector at *EcoRI* and *SalI* sites. The recombinant plasmid stabilized in *E. coli* Top10 strain was then transformed into yeast cells harboring rice cDNA library in pAD-GAL4-2.1 prepared previously by Kumar et al (2012). The transformed cells were screened by plating on 3 DO + 5 mM of 3-AT minimal medium agar plates followed by replica plating on 4 DO + 5 mM of 3-AT minimal medium agar plates. The positive clones growing on 4DO + 5 mM of 3-AT minimal medium were further screened for the activation of the *MEL1* reporter gene encoding α -galactosidase by streaking onto X- α -gal (5-Bromo-4-chloro-3-indolyl α -D-galactopyranoside) containing YPD medium agar plates. Finally, plasmids from the positive clones were isolated and the rice cDNA fragments inserted in pAD-GAL4-2.1 from the clones were PCR amplified using pAD-GAL4-2.1 vector-specific primers. The PCR products were eluted from the agarose gel, purified, and sequenced to identify the putative interacting partners of OsCBSX3.

3.8.3. Interaction analysis of OsCBSX3 with full-length of putative interacting partners by Y2H

The putative genes encoding thioredoxins from rice (*OsTrx y*, *OsTrx f*, *OsTrx m*, *OsTrx o*, *OsTrx 1*, *OsTrx 10* and *OsTrx 23*) were cloned into pGADT7 vector at *EcoRI* and *XhoI* sites and each of these recombinant plasmids stabilized and obtained from *E. coli* Top10 strain was co-transformed along with *OsCBSX3*_pGBKT7 into AH109 cells and plated on 2 DO minimal medium agar plates. The transformed clones were screened on 3 DO and 4 DO minimal medium agar plates. Three of these putative *OsTrxs*, with deletion of the nucleotides corresponding to their signal peptides, which was referred by suffix (-sp) in their name, viz., *OsTrx f(-sp)*, *OsTrx m(-sp)*, *OsTrx o(-sp)*, along with *OsTrx y*, were also cloned into pGBKT7, all at *EcoRI* and *PstI* sites. In addition, genes encoding other putative interacting partners, viz., putative spermidine synthase, putative retrotransposon and putative glycosyl transferase(-sp) at *EcoRI* and *BamHI* sites, initiation factor-2 subunit family domain containing protein(-sp) at *NdeI* and *BamHI* sites, and three *Arabidopsis* AtTrxs {AtTrx y(-sp), AtTrx m(-sp) and AtTrx o(-sp)} at *EcoRI* and *PstI* sites, were also cloned into pGBKT7. All these constructs were

analyzed for auto-activation activity. Thereafter, each of these constructs was co-transformed along with *OsCBSX3_pGADT7* or *OsCBSX3(-sp)_pGBKT7* into AH109 cells, followed by the screening of the transformed clones on 3DO and 4DO minimal medium agar plates.

3.8.4. Transformation of *Saccharomyces cerevisiae*

The yeast transformation was carried out following the lithium acetate (LiAc)-mediated transformation method as described in Yeast Protocols Handbook (Clontech Laboratories, Inc.) with minor modifications. All the steps were performed in sterile conditions. Briefly, several yeast colonies (AH109 strain) of 2–3 mm diameter were inoculated into 50 ml of YPD medium. In the case of library screening, the yeast cells containing the collection of rice cDNA libraries inserted in the pAD-GAL4-2.1 vector were inoculated into 50 ml of SD/-Leu minimal medium. The culture was incubated overnight for 16–18 hr at 30°C with shaking at 200 rpm till stationary phase ($OD_{600} > 1.5$). The next day, about 30 ml of overnight culture was transferred into a flask containing 300 ml of YPD (or SD/-Leu for cDNA library cells) bringing the OD_{600} up to 0.2–0.3, followed by incubation at 30°C with shaking at 200 rpm for 3 hr until OD_{600} reaches 0.4–0.6. Then, the cells were transferred into 50-ml polypropylene tubes and centrifuged at 1,000 x g for 5 min at room temperature. The supernatants were discarded, and the cell pellets were thoroughly resuspended in 40 ml of 1X TE buffer (Appendix 3, Table A3.1). The cells were pooled into a single tube and centrifuged again at 1,000 x g for 5 min at room temperature. The supernatant was decanted, and the yeast competent cells were resuspended in 1.5 ml of freshly prepared TE-LiAc Solution (Appendix 3, Table A3.4, A3.2). Into another 1.5-ml microcentrifuge tube, 1 µg of plasmid DNA (1 µg of each plasmid for co-transformation with two different plasmids) and 0.1 mg of carrier DNA (sonicated salmon sperm DNA, denatured by heating at 95°C for 20 min and rapidly cooled on ice prior to use) were added and mixed. To this, 0.1 ml of yeast competent cells were added and mixed well by vortexing. Thereafter, 0.6 ml of TE-LiAc-PEG Solution (Appendix 3, Table A3.5) solution was added and mixed by vortexing at high speed for 10 sec. The cells were then incubated at 30°C for 30 min with shaking at 200 rpm. Subsequently, 70 µl of DMSO was added and mixed well by gentle inversion, but not by vortexing. Then heat shock was given for 15 min in a 42°C water bath followed by chilling on ice for 10 min. The cells were then pelleted by centrifugation at 3,000 rpm for 10 sec at room temperature. The supernatant was discarded, and the cells were resuspended in 0.5 ml of 1X TE buffer. Finally, 100 µl of the resuspended cells were plated on an appropriate SD agar plate and the plates were incubated upside-down at 30°C until the colonies appear (generally, 3–4 days). For library screening, the entire content of transformed

cells was plated on several numbers of 3 DO (SD/-Leu/-Trp/-His) plates to avoid the loss of interacting partners.

3.8.5. Isolation of plasmid from yeast cells

Yeast plasmid was isolated according to the method detailed in Instruction Manual, HybriZAP-2.1 (Agilent Technologies). Briefly, 2 ml of SD/-Leu/-Trp minimal medium was inoculated with a large yeast colony freshly grown on a 3 DO plate. The culture was incubated overnight at 30°C and was transferred to a 1.5-ml microcentrifuge tube followed by centrifugation at $14,000 \times g$ for 10 sec to pellet the cells. The supernatant was decanted, and the cells were resuspended in 0.2 ml of yeast lysis solution (10 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, 1% SDS, 2% Triton X-100) by vortexing. Onto it, 0.2 ml of phenol-chloroform-isoamyl alcohol solution (25:24:1, v/v/v) and 0.3 g of acid-washed glass beads were added, and the suspension was vortexed for 2 min. The suspension was then centrifuged at $14,000 \times g$ for 5 min at room temperature. The upper aqueous phase containing DNA was transferred to a new microcentrifuge tube and $1/10^{\text{th}}$ volume of 3 M NaOAc (pH 5.2) and 2.5X volume of ethanol was added onto it to precipitate the DNA. The suspension was centrifuged at $14,000 \times g$ for 10 min. The supernatant was decanted, and the DNA pellet was washed with 1 ml of 70% ethanol and centrifuged again at $14,000 \times g$ for 10 min. The supernatant was decanted, and the DNA pellet was air-dried for about 20 min and resuspended in 50 μl of 1X TE buffer. The concentration and purity of isolated DNA were determined using a spectrophotometer. It was then used as a template (50 ng in 50 μl reaction) to PCR amplify the rice cDNA fragments inserted in pAD-GAL4-2.1 using the vector-specific primers.

3.9 Interaction analysis of OsCBSX3 and OsTrx y by Bimolecular Fluorescence Complementation (BiFC) *in planta* in onion epidermal cells

3.9.1. Gene and vector constructs used in BiFC assay

For interaction analysis by BiFC, CDS of *OsCBSX3* (full length) and *OsTrx y* (the clone we used in Y2H) were cloned into split-EYFP (Enhanced Yellow Fluorescent Protein) vectors - pSATN-nEYFP-C1 and pSATN-cEYFP-C1 (Citovsky, 2006). *OsCBSX3* was integrated at *Bgl*III and *Sal*I sites, and *OsTrx y* was cloned at *Bgl*III and *Eco*RI sites, in both the vectors. The recombinant plasmids were stabilized in the *E. coli* Top10 strain and were used to transiently transform onion epidermal cells using biolistic bombardment.

3.9.2. Onion epidermal strips preparation

On the day of bombardment, the bulb of spring onion was cut into quarters, and the epidermis layer on the concave surface of the bulb was gently cut into small rectangular strips of approximately 1 cm² using a razor blade. The epidermal strips were then peeled off with fine forceps and laid on the center on the 0.5X MS plate (0.5X MS medium, 3% Sucrose, 0.3% Phytigel; pH 5.8) with its outer surface facing down the medium. An area of about 3 cm in diameter at the center was covered with the epidermal strips. The plates were incubated in dark at room temperature till ready for biolistic bombardment.

3.9.3. Microcarrier-DNA preparation, and bombardment on onion epidermal cells

To prepare gold microcarriers and its subsequent coating with DNA for bombardment, the method detailed in the instruction manual of Biolistic PDS-1000/He (Bio-Rad) was followed. Briefly, 30 mg of 1 µm gold microcarriers (BioRad, 1652263) was transferred in 1.5 ml microcentrifuge tube and 1 ml of 70% ethanol was added onto it. The tube was vortexed vigorously for 5 min and then the microcarriers were allowed to soak in 70% ethanol for 15 min. The microcarriers were pelleted by briefly spinning the tube in a centrifuge for 5 sec and the supernatant was discarded. The microcarriers were resuspended in 1 ml of sterile water, vortexed vigorously for 1 min, allowed to settle for 1 min, and pelleted by briefly spinning the tube. The supernatant was discarded and the washing of microcarriers with sterile water was repeated two more times. Then, the microcarriers were resuspended in 500 µl of sterile 50% glycerol (microcarrier concentration = 60 mg/ml). Aliquots of the prepared microcarriers were stored at -70°C.

Coating of microcarriers with DNA was carried out on the day of bombardment. The microcarriers prepared in 50% glycerol were vortexed for 5 min to resuspend and disrupt the agglomerated particles. While the tube containing microcarriers was under continuous vortexing, carefully 15 µl of microcarriers were pipetted out and transferred into a new 1.5 ml microcentrifuge tube. Each tube containing 15 µl of microcarriers was then vortexed continuously while adding DNA and other reagents. First, the combination of recombinant plasmids (pSATN-nEYFP-C1 and pSATN-cEYFP-C1; 2.5 µg each) containing the genes of interest to be co-transformed in 40 µl of volume were added. Next, 50 µl of 2.5 M CaCl₂ followed by 20 µl of 0.1 M spermidine were added. The vortexing was continued for another 30 min, but at 4°C. After that, the microcarriers were allowed to settle for 1 min and pelleted by briefly spinning for 2 sec. The supernatant was discarded. The microcarriers were then

washed by adding 100 µl of absolute ethanol, vortex followed by a short spun and the decantation of supernatant. The ethanol washing step was performed thrice. Finally, the microcarriers were resuspended in 20 µl of absolute ethanol, vortexed, and pipetted out onto a sterile microcarrier disk. The ethanol was allowed to evaporate and subsequently, the microcarriers were bombarded into onion epidermal strips placed on MS plate and the plates were incubated overnight in the dark at 28°C. The next day, the peels were mounted on a slide using 30% glycerol and were observed for the complemented EYFP fluorescence under a confocal microscope.

3.10 Generation of transgenic rice plants overexpressing *OsCBSX3* and *OsCBSCLC6* and their subsequent screening

3.10.1 Gene and vector constructs used for transformation of rice

For the overexpression of *OsCBSX3*, its CDS was cloned into pCAMBIA1302 plant binary vector at *Bgl*III and *Spe*I restriction sites with the stop codon included at its 3' end to prevent the formation of GFP fusion protein. Whereas CDS of *OsCBSX3* without the stop codon at its 3' end was also cloned into pCAMBIA1302 at the same *Bgl*III and *Spe*I restriction sites to produce *OsCBSX3*_GFP fusion overexpressing transgenic plant used in protein localization study (section 3.5.2) In the case of *OsCBSCLC6*, its CDS was first cloned into pRT101 vector at *Xho*I and *Kpn*I sites. Then, the gene cassette comprising of 35S CaMV promoter, *OsCBSCLC6*, and Poly(A) signal, was excised from pRT101 by restriction digestion with *Hind*III and inserted into pCAMBIA1300 plant binary vector.

All these constructs were stabilized into *E. coli* Top10 strain and the recombinant plasmids were finally transformed into *A. tumefaciens* LBA4404 strain, which was subsequently used to transform the rice calli.

3.10.2 Rice calli production, its transformation, and regeneration into plants

Rice var. PB-1 was used for the production of transgenic plants. The tissue culture method as detailed by Sahoo et al (2011) was followed. The temperature in the tissue culture room was maintained at 28°C with 14 hr light and 10 hr dark cycle. Approximately 200 dehusked rice seeds were surface sterilized by soaking in 100 ml of sterilization solution {0.1% mercuric chloride (HgCl₂) and 0.02% of Triton X-100} with gentle but constant shaking for 10 min. The solution was decanted, and the seeds were washed with about 200 ml of sterilized water while shaking constantly for 5 min. The solution was decanted, and the washing step was repeated

four more times. Then the seeds were dried for 5 min by spreading on autoclaved 3mm Whatman paper and transferred onto plates containing Callus Induction medium (12-14 seeds per plate). The plates were incubated in the dark for about 14 days. Then, the embryogenic calli were excised, cut into two to four halves using a scalpel, and sub-cultured onto a fresh Callus Induction medium. The calli were incubated for another 5 days in the dark after which they were transformed with *Agrobacterium*. For calli obtained from about 200 rice seeds, 100 ml of YEM broth supplemented with Streptomycin, Rifampicin, and Kanamycin was inoculated with 0.5 ml or 0.5% of fully grown primary culture and was incubated at a 28°C incubator shaking at 180 rpm until the OD₆₀₀ reaches approximately 0.6. After that, the culture was transferred into 50 ml polypropylene tubes and cells were pelleted at 4,000 rpm for 10 min at room temperature. The medium was discarded, and the cells were resuspended in 100 ml of MS Resuspension medium. The calli were added onto *Agrobacterium* resuspension and incubated for 20 min with gentle shaking. The calli were then dried on sterile 3mm Whatman paper for 5 min and transferred onto the Co-cultivation medium followed by incubation in the dark for about 48 hr until *Agrobacterium* growth appears on the periphery of the calli. Then, the calli were washed five times with about 200 ml of Cefotaxime (Taxim, Alkem Laboratories Ltd.) solution (250 mg/l of sterilized water) for about 5 min each with gentle but constant shaking. The calli were then dried on a sterile 3mm Whatman paper and transferred onto a Selection medium and incubated for 12 days in the dark. Subsequently, the healthy calli were sub-cultured two more times onto fresh Selection medium at 10-day intervals, while maintaining in the dark. After the third selection, the proliferated micro-calli were shifted onto Regeneration-I medium and incubated in the dark for the first 7 days after which they were exposed to a regular light cycle of tissue culture room (14 hr light and 10 hr dark) for another 7 days. In the subsequent regeneration and rooting stages as well, the calli were exposed to a regular light cycle of the tissue culture room. Next, the regenerating calli were transferred onto Regeneration-II medium and regenerated shoots were transferred onto Rooting medium in a jam bottle. The rooted plantlets were then transferred to vermiculite for hardening in a greenhouse for 10-12 days. After that, the tissue-cultured plantlets were shifted onto the soil in earthen pots.

3.10.3 PCR screening of putative transgenic plants

For PCR screening of the putative transgenic plants, a rapid genomic DNA isolation was performed using Extraction (Sigma-Aldrich, E7526) and Dilution (Sigma-Aldrich, D5688) solutions and was proceeded directly for PCR. Rice leaf tip (young leaf) of about 1 cm length

was cut with a scissor and put into a 1.5 ml microcentrifuge tube, which was frozen into liquid nitrogen. The tube was taken out from the liquid nitrogen and the frozen leaf tissue inside it was immediately crushed into minute fragments using a 1 ml pipette tip. Onto it, 50 μ l of Extraction solution was added and mixed well by vortexing. Fragments of leaf tissue were ensured to be completely immersed into the buffer which was achieved by giving a short spin. The tube was then incubated at 95°C on a dry bath for 10 min. After that, 50 μ l of Dilution solution was added onto it and the contents were mixed well by vortexing. The tube was then briefly centrifuged to settle down the debris. The supernatant (2 μ l) was used directly as templated in 20 μ l of PCR comprising of 15 μ l nuclease-free water, 2 μ l of 10X PCR buffer, and 0.25 μ l each of 5U/ μ l Taq polymerase, 10 mM dNTPs, 10 μ M forward primer, and 10 μ M reverse primer. The reaction conditions were as follows: initial denaturation at 95°C for 5 min, 40 cycles of denaturation at 95°C for 1 min, annealing at primer specific temperature for 1 min, and extension at 72°C for 1 min (product size < 1 kb), followed by final extension step at 72°C for 7 min. PCR product was analyzed by resolving on an agarose gel.

3.10.4 Analysis of transgene integration by Southern blot

The putative PCR positive transgenic (T_0) plants were further analyzed by Southern blot to confirm transgene integration, transgene copy number, and to identify independent single transgene integrated lines to proceed for later plant growth analysis.

3.10.4.1 Genomic DNA (gDNA) isolation

The gDNA from rice leaves were isolated following the CTAB extraction method detailed by Doyle (1991) with minor modifications. All the steps were performed at room temperature or else specified in the step. Pipette tips cut at the bottom with a razor blade to widen the opening were used for transferring samples. Briefly, 200 mg of frozen leaf tissue was ground to a fine powder in liquid nitrogen using mortar and pestle. A pinch of poly vinyl pyrrolidone (PVP) was added and mixed by further grinding. The homogenized powder (before it thawed) was immediately transferred into a 15 ml polypropylene tube containing 2 ml of CTAB extraction buffer (Appendix 7, Table A7.1) pre-heated at 60°C and the tube was swirled gently to mix the content. The sample was incubated in a 60°C water bath for 30 min with intermittent gentle mixing. Then, an equal volume of chloroform: octanol (24:1) was added and mixed gently but thoroughly. The tube was centrifuged at 6,000 rpm for 15 min, and the upper aqueous phase was pipetted out and transferred to a new polypropylene tube. Next, the aqueous phase was treated with 1 μ l of 10 mg/ml RNase A (Thermo Scientific, EN0531) with incubation at 37°C

for 30 min. Subsequently, the aqueous phase was extracted again with an equal volume of chloroform: octanol (24:1) followed by centrifugation. The upper aqueous phase was transferred into a new polypropylene tube and a double volume of cold absolute ethanol was added. The tube was mixed gently and incubated on ice for an hour to precipitate the DNA. The tube was then centrifuged at 6,000 rpm for 10 min to pellet down the DNA. The supernatant was decanted, and the pellet was washed with 2 ml of 70% ethanol, swirled gently for a minute, and centrifuged at 5,000 rpm for 10 m. The supernatant was decanted, and the DNA pellet was washed again using 70% ethanol. Finally, the DNA pellet was air-dried for about 20-30 min to remove ethanol and dissolved in 100 μ l of 1X TE buffer (Appendix 3, Table A3.1).

The concentration and purity of the isolated DNA were determined by measuring the absorbance of DNA (diluted 250 x to 1 ml) at 260 nm (A_{260}), 280 nm (A_{280}), and 230 nm (A_{230}) using a UV-Vis spectrophotometer. An absorbance of 1 at 260 nm corresponds to approximately 50 ng/ μ l of dsDNA, whereas the ratio of A_{260}/A_{280} and A_{260}/A_{230} should be approximately 1.8 and 2.3–2.4, respectively for pure dsDNA (Koetsier and Cantor, 2019). The integrity of the DNA was further analyzed by resolving it on 0.8% agarose gel.

3.10.4.2 Restriction digestion and gel electrophoresis of gDNA

Restriction enzyme that has only a single restriction site in the T-DNA region of vector and none in the CDS of the gene was selected to determine the transgene integration as well as the transgene integration number in transgenic plants by southern blot. In the case of *OsCBSX3*, *Bam*HI was used, which has a single restriction site in the vector T-DNA region, and none in its CDS as well as in the endogenous gene (including intronic regions). For *OsCBSCLC6*, *Eco*RI was used. However, *Eco*RI has a single restriction site in the intronic region in the endogenous *OsCBSCLC6* which lies between 358/359 bp in its CDS. Therefore, we amplified the probe from the region between 1641 to 2391 bp of *OsCBSCLC6* CDS to be used for the Southern blot.

The restriction digestion reaction comprised of 10 μ g of gDNA, 10 μ l of Restriction enzyme (NEB), 30 μ l of 10X Restriction buffer, and sterile nuclease-free water making up the reaction volume to 300 μ l. The reaction was incubated overnight in a 37°C water bath and terminated by heat inactivation for 20 min at 65°C in a water bath. Importantly, 6 μ l of the enzyme was added to the reaction mixture in the beginning while the remaining 4 μ l was added after 8-10 hr, the next day. Following restriction digestion, the samples were concentrated or

reduced the volume to < 20 µl using SpeedVac (Thermo Scientific). The samples were then mixed with 6X DNA loading dye and loaded in 0.8% agarose gel prepared in 0.5X TBE buffer (Appendix 6, Table A6.3). The gel was run overnight at 35 V in 0.5X TBE buffer until the Bromophenol blue dye migrated to about 3/4th length of the gel.

3.10.4.3 Preparation of gel for transfer

After electrophoresis, the gel was carefully transferred into a glass tray and treated with different solutions as detailed below while keeping on a gently rotating shaker. First, the gel was immersed in 200 ml of EtBr solution (0.5 µg/ml) to stain the DNA for 15 min. DNA in the gel was visualized through gel doc, and the gel was transferred back into glass tray again, followed by de-staining two times with 200 ml of autoclaved water for 15 min each. Next, the gel was treated with 200 ml of Depurination solution (Appendix 8, Table A8.1) for 15 min (Bromophenol blue dye turned yellow in this step). After depurination, the gel was washed with 200 ml of autoclaved water for 5 min. The DNA was then denatured by treating the gel two times with 200 ml of Denaturing solution (Appendix 8, Table A8.2) for 15 min each. The gel was then treated two times with 200 ml of Neutralizing solution (Appendix 8, Table A8.3) for 15 min each and continued to be immersed in Neutralizing solution until proceeded with blotting.

3.10.4.4 Preparation of membrane and blotting papers for transfer

Prior to the transfer, a piece of nylon membrane (Hybond-N⁺, Amersham, GE Healthcare, RPN303B), three pieces of 3mm Whatman paper, and pieces of a large number of blotting papers making a stack of about 7-8 cm thickness were cut corresponding to the size of the gel to be blotted. A piece of 3mm Whatman paper served as a wick. This paper wick should be wider than the gel and long enough to touch the bottom surface of the glass tray while placed over a glass plate supported on a glass tray.

3.10.4.5 Transfer of DNA from gel to membrane

Upward Capillary Transfer method was followed to transfer DNA from agarose gel to the membrane. Half of the glass tray was filled with 20X SSC buffer (Appendix 8, Table A8.4, approximately 1.5 L in a tray) and the glass plate was placed over the glass tray at the center. Over the glass plate, a 3mm Whatman paper wick was placed, with its long ends equally immersed into 20X SSC buffer on the glass tray. The surface of paper wick over the glass tray was thoroughly wetted by pouring 20X SSC buffer onto it and the trapped air bubbles were

removed by rolling a roller or glass pipette over it. Next, the gel was inverted so that its underside was now uppermost and was placed over the center of the wet paper wick on the glass plate while avoiding air bubbles formation. The membrane was dipped in 2X SSC to make it completely wet, and placed exactly over the gel, avoiding air bubbles. Accordingly, three pieces of 3mm Whatman papers were soaked completely in 2X SSC buffer and placed exactly on top of the membrane. The pipette was rolled across the surface of the Whatman paper to remove any air bubbles if entrapped. Then after, precisely over the Whatman paper, a stack of dry blotting papers was placed. Finally, another glass plate was placed over the stack onto which a thick book weighing about 500 g was kept. The transfer was allowed to proceed for about 16 hr.

After completion of the transfer, the paper stacks above the gel were removed. The gel along with the membrane and the three pieces of Whatman papers above were turned with gel side up and Whatman paper at the bottom and laid on a glass plate. Positions of wells on the gel were marked onto the membrane with a pencil. A small piece of membrane on one of the four corners was cut with a scissor to make an identification mark the membrane surface. The gel was then removed from the top, stained with EtBr solution (0.5 µg/ml), and visualized in Gel Doc to ascertain the transfer of the DNA. The membrane along with the Whatman paper underneath was exposed to UV for 2 min (1200 units) for cross-linking of DNA. The membrane was then immersed in 2X SSC buffer on a glass tray and proceeded to the next steps.

3.10.4.6 PCR amplification of Digoxigenin (DIG) labeled probe

The hybridization probe specific to the transgene region was PCR amplified using PCR DIG Probe Synthesis Kit (Roche, 1636090) which included DIG-dUTP in the mixture of nucleotides that was incorporated into the PCR amplified probe. The PCR mixture comprised of 1.25 µl of DIG labeling mix, 0.25 µl of 10 mM dNTPs, 1.5 µl of each 10 µM primer, 0.5 µl of 3.5 U/µl Taq polymerase, 2.5 µl of 10X PCR buffer, 1 ng of plasmid (template) and nuclease-free water making up the reaction volume to 25 µl. A control DIG unlabeled PCR was also performed with essentially the same reaction mixture composition, however instead of the DIG labeling mix, an additional 10 mM dNTPs were added onto it. The reaction conditions were as follows: initial denaturation at 95°C for 2 min, 35 cycles of denaturation at 95°C for 30 sec, annealing at primer specific temperature for 30 sec and extension at 72°C for 1 min (product size < 1 kb), followed by final extension step at 72°C for 7 min. Both DIG-labeled and unlabeled PCR products (3 µl of each) were analyzed by resolving on an agarose gel. Due to multiple

incorporations of DIG-dUTP during PCR, the molecular weight of the DIG-labeled PCR product increases significantly compared to the unlabeled PCR product which was also observed on the agarose gel.

3.10.4.7 Pre-hybridization, hybridization, and immunodetection

Pre-hybridization and hybridization of the DIG-labelled probes were carried out using DIG DNA Labeling and Detection Kit (Roche, 11093657910) following the manufacturer's protocol. The membrane from 2X SSC buffer was placed inside the hybridization bottle and pre-hybridized with 15 ml of pre-warmed pre-hybridization solution {15 ml DIG Easy Hyb buffer + 20 µl of denatured 10 mg/ml sonicated salmon sperm DNA (sssDNA)}. Note that sssDNA was denatured by heating at 95°C for 10 min and rapidly cooled on ice prior to addition into DIG Easy Hyb buffer which was pre-warmed at 60°C. The bottle was incubated in a hybridization oven at 50°C for 2 hr. Then, the prehybridization solution was decanted and 15 ml of pre-warmed hybridization solution {15 ml DIG Easy Hyb buffer + 1 µl of denatured 10 mg/ml sssDNA + 15 µl of denatured DIG-labeled DNA probe (~25 ng/ml final concentration)} was added. Note that both sssDNA and DIG-labeled DNA probes were denatured by heating at 95°C for 10 min and rapidly cooled on ice prior to addition into DIG Easy Hyb buffer which was pre-warmed at 60°C. The bottle was incubated overnight in a hybridization oven at 50°C for ~16 hr.

After hybridization, the membrane was transferred into a glass tray and washed two times with 150 ml Wash solution I (Appendix 8, Table A8.5) for 15 min each at room temperature with gentle shaking. The membrane was then washed twice with 150 ml 65°C pre-warmed Wash solution II (Appendix 8, Table A8.6) for 15 min each at 65°C on a gently shaking water-bath. The membrane was rinsed for 5 min in 100 ml of Wash buffer (Appendix 8, Table A8.7, A8.8). Next, the membrane was incubated in 100 ml of 1X Blocking solution (Appendix 8, Table A8.9) for 30 min followed by incubation with 30 ml of antibody solution {3 µl Anti-DIG-conjugated with Alkaline Phosphatase in 30 ml of 1X Blocking solution (final concentration = 75 mU/ml)} with gentle shaking. Then, the membrane was washed twice again with 100 ml of Wash buffer for 15 min each followed by equilibration in 50 ml of Detection buffer for 5 min (Appendix 8, Table A8.10).

The detection of the DIG-labeled probes was carried out using CSPD, a chemiluminescent substrate (Roche, 11755633001). The equilibrated membrane was placed on a development folder with DNA side facing up and quickly applied ~ 20 drops CSPD out of

the dropper bottle uniformly covering the membrane area. The membrane was immediately covered with the second sheet of the folder, spreading the substrate evenly and avoiding air bubbles over the membrane. It was then incubated for 5 min in the dark at room temperature. Then, the excess liquid was squeezed out by gently rolling a pipette on top of it, and the development folder was sealed completely by a sealing machine. Next, the membrane was incubated at 37°C for 10 min in the dark and then exposed to X-ray film enclosed in an X-ray film cassette at room temperature. The X-ray film was developed first, after 1 hr of exposure by sequentially immersing and gently agitating the film, while holding on its two opposite corners, into the plastic trays containing Developing solution (Appendix A8.1) for 2 min, water for 1 min, fixing solution (Appendix A8.2) for 2 min and lastly in the water again for 1 min. Based on the signal intensity in film developed after 1 hr exposure, a shorter or longer exposure time for another X-ray film, if required, was carried out.

3.10.5 Screening of T₁ and T₂ transgenic lines

The progeny of independent transgenic lines with a single transgene integration number as identified by Southern blot were raised to T₁ and T₂ generations to obtain their homozygous lines. The T₁ seeds were screened by germinating on Hygromycin selection (30 mg/l, prepared in Milli-Q water). Briefly, the seeds were treated with Bavistin as detailed in section 3.1.1, soaked in water for 24 hr and then placed on 3 mm Whatman paper soaked with 5 ml of Hygromycin selection in a petri-plate. Accordingly, the seeds were also placed on only water or no Hygromycin selection plate as a control to check germination percentage. T₁ progeny showing 3:1 transgene segregation on hygromycin were transferred to soil. They were further screened by tissue PCR as performed in section 3.10.2. T₂ seeds were collected from each positive T₁ plant and were screened again on Hygromycin selection along with the no Hygromycin control. T₂ lines with 100% germination on hygromycin were considered homozygous and were further screened by tissue PCR. Accordingly, the T₂ homozygous progenies from three independent transgenic lines were identified and proceeded further for their growth analysis.

3.10.6 Expression analysis of transgene by qRT-PCR

The expressions of the transgene in T₂ homozygous transgenic plants overexpressing either *OsCBSX3* or *OsCBSCLC6* were determined by performing qRT-PCR. Total RNA was isolated from leaves of transgenic plants and the WT controls, first-strand cDNA was synthesized

followed by qRT-PCR, essentially following the same procedure as detailed in sections 3.1.2 to 3.1.4.

3.10.7 Extraction of plant protein

Extraction of plant protein was carried out following the method detailed by Zivy et al (1983). Briefly, 100 mg of leaf tissue in the presence of 10 mg Polyvinylpolypyrrolidone was homogenized to a fine powder form in liquid nitrogen. The powder was resuspended in 2 ml of protein extraction buffer (Appendix 4, Table A4.1) followed by two rounds of centrifugation at 12,000 rpm for 15 min at 4°C. Onto the supernatant, 8 volumes of prechilled acetone containing 10 mM of β -mercaptoethanol was added and incubated overnight at -20°C. The precipitated protein was then pelleted by centrifuging at 8,000 rpm, air-dried, and resuspended in 200 μ l of Laemmli buffer (Appendix 5, Table A5.1) followed by heating at 100 °C for 5 min. The solubilized protein was proceeded for Bradford quantification and immune-blot assay (see section 3.6.3).

3.11 Plant growth analysis under control and stress conditions

3.11.1 Plant growth and stress treatments

The T₂ homozygous independent transgenic lines along with the WT plants were grown on soil pots for their growth studies under controlled and stressed conditions. The study was conducted in a greenhouse maintained at 28 \pm 2°C temperature and 14 h light/ 10h dark cycle. For growth under controlled conditions, the plants were grown till maturity with normal irrigation. In case of stresses, the transgenic plants along with the WT were subjected to drought and salinity at the panicle emergence stage. The salinity stress was imposed by immersing the pot containing plant in 150 mM NaCl solution for 40 days. The salt-water level in the tray was maintained regularly by maintaining the lost volume with water. The drought stress was induced by withholding the irrigation for 14 days (volumetric water content of the soil was reduced to 20%). After respective periods of stress treatment, the stress was discontinued by removing the salt solution followed by normal watering (in case of salinity stress) or by resuming normal irrigation (in case of drought stress), and the plants were allowed to recover.

In salinity stress, the soil electrical conductivity (EC) was determined following the method described by Kargas et al (2020) with minor modification. Soil samples (5 cm depth from the top) were collected, oven-dried at 50°C for 24 h, powdered by mortar and pestle, and passed through a 2 mm sieve. Then, 5 g of the finely powdered soil was transferred into a 50

ml polypropylene tube and added deionized water onto it up to a volume of 25 ml, making a suspension of 1:5 (w/v). The soil suspension was mixed well by vortexing and kept on a shaker for 1 hr. The soil particles were then allowed to settle down for 20 min. The EC (1:5) as well as the temperature of extract (EC increases proportionally with the temperature), without disturbing the settled soil particles, were then determined using TetraCon 325 sensor connected to the Handheld meter Cond 340i (WTW). Note that prior to Soil EC measurement, the instrument was calibrated with 0.01 M KCl which should have an EC of approximately 1.413 dS/m. Finally, the obtained EC (1:5) of soil extract was converted to EC of soil saturated paste extract (EC_e) using the formula, $EC_e = f \cdot EC(1:5)$, where $f = 5.46$ in case of $EC_e < 3$ dS, while it was 6.71 when $EC_e > 3$ dS/m (Kargas et al, 2020). For drought stress, the soil moisture tension was determined by using WaterScout SM100 Soil Moisture Sensor connected to FieldScout Soil Sensor Reader (Spectrum Technologies, Inc.). Prior to measuring soil moisture, the instrument was calibrated in Milli-Q water.

3.11.2 Plant physiological parameter analyses

3.11.2.1 Leaf relative water content

Relative water content (RWC) of the leaf was measured following the method described by Yamasaki and Dillenburg (1999). RWC was taken for the penultimate leaf. Briefly, each leaf strip (approximately 5 cm from the tip) was cut with a scissor and transferred into a pre-weighed 15 ml polypropylene tube. The tube was tightly capped to prevent the loss of water vapor. Immediately after taking samples to the lab, the weight of each tube containing leaf strip was noted from which the fresh weight (FW) of each sample was determined. After that, deionized water was added into each tube to completely immerse the leaf sample, tied the cap and the tube was incubated overnight at room temperature. The next day, the leaf strip was taken out of the tube, gently wiped with tissue paper to remove water from the leaf surface and its turgid weight (TW) was determined. The leaf strip was then transferred into a new dry tube with no cap tied and incubated at 60°C in a dry oven for 48 hr. The dry weight (DW) of the sample was then determined. Finally, through FW, TW, and DW, the RWC of each sample was calculated using the following equation: $RWC (\%) = \{(FW - DW)/(TW - DW)\} \times 100$.

3.11.2.2 Leaf electrolyte leakage

The leaf electrolyte leakage measurement was performed following the method described by Lutts et al (1996) with minor modifications. Briefly, the leaf strip (approximately 5 cm of the

penultimate leaf) was cut and transferred immediately into a pre-weighed 15 ml polypropylene tube to determine its weight. The leaf strip was then rinsed with deionized water, immersed completely in deionized water to a ratio of 1 mg/ml, and incubated overnight at room temperature. After 12 hr, the electrical conductivity (EC) of the solution (Lt) was measured. The samples were then autoclaved, cooled down to room temperature, and the final EC of the solution (L0) was measured. The electrolyte leakage was then calculated as the ratio of Lt and L0 and was expressed in percentage.

3.11.2.3 Chlorophyll-a fluorescence analysis using Handy PEA

Plant photosynthetic performance was determined based on chlorophyll-a fluorescence kinetics of the penultimate leaf by using Handy PEA (Hansatech Instruments). The method as detailed by Wungrampha et al (2019) was followed for optimization of the fluorescence measurement and the subsequent analyses. The leaves were dark-adapted with the accessory clips for 20 min prior to chlorophyll-a fluorescence measurement. First, light intensity to be used for rice (var. PB1) was optimized by measuring the Fv/Fm (a photosynthetic efficiency parameter) from dark-adapted leaves exposed to different intensities of 650 nm actinic light ranging from 2,700 to 3,100 $\mu\text{mol photons m}^{-2} \text{sec}^{-1}$ for 10 sec. The highest Fv/Fm value for rice leaf was obtained at 3,000 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ light intensity, hence the subsequent chlorophyll-a fluorescence measurements were taken at 3,000 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$.

3.11.2.4 Leaf gas exchange analysis using IRGA

Leaf gas exchange parameters, including photosynthesis rate, transpiration rate, and conductance to CO₂ and H₂O, were measured using IRGA (Infra-red Gas Analyzer; LI-6400XT, Licor). All the measurements were taken from the penultimate leaf. The instrument was set to 1000 $\mu\text{mol m}^{-2} \text{sec}^{-1}$ for in-chamber photosynthetically active radiation (PARi), 400 $\mu\text{mol CO}_2 \text{mol}^{-1}$ for reference CO₂ concentration (CO₂R), 500 $\mu\text{mol sec}^{-1}$ for flow rate (Flow), for all the measurements taken.

3.12 Generation and growth analysis of transgenic *Arabidopsis* overexpressing *OsCBSX3*

OsCBSX3 overexpressing transgenic *Arabidopsis thaliana* {Columbia-0 (Col-0) ecotype} was generated by *Agrobacterium*-mediated floral dip transformation method as described by Zhang et al (2006) with few changes. The same *OsCBSX3*_pCAMBIA1302 (without GFP fusion) that was used for rice transformation (section 3.10.1) was transformed into *Agrobacterium* GV3101 strain, which was subsequently used for the transformation of *Arabidopsis*. For growing the

Arabidopsis plants, seeds were transferred into a microcentrifuge tube and treated with sterilization solution {0.1% mercuric chloride (HgCl₂) and 0.02% of Triton X-100} for 10 min, followed by several washing steps with sterile water until the froth stopped forming. The seeds, while still immersed in water, were then stratified at 4°C in the dark for 3 days. After stratification, the seeds were placed onto a 0.5X MS medium (0.5X MS medium, 1% Sucrose, 0.3% Phytigel; pH 5.7) using a 20 µl pipette with the cut tip. The plates were then incubated in an *Arabidopsis* growth room maintained at 18°C temperature and 14 hr light and 10 hr dark cycle. After 2 weeks, the seedlings were transferred onto potting mixture {vermiculite and vermicompost (2:1)}, sowing 5 seedlings per pot (3 x 3 inches). The pots were placed on a plastic tray and for the first week of transplanting the entire plants in the tray were covered with a transparent cover. When the plants attained the reproductive stage, the first bolt was clipped off to induce the proliferation of more secondary inflorescences, and about a week later, the *Agrobacterium* transformation was performed.

As detailed in section 3.4.10.1, the *Agrobacterium* GV3101 cells harboring recombinant plasmid were cultured in YEM broth supplemented with antibiotics (Gentamycin, Rifampicin, and Kanamycin). However, the secondary culture was scaled up to 1 L for the transformation of 12-15 pots of *Arabidopsis*, and the cells were cultured till the OD₆₀₀ reached about 1.2 - 1.5. The pelleted cells were resuspended in 1 L of infiltration medium (0.5X MS, 5% Sucrose, and 0.02% Silwet L-77; pH 5.7). The *Agrobacterium* suspension was added in a magenta box placed in a vacuum desiccator and the *Arabidopsis* pots were inverted with the inflorescences dipped into the suspension, followed by vacuum infiltration for 3 min. The *Agrobacterium* suspension was changed after infiltration of 5-6 pots on average. The plants were then wrapped with a plastic film and laid in a horizontal position for 24 hr. Then the plastic film was removed, and plants were grown normally.

Seeds obtained from the transformed plants were sterilized and stratified as mentioned above and grown on a 0.5X MS plate containing 25 mg/l Hygromycin selection. The Hygromycin positive putative transformed plants were screened by tissue PCR (as described in section 3.10.3). Hygromycin and tissue PCR screening was followed till T₃ generation to obtain homozygous transgenic lines. The expression of the transgene was confirmed by qRT-PCR. The growth and development of T₃ homozygous transgenic lines were analyzed.

For *in vitro* growth analysis of the plants at the seedling stage, the five-day-old seedlings germinated on 0.5X MS medium were transferred onto 0.5X MS medium

supplemented with 100 mM NaCl, 100 mM Mannitol, and 1 mM H₂O₂ for imposing salinity, osmotic, and oxidative stresses, respectively in square plates. The plants grown in vermiculite pots were imposed to salinity and drought stress at the transition of vegetative to reproductive phase of the plants by irrigating with 150 mM NaCl solution and by withholding irrigation, respectively.

3.13 Primer designing

All the primers used in the study were designed using the Primer3Plus online tool (<https://primer3plus.com/cgi-bin/dev/primer3plus.cgi>) and the specificity of the primer pair was analyzed by the Primer-Blast online tool (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi>). Sequences of the primers are listed in Appendix 9, Table A9.1.

3.14 Heat map preparation

All the heatmaps produced in the study were prepared using MultiExperiment Viewer (MeV) 4.9.0 tool (Saeed et al, 2003).

3.15 Multiple alignment and phylogenetic tree

Multiple alignments of the sequences and the phylogenetic analyses were performed using the MEGA X tool (Kumar et al, 2018).

3.16 Statistical analysis

The statistical significance in the data was calculated by unpaired, two-tailed Student's *t*-test using GraphPad Prism version 5.02.

4. Results

To understand the basic cellular functions of two distinct members of a CDCP superfamily, *OsCBSX3* (a polypeptide containing only a single pair of CBS domains) and *OsCBSCLC6* (a member of anion transport proteins containing a pair of CBS domains), we began our study with their gene expression profiling, gene amplification, and subcellular localization of the respective proteins. Next, we attempted to understand their biological functions through expression in a bacterial system, sequence analyses, yeast two-hybrid assays, and morphological and physiological analyses of the transgenic plants overexpressing these two CDCPs. This chapter presents the results obtained from these studies on *OsCBSX3* and *OsCBSCLC6* in two major sections, 4.1 and 4.2, respectively.

4.1 Functional characterization of *OsCBSX3*

4.1.1 Expression analysis of *OsCBSX3*

The expression of *OsCBSX3* (LOC_Os02g57280.1) based on qRT-PCR was analyzed in two contrasting rice genotypes, viz., IR64 (salt-sensitive) and Pokkali (salt-tolerant), in different tissues at different developmental stages. Its expression was also studied in these two rice genotypes under various stress treatments at the seedling stage. Interestingly, *OsCBSX3* completely overlaps with the 3' region of *OsCYTOCHROME P450* (*OsCYP450*; LOC_Os02g57290; Fig. 4.1) at the opposite strand in an antiparallel overlapping arrangement. Transcripts from such antiparallel overlapping genes (referred to as Natural Antisense Transcripts / NATs) potentially form RNA duplexes, and the transcriptional and post-transcriptional functions of such overlapping genes are known to be regulated by one another through various modes of interferences (Boi et al, 2004; Latgé et al, 2018). Therefore, to assess if such regulation exists between *OsCBSX3* and *OsCYP450* at the transcript level, we also analyzed the expression of *OsCYP450* in this study. For qRT-PCR amplification of the products specific to *OsCBSX3* transcripts, the exon-exon junction spanning primers (both forward and reverse) were designed (specific to LOC_Os02g57280.1 isoform of *OsCBSX3*; Fig. 4.1). Whereas qRT-PCR primers for *OsCYP450* transcripts were designed from two different exonic regions which do not overlap with *OsCBSX3*. The specificity of each primer set was confirmed by a single melt curve peak in qRT-PCR and a single band of the appropriate size on an agarose gel.

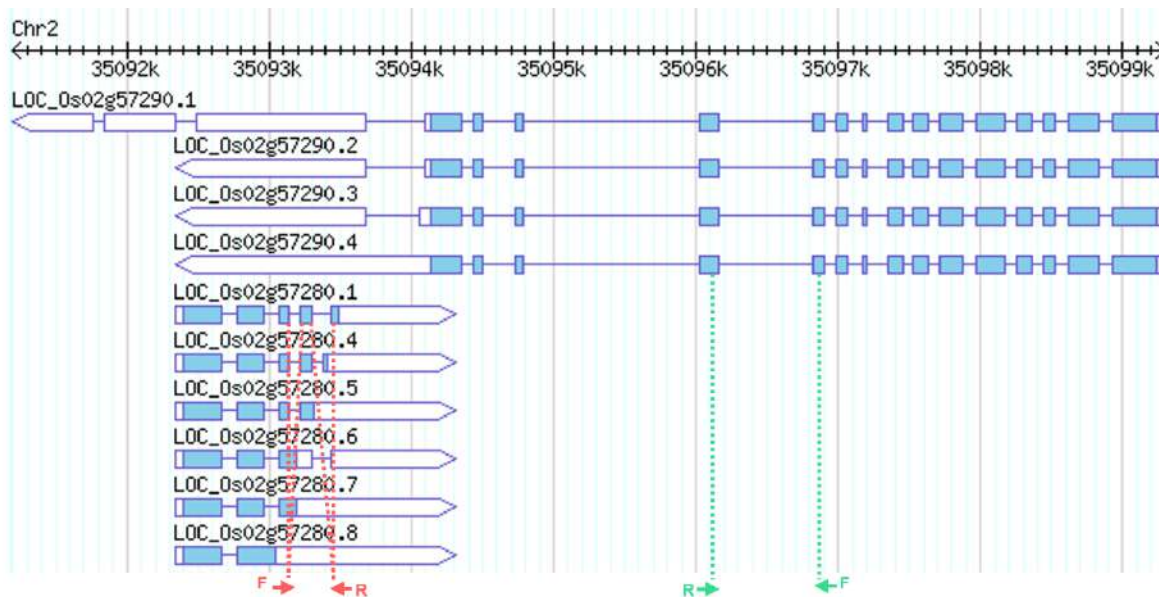


Fig. 4.1: Genomic organization of *OsCBSX3* and *OsCYP450*, and position of the primers designed specific to *OsCBSX3* and *OsCYP450* transcripts for their expression analysis by qRT-PCR. *OsCBSX3* (LOC_Os02g57280) overlaps with *OsCYP450* (LOC_Os02g57290) in an anti-parallel arrangement. Blue coloured boxes represent the exons connected by blue lines representing introns. White boxes with blunt ends on both sides represent 5' UTR, whereas that with pointed end at one side and blunt on the other represent 3'UTR. The exonic region of *OsCBSX3* corresponding to its exon-exon spanning qRT-PCR primers are presented with red dotted lines. Similarly, the exonic region of *OsCYP450* corresponding to its qRT-PCR primers are presented with green dotted lines. 'F' and 'R' represent forward and reverse primers.

During all three developmental stages, viz., seedling, tillering, and early flowering, the expression of both *OsCBSX3* and *OsCYP450* was observed to be significantly higher in shoots than in roots (Fig. 4.2 a, b). The transcript level of both the genes elevated in the shoots as plants progressed from seedling to tillering and flowering stages of development. The panicle at the early flowering stage also exhibited expression of these two genes, although less than their levels in leaf, but higher than in the root tissues. The transcript level of *OsCBSX3* appeared to be similar in both Pokkali and IR64; however, *OsCYP450* transcripts appeared to be higher in the shoots of Pokkali than in IR64 at tillering and flowering stages. Although the overall expression pattern of *OsCBSX3* and *OsCYP450* was observed to be similar across different tissues, the transcript level of *OsCBSX3* was significantly higher than that of *OsCYP450* in the shoot tissues at all three developmental stages. The *in-silico* expression analysis based on mRNA-seq datasets obtained from the Genevestigator tool also exhibited higher transcript levels of *OsCBSX3* and *OsCYP450* in the shoots with transcripts of the former being significantly higher than that of the latter (Fig. 4.2 c). In addition, our qRT-PCR based expression analysis in the shoots of both Pokkali and IR64 seedlings showed expression of *OsCBSX3* to be diurnally regulated with significantly higher expression during the day (Fig. 4.2 d). While *OsCYP450* transcripts also appeared to be higher during the day, the difference

was not statistically significant (Fig. 4.2 e). Consistently, the available Affymetrix GeneChip datasets from Genevestigator also justified the higher expression of *OsCBSX3* during the day than at night (Fig. 4.2 f).

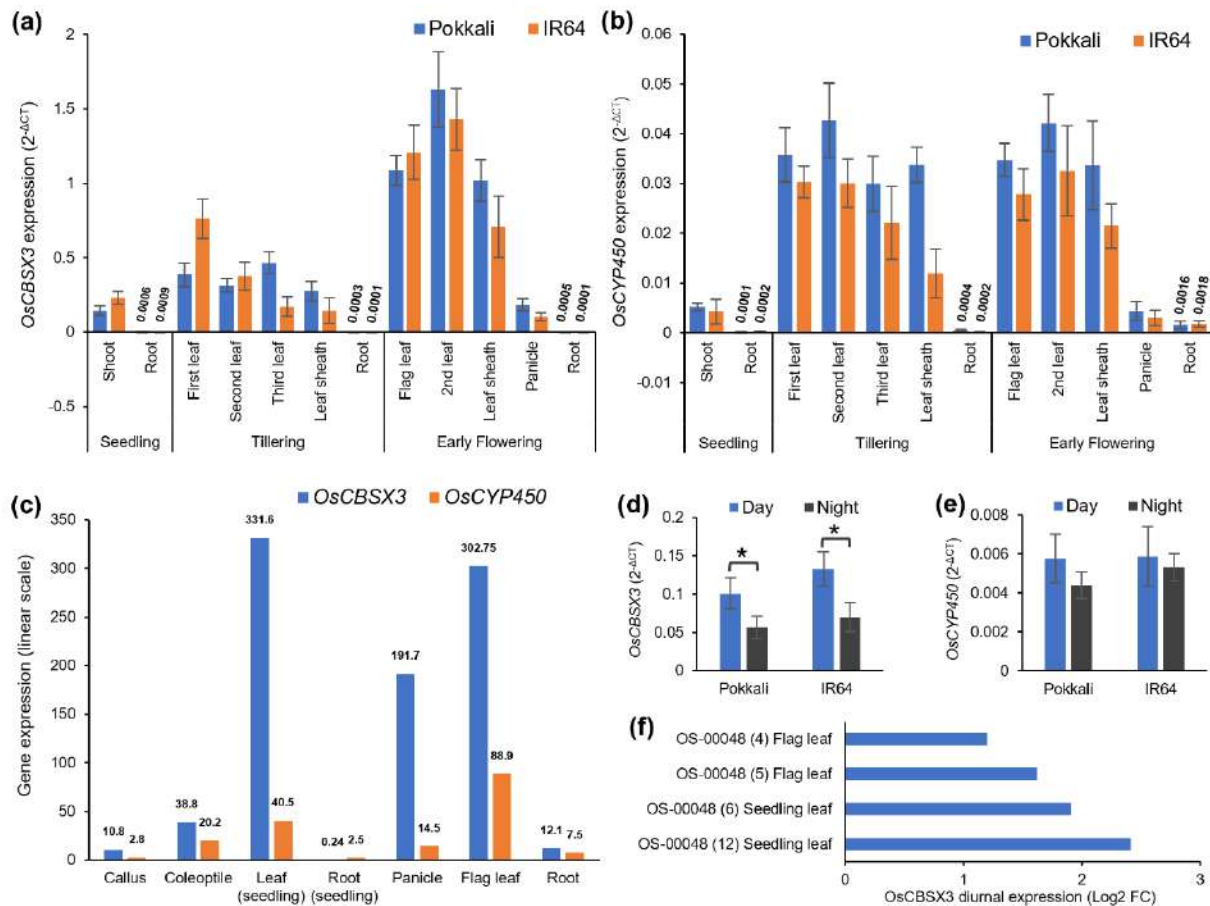


Fig. 4.2: Transcript abundance of *OsCBSX3* and *OsCYP450* in different tissues at the seedling, tillering, and flowering stages of rice. Bar graphs in (a) and (b) represent the qRT-PCR based expression of *OsCBSX3* and *OsCYP450*, respectively, in IR64 and Pokkali rice plants, in various tissues and at different developmental stages, viz., shoot and root tissues of 2-week-old seedlings; top three leaves, leaf sheath of the third leaf and root tissues at tillering stage; and the flag leaf, penultimate leaf, and leaf sheath of the penultimate leaf, the panicle and the root tissues at flowering stage. The expression level is expressed as $2^{-\Delta CT}$. (c) Normalized transcript level of *OsCBSX3* and *OsCYP450* in different tissues of rice based on available mRNA-seq datasets from Genevestigator tool. The transcript level is expressed in linear scale as it is presented in the Genevestigator. Bar graphs in (d) and (e) represent the qRT-PCR based diurnal expression pattern of *OsCBSX3* and *OsCYP450*, respectively, in shoot tissues of 2-week-old seedlings of Pokkali and IR64 (expressed as $2^{-\Delta CT}$). *: $P \leq 0.05$ (calculated by unpaired, two-tailed Student's *t*-test). (f) Expression of *OsCBSX3* during the day relative to its expression at night based on the available Affymetrix GeneChip datasets from the Genevestigator ($P < 0.05$). Note that no data was available for diurnal expression of *OsCYP450* at $P < 0.05$. In (a), (b), (d), and (e), the bar represents mean \pm SD, $n=3$.

Expression analysis based on qRT-PCR under different stress treatments in shoot tissues of rice seedlings (Fig. 4.3 a) exhibited *OsCBSX3* to be inducible to desiccation (air-dry), salinity (200 mM NaCl), cold (4°C), and heat (42°C) treatments in Pokkali. It should be noted that the gene expression under stress treatments are expressed as Log₂ of $2^{-\Delta\Delta CT}$ or Log₂ of Fold Change (FC) relative to their expression under control condition, and only the Log₂ FC values of >2.0 and <2.0 are considered as significantly upregulated and downregulated, respectively in this study. The *OsCBSX3* expression was also upregulated, however later after 12 hr, by oxidative (10 mM H₂O₂) and salicylic acid (100 μM) treatments, whereas no significant change was noticed under heavy metal (100 μM CuCl₂) treatment. The upregulation of *OsCBSX3* was observed in IR64 as well, however only in response to desiccation and at later time-points in cold stress. In contrast, the expression of *OsCYP450* was downregulated in response to desiccation and heat treatments in the shoots of IR64 seedlings, and at the early time point (2 hr) of desiccation and salinity stress in Pokkali. Whereas in response to cold, heavy metal, and salicylic acid treatments, the *OsCYP450* showed upregulation in Pokkali but not in IR64 shoots.

We also performed the *in-silico* expression analysis of *OsCBSX3* and *OsCYP450* under various stress treatments using Affymetrix GeneChip and mRNA-seq datasets available in Genevestigator ($P < 0.05$; Fig. 4.3 b). Consistent with our qRT-PCR based expression profile, *OsCBSX3* was identified to be upregulated by salinity, desiccation, and heat stresses, except for the cold treatment that its expression was found to be downregulated instead. This analysis additionally identified *OsCBSX3* expression to be induced under submergence and anoxic conditions, while downregulated by inoculation of the fungal pathogen. Interestingly, the expression of *OsCBSX3* and *OsCYP450* in these expression datasets also exhibited an inverse expression pattern between these two genes with the downregulation of *OsCYP450* upon upregulation of *OsCBSX3*, specifically in response to drought, salinity, and submergence treatments.

Taking together these results from expression studies, we infer that *OsCBSX3* is predominantly expressed in shoot tissues and is responsive to various stresses. In relation to *OsCYP450*, overlapping with *OsCBSX3* on the opposite strand, the higher expression of *OsCBSX3* appears to have a negative effect on the expression of *OsCYP450*.

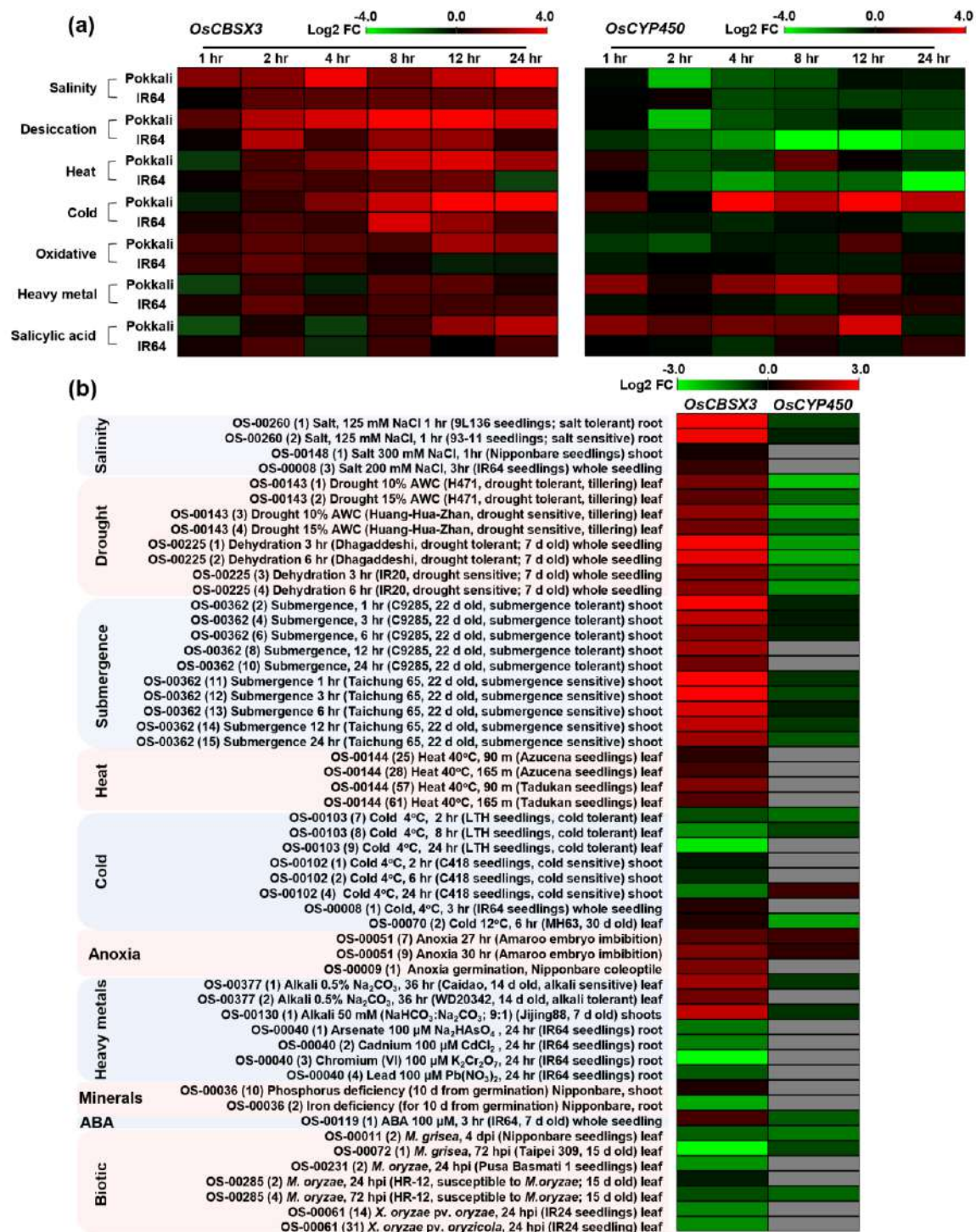


Fig. 4.3: Expression pattern of *OsCBSX3* and *OsCYP450* in response to various stress treatments in rice. (a) Heat-map representing the qRT-PCR based transcript abundance of *OsCBSX3* (left panel) and *OsCYP450* (right panel) in the shoot tissues from 2-week-old seedlings of IR64 and Pokkali, subjected to different stress treatments and at different time points. Expression level is expressed as an average of Log₂ of fold change (Log₂ FC) from three replicates (n=3). (b) Expression level of *OsCBSX3* and *OsCYP450* in different tissues of rice plants subjected to various stress treatments, based on both the Affymetrix GeneChip and the mRNA-seq datasets ($P < 0.05$) retrieved from Genevestigator tool. Grey boxes denote the insignificant value ($P > 0.05$). The transcript level is expressed as Log₂ FC. The Heat-maps were generated using MeV 4.9.0 tool.

4.1.2 Subcellular localization of OsCBSX3

The *in-silico* protein subcellular localization prediction by TargetP-1.1 server (Emanuelsson et al, 2000) anticipated OsCBSX3 to possess a chloroplast transit peptide and localize in the chloroplast (Fig. 4.4 a). It also predicted a potential cleavage site of the transit peptide at N-terminal 35th amino acid of the protein. To validate it *in-vivo*, the coding sequence (CDS) of *OsCBSX3* was cloned into the pCAMBIA1302 vector in-frame with the reporter gene for GFP, under CaMV 35S promoter. The OsCBSX3-GFP fusion protein was transiently expressed in tobacco leaf epidermal cells (*Nicotiana benthamiana*) through *Agrobacterium* (*A. tumefaciens*, GV3101 strain) infiltration method (Sparkes et al, 2006). Under a confocal microscope, the GFP signal was observed distinctly in the chloroplast for OsCBSX3-GFP fusion, which merged well with the autofluorescence signal from chloroplast (Fig. 4.4 b). However, the leaf infiltrated with the empty pCAMBIA1302 vector emitted a GFP signal in the cytoplasm.

The chloroplastic localization of OsCBSX3 was further confirmed in the protoplast isolated from stem and sheath region of rice seedlings (var. PB-1) stably overexpressing OsCBSX3-GFP fusion under CaMV 35S promoter. These stable transgenic plants were generated through *Agrobacterium tumefaciens* (LBA4404 strain) mediated transformation method (detailed in section 4.1.4). Observation under the confocal microscope detected significant fluorescence in the OsCBSX3-GFP overexpressing rice protoplast, that merged with the chloroplast autofluorescence (Fig. 4.4 c). However, no significant fluorescence emission was observed in the protoplast isolated from wild-type (WT) plants. Taken together, these experiments confirmed that OsCBSX3 is a chloroplast localized protein.

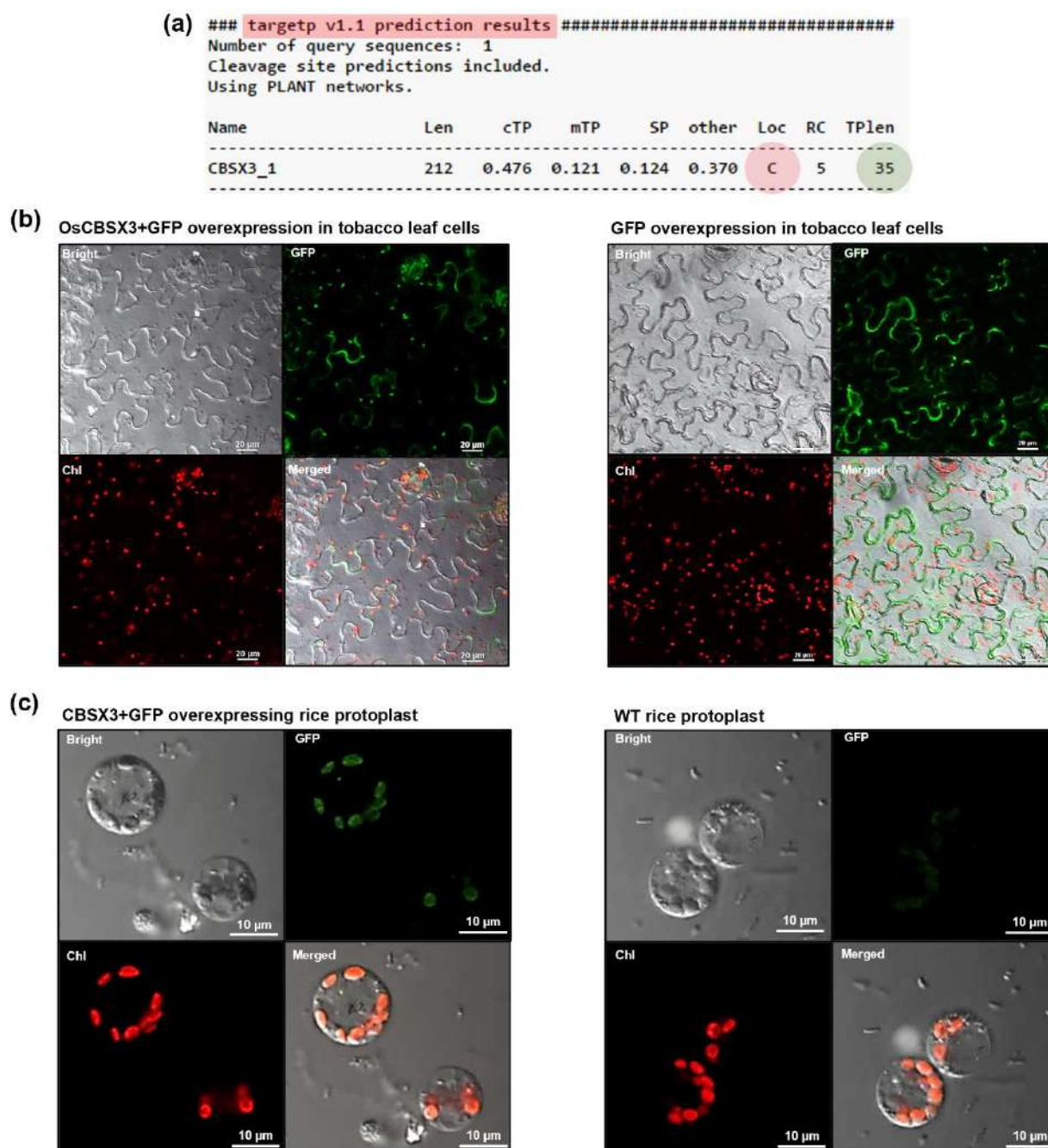


Fig. 4.4: Subcellular localization of OsCBSX3. (a) *In silico* sequence analysis using TargetP-1.1 predicted OsCBSX3 to possess a chloroplast transit peptide (cTP) and localize in the chloroplast (C; highlighted in red), with a potential cleavage site of the transit peptide at N-terminal 35th amino acid of the protein (highlighted in green). (b) Confocal microscopic images depicting the emission of GFP fluorescence from the tobacco leaf epidermal cells (*N. benthamiana*) transiently overexpressing OsCBSX3+GFP fusion protein (left panel) and only GFP (right panel). OsCBSX3, cloned in-frame with downstream CDS for GFP in pCAMBIA1302 vector was used for the overexpression of OsCBSX3+GFP fusion, while empty pCAMBIA1302 vector was used as a control overexpressing only GFP. Scale bar = 20 μ m. (c) Confocal microscopic images exhibiting the GFP fluorescence emission from rice protoplasts isolated from transgenic plants stably overexpressing OsCBSX3+GFP fusion protein (left panel) and the WT plants (right panel). Scale bar = 10 μ m. In both (b) and (c), 'Chl' denotes the chlorophyll autofluorescence signal.

4.1.3 Biochemical characterization of OsCBSX3 and identification of its interacting proteins

Two Bateman modules constituting a pair of CBS domains associate to form a CBS module, resulting in the formation of a homodimer in the case of CDCPs with a single pair of CBS domains. In each CBS module, its four CBS domains are known to establish four potential ligand binding sites, which most commonly accommodate adenosine nucleotides (AMP/ADP/ATP). To characterize such homo-oligomeric status and ligand binding affinity of OsCBSX3, we expressed it in *E. coli*, purified it, and assayed for these properties. We also carried out screening of the rice cDNA library through yeast-two hybrid assay to find the proteins interacting with OsCBSX3.

4.1.3.1 Expression and purification of OsCBSX3 in *E. coli*

For the heterologous expression of OsCBSX3 in *E. coli*, the CDS of *OsCBSX3* was cloned into pET28a (bacterial expression vector) at *Bam*HI and *Eco*RI restriction sites, to produce OsCBSX3 with the Histidine tag at its N-terminus (6X-His_OsCBSX3; Fig. 4.5). The recombinant *OsCBSX3_pET28a* plasmid was stabilized in *E. coli* (Top10) cloning strain and subsequently transformed into *E. coli* BL21 (DE3) protein expression strain. The protein of desired size (6X-His_OsCBSX3 = 26.75 kDa) was identified to be expressed on induction of the *E. coli* culture with IPTG (0.1 mM) for 4 hr at 37°C (Fig. 4.6 a). However, SDS-PAGE and western blot analysis of insoluble and soluble fractions of the lysed cells showed 6X-His_OsCBSX3 to aggregate in the insoluble fraction (Fig. 4.6 b, c). To obtain 6X-His_OsCBSX3 in the soluble fraction, induction at lower incubation temperatures (18 °C and 28 °C) was attempted, nonetheless the 6X-His_OsCBSX3 still aggregated into the inclusion bodies. Alternatively, we followed *in vitro* re-folding approach by solubilizing the inclusion bodies in a buffer containing a high concentration of chaotropic agents (Guanidine hydrochloride and Urea), followed by purification of the recombinant 6X-His_OsCBSX3 by Ni-NTA affinity chromatography (Fig. 4.6 d) and its subsequent refolding by gradual removal of chaotropes through dialysis. However, 6X-His_OsCBSX3 started to aggregate as we gradually reduced the concentration of urea to 2 M in the dialysis buffer.

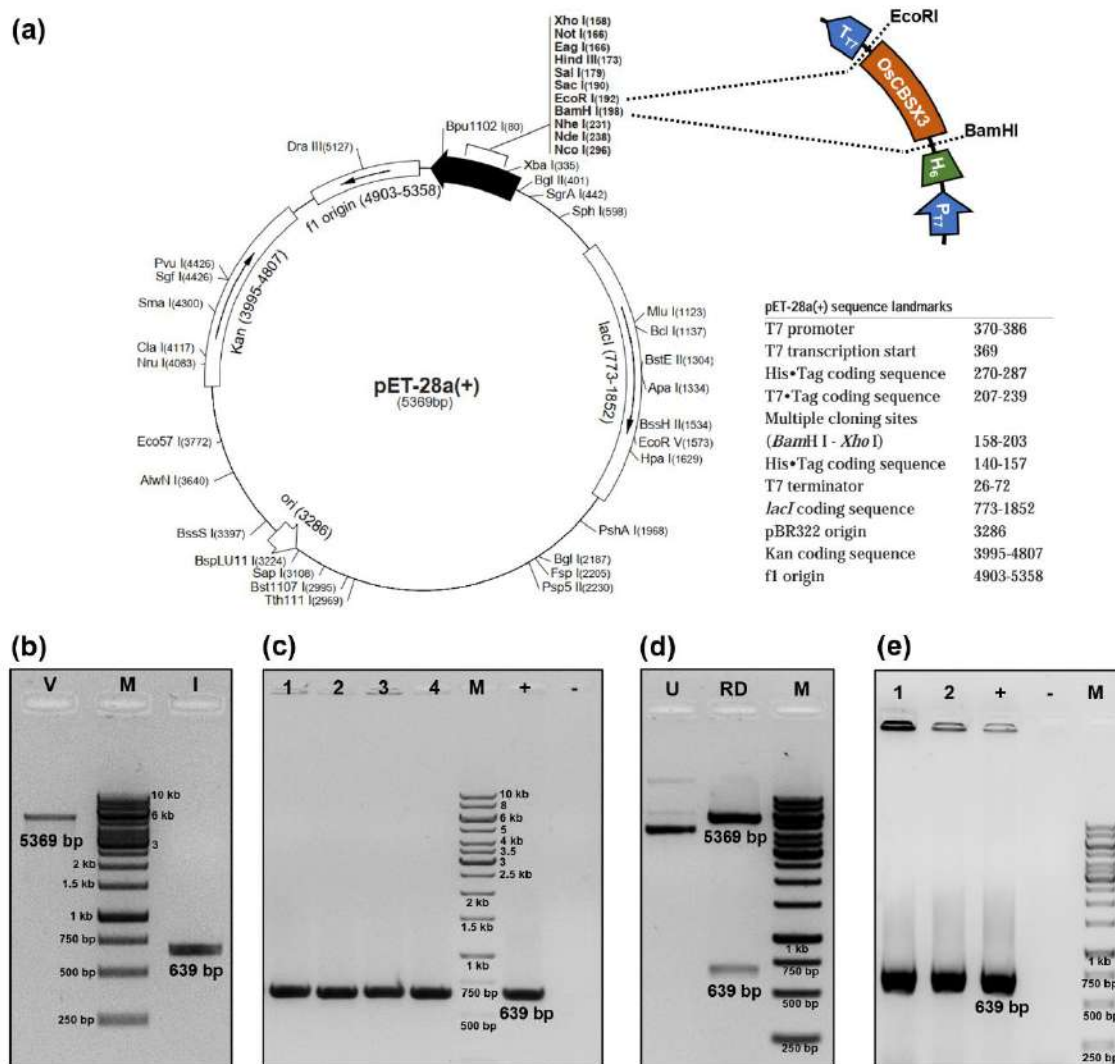


Fig. 4.5: Cloning of *OsCBSX3* into pET28a bacterial expression vector. (a) Schematic representation of pET28a vector (Vector map is acquired from Nova gen, TB074 12/98). The *OsCBSX3* insertion site (in-frame with the upstream CDS for 6X His tag; H₆) and the final expression cassette is depicted as an extension from the vector map (P_{T7}: T7 promoter; T_{T7}: T7 terminator). (b) Vector (V; pET28a) and insert (I; *OsCBSX3*), examined on an agarose gel after restriction digestion and purification, but prior to ligation reaction. Note that the vector is digested sequentially, first with *Bam*HI, followed by purification and second digestion with *Eco*RI, whereas the insert is digested simultaneously with *Bam*HI and *Eco*RI (double digestion). The vector and insert are then ligated and transformed into the competent cells of *E. coli* (Top10). (c) Colony PCR screening of *OsCBSX3*_pET28a transformed Top10 clones. (d) Confirmation of the PCR positive clone by restriction digestion of the isolated plasmid (U: undigested plasmid; RD: plasmid digested with *Bam*HI and *Eco*RI). The recombinant *OsCBSX3*_pET28a plasmid was then transformed into the competent cells of *E. coli* BL21 (DE3). (e) Colony PCR screening of *OsCBSX3*_pET28a transformed BL21 (DE3) clones. PCR positive clone was proceeded for protein expression. Transformed clones are marked numerically. 'M': DNA ladder; '+': positive control; '-': no template control.

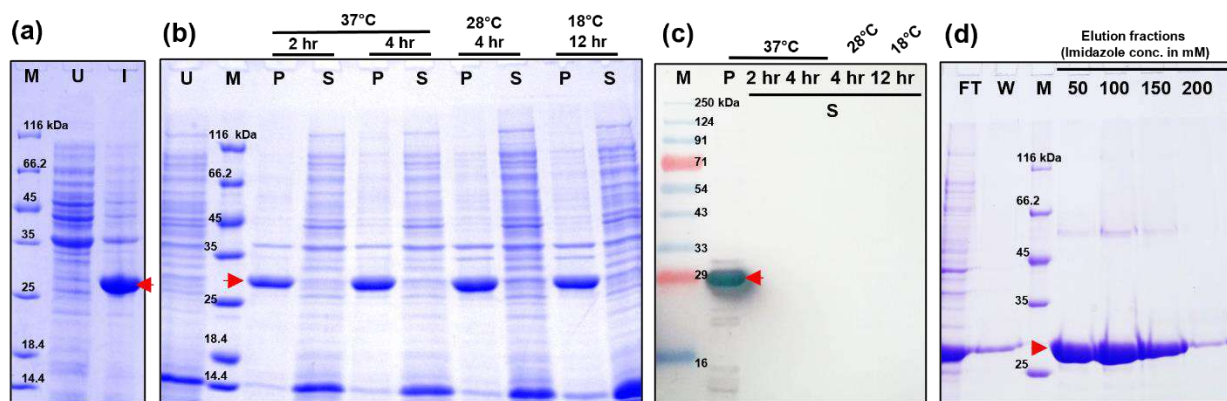


Fig. 4.6: Expression of 6XHis-OsCBSX3 in *E. coli* BL21 (DE3). (a) The total cell lysate of uninduced (U) and induced (I) cells resolved in 12% SDS-PAGE gel. (b) Total cell lysate of the uninduced (U) cells, and the insoluble (P) and soluble (S) fractions of cell lysate from induced cells resolved in SDS-PAGE gel. The cultures were induced and incubated at different temperatures and time points to improve the solubility of 6XHis-OsCBSX3. (c) Western blot of the proteins from soluble (S) fractions of cell lysate {from all four corresponding samples shown in (b)} using anti-His antibody conjugated with alkaline phosphatase. Insoluble (P) fraction from induced cells grown at 37°C for 2 hr was used as a positive control. Expression of 6XHis-OsCBSX3 was confirmed, however the protein aggregated in the insoluble fraction. (d) Recombinant 6XHis-OsCBSX3 solubilized from the insoluble fraction using Guanidine HCl and Urea, purified by Ni-NTA affinity chromatography, and resolved in SDS-PAGE gel. Proteins were eluted in an elution buffer with increasing concentrations of imidazole. The purified 6XHis-OsCBSX3 was proceeded for refolding by dialysis. ‘FT’: Flow through; ‘W’: Wash fraction. The position of induced protein band corresponding to 6XHis-OsCBSX3 (= 26.75 kDa) is shown by a red arrow. ‘M’: protein marker.

Continuing our attempts to bring OsCBSX3 in the soluble fraction, we also cloned CDS of *OsCBSX3* into pGEX-4T-1 (bacterial expression vector) at *Bam*HI and *Eco*RI site to obtain OsCBSX3 in-fusion with the Glutathione S-transferases (GST) tag at its N-terminus and expressed it in *E. coli* BL21 (DE3) (Fig. 4.7 a-d). However, this approach also did not help to obtain OsCBSX3 in the soluble fraction despite testing different induction temperatures and varied time points for bacterial culture incubation (Fig. 4.7 e).

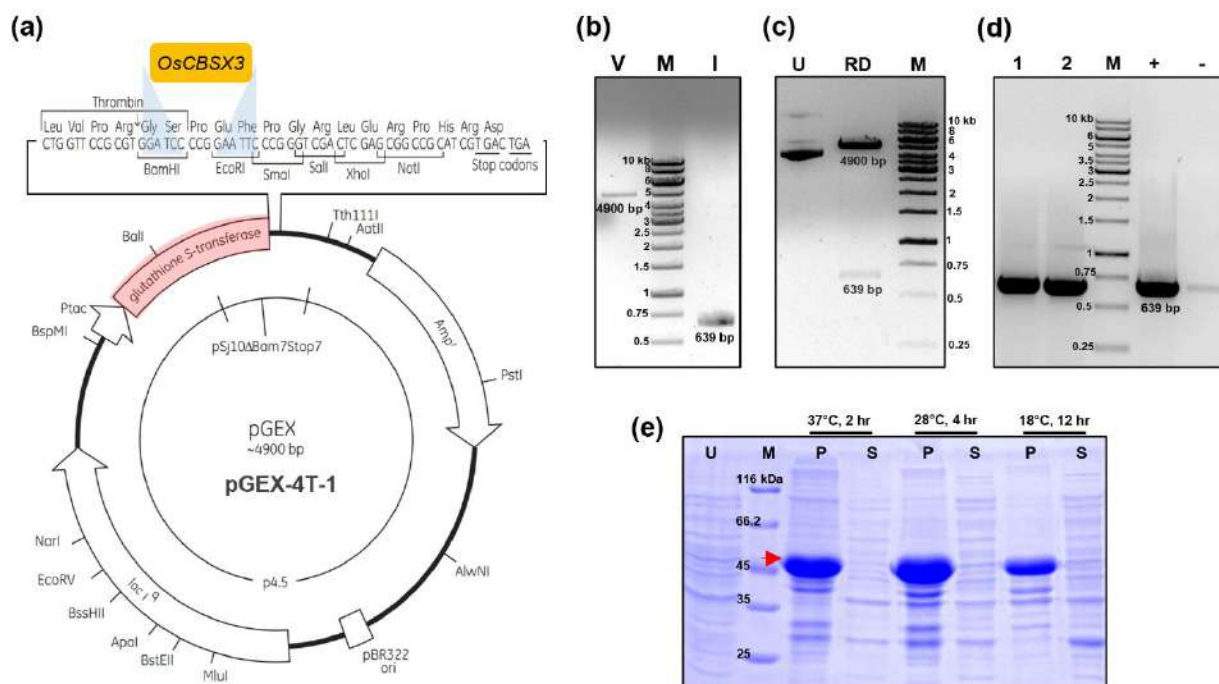


Fig. 4.7: Cloning of *OsCBSX3* into pGEX-4T-1 and expression in *E. coli* BL21 (DE3). (a) Schematic representation of pGEX-4T-1 (vector map is acquired from GE Life Sciences, 28-9191-62 AC 07/2009). The *OsCBSX3* (cloned in-frame with the upstream CDS for GST tag) and its insertion sites are depicted as an extension. (b) Vector (V; pGEX-4T-1) and insert (I; *OsCBSX3*) examined on an agarose gel after restriction digestion (*Bam*HI and *Eco*RI) and purification, but prior to ligation reaction. The purified vector and insert were ligated and transformed into the competent cells of *E. coli* (Top10) cells. (c) Screening of the transformed Top10 clone by restriction digestion of the isolated plasmid (U: undigested plasmid; RD: plasmid digested with *Bam*HI and *Eco*RI). The recombinant *OsCBSX3*_pGEX-4T-1 plasmid was then transformed into *E. coli* BL21 (DE3) cells. (d) Colony PCR screening of *OsCBSX3*_pGEX-4T-1 transformed BL21 (DE3) clones. Transformed clones are marked numerically. The positive clone was proceeded for protein expression. 'M': DNA ladder; '+': positive control; '-': no template control. (e) The total cell lysate from uninduced cells (U), and the insoluble (P) and the soluble fraction (S) from induced cells resolved in 12% SDS-PAGE gel. The induced protein corresponding to GST-*OsCBSX3* (49 kDa) is shown by a red arrow. The cultures were induced and incubated at different temperatures and for time points to improve the solubility of GST-*OsCBSX3*. 'M': Protein marker.

Next, we opted to use *E. coli* ArcticExpress (DE3) expression strain, which is known to address the hurdle of protein solubility. This strain expresses cold-adapted chaperonins 'Cpn10' and 'Cpn60' from the psychrophilic bacterium (*Oleispira antarctica*), which possess high protein refolding activity at temperatures of 4-12°C (<https://www.agilent.com/cs/library/usermanuals/Public/230191.pdf>). Therefore, we transformed competent cells of ArcticExpress (DE3) with *OsCBSX3*_pET28a (Fig. 4.8 a). Expression in this strain yielded some amount of *OsCBSX3* in the soluble fraction, however, its subsequent purification using Ni-NTA affinity chromatography co-eluted other undesirable proteins as well (Fig. 4.8 b-e). Although increasing NaCl concentration to 400 mM in the buffer significantly reduced the co-elution of other non-specific proteins during purification, the

purified protein was highly prone to aggregation, particularly when attempted to concentrate using Centrifugal Concentrators (Vivaspin 6, GE Healthcare).

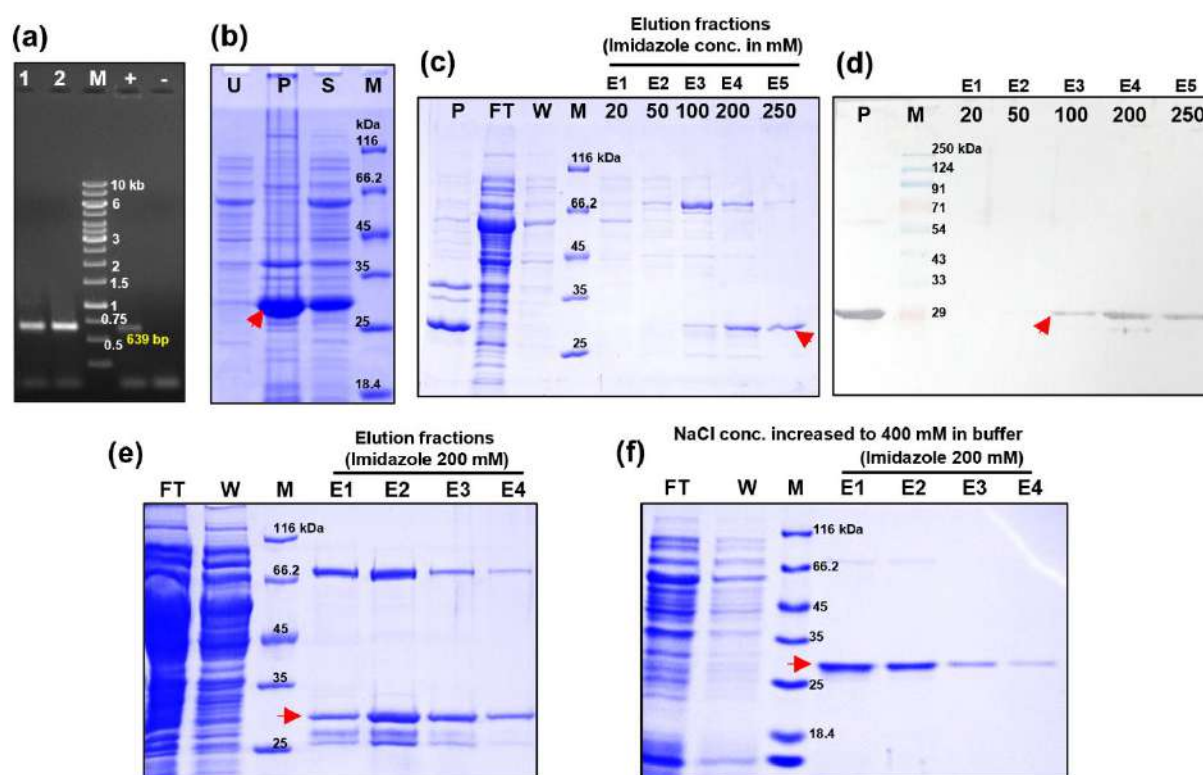


Fig. 4.8: Expression of 6XHis-OsCBSX3 in *E. coli* ArcticExpress (DE3). (a) Colony PCR screening of *E. coli* ArcticExpress (DE3) transformed with *OsCBSX3_pET28a*. Transformed clones are marked numerically; 'M': DNA ladder; '+' and '-' denotes positive and no template control, respectively. (b) Total cell lysate of uninduced cells (U), and the insoluble (P) and soluble (S) fraction of induced cells resolved in 12% SDS-PAGE gel. (c) 6XHis-OsCBSX3 purified from the soluble fraction (S) by Ni-NTA affinity chromatography along with the insoluble fraction (P) from induced cells resolved in SDS-PAGE gel. (d) Western blot of samples corresponding to the ones shown in (c) using anti-His antibody conjugated with alkaline phosphatase which confirmed the presence of 6XHis-OsCBSX3 in soluble fraction as well. SDS-PAGE gel images in (e) and (f) represents 6XHis-OsCBSX3 that is lysed, Ni-NTA affinity purified and eluted in buffer containing 150 mM and 400 mM of NaCl, respectively. Low NaCl concentration in the buffer resulted in co-purification of 6XHis-OsCBSX3 along with some undesirable proteins, while increasing NaCl to 400 mM in the buffer significantly reduced the contamination of non-specific proteins. Position of the induced protein corresponding to 6XHis-OsCBSX3 (26.75 kDa) is shown by a red arrow. In (c) to (f), different Ni-NTA affinity chromatography purification fractions are denoted as 'FT': Flow through; 'W': Wash; 'E': Elution. 'M': Protein marker.

We finally attempted to express OsCBSX3 that has a deletion of 35 residues from the N-terminal region which were predicted to constitute chloroplast signal peptide (sp) by TargetP 1.1 server (see Fig. 4.4 a). Hence, we referred to this modified protein as OsCBSX3(-sp). The construct was prepared by cloning the CDS of *OsCBSX3* lacking '105 nucleotides' from the 5' region that corresponds to signal peptide {*OsCBSX3(-sp)*} into pET28a at the same cloning sites used for cloning the full-length gene as mentioned above. The recombinant plasmid was then transformed into *E. coli* BL21 (DE3) to express 6XHis_OsCBSX3(-sp) (Fig. 4.9 a-c). This

approach resulted in an improved yield of 6X-His_OsCBSX3(-sp) in the soluble fraction (23.11 kDa; Fig. 4.9 d, e) which was used subsequently for the oligomeric and ATP binding studies. However, the Ni-NTA affinity chromatography purified OsCBSX3(-sp) still remained prone to aggregation, despite being screened in variable buffer compositions.

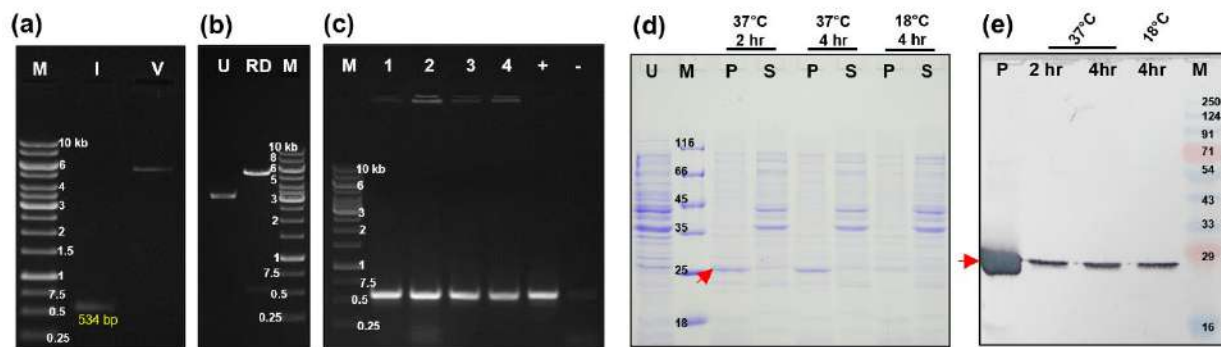


Fig. 4.9: Cloning of *OsCBSX3(-sp)* in pET28a and expression of 6XHis-*OsCBSX3(-sp)* in BL21(DE3). (a) Vector (pET28a; V) and insert (*OsCBSX3(-sp)*; I) examined on an agarose gel prior to ligation. (b) Confirmation of *E. coli* Top10 positive transformants by restriction digestion of the isolated plasmid (U: undigested plasmid; RD: plasmid digested with *Bam*HI and *Eco*RI). (c) Colony PCR screening of *OsCBSX3(-sp)*_pET28a transformed *E. coli* BL21 (DE3) clones. ‘M’: DNA ladder; ‘+’: positive control; ‘-’: no template control. (d) Total cell lysate from uninduced cells (U), and the insoluble (P) and soluble (S) fractions from induced cells expressing 6XHis-*OsCBSX3(-sp)* resolved in 12% SDS-PAGE gel. The induced cultures were incubated at 18°C and 37°C. (e) Confirmation of 6XHis-*OsCBSX3(-sp)* expression in all soluble and one insoluble fraction corresponding to the ones shown in (d) by western blot using anti-His antibody conjugated with alkaline phosphatase. In (d) and (e), the expected position of 6XHis-*OsCBSX3(-sp)* (23.11 kDa) is shown by red arrow; ‘M’: protein marker.

4.1.3.2 Oligomeric status of *OsCBSX3*

We performed Gel Permeation Chromatography (GPC; Superdex 75 column) to analyze the oligomeric status of modified *OsCBSX3* {6X-His-*OsCBSX3(-sp)*} pre-purified by Ni-NTA affinity chromatography. The protein molecular weight calculated based on the elution volume of *OsCBSX3(-sp)* from the GPC column implied that the protein is in a monomeric state (Fig. 4.10). However, owing to the less solubility and aggregating nature of *OsCBSX3(-sp)* observed *in vitro*, we cannot exclude the possibility of this expressed protein being in misfolded or non-native conformation. Therefore, we also proceeded to determine its oligomeric status *in vivo* through yeast two-hybrid (Y2H) assay.

For Y2H, the CDS of *OsCBSX3* (full length) as well as *OsCBSX3(-sp)* was cloned into both pGADT7 and pGBKT7 Y2H vectors in-frame with the upstream CDS encoding activation domain (AD) and binding domain (BD), respectively (Fig. 4.11 a-d). The recombinant

plasmids (*OsCBSX3_pGADT7* and *OsCBSX3_pGBKT7*) stabilized in *E. coli* (Top10) were co-transformed into yeast cells (*Saccharomyces cerevisiae*, AH109 strain), followed by screening on 3 Synthetic Drop-out (SD/- Leu/-Trp/-His or 3 DO) and 4 DO (SD/-Trp/-Leu/-His/-Ade) minimal medium supplemented with 5 mM of 3-Amino-1,2,4-triazole (3-AT, a Histidine anti-metabolite). *OsCBSX3* showed no autoactivation of selection genes in yeast, while the yeast clones co-transformed with *OsCBSX3_pGADT7* and *OsCBSX3_pGBKT7* showed growth on both 3 DO and 4 DO minimal medium (Fig. 4.11 e).

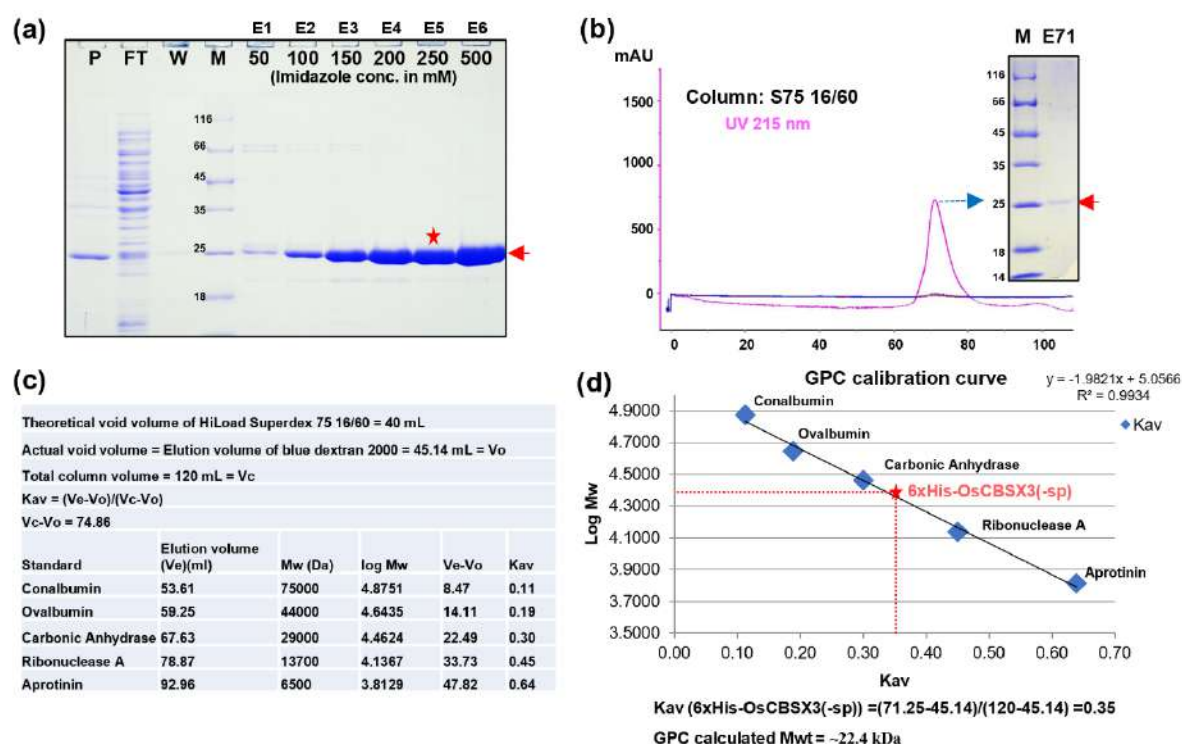


Fig. 4.10: Gel permeation chromatography (GPC) of 6XHis-OsCBSX3(-sp) for determination of its oligomeric status. (a) Ni-NTA affinity chromatography purified 6XHis-OsCBSX3(-sp) analyzed in SDS-PAGE gel. Different purification fractions are indicated as- ‘FT’: Flow through; ‘W’: Wash; ‘E’: Elution; while ‘P’ represents the insoluble protein fraction; ‘M’: protein marker. 6XHis-OsCBSX3(-sp) is shown by a red arrow. The ‘E5’ fraction eluted with 250 mM Imidazole containing elution buffer (marked with a red star) was used for the downstream GPC analysis. (b) GPC chromatogram/elution profile of 6XHis-OsCBSX3(-sp) passed through Superdex 75 column. X-axis represents the elution volume (Ve) and Y-axis represents the absorbance at 215 nm (purple) as milli Absorbance Units (mAU). The GPC eluted fraction {E71; elution volume (Ve) = 71.25 ml} was confirmed to be 6XHis-OsCBSX3(-sp) by SDS-PAGE {shown in the inset, marked by a red arrow}; ‘M’: protein marker. (c) Partition co-efficients (Kav) of five different protein standards of known native molecular weights (Mw) obtained based on their Ve and properties of the column (Vo: void volume; Vc: column volume), which were used to prepare the calibration curve. (d) The approximate Mw of 6XHis-OsCBSX3(-sp) determined by plotting its calculated Kav on the calibration curve (marked by red dotted lines and red star). The Mw was observed to be approximately 22.4 kDa. (Theoretical Mw = 23.11; predicted by Compute pI/Mw tool, ExPASy).

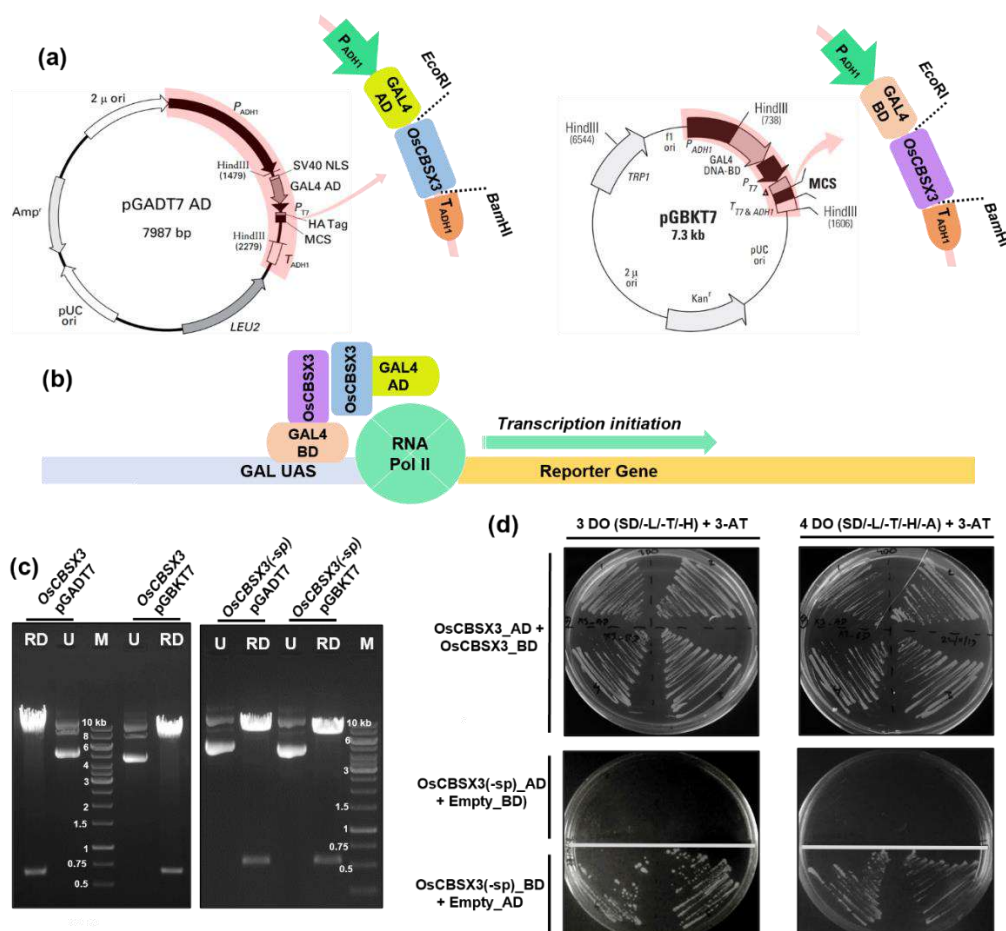


Fig. 4.11: Determination of *OsCBSX3* homo-oligomerization by yeast two-hybrid (Y2H) assay. (a) Vector map of pGADT7 (left panel, acquired from Clontech, PT3249-5) and pGBKT7 (right panel, acquired from Clontech, PT3248-5). Expression cassette in each vector map is highlighted in red, and the final cassette is depicted on top-right corner of each vector map. The CDS of both *OsCBSX3* and *OsCBSX3(-sp)* was inserted in-frame with the upstream region encoding GAL4 Activation Domain (GAL4AD) and GAL4 DNA Binding Domain (GAL4BD) in pGADT7 and pGBKT7, respectively. Transcription of insert in both the vectors is regulated by the upstream alcohol dehydrogenase 1 (ADH1) promoter (P_{ADH1}) and terminated by the downstream ADH1 terminator (T_{ADH1}). (b) Schematic representation of the working principle of Y2H assay. The GAL4BD fusion protein binds to GAL UAS (Upstream Activating Sequences). *OsCBSX3*, if self-associate, brings GAL4AD in proximity to GAL4BD, that subsequently drives transcription of the reporter gene. (c) Confirmation of the cloning of *OsCBSX3* and *OsCBSX3(-sp)* in pGADT7 and pGBKT7 by restriction digestion of the plasmids isolated from transformed *E. coli* (Top10) clones. ‘M’: DNA ladder; U: undigested plasmid; RD: plasmid digested with *EcoRI* and *BamHI*. (d) *OsCBSX3_pGADT7* and *OsCBSX3_pGBKT7* co-transformed yeast (AH109) clones survived on 3 DO and 4 DO minimal medium, implying the homo-oligomerization of *OsCBSX3* (top panel). In case of *OsCBSX3(-sp)*, the yeast clones co-transformed with *OsCBSX3(-sp)_pGBKT7* and Empty pGADT7 proliferated on both 3 DO and 4 DO minimal medium, implying strong auto-activation by *OsCBSX3(-sp)* (bottom panel).

This result, in contrast to the GPC analysis mentioned above, suggests that OsCBSX3 proteins interact or associate with each other *in vivo* into a higher oligomeric structure (possibly a dimer). However, in the case of *OsCBSX3(-sp)*, co-transformation of yeast cells with *OsCBSX3(-sp)_pGBKT7* and pGADT7 (empty vector) exhibited strong auto-activation of the selection marker genes, enabling transformed yeast cells to proliferate on both 3 DO and 4 DO minimal medium (Fig. 4.11 e). Surprisingly, the yeast clones co-transformed with *OsCBSX3_pGBKT7* and *OsCBSX3(-sp)_pGADT7* showed no growth on 3 DO and 4 DO minimal medium, which implies the probable involvement of deleted N-terminal region in forming associations among OsCBSX3 proteins.

Therefore, based on these observations, we propose that OsCBSX3 exists in a homo-oligomeric state *in vivo*, and its N-terminal region that is predicted as a signal peptide in this study is essential for its homo-oligomeric state. Accordingly, in addition to non-native protein conformation, the deletion of that corresponding signal peptide region in 6X-His_*OsCBSX3(-sp)* may also account for its existence as monomeric protein *in vitro* as observed by GPC. Other than its oligomeric state, the autoactivation activity of *OsCBSX3(-sp)* in Y2H assay implies its likely role as a transcription activator in the chloroplast where it was found to be localized (see section 4.1.2).

4.1.3.3 Sequence analysis for ligand binding residues and ATP binding assay for OsCBSX3

Several structural studies on CBS domain-ligand complexes have identified conserved residues that form the canonical ligand-binding sites in a Bateman module (Ereño-Orbea et al, 2013a). Accordingly, based on some of these available structural studies and the respective CBS domain sequences from these CDCPs, we analyzed the conservation of potential ligand-binding residues in CBS domain sequences from *OsCBSX3* by sequence alignment (Fig. 4.12). We observed conserved Thr/Ser and Leu/Ile in Block I, and hydrophobic and the Pro (P) residues at respective positions in the conserved “h-y-y'-h'-P” motif (‘h’: hydrophobic, ‘y’: any residue) of Block II in both CBS domains (CBS1 and CBS2) of *OsCBSX3*. The two CBS domains were also observed to have conserved Asp (D) residue in “G-h-h'-T/S-y-y'-D/N” motif of Block III, however only CBS2 contained conserved Ser (S) in this motif, while it was substituted by hydrophobic Val (V) in CBS1, which may interfere with adenosine nucleotide binding. Therefore, based on this sequence alignment, we anticipate that *OsCBSX3* has a

potential binding site for only one adenosine nucleotide per each Bateman module or two per each CBS module.

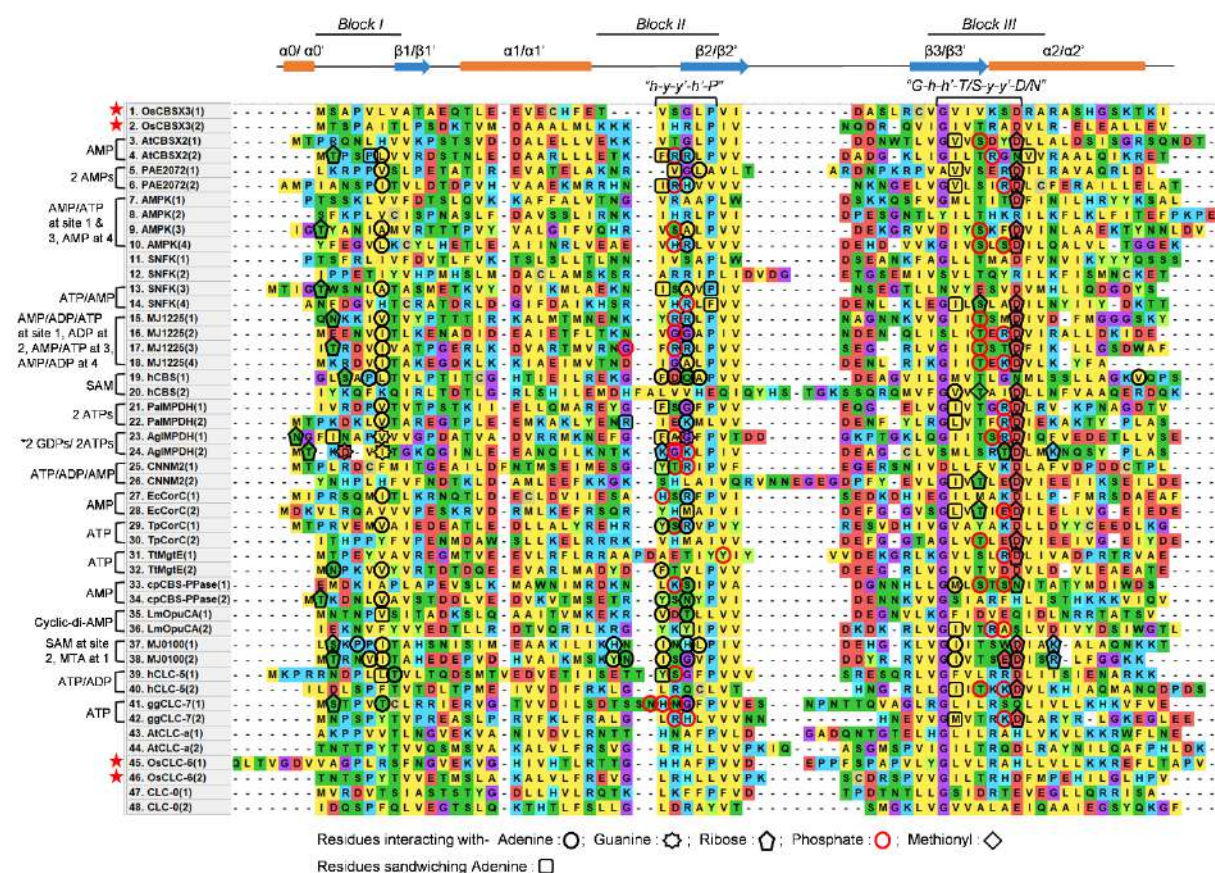


Fig. 4.12: Analysis of ligand binding residues in the sequences of CBS domains. Sequence alignment of the CBS domains from OsCBSX3 and OsCBSX3(2) along with the CBS domains from different CDCPs wherein the residues interacting with ligands have been structurally studied (see Table. 2 for details). CBS domains from each protein is numbered as (1), (2), (3) and (4), according to their arrangement from N-terminus to C-terminus. Three blocks in CBS domains that form the ligand binding site and the two conserved motifs “h-y-y’-h’-P” and “G-h-h’-T/S-y-y’-D/N” (‘h’: hydrophobic, ‘y’: any residue) present in Block II and Block III, respectively are highlighted on the top. Ligands that bind to CBS domains in these CDCPs are shown on the left, while the residues involved in interaction with the ligands are marked by enclosing in different shapes. Each CBS domain sequence from OsCBSX3 and OsCBSX3(2) is highlighted by red star. In case of CLCs, the CBS domain sequences from AtCLC-a which has been reported to be regulated by ATP/AMP (De Angeli et al, 2009), but not structurally shown as well as the CBS domains from CLC-0 that is reported to show no ATP binding (Meyer and Dutzler, 2006), are also analyzed for comparisons. Note that the residues from both CBS domains in a Bateman module as well as the residues from adjacent Bateman module interacts with the ligand in each binding site. For e.g., with one ATP binding to *Arabidopsis* AtCBSX2, the residues from both CBS1 and CBS2 domains and the residues from other Bateman module in a dimer interacts to it. In case of AglMPDH, similar residues are involved in interaction with both GDP and ATP, and the residues interacting with GDP are highlighted in the figure. The sequence alignment is performed by ClustalW using Mega X program.

After this *in silico* analysis, we proceeded to study the binding of ATP to OsCBSX3 *in vitro* using the recombinant 6X-His_OsCBSX3(-sp) expressed in *E. coli* and the radiolabeled ATP [α - 32 P]. However, contrary to the *in-silico* sequence analysis, we observed no ATP binding to 6X-His_OsCBSX3(-sp) (Fig. 4.13). Nevertheless, considering the *in vitro* instability

and the deletion of predicted N-terminal sequence peptide from the recombinant 6X-His_OsCBSX3(-sp) as mentioned above, we hypothesize that the CBS module of CBSX3 has at least two potential binding sites for adenosine nucleotides.



Fig. 4.13: ATP binding assay of 6xHis-OsCBSX3(-sp). Phosphorimager-derived image of purified 6xHis-OsCBSX3(-sp) resolved on SDS-PAGE gel after incubation with [α - 32 P]ATP followed by UV crosslinking. ‘X3(-sp)’: 6xHis-OsCBSX3(-sp), ‘-’: BSA as a negative control; ‘+’: CBSX member from rice that is identified to bind ATP (marked with red arrow; unpublished data) as a positive control. No ATP binding to OsCBSX3(-sp) was observed.

4.1.3.4 Identification of OsCBSX3 interacting partners by yeast two-hybrid assay

To understand the cellular function of OsCBSX3, we attempted to find its interacting proteins by employing the yeast two-hybrid (Y2H) approach. A rice cDNA library (prepared from the leaves of 10 days old Pokkali seedlings) that was previously inserted into pAD-GAL4-2.1 vector using “HybriZAP-2.1 Two Hybrid kit” (Stratagene) and transformed into AH109 yeast cells by Kumar et al (2012) (Fig.4.14 a), were used for the screening. This library of rice cDNAs inserted in the pAD-GAL4-2.1 vector expresses their respective proteins in-fusion with the transcription activation domain (AD) of the GAL4 transcription factor at the N-terminus. Therefore, the CDS of *OsCBSX3* was cloned in the pBD-GAL4-Cam vector to produce OsCBSX3 in-fusion with the DNA binding domain (BD) of the GAL4 transcription factor at N-terminus (Fig. 4.14 b-d). This GAL4BD-OsCBSX3 protein was used as a ‘bait’ to screen for its interacting partners from the library of rice proteins fused to GAL4AD (i.e., ‘prey’), the interaction of which would bring AD and BD of GAL4 transcription factor in close proximity, resulting in activation of the selection or reporter genes in the Y2H assay.

The yeast clones harboring rice cDNA library in pAD-GAL4-2.1 were transformed with *OsCBSX3*_pBD-GAL4-Cam and plated onto 3DO minimal medium (SD/- Leu/-Trp/-His) supplemented with 5 mM of 3-AT. The colonies growing on a 3DO minimal medium were streaked onto a more stringent 4DO + 5 mM 3-AT minimal medium (SD/-Trp/-Leu/-His/-Ade) (Fig. 4.14 e). The clones growing on 4DO medium were screened further for the activation of *MEL1* (a reporter gene that encodes α -galactosidase) by streaking onto X- α -gal (5-Bromo-4-chloro-3-indolyl α -D-galactopyranoside) containing YPD medium (Fig. 4.14 f).

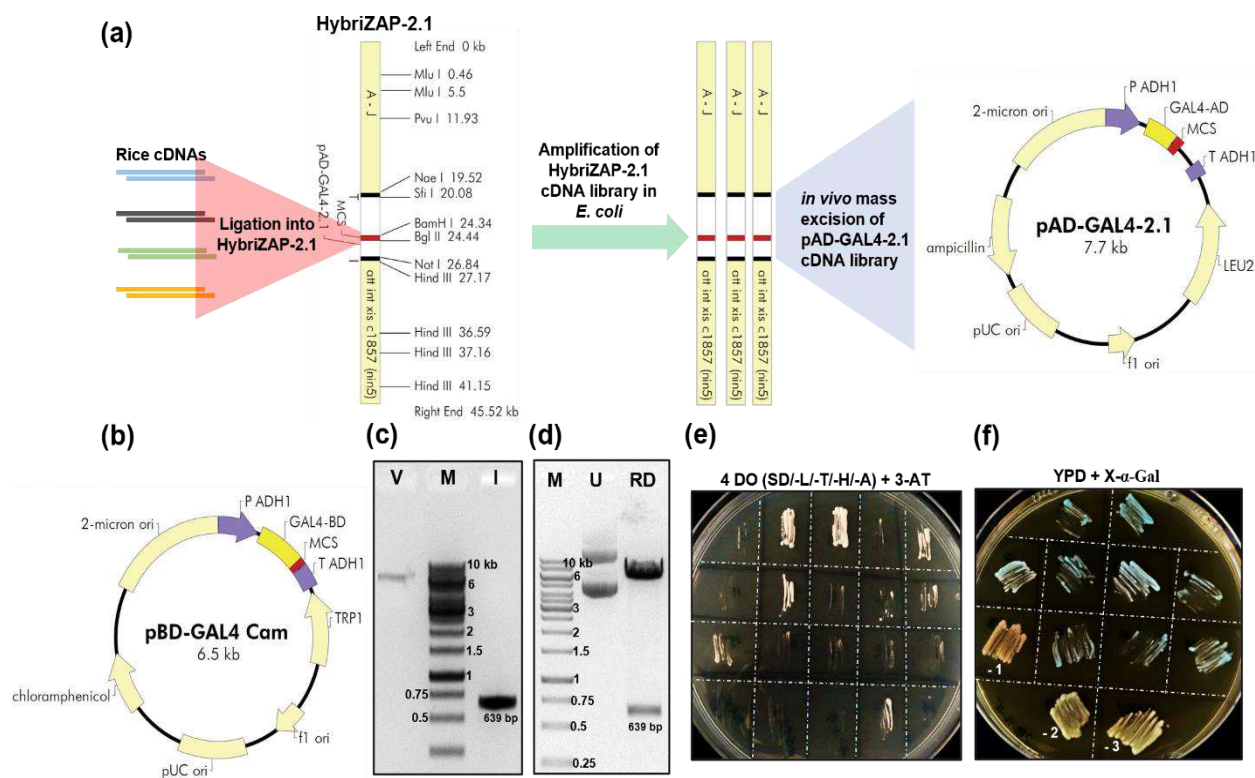


Fig. 4.14: Cloning of *OsCBSX3* in pBD-GAL4 Cam and screening of the transformed yeast clones by Y2H assay. (a) Schematic representation of the preparation of rice cDNA library in pGADT7 phagemid vector through HybridZAP-2.1 lambda phage vector system {prepared previously by Kumar et al (2012)}. The rice cDNA inserts were ligated into HybriZAP-2.1 vector, packaged to lambda phage, and then amplified in *E. coli*. The amplified phage cDNA library was eventually converted to pAD-GAL4-2.1 phagemid cDNA library by co-infecting the *E. coli* cells with lambda phage and the helper phage. The excised pAD-GAL4-2.1_cDNA library plasmids were subsequently transformed into yeast AH109 strain where they produce proteins encoded by rice cDNA in fusion with GAL4 Activation domain (GAL4 BD) at the N-terminus. (b) Vector map of pBD-GAL4 Cam onto which CDS of *OsCBSX3* was ligated in-frame with the upstream CDS for GAL4 DNA Binding Domain (GAL4 BD). (c) Vector (V, pBD-GAL4 Cam) and insert (I, *OsCBSX3*) analyzed on an agarose gel prior to ligation. (d) Confirmation of *E. coli* (Top10) clone transformed with *OsCBSX3*-pBD-GAL4 Cam by restriction digestion (U: undigested plasmid; RD: plasmid digested with *EcoRI* and *SalI*). The recombinant *OsCBSX3*-pBD-GAL4 Cam plasmid was transformed into AH109 cells pre-transformed with pAD-GAL4-2.1 rice cDNA library. (e) Screening of the yeast (AH109) clones harbouring rice cDNA inserted pAD-GAL4-2.1 and *OsCBSX3*_pBD-GAL4 Cam on 4 DO + 5mM of 3-AT minimal medium (SD/-Trp/-Leu/-His/-Ade). (f) Screening of 4 DO positive yeast clones for the activation of *MEL1* on YPD medium containing X- α -Gal. The positive clones developed blue colour, whereas the negative controls (-1, -2 and -3) remained white. '-1': AH109 harbouring pBD-GAL4 + rice cDNA pAD-GAL4; '-2': untransformed AH109; '-3': AH109 harbouring only rice cDNA pAD-GAL4. Images in (a) and (b) are acquired from Agilent Technologies (HybriZAP-2.1 Two-Hybrid Libraries Instruction manual #977519-12); In (c) and (d), 'M': DNA ladder.

From those clones that proliferated on 4DO medium and turned blue on X- α -gal containing YPD medium, the inserts corresponding to rice cDNA fragments were PCR amplified using pAD-GAL4-2.1 vector-specific primers flanking these inserts and the isolated plasmids as templates (Fig. 4.15 a-c). PCR products were eluted from the agarose gel, purified, and sequenced to identify the putative interacting partners of OsCBSX3. Although we used plasmid isolated from yeast cells that were lysed in lysis buffer with the aid of glass beads for PCR amplification of the inserts, we found yeast cell lysates obtained using lyticase as well as the direct colony PCR by pre-boiling the cell suspension, to be equally efficient for PCR amplification from yeast (Fig. 4.15 d).

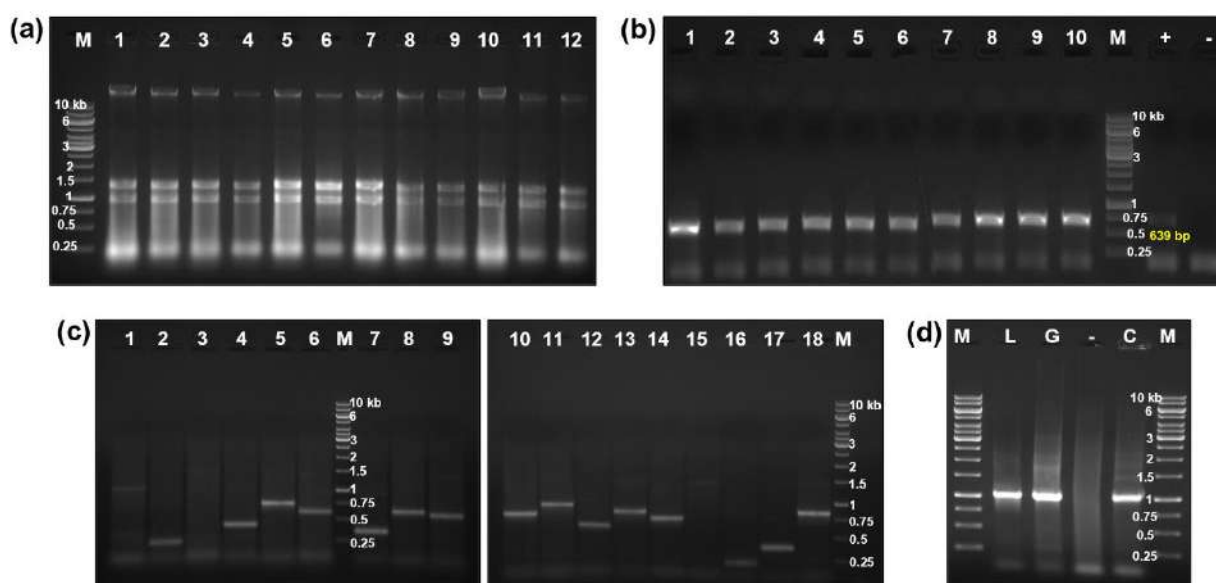


Fig. 4.15: Plasmid isolation from yeast and PCR amplification of the DNA inserts corresponding to rice cDNA in pAD-GAL4-2.1 from clones identified to be positive in Y2H assay. (a) Representative agarose gel image of crude plasmids isolated from yeast clones that were positive on 4 DO minimal medium and X- α -Gal chromogenic assay. (b) PCR screening of Y2H positive clones for the presence of *OsCBSX3*-pBD-GAL4 Cam using *OsCBSX3* specific primers. (c) PCR amplified DNA fragments corresponding to rice cDNA inserted into pAD-GAL4-2.1 from Y2H positive clones. PCR was performed using pAD-GAL4-2.1 specific primers (flanking the inserts). PCR products were eluted from the gel, purified, and sequenced. (d) PCR efficiency from yeast cells lysed by three different methods. (1) using Lyticase 'L' for cell lysis and the lysate was used for PCR; (2) using Glass Beads 'G' to lyse cells in SDS buffer for plasmid isolation {as in (a)} which is used as template for PCR; (3) by direct Colony PCR 'C' in which cells were picked from the culture plate with small pipette tip, resuspended in 20 μ l water, heated at 98°C for 5 min, vortexed, and finally 2 μ l of that cell suspension was as template for PCR. A Y2H positive clone containing rice cDNA insert in pAD-GAL4 was used to isolate template for PCR and the insert was amplified using pAD-GAL4 specific primer. Clones are labelled numerically. 'M': DNA ladder; '+': positive control; '-': no template control.

The analysis of insert sequences from Y2H positive clones identified putative thioredoxin (LOC_Os01g73234), putative spermidine synthase (LOC_Os06g33710), putative glycosyl transferase (LOC_Os07g48370), putative retrotransposon (LOC_Os10g29650), and initiation factor-2 subunit family domain-containing protein (LOC_Os11g11050) to be possible interacting partners of OsCBXS3. Whereas many other inserts obtained from Y2H library screening were found to be either noticeably short or out of frame with the N-terminus AD, hence, have not proceeded for the downstream analysis.

The thioredoxins (Trxs) have been reported to interact with CBSX homologs (AtCBSX1, 2, and 3) in *Arabidopsis* (Yoo et al, 2011; Shin et al, 2020). Accordingly, the putative thioredoxin that we obtained from Y2H library screening is found to share high sequence identity (56.89%) with AtTrx y, which is among the AtTrxs reported to interact with AtCBSX1 and AtCBSX2. Thus, we termed this putative Trx from rice as OsTrx y in this study. To further validate the interaction of OsTrx y (full length) with OsCBSX3 by Y2H assay, we planned to amplify the full-length CDS of putative *OsTrx y*. Intriguingly, our attempt to PCR-amplify the full-length CDS of *OsTrx y* consistently obtained the product of smaller size (in accordance with the gene sequence available in Rice Genome Annotation Project Database). Sequencing of the product identified a deletion of 28-141 nucleotides in *OsTrx y*, which corresponded to the chloroplast signal peptide in the protein (analyzed by TargetP 1.1, Emanuelsson et al, 2000). However, this deletion caused no frameshift or loss of residues from Trx domain in the protein. Hence, we proceeded with this isoform of *OsTrx y* for downstream analysis. We amplified two more genes encoding putative chloroplast localized Trxs from rice {termed in this study as OsTrx f (LOC_Os01g68480) and OsTrx m (LOC_Os04g44830)} whose orthologs in *Arabidopsis* are known to interact with AtCBSX1 and AtCBSX2. In addition, we also amplified putative mitochondrial OsTrx o (LOC_Os06g45510), and three putative cytosol Trx h members: OsTrx 1 (LOC_Os01g07376), 10 (LOC_Os03g58630), and 23 (LOC_Os07g08840). We cloned CDS of all these putative *OsTrxs* into pGADT7 vector, in-frame with the upstream CDS for GAL4 AD, whereas *OsCBSX3* was cloned into pGBKT7 vector, in-frame with upstream CDS for GAL4 BD (Fig. 4.16 a). Subsequently, *OsCBSX3* in pGBKT7 along with each *OsTrx* in pGADT7 were co-transformed into AH109 yeast cells and the transformants were analyzed for growth on 3DO and 4DO minimal medium plates. We found that clones co-transformed only with *OsTrx y*_pGADT7 and *OsCBSX3*_pGBKT7 proliferated on both 3DO and 4DO medium, implying that OsCBSX3 interacts specifically with OsTrx y (Fig. 4.16 b).

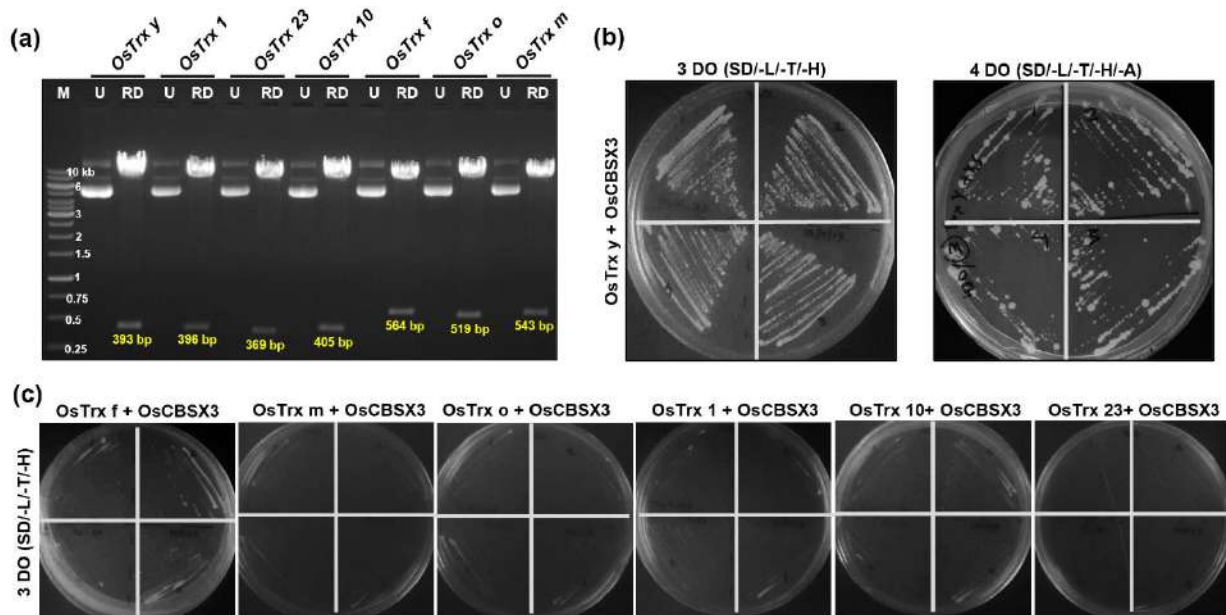


Fig. 4.16: Cloning of genes encoding putative *OsTrx* members from rice in pGADT7 and analysis of their interaction with *OsCBSX3* by Y2H assay. (a) Screening of *E. coli* (Top10) clones transformed with putative members of *OsTrx* cloned into pGADT7 at *EcoRI* and *XhoI* sites. ‘U’: Undigested plasmid; ‘RD’: plasmid digested with *EcoRI* and *XhoI*; ‘M’: DNA Ladder. Each recombinant *OsTrx*_pGADT7 plasmid along with *OsCBSX3*_pGBKT7 were co-transformed into yeast cells (AH109) and grown on 2 DO minimal medium (SD/-L/-T) from where the transformants were streaked onto 3 DO (SD/-L/-T/-H) and 4 DO (SD/-L/-T/-H/-A) minimal medium. (b) Growth of yeast clones co-transformed with *OsTrx* *y*_pGADT7 and *OsCBSX3*_pGBKT7 on both 3DO (SD/-L/-T/-H) and 4DO (SD/-L/-T/-H/-A) minimal medium, indicating the interaction between *OsCBSX3* and *OsTrx* *y*. (c) 3 DO minimal medium plates showing no growth of the yeast clones co-transformed with *OsCBSX3* and each of the remaining *OsTrx* members.

Since *OsTrx* *y*, which we identified to interact with *OsCBSX3*, has a deletion of a region that corresponds to the predicted signal peptide residues, we speculated the possible interference of such signal peptides present in other *OsTrxs* in their interaction with *OsCBSX3*. Therefore, we considered to further analyze the interaction of *OsCBSX3* with these *OsTrxs*, but in their signal peptide deleted state by Y2H assay. Note that as in the case of *OsCBSX3*(-sp), all the genes with suffix (-sp) in this study represent the deletion of a region corresponding to the signal peptide predicted by TargetP 1.1 tool. However, as we mentioned in the previous section (4.1.3.2) that deletion of the predicted signal peptide residues results in strong autoactivation by *OsCBSX3* in the Y2H assay, we used *OsCBSX3* and *OsCBSX3*(-sp) cloned in pGADT7 for their interaction analysis by Y2H. Accordingly, we cloned putative *OsTrxs*, viz., *OsTrx* *f*(-sp), *OsTrx* *m*(-sp) and *OsTrx* *o*(-sp) as well as *OsTrx* *y* into pGBDT7. In addition, we also amplified and cloned other four genes identified through Y2H cDNA library screening which encodes for putative spermidine synthase, putative retrotransposon, putative glycosyl transferase (-sp), and initiation factor-2 subunit family domain-containing protein (-sp), and three *AtTrxs* {*AtTrx* *y*(-sp) (At1g76760), *AtTrx* *m*(-sp) (At4g03520) and *AtTrx* *o*(-sp)

(At2g35010)} whose products are reported to interact with AtCBSX1 and AtCBSX2 (Yoo et al, 2011) into pGBKT7 vector for their subsequent interaction analysis with OsCBSX3 (Fig. 4.17).

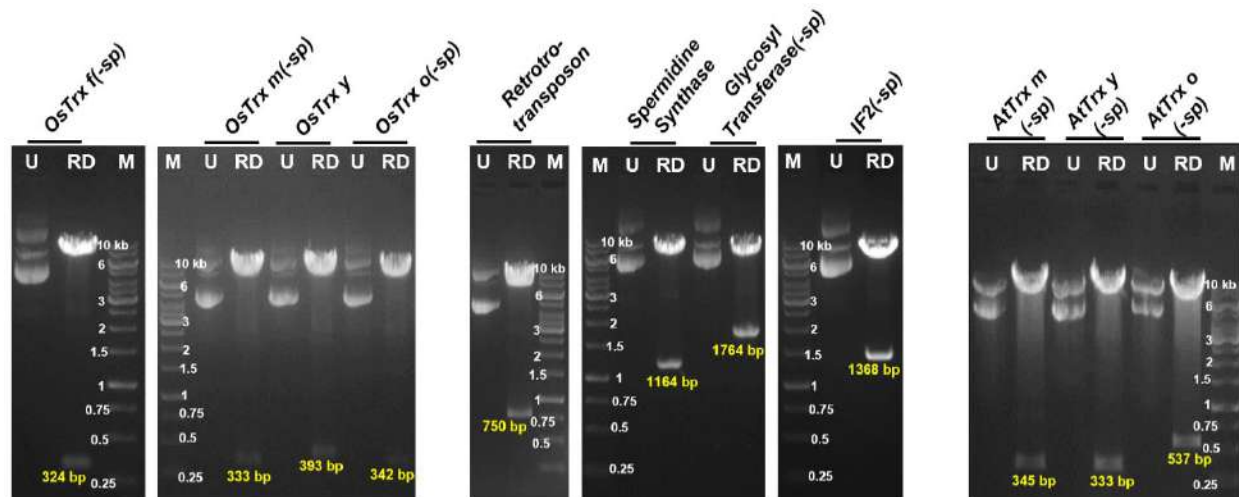


Fig. 4.17: Cloning of genes encoding putative interacting partners of OsCBSX3 from rice and *Arabidopsis* into pGBKT7 for Y2H assay. Confirmation of the cloning of different genes in pGBKT7 by restriction digestion of the respective plasmids isolated from the *E. coli* (Top10) transformed clones. The cloned genes include putative *OsTrx* members from rice {*OsTrx y*, *OsTrx m(-sp)*, *OsTrx f(-sp)* and *OsTrx o(-sp)*}, and *AtTrx* members from *Arabidopsis* {*AtTrx y(-sp)*, *AtTrx m(-sp)* and *AtTrx o(-sp)*} and four other genes from rice that encodes for retrotransposon, spermidine synthase, glycosyl transferase and IF2 subunit family domain containing protein. All *Trx* genes from both rice and *Arabidopsis* were cloned at *EcoRI* and *PstI* restriction sites; genes for Glycosyl transferase and Spermidine synthase at *EcoRI* and *BamHI* sites; IF2 at *NdeI* and *BamHI* sites; Retrotransposon at *NdeI* and *PstI* sites. Accordingly, cloning of these genes in pGBKT7 were confirmed by restriction digestion of the isolated plasmids with the respective restriction enzymes used for cloning. U: undigested plasmid; RD: plasmid digested with restriction enzymes; 'M': DNA ladder. Genes with deletion of sequences corresponding to the predicted signal peptide are labelled with suffix (-sp).

None of these genes cloned into pGBKT7 showed autoactivation of Y2H reporter genes. Therefore, we co-transformed AH109 yeast cells with either *OsCBSX3* or *OsCBSX3(-sp)* cloned into pGADT7 and each gene (encoding a putative interacting partner) cloned into pGBKT7 and analyzed the growth of the transformed clones on 3DO and 4DO minimal medium. However, despite the removal of signal peptides, none of these putative *OsTrxs* exhibited interaction with *OsCBSX3* or *OsCBSX3(-sp)*, except for *OsTrx y* which consistently showed interaction with *OsCBSX3* (Fig. 4.18). Surprisingly, we observed that the interaction with *OsTrx y* was hampered when the predicted signal peptide was deleted from *OsCBSX3* {i.e., when *OsCBSX3(-sp)* was co-transformed with *OsTrx y*}. Besides, the other four putative interacting proteins identified from Y2H library screening as well as the three *AtTrxs* known to interact with AtCBSX1 and AtCBSX2 also failed to restore yeast growth on 3DO and 4DO minimal medium, indicating no interaction with both *OsCBSX3* and *OsCBSX3(-sp)*.

Therefore, based on our Y2H protein interaction analysis, we suggest that OsCBSX3 interacts specifically with OsTrx y among other OsTrx members. Furthermore, we propose that the predicted N-terminal signal peptide region of OsCBSX3 is critical for the interaction, as its deletion affected the interaction of OsCBSX3 with OsTrx y. It is noteworthy that this predicted signal peptide region of OsCBSX3 was also identified to be essential for OsCBSX3 homo-oligomerization, while its deletion exhibited auto-activation of Y2H selection markers (section 4.1.3.2).

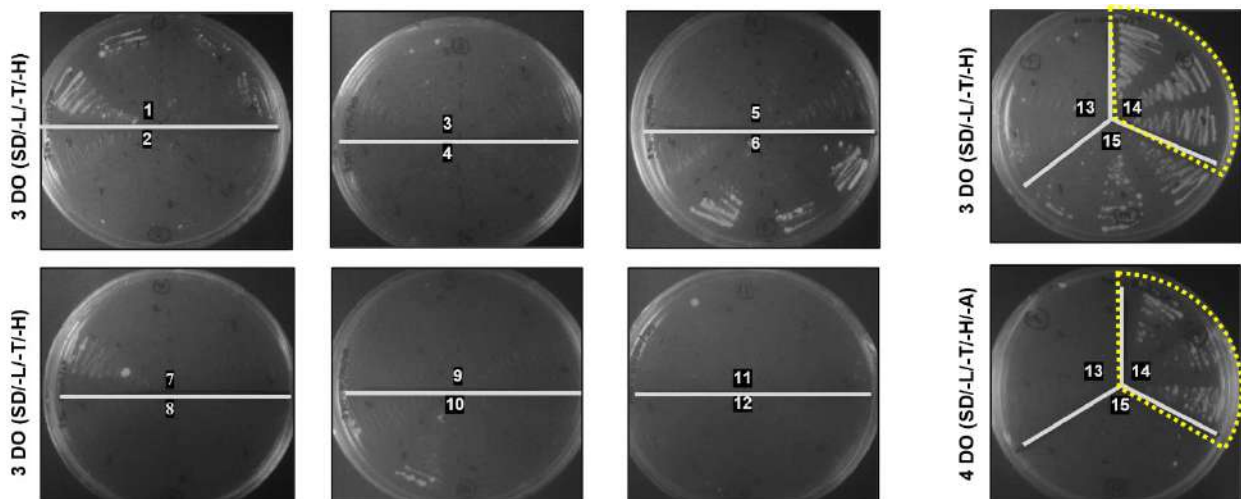


Fig. 4.18: Interaction of OsCBSX3 with different putative interacting partners by Y2H assay. Images on the left panel show the growth of yeast clones co-transformed with ‘1’: *AtTrx m(-sp) + OsCBSX3(-sp)*; ‘2’: *AtTrx y(-sp) + OsCBSX3(-sp)*; ‘3’: *AtTrx o(-sp) + OsCBSX3(-sp)*; ‘4’: *OsTrx f(-sp) + OsCBSX3(-sp)*; ‘5’: *OsTrx m(-sp) + OsCBSX3(-sp)*; ‘6’: *OsTrx y(-sp) + OsCBSX3(-sp)*; ‘7’: *OsTrx o(-sp) + OsCBSX3(-sp)*; ‘8’: *Glycosyl transferase + OsCBSX3(-sp)*; ‘9’: *IF2 + OsCBSX3(-sp)*; ‘10’: *Retrotransposon + OsCBSX3(-sp)*; ‘11’: *Spermidine synthase + OsCBSX3(-sp)* and ‘12’: *Empty_BD + OsCBSX3(-sp)_AD*, on 3 DO minimal medium. That on the right panel show the growth of yeast clones co-transformed with ‘13’: *AtTrx y + OsCBSX3*; ‘14’: *OsTrx y + OsCBSX3* and ‘15’: *AtTrx m + OsCBSX3* on 3 DO and 4 DO minimal medium. Only the clones co-transformed with *OsTrx y_pGBKT7 + OsCBSX3_pGADT7* (14) proliferated on 3 DO and 4 DO minimal medium.

4.1.3.4 Interaction analysis *OsCBSX3* and *OsTrx y* by BiFC assay

To re-validate the interaction between *OsCBSX3* and *OsTrx y*, we performed Bimolecular Fluorescence Complementation (BiFC) assay. The CDS of *OsCBSX3* and *OsTrx y* were cloned into both pSATN-nEYFP-C1 and pSATN-cEYFP-C1 vectors in-frame with the upstream region encoding N-terminal and C-terminal fragment of yellow fluorescent protein (YFP), respectively (Fig. 4.19 a-d). These constructs in two combinations, (a) *OsCBSX3-nEYFP* and *Trx y-cEYFP* and (b) *Trx y-nEYFP* and *OsCBSX3-cEYFP*, were bombarded into onion epidermal cells. Under a confocal microscope the reconstituted YFP fluorescence was detected in the epidermal cells co-expressing *OsCBSX3* and *OsTrx y* in both the combinations of split YFP fusion (Fig. 4.19 e). This analysis hence confirmed the interaction between *OsCBSX3* and *OsTrx y* in *planta*.

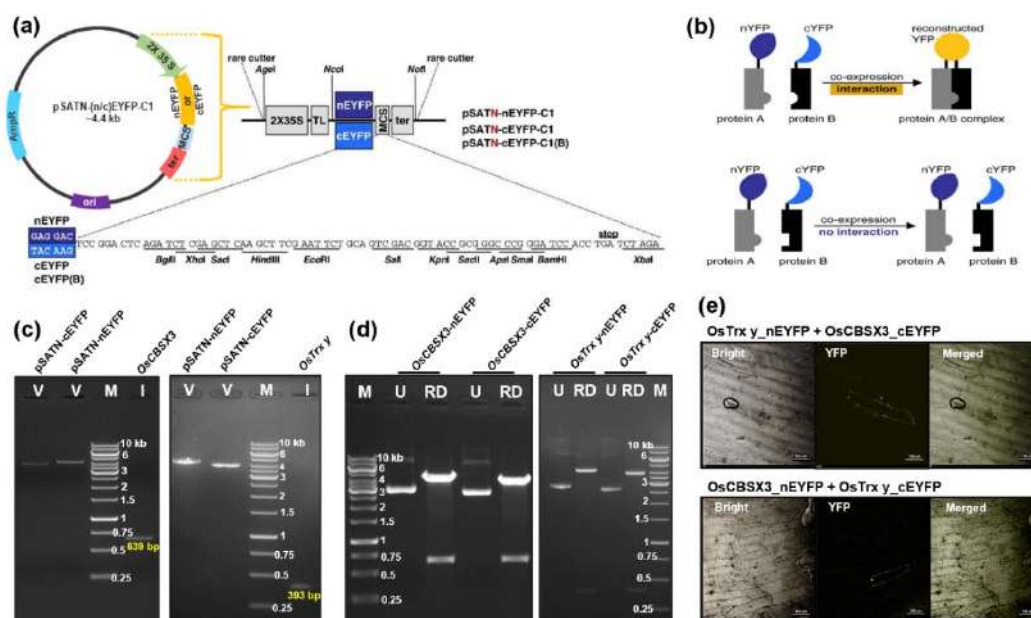


Fig. 4.19: Cloning of *OsCBSX3* and *OsTrx y* into pSATN_(n/c)EYFP-C1 and BiFC assay. (a) Schematic representation of pSATN_(n/c)EYFP-C1 vector that encodes either N-terminal (nEYFP) or C-terminal (cEYFP) fragment of the split EYFP driven by 2X CaMV 35 S promoter (2X 35 S). The CDS of *OsCBSX3* and *OsTrx y* were cloned in frame with the upstream CDS for both nEYFP and cEYFP. ‘TL’: TEV translation leader; ‘MCS’: Multiple cloning sites; ‘ter’: CaMV 35S poly(A) transcriptional terminator. (b) The working principle of BiFC assay which is based on EYFP fluorescence emission from the reconstituted nEYFP and cEYFP upon interaction between the two test proteins. No fluorescence is emitted if the test proteins have no interactions. (c) Vector (V) and insert (I) examined on an agarose gel prior to ligation. *OsCBSX3* was cloned at *BglIII* and *SalI* sites, while *OsTrx y* at *BglIII* and *EcoRI* sites; ‘M’: DNA ladder (d) Screening of the transformed *E. coli* (Top10) clones by restriction digestion of the isolated plasmids (U: undigested plasmid; RD: plasmid digested with respective restriction enzymes used for cloning); ‘M’: DNA ladder. (e) Confocal microscopy images exhibiting the reconstituted EYFP fluorescence in onion epidermal cells co-expressing *OsCBSX3* and *Trx y* in-fusion with split nEYFP and cEYFP in both the combinations. Scale = 100 μ m. Structural feature of expression cassette of the vector in (a) and the BiFC assay representation in (b) are acquired from Citovsky et al, 2006.

4.1.4 Generation and growth assessment of transgenic rice plants overexpressing *OsCBSX3*

To study the physiological function of *OsCBSX3* in rice plants, a transgenic approach was followed through the overexpression *OsCBSX3*. The full-length CDS of *OsCBSX3* was cloned into the pCAMBIA 1302 plant expression binary vector downstream of 35S CaMV promoter at *Bgl*III and *Spe*I sites (Fig. 4.20). To produce transgenic plants overexpressing *OsCBSX3*-GFP fusion (used for subcellular localization study; section 4.1.2), the stop codon of the *OsCBSX3* was excluded during the cloning process. The recombinant *OsCBSX3_pCAMBIA1302* plasmid was transformed into *Agrobacterium tumefaciens* (LBA4404 strain), which was subsequently used for the transformation of embryonic calli of rice (PB-1 variety). The transformed calli were selected on Hygromycin-containing MS medium, regenerated to shoots and roots, hardened on vermiculite, and finally transferred onto the soil (Fig. 4.21). The transgenic plants were screened by PCR, whereas the transgene copy number and individual transgenic events were identified by Southern blot (Fig. 4.22).

Independent transgenic lines with a single transgene integration number were raised to T₁ and T₂ generations to obtain their homozygous lines. The T₁ seeds were screened by germinating on Hygromycin selection (30 mg/L water). They showed 3:1 segregation of transgene and the positive lines were screened further by PCR and grown to harvest the T₂ seeds (Fig. 4.23 a, b). Similarly, the T₂ seeds from independent T₁ lines were again germinated on Hygromycin selection (30 mg/L water) and only those independent T₂ lines that showed 100% germination on Hygromycin (the homozygous lines) were used further for the physiological studies. The expression analysis by qRT-PCR confirmed a significantly higher accumulation of *OsCBSX3* transcripts in the transgenic plants (Fig. 4.23 c). As we observed *OsCYP450* be downregulated upon the upregulation of *OsCBSX3* in gene expression study (section 4.1.1), we analyzed the expression of *OsCYP450* in *OsCBSX3* overexpressing transgenic plants. We also analyzed the expression of *OsTrx y* (whose product was identified to interact with *OsCBSX3*) in these transgenic lines. However, no alteration in the expression of these two genes was observed in the transgenic plants overexpressing *OsCBSX3* (Fig. 4.23 c).

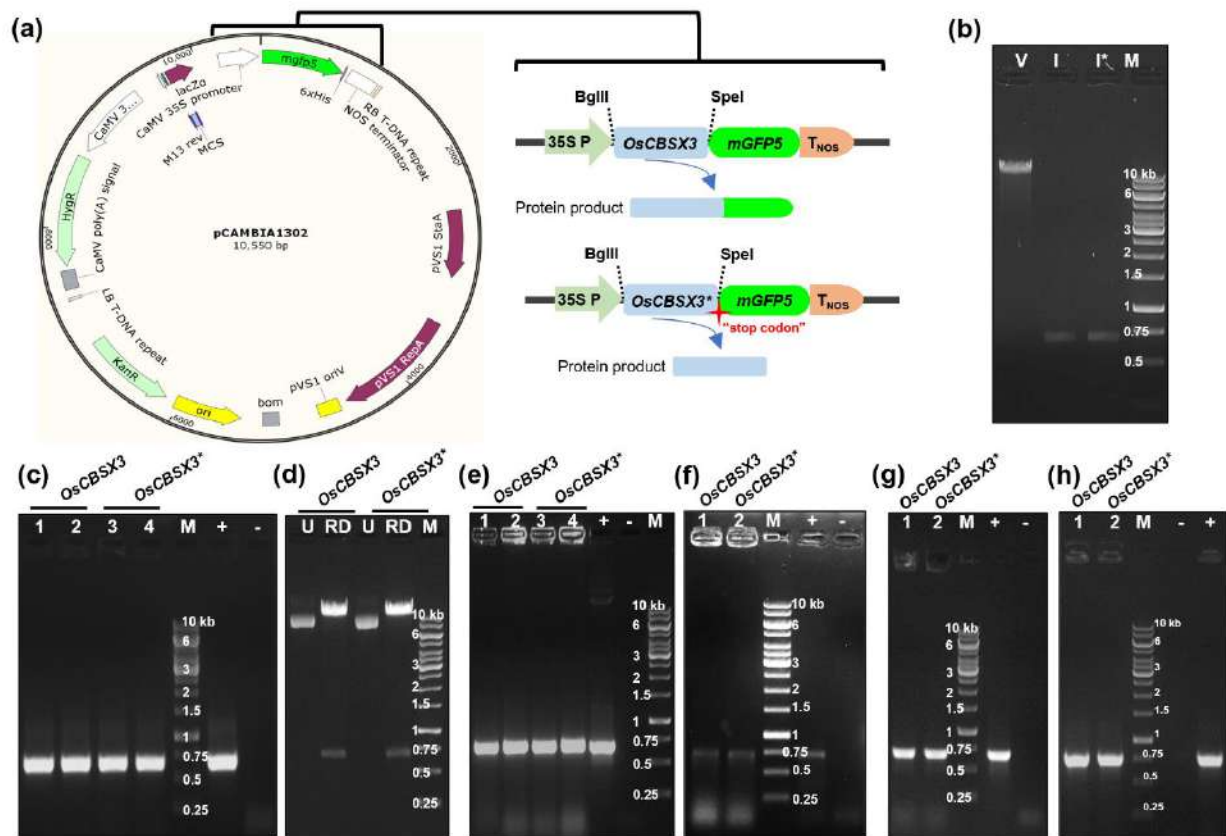


Fig. 4.20: Cloning of *OsCBSX3* into pCambia1302 plant transformation binary vector. (a) Representation of pCambia1302 vector map (prepared using SnapGene, snapgene.com) and two different gene constructs used for plant transformation, one cloned with CDS of *OsCBSX3* excluding stop codon (*OsCBSX3*) that expresses *OsCBSX3* in-fusion with GFP, while another including stop codon (*OsCBSX3**, stop codon is marked as red star) that expresses only *OsCBSX3*. In both the constructs, *OsCBSX3* was inserted at *Bgl*III and *Spe*I sites, with upstream CaMV 35S promoter (35S P) and downstream NOS terminator (T_{NOS}). Note that a single nucleotide is incorporated at the 5' end of *OsCBSX3* CDS to maintain translational frame with the start codon present upstream of *Bgl*III site. (b) Vector (V: pCambia1302) and insert (I: *OsCBSX3*; I*: *OsCBSX3**) examined on an agarose gel prior to ligation reaction. Agarose gel images in (c) and (d) depicts the screening of transformed *E. coli* (Top10) clones by colony PCR and restriction digestion of the isolated plasmids, respectively. The recombinant plasmids were transformed into *Agrobacterium tumefaciens* (both LBA4404 and GV301 strains) for plant transformation. Agarose gel images in (e) and (f) depicts the screening of transformed *Agrobacterium* (LBA4404) clones for the presence of recombinant plasmid and helper plasmid by colony PCR with *OsCBSX3* and *Vir* gene specific primers, respectively. Similarly, the agarose gel images in (g) and (h) depicts the screening of transformed *Agrobacterium* (GV301) clones for the presence of recombinant plasmid and the helper plasmid, respectively by colony PCR. Transformed clones are labelled numerically. 'U': undigested plasmid; 'RD': plasmid digested with *Bgl*III and *Spe*I; 'M': DNA ladder; '+': positive control; '-': no template control.

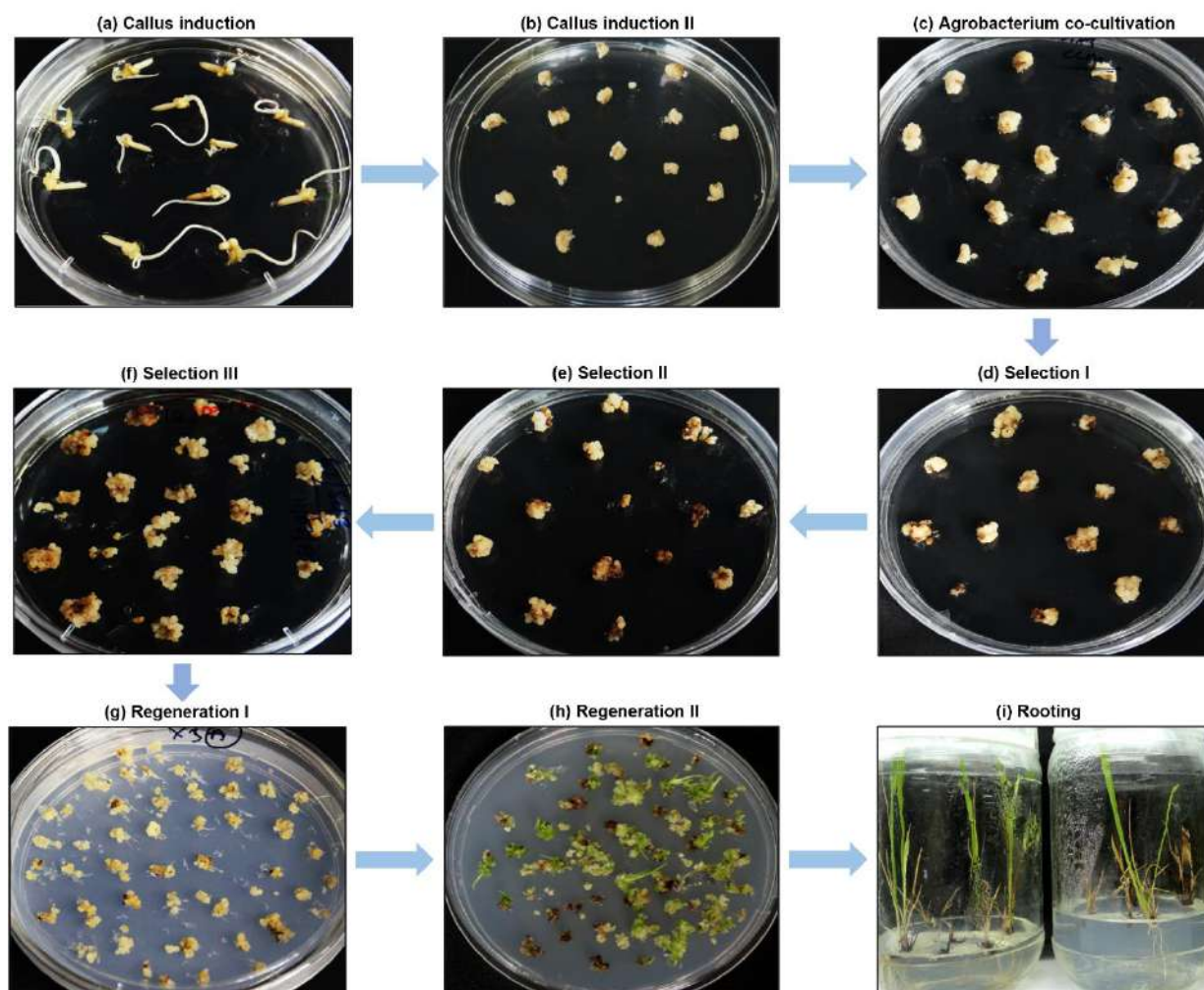


Fig. 4.21: Generation of transgenic rice plants by *Agrobacterium* mediated transformation. Schematic representation of various steps involved in the generation of transgenic rice by *Agrobacterium*-mediated transformation. **(a)** The dehusked seeds of rice (var. PB1) were sterilized and induced to produce calli on Callus Induction (CI) medium. **(b)** The induced calli were separated, cut into smaller pieces and sub-cultured on fresh CI medium. **(c)** The calli were collected and incubated with *Agrobacterium tumefaciens* (LBA4404 strain, harboring the recombinant plasmid) in Resuspension medium for 15 minutes with gentle shaking and transferred onto Co-cultivation medium (CCM) for about 2 days (until slight growth of *Agrobacterium* appears on the base of calli). **(d)** The calli were collected, washed with Cefotaxime solution (250 mg/L) to eliminate the *Agrobacterium*, and transferred onto Selection medium containing 50 mg/L Hygromycin (to select the transformed calli) and 250 mg/L Cefotaxime (to inhibit *Agrobacterium* growth). **(e, f)** The calli were subsequently sub-cultured for two times onto fresh Selection medium after a period of about every 10 days (Selection II and Selection III). **(g)** The calli were then transferred onto Regeneration I medium. After a week, the calli were exposed to light (14hr ON, 10hr OFF) for shoot regeneration. **(h)** The calli were then transferred onto Regeneration II medium for enhancing shoot generation. **(i)** The regenerated shoots were transferred onto Rooting medium containing 30 mg/L Hygromycin in Jam bottles for root regeneration. The regenerated plants were subsequently transferred onto vermiculite in green house for hardening process and finally transferred onto soil. Note that the calli were incubated at 28°C temperature throughout the tissue culture steps and in the dark till first week of the Regeneration I step.

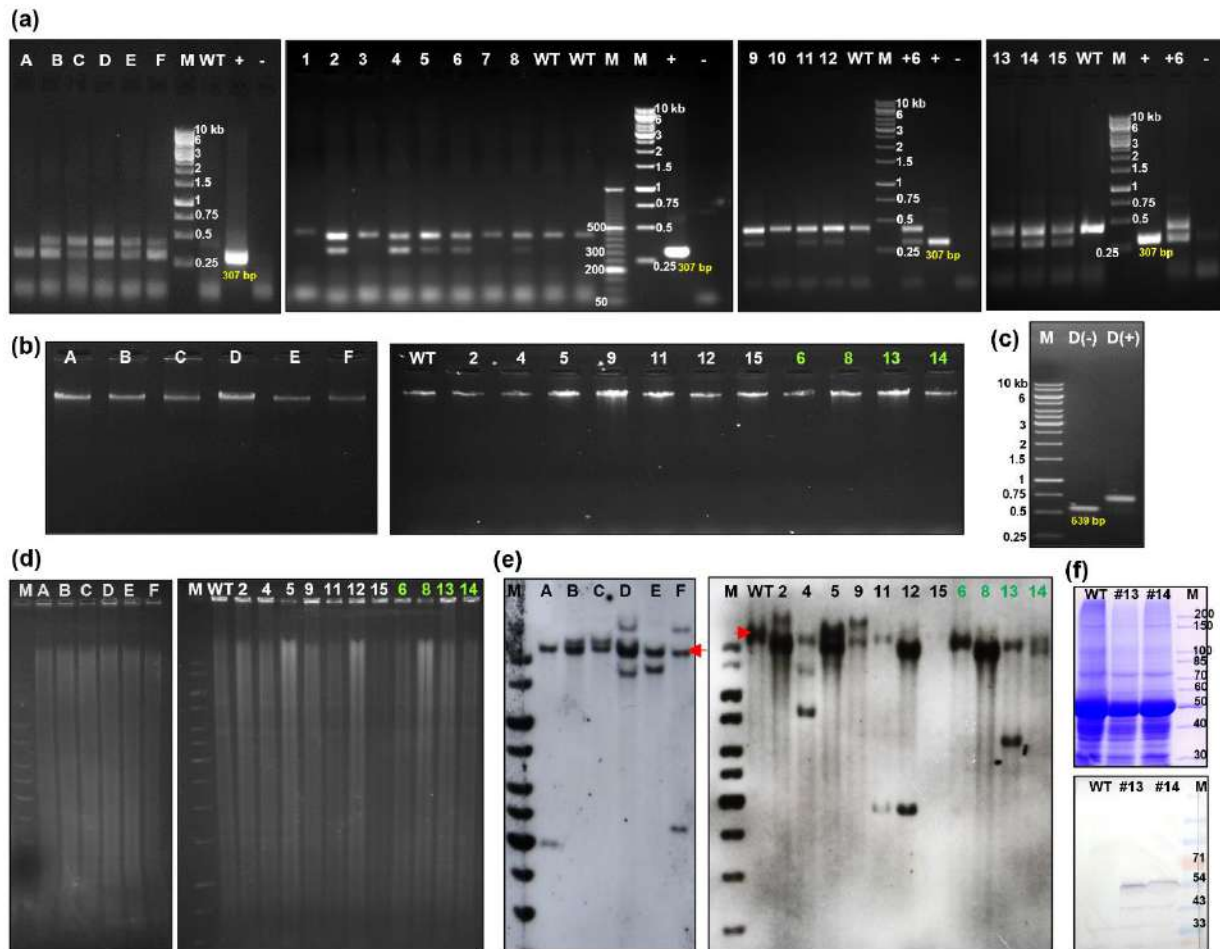


Fig. 4.22: Screening of *OsCBSX3* overexpressing (T_0) transgenic rice plants by PCR, Southern blot, and Western blot. (a) Screening of tissue culture generated transformed plants by PCR using *OsCBSX3* specific internal primers. A transgene specific product of 307 bp was obtained in transgenic plants. (b-e) Screening of the transgenic plants for transgene integration and its copy number by Southern blot. (b) Genomic DNA isolated from the PCR positive transgenic plants resolved on a agarose gel. Note that the four lines labelled in green font represents the gDNA from transgenic plants overexpressing *OsCBSX3+GFP* fusion. (c) PCR amplification of hybridization probe labelled with digoxigenin (DIG)-dUTP using full-length *OsCBSX3* primer sets. The incorporation of DIG-dUTP resulted in higher molecular weight of the PCR product {D(+): DIG-dUTP used in PCR; D(-): PCR without DIG-dUTP}. (d) Plant gDNA resolved on 0.8% agarose gel after digestion with *Bam*HI which has a single restriction site in the vector T-DNA but none in *OsCBSX3*. (e) Immuno-detection of the DIG-dUTP labelled hybridized probe or the transgene by exposing the membrane to X-ray film. (f) Detection of *OsCBSX3+GFP* overexpression in the transgenic plants by Western blot using anti-GFP antibody. The SDS-PAGE gel on the top panel depicts the loading control of the soluble proteins extracted from the plants. The membrane on the lower panel shows the immuno-detection of *OsCBSX3+GFP* (51 kDa) in transgenic plants. 'WT': Wild-type; '#13' and '#14': *OsCBSX3+GFP* transgenics; 'M': Protein marker. (a-e) 'gDNA from transgenic plants' are labelled both alphabetically (A-F) and numerically (1-15), while that from 'Wild-type plant' is labelled as 'WT'; 'M': DNA ladder; '+' and '-' are positive and no template control, respectively (in PCR). The position of endogenous *OsCBSX3* in Southern blot is marked by a red arrow.

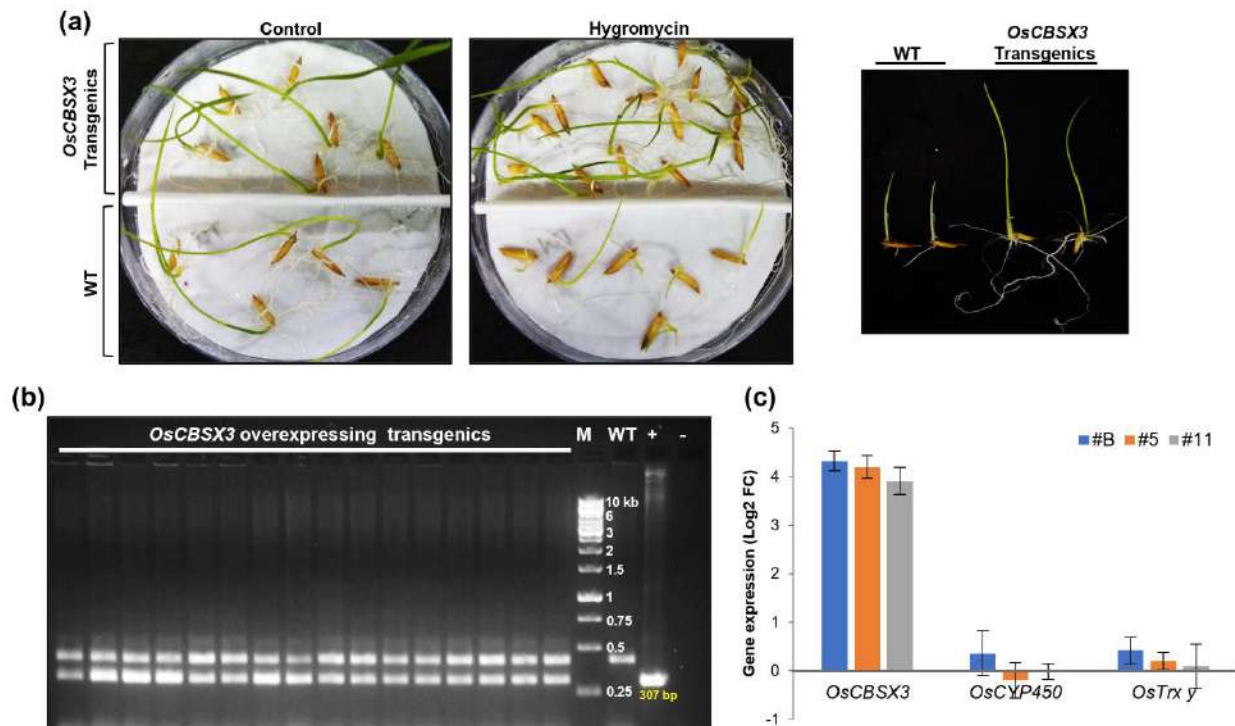


Fig. 4.23: Screening of *OsCBSX3* overexpressing T₁ plants and expression analysis of the transgene. (a) Representative image showing the screening of T₁ generation seeds from *OsCBSX3* overexpressing transgenic lines after a week of germination on hygromycin solution (30 mg per L water). The closer view of phenotypic difference between the corresponding WT and transgenic seedlings germinated on hygromycin is shown in the right panel. (b) The hygromycin tested positive seedlings were further confirmed by PCR using *OsCBSX3* specific internal primers. A transgene specific product of 307 bp was obtained from transgenic plants. ‘M’: DNA ladder; ‘WT’: Wild type; ‘+’: positive control; ‘-’: no template control. (c) Bar graph representing the qRT-PCR based expression of *OsCBSX3* along with two associated genes, *OsCYP450* and *OsTrx γ*, in shoot tissues of *OsCBSX3* overexpressing T₂-homozygous transgenic lines (#B, #11 and #5) at the seedling stage. Gene expression in these transgenic plants are plotted relative to their expression in WT plants, and the values are expressed as Log₂ of 2^{-ΔΔCT} or Log₂ of Fold Change (FC). Bar represents mean ± SD, n=3.

The overexpression of two *CBSX* homologs from *Arabidopsis*, *AtCBSX1* and *AtCBSX3*, have been reported to result in severe sterility and retarded plant growth, respectively in transgenic *Arabidopsis* lines (Yoo et al, 2011; Shin et al, 2020). However, we observed normal growth and development of *OsCBSX3* overexpressing transgenic rice lines under normal growth conditions. Although we obtained two transgenic lines with significantly taller phenotypes, we considered them as an effect of gene disruption due to transgene integration in the genome and continued our study with other transgenic lines that showed normal growth phenotypes. As our expression study identified *OsCBSX3* to be induced by stresses, particularly drought and salinity, we analyzed the growth and yield performances of three independent *OsCBSX3* overexpressing homozygous T₂ transgenic lines under these two stresses. The plants grown in pots were subjected to stress treatment at panicle emergence stage (90 days post-germination) by immersing the pot in 150 mM NaCl solution for salinity stress, and by

withholding the irrigation for drought stress (Fig. 4.24). Salinity was maintained regularly by leveling the volume of NaCl solution lost by evaporation with water. The soil electrical conductivity (EC) on the final day of the salinity stress (40 days of stress) was observed to be 14 ± 1 dS/m (soil solution temperature during EC measurement was 19°C). The drought stress was imposed for 2 weeks until the volumetric water content of the soil reached 20%, as measured by WaterScout SM100 Soil Moisture Sensor (Spectrum Technologies, Inc.).

The physiological performance of the plants analyzed based on the chlorophyll-a fluorescence emission (using Handy PEA⁺, Hansatech Instruments) and the conduction of CO₂ and H₂O through stomata in the leaves {using Infra-red Gas Analyzer (IRGA), Licor-xT} indicated better endurance of *OsCBSX3* transgenic lines to both salinity and drought stresses than the WT counterparts. The transgenic plants exhibited a significantly higher range of Fv/Fm and performance index which represents the chlorophyll-a fluorescence-based maximum quantum efficiency of Photosystem II (PSII) and the conservation of energy from excitation of PSII to a reduction of PSI, respectively (Fig. 4.25 a, b). Accordingly, the photosynthesis rate, the transpiration rate, and the stomatal conductance on account of the exchange of CO₂ and H₂O in the leaves were also observed to be significantly higher in *OsCBSX3* overexpressing transgenic lines than in the WT plants (Fig. 4.25 c-e). In addition to stress, the transgenic plants were observed to exhibit better physiological performance than the WT plants in control conditions as well. And consistent with these physiological performances, the *OsCBSX3* transgenic plants also produced slightly but significantly higher grain yield than the WT plants under both salinity and drought stress conditions (Fig. 4.26).

Therefore, based on these results, we suggest that the overexpression of stress-inducible *OsCBSX3* renders better physiological performances and higher yield per plant under salinity and drought stresses. Besides, we also infer that the overexpression of *OsCBSX3* CDS through transgenic approach has no effect on the expression of *OsCYP450* (a gene that overlaps with *OsCBSX3* at its opposite strand in the genome) and *OsTrx y* (a gene whose product interacts with *OsCBSX3*) in the transgenic plants.

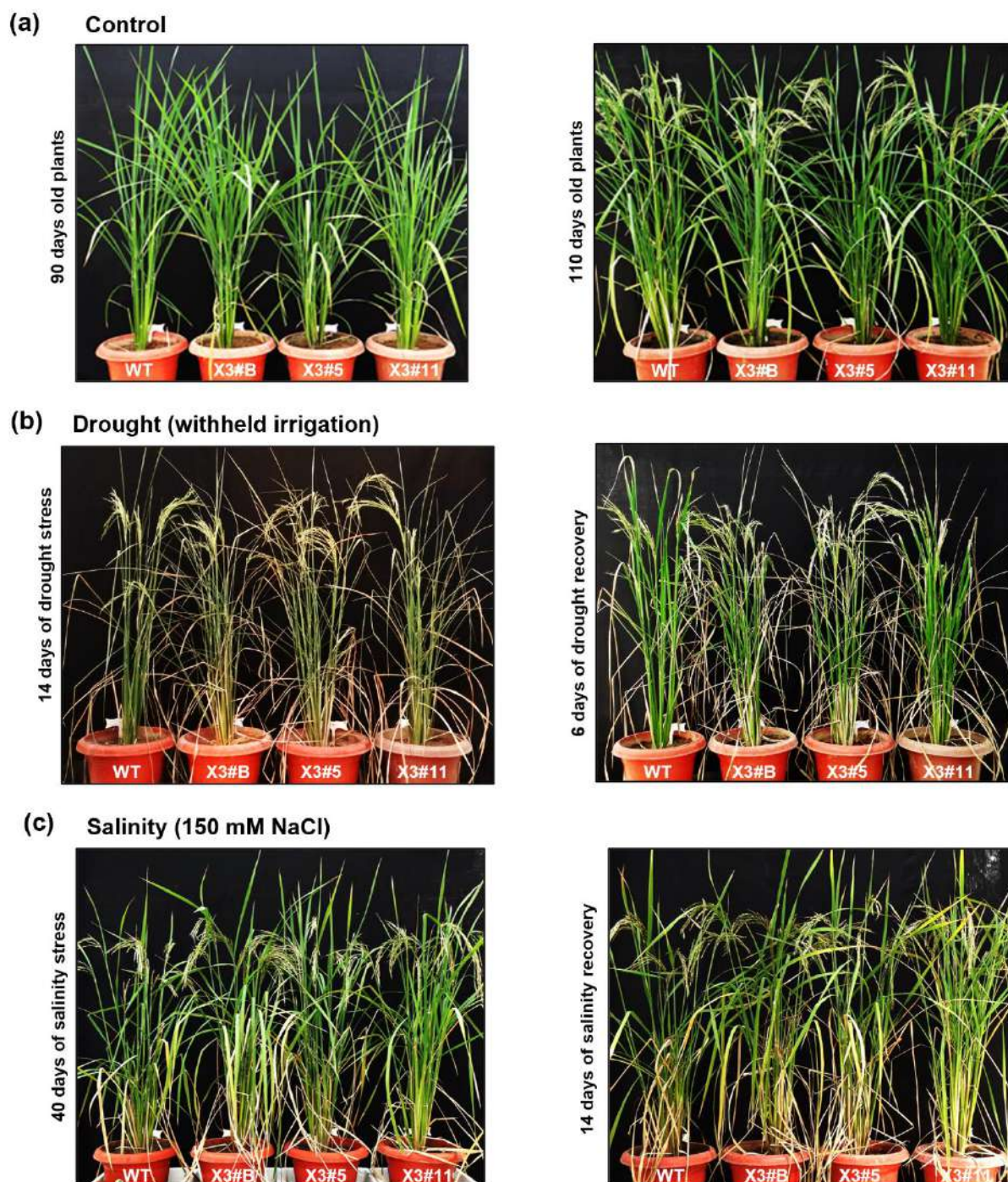


Fig. 4.24: *OsCBSX3* overexpressing T₂ transgenic rice lines under control, drought, and salinity stressed conditions. (a) Plants grown under control conditions at 90 days (left panel) and 110 days (right panel) after sowing (DAS). (b) Plants subjected to drought stress by withholding irrigation for 14 days (left panel; 104 DAS) and 6 days after the drought stress recovery (right panel; 110 DAS). (c) Plants subjected to salinity stress by keeping the pots in 150 mM NaCl solution for 40 days (left panel; 130 DAS) and 14 days after the salinity stress recovery (right panel; 144 DAS). Plants were subjected to drought or salinity stress on at 90 DAS. ‘WT’: Wild type plant; ‘X3#B’, ‘X3#11’ and ‘X3#5’ are the three independent single copy integrated *OsCBSX3* overexpressing homozygous T₂ lines. ‘DAS’: days after sowing.

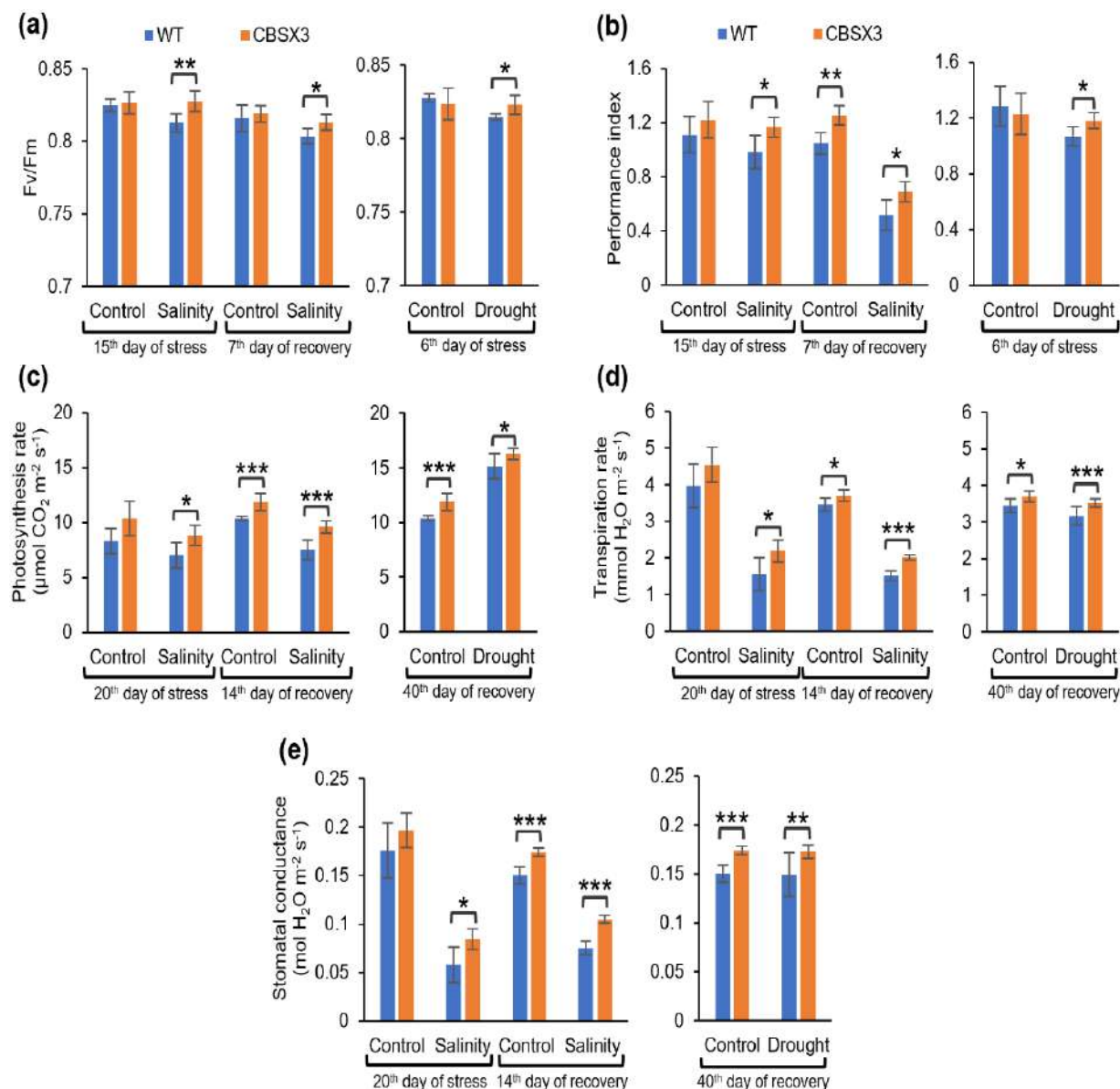


Fig. 4.25. Physiological performance of *OsCBSX3* overexpressing transgenic plants. Bar graphs in (a) and (b) represent two photosynthesis parameters: Fv/Fm and the Performance Index, respectively, determined based on chlorophyll-a fluorescence emission from a dark-adapted penultimate leaf using Handy PEA (Hansatech Instruments). The fluorescence parameters from salinity treated plants were measured on 15th day of stress treatment and on 7th day post stress recovery, while that from drought treated plants were measured on 6th day of stress treatment. The graphs in (c), (d), and (e) represent photosynthesis rate, transpiration rate, and the stomatal conductance, respectively, determined based on CO₂ and H₂O exchange in a penultimate leaf using IRGA (LI-6400XT, Licor). The gas exchange parameters from salinity treated plants were measured on 20th day of stress treatment and on 14th day post stress recovery, while that from drought plants were measured on 40th day post stress recovery. Bar represents mean \pm SD, n=6. *: P \leq 0.05; **: P \leq 0.01; ***: P \leq 0.001 (calculated by unpaired, two-tailed Student's *t*-test).

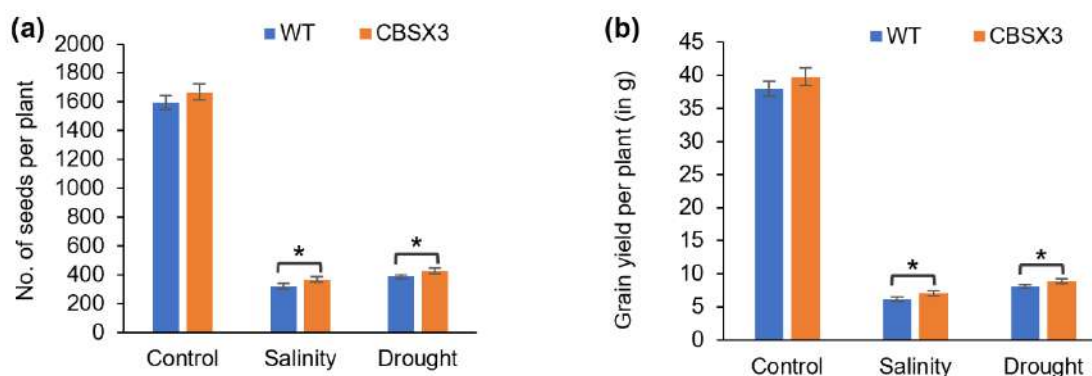


Fig. 4.26. Yield of *OsCBSX3* overexpressing transgenic plants under control and stress conditions. Bargraphs in (a) and (b) depict the average number of filled seeds and average weight of filled grains per plant, respectively in *OsCBSX3* transgenic and the WT plants under control as well as salinity and drought stress conditions. Bar represents mean \pm SD, n=3. *: $P \leq 0.05$ (calculated by unpaired, two-tailed Student's *t*-test).

4.1.4 Growth assessment of transgenic *Arabidopsis* overexpressing *OsCBSX3*

The overexpression of two *CBSX* homologs from *Arabidopsis*, *AtCBSX1* and *AtCBSX3*, have been reported to cause sterility and growth reduction, respectively in transgenic *Arabidopsis* plants (Yoo et al, 2011; Shin et al, 2020). To analyze the effect of *OsCBSX3* in the growth and development of *Arabidopsis*, we also generated *OsCBSX3* overexpressing *Arabidopsis thaliana* (Ecotype Col-0) by *Agrobacterium*-mediated floral-dip transformation method. *OsCBSX3* cloned into pCAMBIA1302 (the one used to produce rice transgenics) was transformed into *A. tumefaciens* (GV3101 strain), which was subsequently used for *Arabidopsis* transformation. The positive T₁ seeds from transformed plants were screened by germinating on half-strength MS medium supplemented with Hygromycin (25 mg/L), and the hygromycin positive plants were subsequently screened for the transgene by PCR (Fig. 4.27 a, b). The T₂ and T₃ generation seeds from each T₁ line were further screened on Hygromycin (25 mg/L) containing half-strength MS medium to obtain the homozygous T₃ plants, which were subsequently used for the plant growth analysis (Fig. 4.27 c). The expression of *OsCBSX3* in *Arabidopsis* T₃ transgenic lines was confirmed by qRT-PCR using *OsCBSX3* specific primers (Fig. 4.27 d).

The T₃ homozygous transgenic lines of *Arabidopsis* overexpressing *OsCBSX3* exhibited normal fertilization and growth as WT plants (Fig. 4.27 e, f), which is contrasting to *AtCBSX1* and *AtCBSX3* overexpression phenotype. Therefore, based on this observation, we infer the function of *OsCBSX3* is diverged and distinct from that of *Arabidopsis AtCBSX1* and *AtCBSX3* homologs.

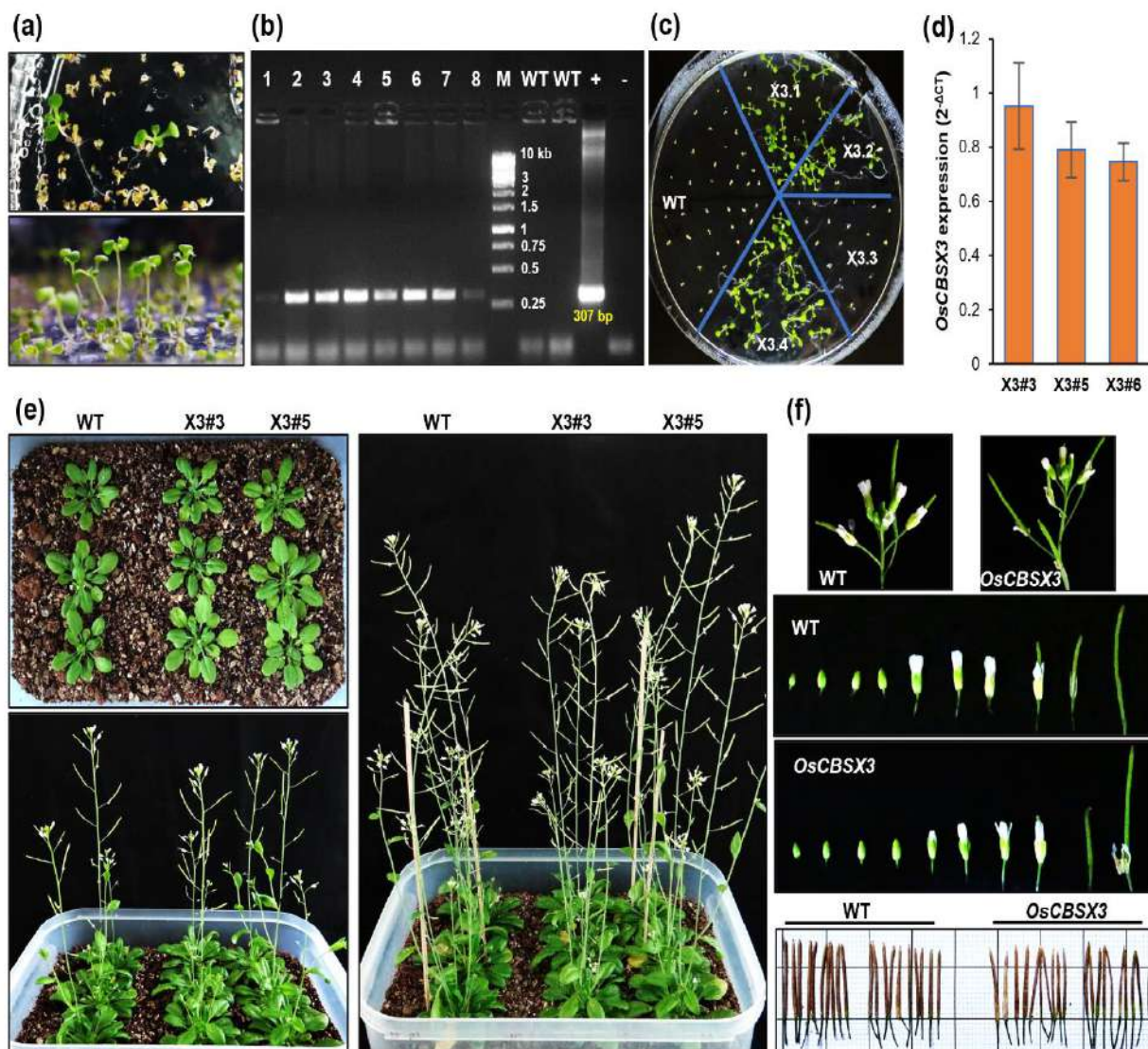


Fig. 4.27: Screening and growth analysis of *OsCBSX3* overexpressing transgenic *Arabidopsis* plants. (a) Screening of *Arabidopsis* T₁ seeds overexpressing *OsCBSX3* obtained from *Agrobacterium* transformed plants by germinating on half-strength MS medium containing Hygromycin (25 mg per L). (b) Screening of the Hygromycin positive plants for the presence of transgene by PCR using *OsCBSX3* specific internal set of primers. Transgenic lines are labelled numerically; 'M': DNA ladder; 'WT': Wild type; '+': positive control; '-': no template control (c) Screening of T₃ seeds by germinating on half-strength MS medium containing Hygromycin (25 mg per L) to identify the homozygous lines while eliminating the segregating population. (d) Confirmation of *OsCBSX3* expression in T₃ homozygous lines by qRT-PCR. The transgene expression is shown as 2^{-ΔCT}. Bar represents mean ± SD, n=3. (e) Growth phenotype of *OsCBSX3* overexpressing T₃ transgenic *Arabidopsis* plants and the WT at vegetative, early flowering and late flowering stages. (f) Comparison of flowers at its different development stages and the matured siliques from *OsCBSX3* overexpressing transgenic and the WT *Arabidopsis* plants.

When subjected to salinity and drought (the conditions that induce *OsCBSX3* expression), as well as the oxidative stress (as *OsCBSX3* was identified to interact with *OsTrx y*, a redox protein), the *OsCBSX3* overexpressing T₃ transgenic seedlings of *Arabidopsis* showed better growth (mainly the shoots) than the WT plants on all these treatments (Fig. 4.28). Moreover, the transgenic plants also exhibited better growth than WT plants under the control conditions.

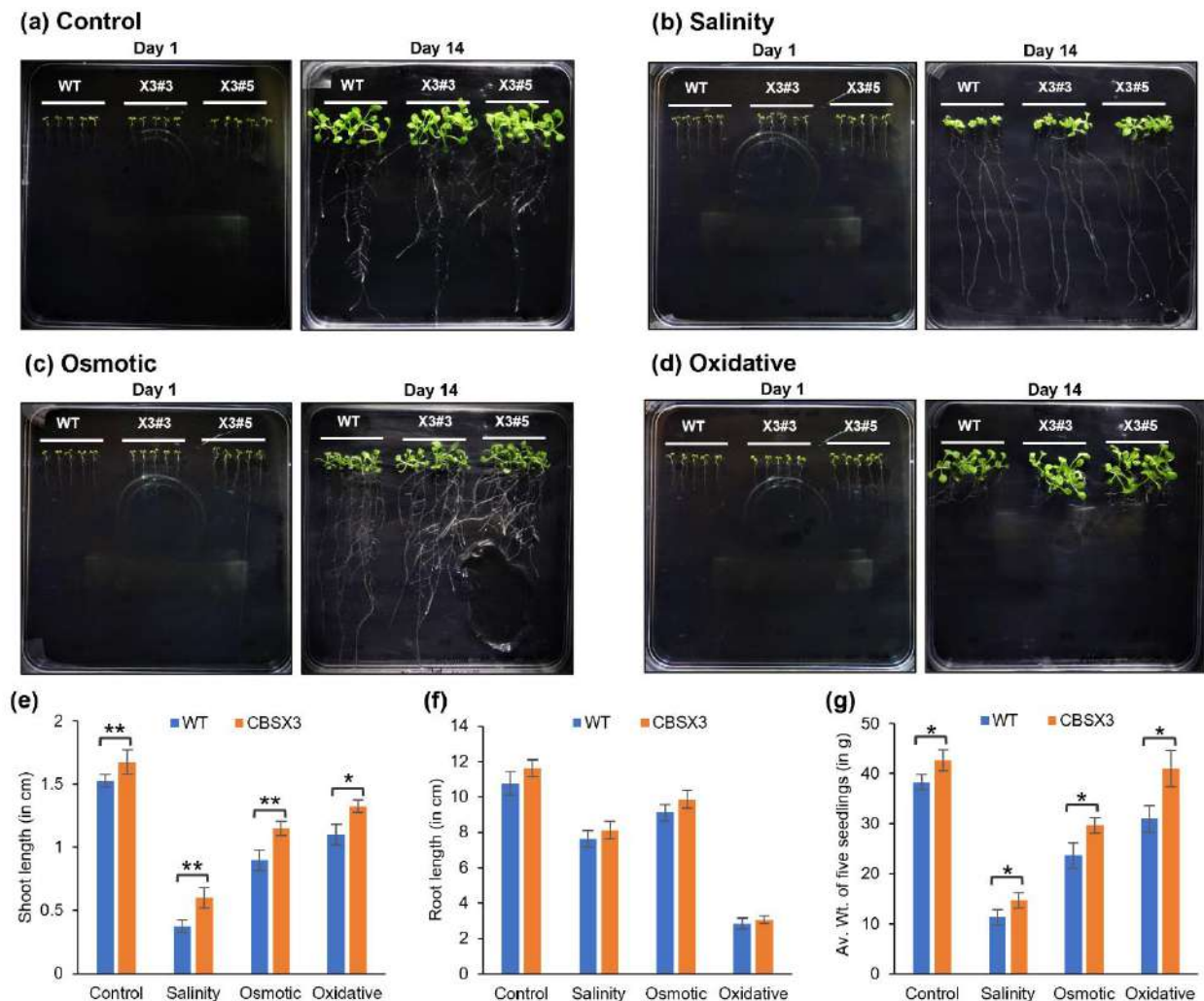


Fig. 4.28: Growth analysis of *OsCBSX3* overexpressing transgenic *Arabidopsis* seedlings. Images in (a) to (d) depicts the growth of *Arabidopsis* seedlings on half-strength MS medium under control, salinity (100 mM NaCl), osmotic (100 mM Mannitol), and oxidative (1 mM H₂O₂) stress conditions, respectively. The seedlings were transferred onto the given plates on 5th day of germination on Half-strength MS medium, which is referred as day one of the treatment (on the left panel). These seedlings on 14th day of treatment are shown on the right panels. The bar graphs in (e), (f), and (g) show shoot length, root length, and an average weight of five seedlings, respectively under control as well as salinity, osmotic, and oxidative stress conditions. Bar represents mean \pm SD; n=4 in (e) and (f), and n=3 in (g). *: P \leq 0.05; **: P \leq 0.01 (calculated by unpaired, two-tailed Student's *t*-test).

The *OsCBSX3* overexpressing T₃ transgenic *Arabidopsis* also exhibited significantly superior growth over WT plants when subjected to salinity (irrigation with 150 mM of NaCl solution) and drought (irrigation withheld) stress treatments at the transition of vegetative to reproductive phase till the maturity of the plants. The transgenic plants produced longer inflorescence stem with a greater number of siliques in both the stresses (Fig. 4.29).

Therefore, with all these observations on the growth performance of the *OsCBSX3* overexpressing transgenic *Arabidopsis*, it can be inferred that *OsCBSX3* functions in stress tolerance in plants.

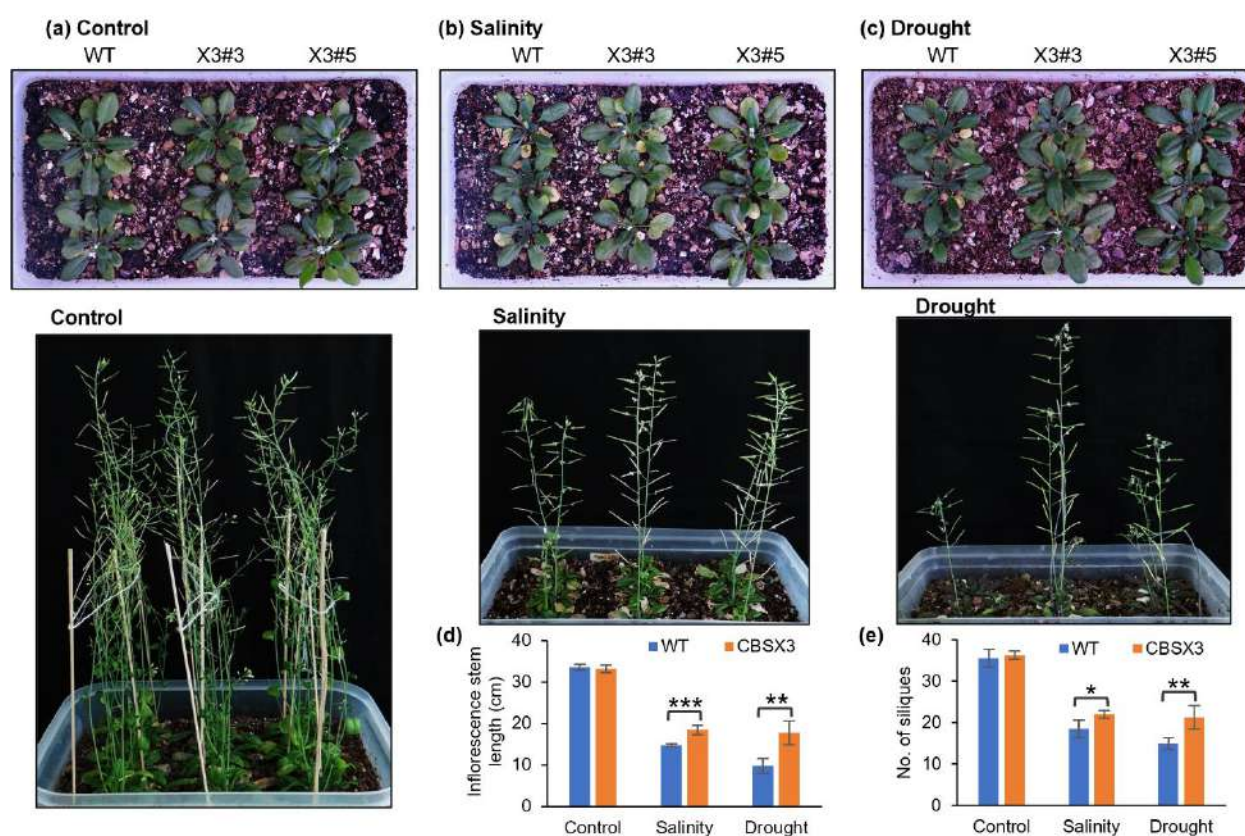


Fig. 4.29: Performance of *OsCBSX3* overexpressing transgenic *Arabidopsis* plants under salinity and drought stress. Images in (a), (b), and (c) on the upper panel depicts *Arabidopsis* plants at the transition of vegetative to reproductive stage, which were then imposed with the control (normal irrigation), salinity (irrigating with 150 mM NaCl solution), and drought (irrigation withheld) conditions, respectively till the maturity phase. The corresponding images in the lower panel depict the respective plants at maturity phase. The bar graphs in (d) and (e) show inflorescence stem length and the number of siliques per plant, respectively. Bar represents mean \pm SD; n=4. *: $P \leq 0.05$; **: $P \leq 0.01$; ***: $P \leq 0.001$ (calculated by unpaired, two-tailed Student's *t*-test).

4.2 Functional characterization of OsCBSCLC6

4.2.1 Gene expression analysis of *OsCBSCLC6*

The qRT-PCR based transcript abundance of *OsCBSCLC6* (LOC_Os08g20570) was analyzed in two contrasting rice genotypes, viz., IR64 (salinity-sensitive) and Pokkali (salinity-tolerant), in different tissues at different developmental stages as well as in the shoots of two-week-old seedlings subjected to various stress treatments. *OsCBSCLC6* was observed to be expressed in all tissues analyzed; however, its expression was predominantly higher in the roots during all three developmental stages, viz., seedling, tillering, and early flowering (Fig.4.30 a). Both Pokkali and IR64 exhibited a similar expression pattern of *OsCBSCLC6* in each tissue at a given developmental stages. Corresponding to our qRT-PCR based expression data, the *in-silico* expression analysis using mRNA-seq datasets obtained from the Genevestigator tool also exhibited significantly higher expression of *OsCBSCLC6* in roots than in the shoot tissues (Fig. 4.30 b).

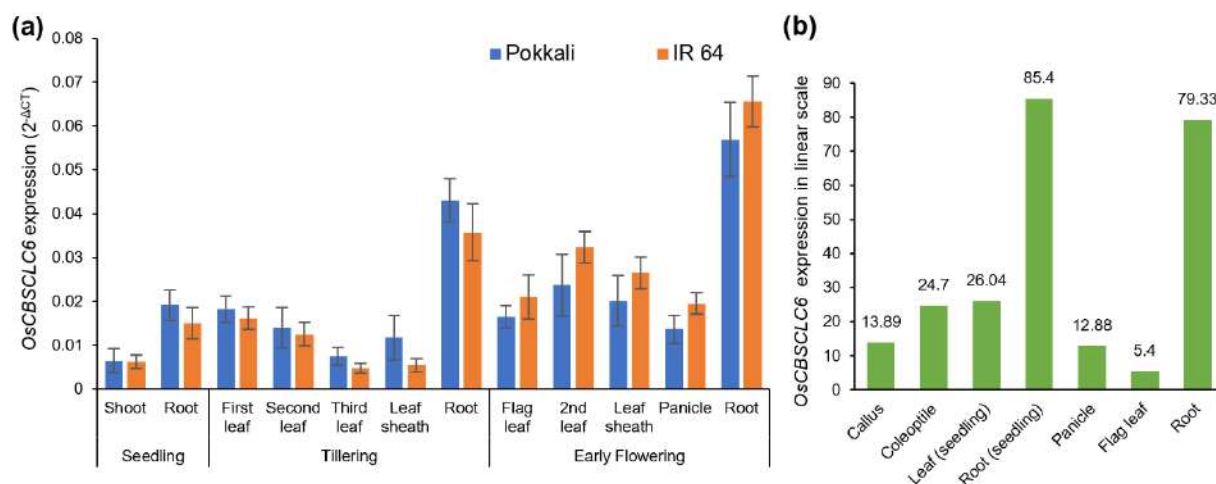


Fig. 4.30: Transcript abundance of *OsCBSCLC6* in different tissues at the seeding, tillering, and flowering stages of rice. (a) The qRT-PCR based expression of *OsCBSCLC6* in two contrasting rice genotypes, IR64 and Pokkali, in various tissues at different developmental stages, viz., shoot and root tissues of 2-week-old seedlings; top three leaves, leaf sheath of the third leaf and root tissues at tillering stage; and the flag leaf, penultimate leaf, leaf sheath of the penultimate leaf, panicle and root tissues at flowering stage. The expression level is expressed as 2^{-ΔCT}. Bar represents mean ± SD, n=3. (b) Transcript level of *OsCBSCLC6* in different tissues of rice based on mRNA-seq datasets retrieved from Genevestigator tool. The transcript abundance is expressed in a linear scale as presented in the Genevestigator.

Among various stress treatments imposed on the rice seedlings, the expression of *OsCBSCLC6*, as analyzed by qRT-PCR, was observed to be induced only under desiccation in the shoot tissues of both Pokkali and IR64 {Log₂ of Fold Change (FC) or Log₂ of $2^{-\Delta\Delta CT}$ > 2.0; Fig. 4.31 a). Its expression was also observed to be upregulated at some time points in response to salinity, cold, and heavy metal treatments in Pokkali; however, it was not altered significantly (Log₂ FC >2.0 or <-2.0) in response to other stress treatments in both the rice genotypes. The *in-silico* expression analysis using Affymetrix GeneChip and mRNA-seq datasets retrieved from Genevestigator tool (P < 0.05) also exhibited the induction of *OsCBSCLC6* in response to drought, salinity as well as to submergence (Fig. 4.31 b).

Altogether, the expression study showed that *OsCBSCLC6* is expressed mostly in root tissues of rice and its expression is upregulated in the shoots of rice seedlings mainly in response to drought stress.

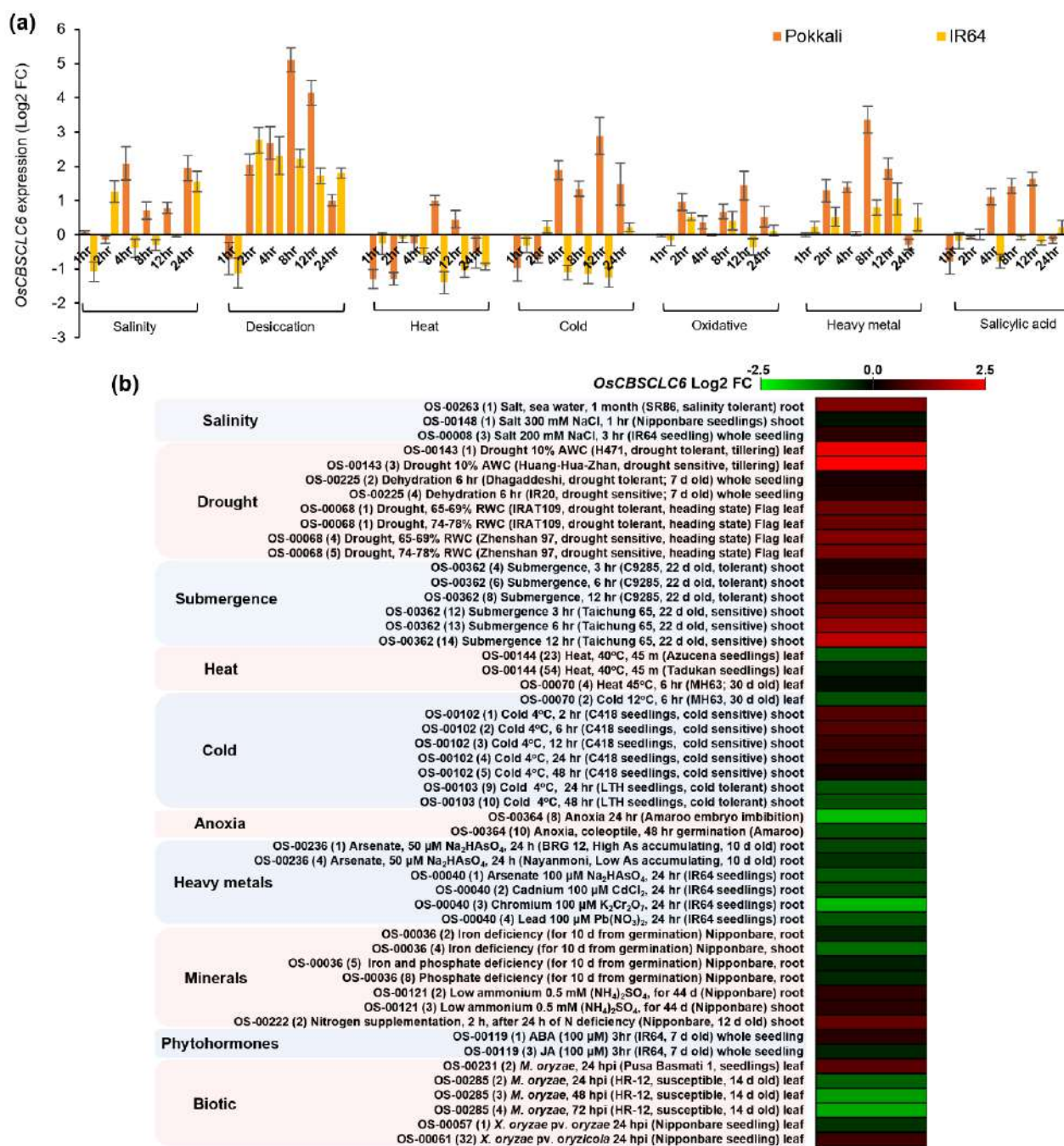


Fig. 4.31: Expression pattern of *OsCBSCLC6* in response to various stress treatments in rice. (a) Bar graph depicting the qRT-PCR based transcript abundance of *OsCBSCLC6* in shoot tissues from 2-week-old seedlings of IR64 and Pokkali, subjected to various stress treatments for different time points. The expression levels are expressed as Log₂ of fold change (Log₂ FC or Log₂ of 2^{- $\Delta\Delta$ CT} relative to the expression in control samples). Bar represents mean \pm SD, n=3. (b) Heat-map depicting the transcript level of *OsCBSCLC6* in different tissues of rice under different stress treatments based on both the Affymetrix GeneChip and the mRNA-seq datasets ($P < 0.05$) retrieved from Genevestigator tool. The transcript levels are expressed as Log₂ FC relative to its expression in control samples. Heat-map was generated using MeV 4.9.0 tool.

4.2.2 Subcellular localization of OsCBSCLC6

To determine the subcellular localization of OsCBSCLC6, its CDS was cloned into pCAMBIA1302 vector under CaMV 35S promoter and in-frame with the CDS for GFP to express OsCBSCLC6+GFP fusion protein (Fig. 4.32 a-d). The transient overexpression of OsCBSCLC6+GFP fusion protein in the leaf epidermal cells of tobacco (*N. benthamiana*) signaled the probable localization of the fusion protein in the plasma membrane; however, it was not very distinct from the fluorescence signal obtained from only GFP that localizes in the cytoplasm (Fig. 4.32 e, f). Therefore, to further confirm its subcellular localization, the OsCBSCLC6+GFP was overexpressed transiently in rice protoplasts which clearly showed the emission of GFP fluorescence in the tonoplast (Fig. 4.32 g). Whereas the protoplast transiently overexpressing only GFP emitted fluorescence in the cytoplasm (Fig. 4.32 h). The non-transfected protoplasts showed only the basal emission of the chloroplast autofluorescence (Fig. 4.32 i). Therefore, with this result, it is inferred that OsCBSCLC6 is localized in the tonoplast of the vacuole.

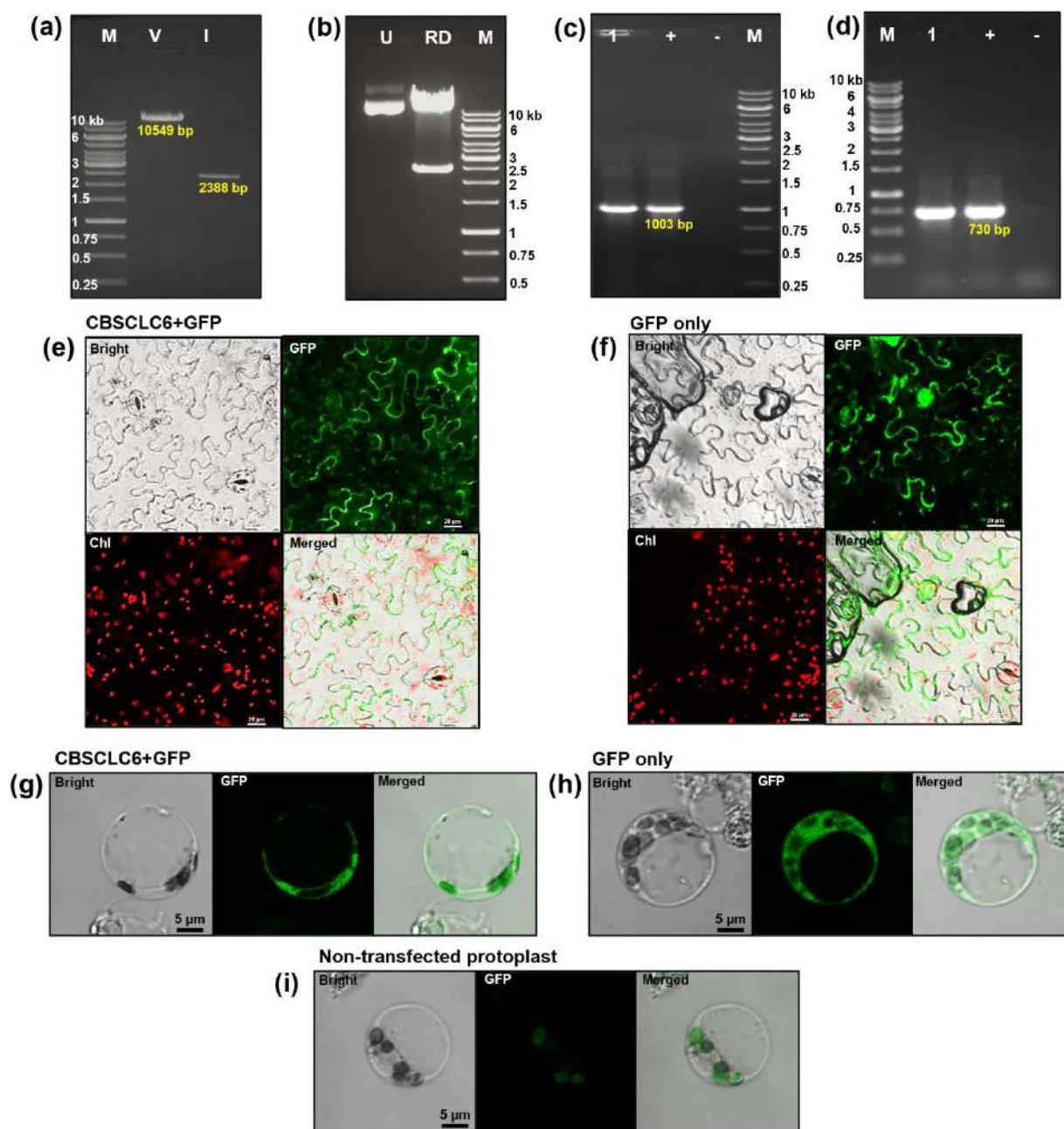


Fig. 4.32: Subcellular localization of *OsCBSCLC6*. Gel images from (a) to (d) represent the cloning of *OsCBSCLC6* in-frame with *GFP* in pCAMBIA1302. (a) Vector (V: pCAMBIA1302) and insert (I: *OsCBSCLC6*) examined on agarose gel prior to ligation reaction. (b) Screening of the transformed *E. coli* Top10 clone by restriction digestion of the isolated plasmid (U: undigested plasmid; RD: plasmid digested with *Bgl*III and *Spe*I). Agarose gel images in (c) and (d) depict the screening of the transformed *A. tumefaciens* (GV301) clone by PCR with *OsCBSCLC6* specific internal set of primers and *Vir* gene specific primers (for helper plasmid), respectively. '1': transformed clone; 'M': DNA ladder; '+': positive control; '-': no template control. Images in (e) and (f) depict the GFP fluorescence emission from tobacco leaf epidermal cells transiently overexpressing *OsCBSCLC6*+*GFP* and only *GFP*, respectively under a confocal microscope (Scale bar 20 μ m). Images in (g), (h), and (i) depict the GFP fluorescence emission from rice protoplasts transiently overexpressing *OsCBSCLC6*+*GFP*, only *GFP*, and the chlorophyll autofluorescence from untransformed protoplast, respectively under a confocal microscope (Scale bar 5 μ m).

4.2.3 CLCs functional motif sequence analysis in OsCBSCLC6

Chloride channels (CLCs) have conserved motifs or high sequence conservation in residues that form an ion conduction pathway in the transmembrane domain. Within these conserved motifs, the presence of two conserved glutamates, gating glutamate, and proton glutamate, have been identified to be critical for CLCs with the exchanger function (Accardi and Miller, 2004; Bergsdorf et al, 2009). Whereas the presence of 'Proline' or 'Serine' (in the position of 'X') in the conserved "G'X'GIP" motif has been identified to impart selectivity for NO_3^- or Cl^- transport function, respectively to CLC members (Bergsdorf et al, 2009; Zifarelli and Pusch, 2009). To gain the putative transport function of OsCBSCLC6, we analyzed its transmembrane domain sequence with the sequences from well-characterized CLCs from different kingdoms, and accordingly, the phylogenetic tree was built as well. We identified OsCBSCLC6 to group with *Arabidopsis* AtCLC-c and AtCLC-g in the phylogenetic tree, sharing 'Serine' in the selectivity filter motif (Fig. 4.33). 'Alanine' was observed at a position corresponding to the 'gating glutamate', while 'serine' was found to occur at a position corresponding to the 'proton glutamate' in OsCBSCLC6. Therefore, based on transmembrane sequence analysis, we inferred OsCBSCLC6 as a channel protein with Cl^- selectivity.

We also analyzed the cytosolic C-terminal region of OsCBSCLC6 consisting of a pair of CBS domains for the conservation of ligand-binding residues in relation to CBS domains from various structurally characterized CDCPs (see Fig. 4.12 in section 4.1.3.3). We identified its CBS2 to have conserved Thr (T) and Asp (D) in "G-h-h'-T/S-y-y'-D/N" motif, while in CBS1, Leu and His were found to occupy the respective positions. Likewise, we observed the conserved Thr residue to be replaced by Leu in CBS1 of both hCLC-5 and ggCLC-7, while their CBS2 possessed both Thr and Asp in "G-h-h'-T/S-y-y'-D/N" motif, and the Bateman module of these two CLCs are reported to bind only one ATP (Meyer et al, 2007; Schrecker et al, 2020). We also analyzed AtCLC-a CBS domains that are reported to be ATP/AMP regulated through electrophysiological study (De Angeli et al, 2009) and CLC-0 CBS domains that showed no ATP binding in the crystal structure (Meyer and Dutzler, 2006). Interestingly, we observed the presence of both conserved Thr and Asp in "G-h-h'-T/S-y-y'-D/N" motif in CBS2 domains of AtCLC-a, whereas both CBS1 and CBS2 from CLC-0 lacked the conserved T/S and D/N residues in the motif, which may account for binding of one ATP per Bateman module in AtCLC-a, while no binding in CLC-0. Therefore, pertaining to the occurrence of Thr and Asp in "G-h-h'-T/S-y-y'-D/N" motif in its CBS2 domain, we propose the Bateman module of OsCBSCLC6 to have a binding affinity for one adenosyl nucleotide ligand.

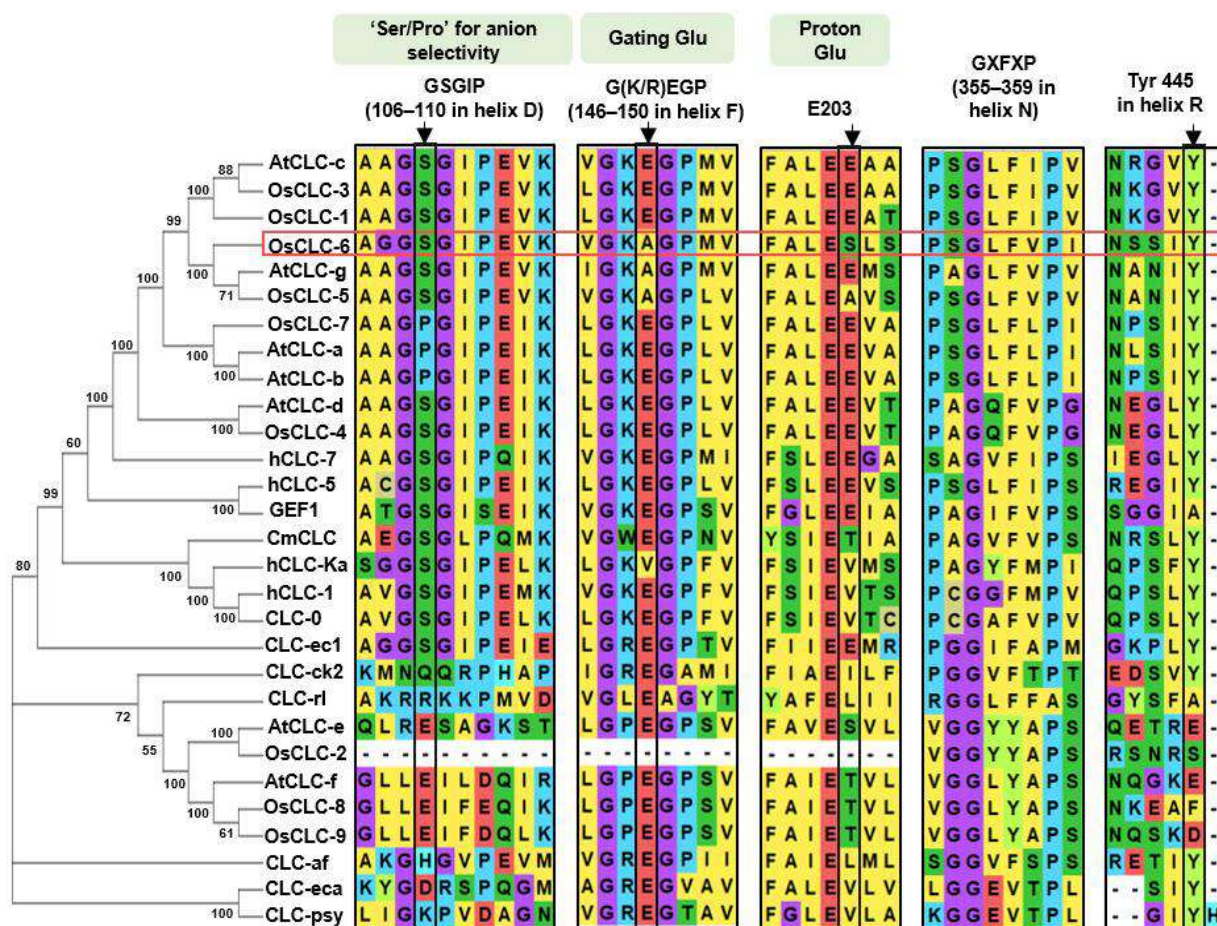


Fig. 4.33: Phylogenetic tree of different Chloride Channel proteins (CLCs) and sequence alignment of the conserved motifs that form ion conduction pore in these proteins. Image depicting the phylogenetic relation and sequence alignment of the transmembrane domain of CLCs from *Arabidopsis*, rice, and other well characterized CLCs from both eukaryotic and prokaryotic origin. The residues from four conserved motifs that form the ion conduction pore and the proton glutamate is presented in the sequence alignment. The sequences were aligned following ClustalW method (except for last block representing Tyr 445 which was prepared following the Muscle alignment method), and the phylogenetic tree was built following Neighbour-Joining method (pairwise deletion; 1000 bootstraps). The numbering of residues is in accordance with the sequence of CLC-ec1 from *E. coli*. 'OsCLC-1 to -9': OsCBSCLCs from Rice; 'AtCLC-a to -g': from *Arabidopsis*; 'hCLC-1, -5, -7, -Ka': from human; 'CLC-0': from *Torpedo marmorata*; 'GEF1': from yeast; 'CmCLC': from *Cyanidioschyzon merolae*; 'CLC-eca': from *Enterococcus casseliflavus*; 'CLC-ck2': from *Citrobacter koseri*; 'CLC-psy': from *Pseudomonas syringae*; 'CLC-af': from *Acidithiobacillus ferrooxidans*; 'CLC-rl': *Rhizobium leguminosarum*. Note that CLC-af and CLC-rl are prokaryotic CLCs that possess cytosolic CBS domains.

4.2.4 Expression of OsCBSCLC6 in *E. coli*

For heterologous expression of OsCBSCLC6 in *E. coli*, its full-length CDS was cloned into the pET28a expression vector (Fig. 4.34 a-d). The recombinant *OsCBSCLC6*_pET28a was transformed into BL21 (DE3) and C43 (DE3) strains of *E. coli*. Protein induction was carried out at 37°C and 18°C, and in three different broths ('LB': Lysogeny Broth; '2X YT': Yeast extract Tryptone; 'TB'; Terrific Broth). However, the expression of OsCBSCLC6 (expected size = 90.48 kDa) was not observed in any of the conditions (Fig. 4.34 e, f). The inefficiency of both BL21 (DE3) and C43 (DE3) strains of *E. coli* to express OsCBSCLC6 urged us to analyze the codon of *OsCBSCLC6* in relation to the rare codons in *E. coli*. We observed a high frequency of codon 'CGG' for 'Arg', followed by 'CCC', 'AGG' and 'AUA' for 'Pro', 'Arg' and 'Iso', respectively in *OsCBSCLC6* which are rare in *E. coli*. Hence, we attempted to express OsCBSCLC6 in *E. coli* Rosetta (DE3), a strain optimized for rare codons (Fig. 4.34 g). However, the induction of OsCBSCLC6 was not observed in Rosetta (DE3) as well, despite using different incubation conditions and culture media (Fig. 4.34 h).

Because of the inability to express full-length OsCBSCLC6 in *E. coli*, we next tried to express only the C-terminal cytosolic region of OsCBSCLC6 containing a pair of CBS domains by cloning 5' end deleted CDS of *OsCBSCLC6* into pET28a vector {two constructs: (a) 1716 nucleotides deleted (CLCΔ1716) and (b) 1752 nucleotides deleted (CLCΔ1752)} (Fig. 4.35 a-c). The transformed *E. coli* BL21 (DE3) cells showed expression for both the deletion constructs, however, the proteins were observed to aggregate in the insoluble fraction (Fig. 4.35 d, e). Solubilization by Guanidine HCl and Urea, and subsequent purification by Ni-NTA affinity chromatography yielded purified OsCBSCLC6, however refolding of the purified protein by dialysis resulted in protein aggregations. Therefore, due to the highly insoluble nature of the OsCBSCLC6 cytosolic fraction, we could not commence any *in vitro* studies on this protein.

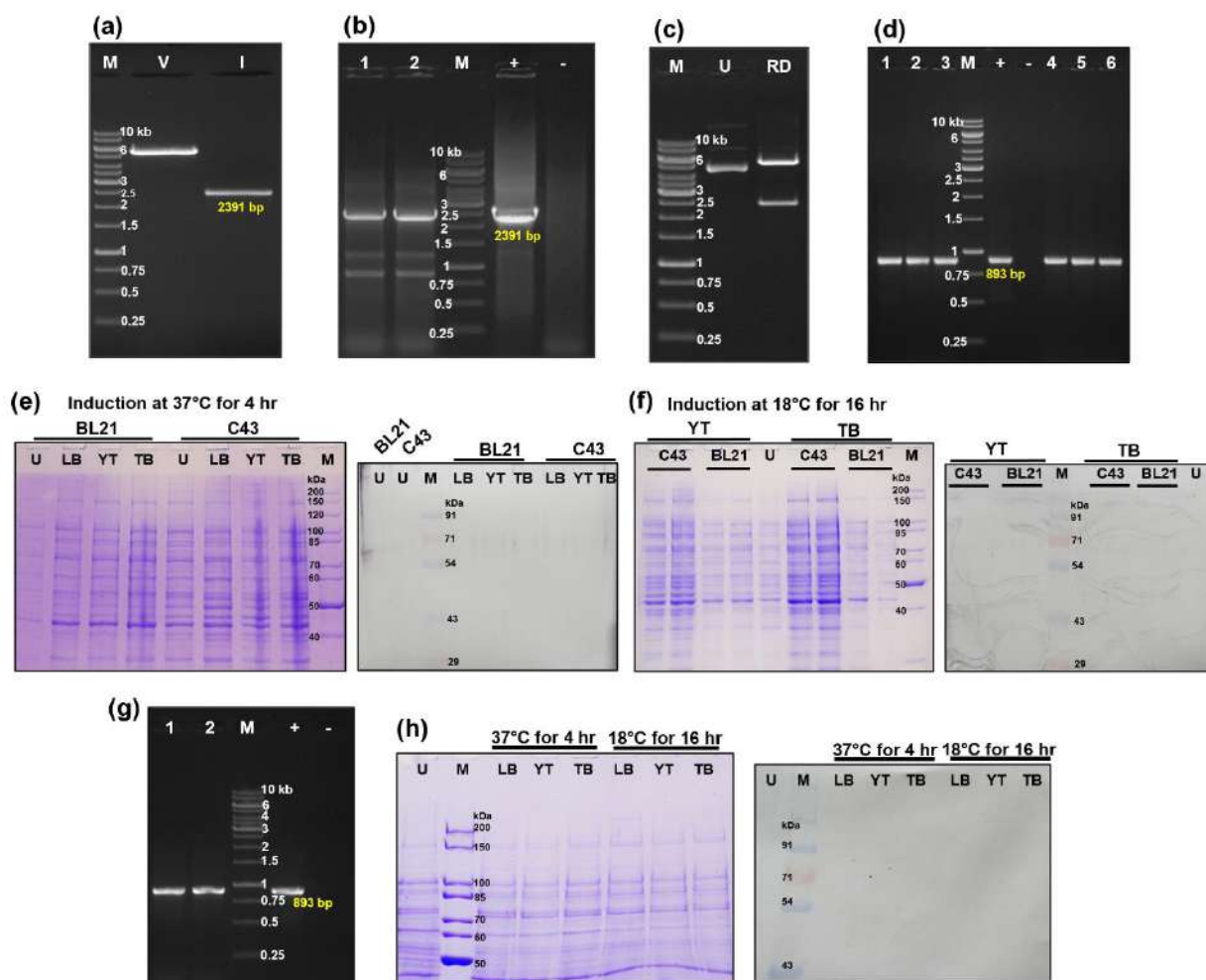


Fig. 4.34: Cloning and expression of *OsCBSCLC6* in *E. coli*. Agarose gel images in (a) to (d) illustrates the cloning of full length CDS for *OsCBSCLC6* into pET28a vector at *EcoRI* and *XhoI* restriction sites, in-frame with the upstream CDS for 6X His tag, and transformation into C43 (DE3) and BL21 (DE3) strains of *E. coli*. (a) Vector (V; pET28a) and insert (I; *OsCBSCLC6*) analyzed on agarose gel prior to ligation reaction. (b) Colony PCR screening of *OsCBSCLC6*_pET28a transformed *E. coli* Top10 clones. (c) Confirmation of PCR positive clone by restriction digestion of the isolated plasmid (U: undigested plasmid; RD: plasmid digested with *EcoRI* and *XhoI*). (d) Colony PCR screening of *OsCBSCLC6*_pET28a transformed *E. coli* C43 (DE3) and BL21 (DE3) clones labelled as '1-3' and '4-6', respectively, using vector specific T7 forward primer and *OsCBSCLC6* internal reverse primer. The positive clones were proceeded for protein expression. (e, f) Expression of 6xHis-*OsCBSCLC6* (expected size = 90.48 kDa) was analyzed by SDS-PAGE (left panel) and western blot (right panel). The total cell lysate from *OsCBSCLC6* transformed uninduced (U) and induced BL21 (DE3) and C43 (DE3) cells incubated at (e) 37°C and (f) 18°C were analyzed. (g, h) Transformation of *OsCBSCLC6*_pET28a recombinant plasmid into *E. coli* Rosetta (DE3) and its expression analysis. (g) Colony PCR screening of *OsCBSCLC6*_pET28a transformed *E. coli* Rosetta (DE3) clones using vector specific T7 forward primer and *OsCBSCLC6* internal reverse primer. (h) SDS-PAGE (left panel) and western blot (right panel) of the total cell lysate from *OsCBSCLC6* transformed uninduced (U) and induced Rosetta (DE3) cells incubated at 37°C and 18°C. Agarose gel: Bacterial clones are marked numerically; 'M': DNA ladder; '+': positive control; '-': no template control. SDS-PAGE: 'M': protein marker; 'LB': Lysogeny Broth; 'YT': 2X Yeast extract Tryptone; 'TB': Terrific Broth. Western blot is carried out using anti-His antibody conjugated with alkaline phosphatase. Note that the full-length *OsCBSCLC6* could not be expressed.

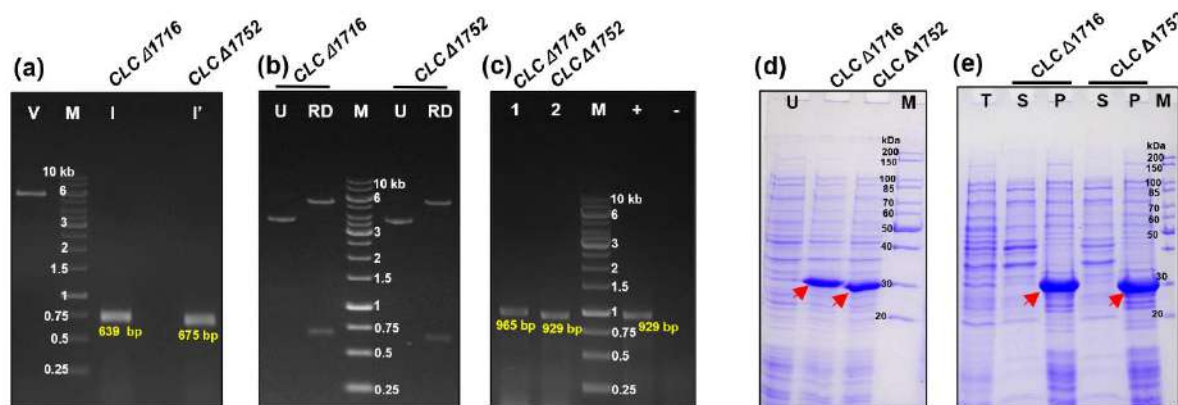


Fig. 4.35: Cloning and expression of truncated *OsCBSCLC6* that codes for its cytosolic C-terminal fraction in *E. coli*. To express the cytosolic region of *OsCBSCLC6* containing two CBS domains, the nucleotide region that corresponds to transmembrane domain was deleted from the CDS of *OsCBSCLC6*. Two different deletion fragments of *OsCBSCLC6*, deletion of (1) 1716 nucleotides (*CLCA1716*) and (2) 1752 nucleotides (*CLCA1752*) from 5' region, were cloned into pET28a. (a) Vector (V; pET28a) and inserts (I; *CLCA1716*; I': *CLCA1752*) examined on agarose gel prior to ligation reaction. (b) Screening of the transformed *E. coli* (Top10) clones by restriction digestion of the isolated plasmids (U: undigested plasmid; RD: plasmid digested with *EcoRI* and *XhoI*). (c) Colony PCR screening of transformed *E. coli* BL21 (DE3) clones using vector specific T7 forward and T7 reverse primers. (d) SDS-PAGE of the total cell lysate from uninduced (U) and induced transformed BL21 (DE3) cells. The induced protein bands of about 28.8 kDa and 27.6 kDa corresponding to 6xHis:CLC Δ 1716 and 6xHis:CLC Δ 1752, respectively are marked by red arrows. (e) SDS-PAGE of insoluble (P) and soluble (S) fractions of the cell lysate from induced cells and the total cell lysate from uninduced cells (T). Both 6xHis:CLC Δ 1716 and 6xHis:CLC Δ 1752 were observed only in the insoluble fraction (shown by red arrows). Agarose gel: 'M': DNA ladder; '+': positive control; '-': no template control. SDS-PAGE: 'M': Protein marker

4.2.5 Homo-oligomeric status of *OsCBSCLC6*

As in many CDCPs, CBS domains pairs (Bateman modules) from CLCs are also known to form homodimers. The insolubility of *E. coli* expressed cytosolic fragment of *OsCBSCLC6* impeded its homo-oligomeric study *in vitro*. Therefore, we used the yeast two-hybrid (Y2H) approach to analyze the homo-oligomerization of the cytosolic region containing a pair of CBS domains of *OsCBSCLC6*. The CDS of *OsCBSCLC6* with 1716 nucleotides deleted at 5' end (*CLCA1716*) was cloned into both pGADT7 and pGBKT7 vectors in-frame with the upstream CDS encoding activation domain (AD) and binding domain (BD), respectively (Fig. 4.36 a, b). The recombinant plasmids (*CLCA1716_pGADT7* and *CLCA1716_pGBKT7*) were co-transformed into the AH109 strain of yeast and screened on 3 DO (SD/-L/-T/-H) minimal medium. The co-transformed yeast clones, however, showed no growth on 3 DO minimal medium (Fig. 4.36 c). This result implies that Bateman modules of *OsCBSCLC6* exist as homo-monomer. However, we also consider the possible affinity of the cytosolic region of *OsCBSCLC6* to localize towards the surface membrane, as has been observed in the case of human CLC-5 (Mo et al, 2004), which can hinder the oligomeric analysis by Y2H.

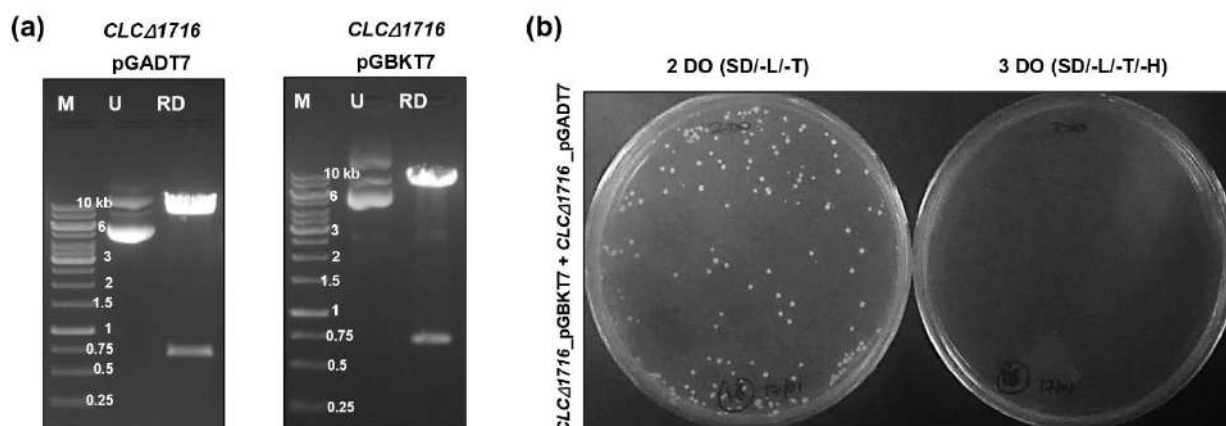


Fig. 4.36: Determination of homo-oligomerization among cytosolic Bateman modules of OsCBSCLC6 by Y2H assay. To determine the homo-oligomerization among cytosolic Bateman modules of OsCBSCLC6 by Y2H assay, its 5' 1716 nucleotide deleted CDS (*CLCA1716*) was cloned into pGADT7 (at *EcoRI* and *XhoI* site) and pGBKT7 (at *EcoRI* and *PstI* site). Agarose gel images in (a) and (b) depict the screening of *E. coli* (Top10) clones transformed with *CLCA1716_pGADT7* and *CLCA1716_pGBKT7*, respectively by restriction digestion of their plasmids (U: undigested plasmid; RD: plasmid digested with restriction enzymes; 'M': DNA ladder). (c) Growth of AH109 yeast clones co-transformed with *CLCA1716_pGADT7* and *CLCA1716_pGBKT7* on 2 DO and 3 DO minimal medium. Note that no growth of the transformed yeast cells was observed on 3 DO minimal medium.

4.2.6 Generation and growth assessment of transgenic rice plants overexpressing *OsCBSCLC6*

To assess the physiological function of *OsCBSCLC6* in rice, transgenic rice plants overexpressing *OsCBSCLC6* were generated. The full-length CDS of *OsCBSCLC6* was cloned into the pRT101 vector downstream of 35S CaMV promoter at *XhoI* and *KpnI* sites. The *OsCBSCLC6* expression cassette was then excised from pRT101 by restriction digestion with *HindIII* and ligated into pCAMBIA1300 plant binary vector (Fig. 4.37 a-c). The recombinant plasmid was transformed into *Agrobacterium tumefaciens* (LBA4404 strain; Fig. 4.37 d) which was subsequently used for transformation of the embryonic calli of rice (PB-1 variety). The transformed calli were selected on Hygromycin-containing MS medium, regenerated to shoots and roots, transferred to vermiculite, and then to soil (tissue culture steps as presented in Fig. 4.21 were followed). The transgenic plants were screened by PCR and the independent transgenic events and the transgene copy number were confirmed by Southern blot (Fig. 4.38). The T₁ seeds of independent transgenic lines with single transgene integration were screened by germination on Hygromycin selection (30 mg/L) followed by PCR (Fig. 4.39 a, b). T₂ seeds from the positive lines were collected and were screened again on Hygromycin selection (30 mg/L). The T₂ seeds from only those T₁ lines that showed 100% germination on Hygromycin (considered to be the homozygous lines) were used further for the plant growth study. The

expression analysis by qRT-PCR confirmed significantly higher expression of *OsCBSCLC6* in the homozygous T₂ transgenic plants (Fig. 4.39 c).

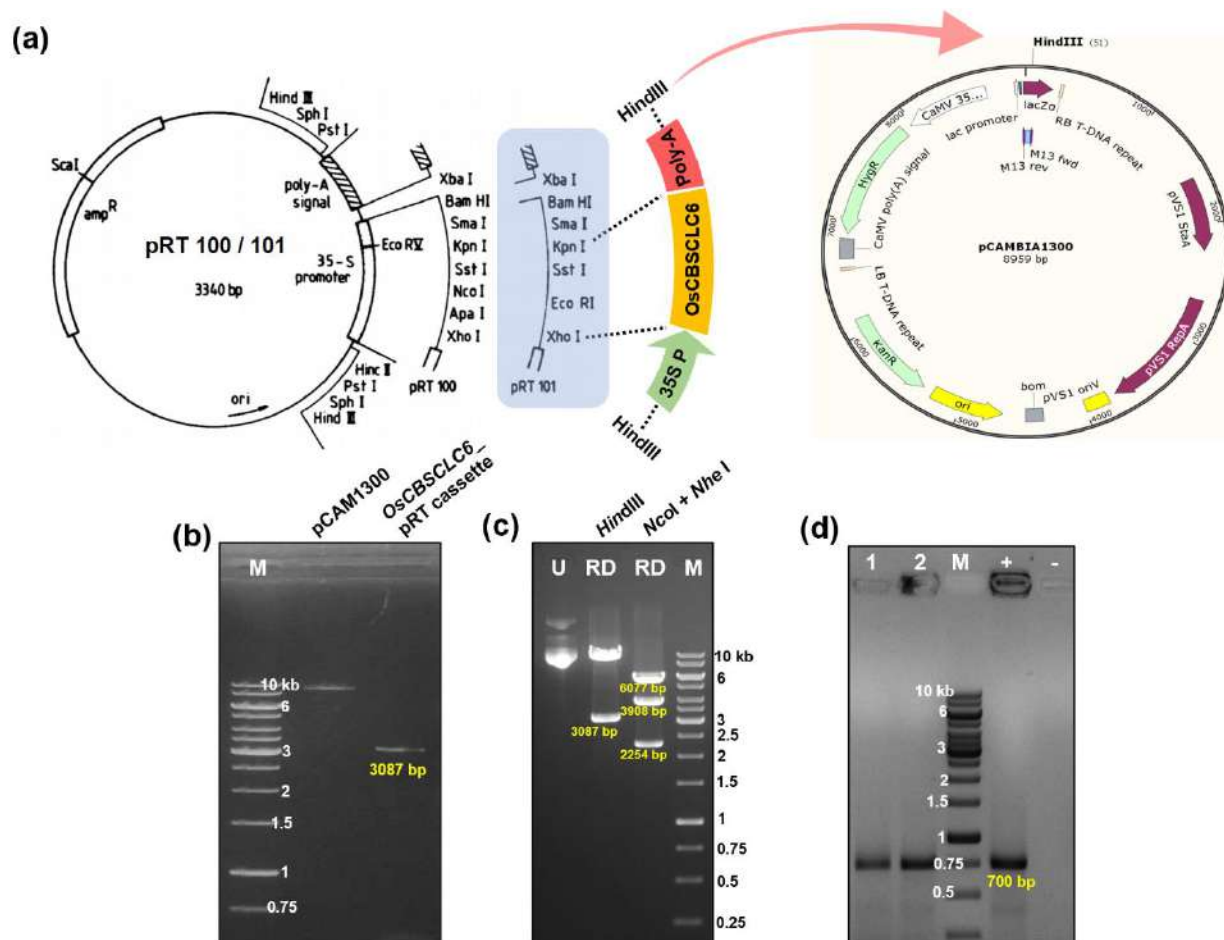


Fig. 4.37: Cloning of *OsCBSCLC6* into pCambia1300 plant transformation binary vector through pRT101. (a) Schematic representation of pRT vector map (acquired from Töpfer et al, 1987), the available cloning sites, and *OsCBSCLC6* expression cassette, and the pCambia1300 vector map (prepared using SnapGene; snapgene.com). The CDS of *OsCBSCLC6* was inserted downstream of CaMV 35S promoter (35S P) and upstream of CaMV polyadenylation signal (Poly-A) in pRT101, and this cassette was excised by restriction digestion with *HindIII* and inserted into pCambia1300 at *HindIII* site. (b) pCambia1300 and 35S P-*OsCBSCLC6*-poly-A cassette examined on agarose gel prior to ligation reaction. (c) Screening of the transformed *E. coli* Top10 clone by restriction digestion of the isolated plasmid (U: undigested plasmid; RD: plasmid digested restriction enzyme). (d) Colony PCR screening of the transformed *A. tumefaciens* (LBA4404) clones with *OsCBSCLC6* specific internal set of primers. Clones are marked numerically; 'M': DNA ladder; '+': positive control; '-': no template control.

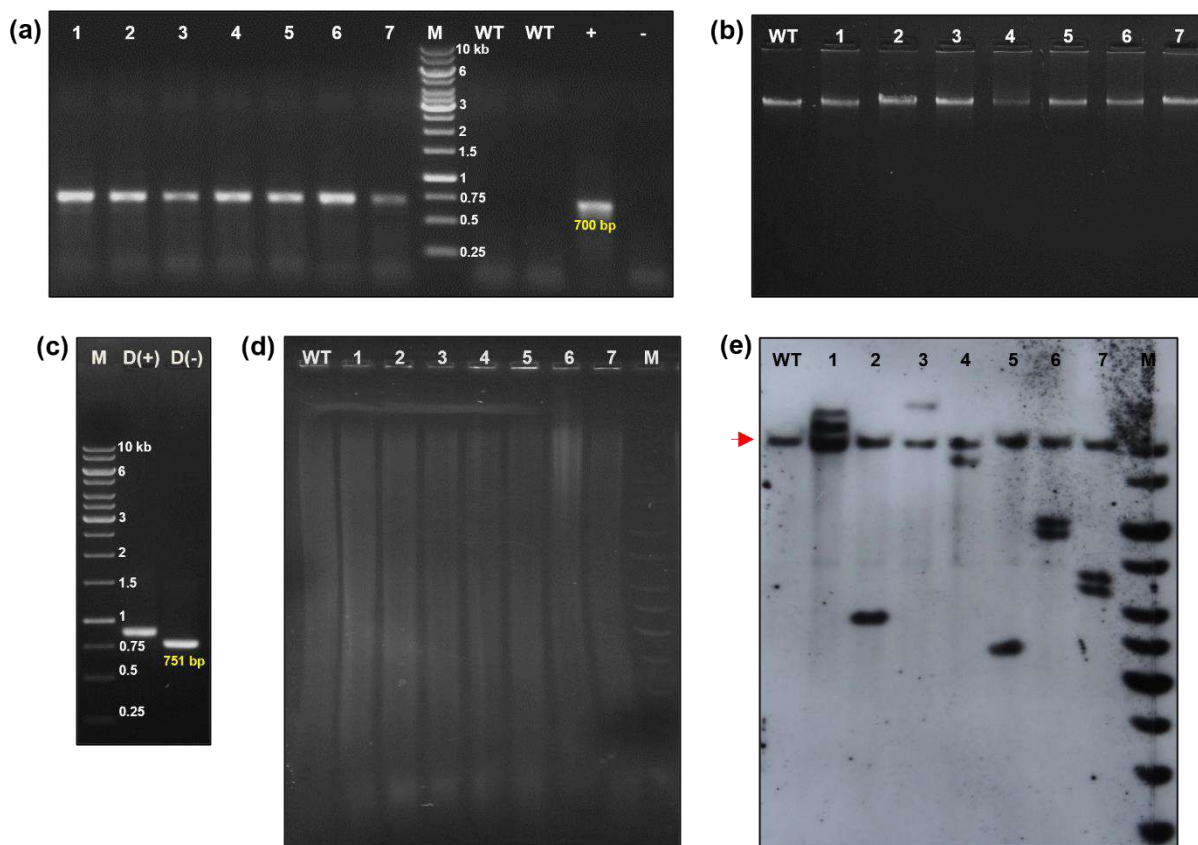


Fig. 4.38: Screening of *OsCBSCLC6* overexpressing T_0 transgenic rice plants by PCR followed by Southern blot. (a) PCR screening of the tissue culture generated putative transgenic plants using *OsCBCLC6* CDS specific internal set of screening primers. A transgene specific product of 700 bp was obtained in transgenic plants. (b) Genomic DNA isolated from the PCR positive transgenic plants. (c) PCR amplification of the hybridization probe labelled with digoxigenin (DIG)-dUTP using *OsCBSCLC6* internal set of primers (amplifies 1641 to 2391 nucleotide region of *OsCBSCLC6*). The incorporation of DIG-dUTP resulted in higher molecular weight of the PCR product {D(+): PCR with DIG-dUTP used; D(-): PCR without DIG-dUTP}. (d) Plant gDNA digested with *EcoRI* restriction enzyme and resolved on 0.8% agarose gel. *EcoRI* has only a single restriction site in T-DNA region of the vector and none in the CDS of *OsCBSCLC6*. In endogenous *OsCBSCLC6*, *EcoRI* has a single cutting site at 1089 nucleotide, which lies in the intronic region and before 1641 to 2391 nucleotide region of *CBSCLC6* CDS that we used for hybridization probe synthesis. (e) Determination of the *OsCBSCLC6* transgene integration and copy number in transgenic plants by immuno-detection of the DIG-dUTP labelled hybridized probe using anti-DIG antibody conjugated with Alkaline Phosphatase. Disodium 3-(4-methoxyspiro {1,2-dioxetane-3,2'-(5'-chloro) tricyclo decan}-4-yl) phenyl phosphate (CSPD) was used as a chemiluminescent substrate for alkaline phosphatase and the membrane was exposed to X-ray film. The position of endogenous *OsCBSCLC6* is marked by a red arrow. Transgenic lines: 2, 3, 4 and 5 have single copy of transgene, while 1, 6 and 7 have two copies of transgene. 'Transgenic plants' are labelled numerically; 'WT': Wild type plant; 'M': DNA ladder; '+' and '-' are positive and no template control, respectively (in PCR).

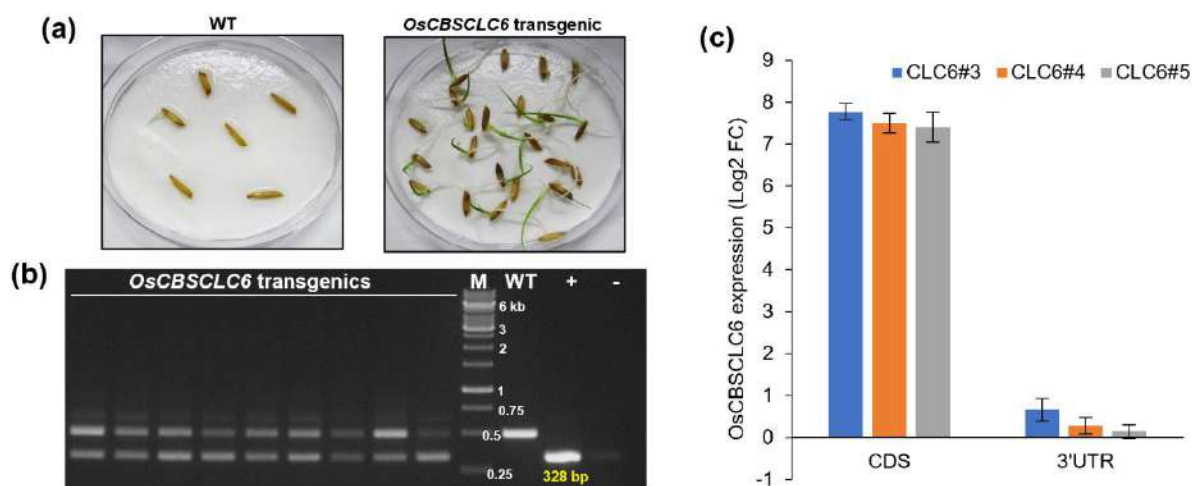


Fig. 4.39: Screening of *OsCBSCLC6* overexpressing T₁ plants and expression analysis of the transgene in T₂ plants. (a) Representative image showing the screening of T₁ generation seeds from *OsCBSCLC6* overexpressing transgenic lines on hygromycin solution (30 mg per L water). (b) Screening of the hygromycin positive seedlings by PCR using *OsCBSCLC6* specific internal primers. ‘M’: DNA ladder; ‘WT’: Wild-type; ‘+’: positive control; ‘-’: no template control. (c) Bar graph representing the qRT-PCR based expression of *OsCBSCLC6* in the youngest leaf at tillering stage in *OsCBSCLC6* overexpressing T₂ homozygous transgenic lines in relation to the transcript level of endogenous *OsCBSCLC6* in WT plants. The expression of the transgene was determined by using the primer sets corresponding to the exonic region that amplifies *OsCBSCLC6* transcripts from both transgene and endogenous gene. Whereas the expression of only endogenous *OsCBSCLC6* was determined using the primer sets corresponding to the 3’UTR region that do not amplify the transcripts from transgene. The expression level is expressed as Log₂ of 2^{-ΔΔCT} or Log₂ of Fold Change (FC). Bar represents mean ±SD, n=3.

In relation to our gene expression study that identified the upregulation of *OsCBSCLC6* under drought stress, we analyzed the growth performance of three independent homozygous T₂ transgenic lines constitutively overexpressing *OsCBSCLC6* under drought stress conditions. As the homologs of *OsCBSCLC6* in *Arabidopsis* are involved in Cl⁻ sequestration and tolerance function in plants (Nguyen et al, 2016; Hu et al, 2017), we also analyzed *OsCBSCLC6* transgenic plants under salinity (150 mM NaCl) treated conditions. The stress treatments were imposed at the pre-flowering stage (81 days after sowing) by withholding the irrigation for drought stress and by irrigating the pots with NaCl solution for salinity stress (Fig. 4.40) The drought stress was imposed for 2 weeks with the soil volumetric water content reaching to 20% on the final day of stress {measured using WaterScout SM100 Soil Moisture Sensor (Spectrum Technologies, Inc.)}. The salinity stress was imposed for 40 days with the soil electrical conductivity (EC) observed to be 14±1 dS/m on the final day of stress (soil solution temperature during EC measurement was 20°C).

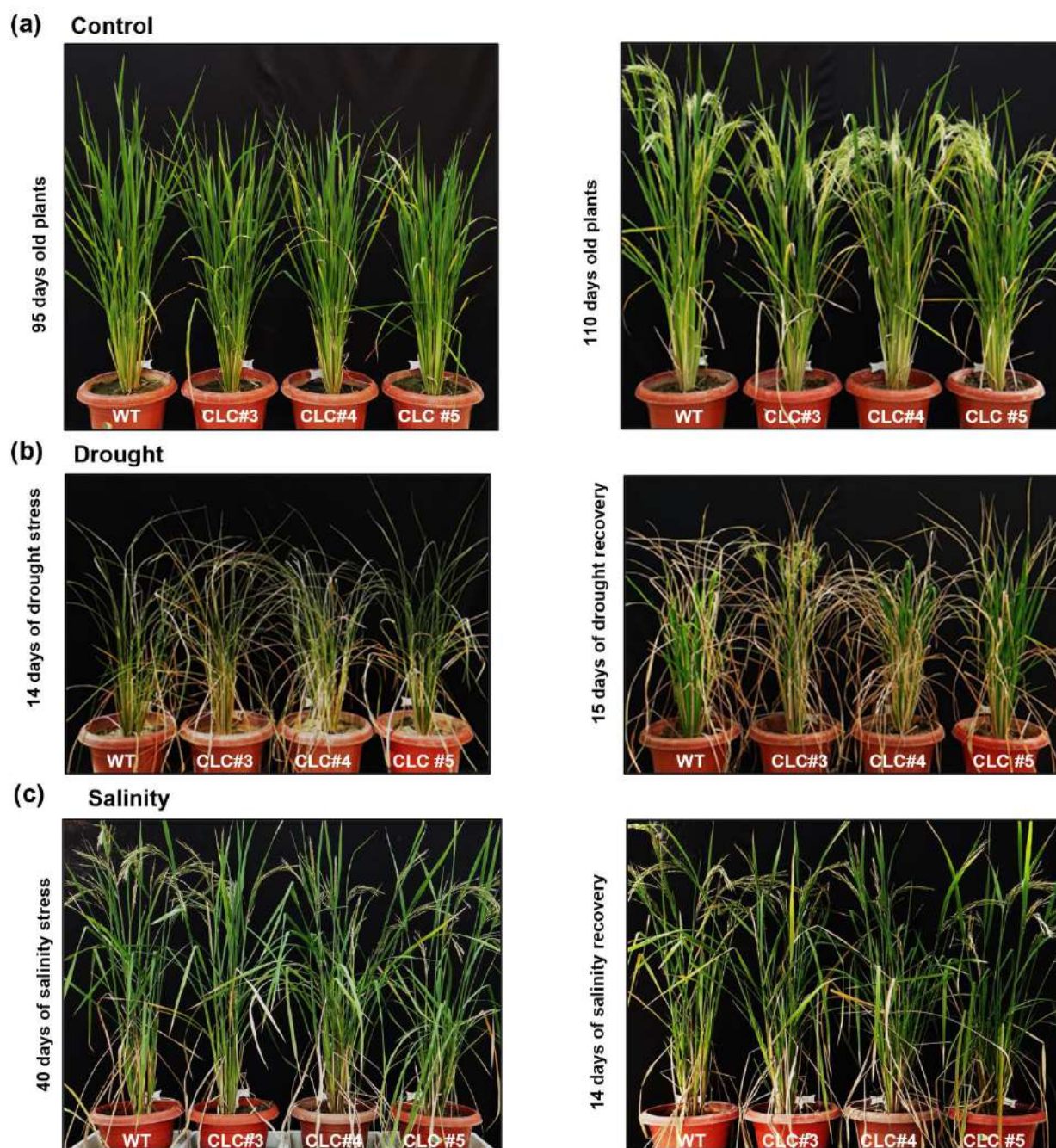


Fig. 4.40. Growth of *OsCBSCLC6* overexpressing T₂ transgenic rice plants under control, drought, and salinity stress conditions. (a) Transgenic and WT plants grown under controlled condition at 95 (left panel) and 110 (right panel) days after sowing (DAS). (b) Plants subjected to drought stress by withholding irrigation for 14 days (left panel; at 95 DAS) and 15 days after the stress recovery (right panel; at 110 DAS). (c) Plants subjected to salinity stress by keeping the pots in 150 mM NaCl solution for 40 days (left panel; at 121 DAS) and 14 days after the stress recovery (right panel; at 135 DAS). Plants were subjected to drought and salinity stresses at 81 DAS. 'WT': Wild-type plant; 'CLC#3', 'CLC#4' and 'CLC#5' are the three independent *OsCBSCLC6* overexpressing T₂ homozygous transgenic lines.

On the final day of stress treatment, the relative water content (RWC) and the electrolyte leakage in the penultimate leaf (leaf below the flag leaf) were observed, which depicts the plant water status and the cellular membrane integrity, respectively. Under drought stress, the *OsCBSCLC6* overexpressing transgenic lines were found to maintain higher water status than the WT plants, while no significant difference was observed in leaf electrolyte leakage between the transgenic and the WT plants (Fig. 4.41 a). In case of salt stress, *OsCBSCLC6* overexpressing transgenic lines exhibited significantly higher leaf RWC and lower electrolyte leakage in relation to the WT plants (Fig. 4.41 b). These two parameters indicated the better tolerance of stress by *OsCBSCLC6* overexpressing transgenic lines.

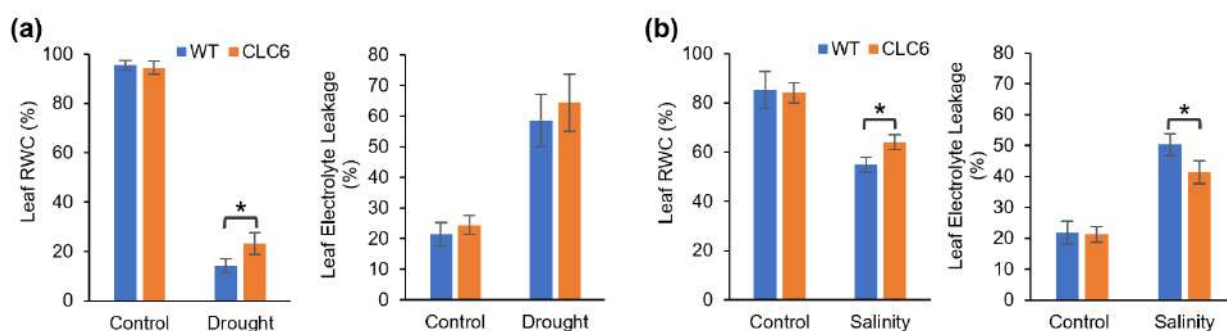


Fig 4.41. Leaf relative water content and electrolyte leakage in plants under controlled, salinity, and drought stressed conditions. (a) Bar graphs representing the leaf relative water content (in left panel) and leaf electrolyte leakage (right panel) in *OsCBSCLC6* overexpressing transgenics (CLC6) and the WT plants on 14th day of drought stress treatment. (b) Bar graphs representing the leaf relative water content (in left panel) and leaf electrolyte leakage (right panel) in *OsCBSCLC6* overexpressing transgenics (CLC6) and the WT plants on 40th day of salinity stress treatment. Bar represents mean \pm SD, n=3. ‘*’: $P \leq 0.05$ (calculated by unpaired, two-tailed Student’s *t*-test).

In addition to leaf RWC and the electrolyte leakage, we also observed healthier physiological performance of *OsCBSCLC6* overexpressing transgenic plants with respect to the chlorophyll-a fluorescence emission (using Handy PEA⁺, Hansatech Instruments) and the conduction of CO₂ and H₂O in the leaves {using Infra-red Gas Analyzer (IRGA), Licor-xT}. The *OsCBSCLC6* transgenic plants exhibited significantly higher Fv/Fm {maximum quantum efficiency of Photosystem II (PSII)} and performance index (conservation of energy from excitation of PSII to reduction of PSI) than the WT plants in salinity stress as well as during salinity and drought stress recovery phases (Fig. 4.42 a, b). Accordingly, the photosynthesis rate, the transpiration rate, and the stomatal conductance based on the exchange of CO₂ and H₂O in the leaves were also observed to be better in *OsCBSCLC6* overexpressing transgenic lines than the WT under both salinity and drought stress (Fig. 4.42 c-e). Moreover, the

OsCBSCLC6 transgenic plants exhibited a higher transpiration rate and stomatal conductance than the WT plants in controlled conditions as well towards the late matured stage.

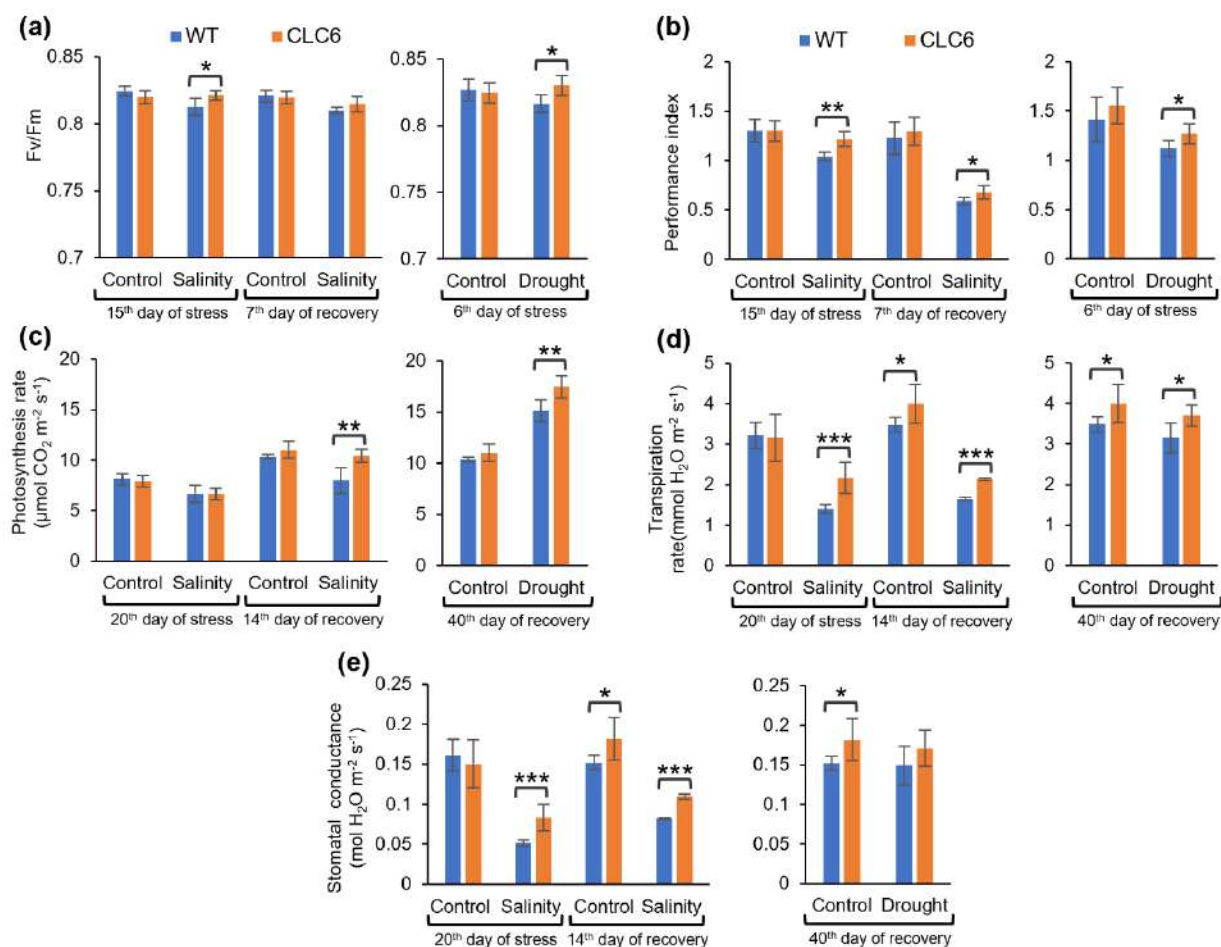


Fig. 4.42. Physiological performance of *OsCBSCLC6* overexpressing transgenic plants. Bar graphs in (a) and (b) represent two photosynthesis parameters: Fv/Fm and the Performance Index, respectively, determined based on chlorophyll-a fluorescence emission from a dark-adapted penultimate leaf using Handy PEA (Hansatech Instruments). The fluorescence parameters from salinity treated plants were measured on 15th day of stress treatment and on 7th day post stress recovery, while that from drought treated plants were measured on 6th day of stress treatment. The graphs in (c), (d), and (e) represent photosynthesis rate, transpiration rate, and the stomatal conductance, respectively, determined based on CO₂ and H₂O exchange in the penultimate leaf using IRGA (LI-6400XT, Licor). The gas exchange parameters from salinity treated plants were measured on 20th day of stress treatment and on 14th day post stress recovery, while that from drought plants were measured on 40th day post stress recovery. Bar represents mean ± SD, n=6. *: P ≤ 0.05; **: P ≤ 0.01; ***: P ≤ 0.001 (calculated by unpaired, two-tailed Student's *t*-test).

Phenotypically, the *OsCBSCLC6* overexpressing transgenic lines exhibited a significant reduction in plant height under controlled growth conditions, which was apparent during the reproductive stage (Fig. 4.40 a; Fig. 4.43 a). Although the difference in plant height between the transgenic and the WT plants was also observed in salinity stressed condition, no significant difference was existed in drought stressed plants. In the case of plant yield, consistent with better physiological performances exhibited by *OsCBSCLC6* overexpressing transgenic plants over the WT, these plants produced significantly higher number of filled grains and thus higher yield per plant under both salinity and drought stressed conditions (Fig. 4.43 b, c).

Therefore, based on the overall performances of the *OsCBSCLC6* overexpressing transgenic lines, we infer that *OsCBSCLC6* has an important role in maintaining cellular homeostasis for normal plant growth and development as well as in imparting stress tolerance in rice.

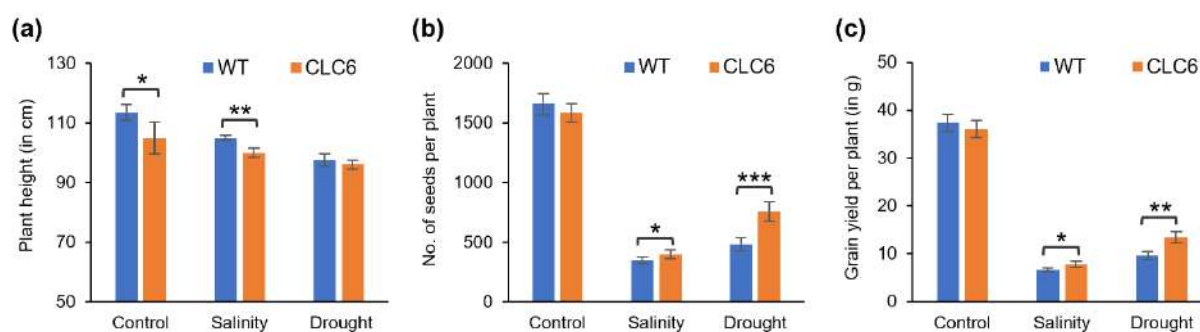


Fig. 4.43. Plant height and yield of *OsCBSCLC6* overexpressing T₂ transgenic rice plants. Bar graphs representing (a) the plant height measured from base of the shoot to tip of the main panicle (n=4), (b) number of filled grains per plant (n=3), and (c) weight of filled grains (in g) per plant (n=3 for WT, n=6 for CLC6), under control as well as salinity, and drought stressed conditions. ‘WT’: wild-type plants; ‘CLC6’: *OsCBSCLC6* overexpressing transgenic lines. Bar represents mean \pm SD. ‘*’: $P \leq 0.05$; ‘**’: $P \leq 0.01$; ‘***’: $P \leq 0.001$ (calculated by unpaired, two-tailed Student’s *t*-test).

5. Discussion

Since its discovery by Bateman (1997), CBS domains have been identified in various proteins in all living organisms, from prokaryotes to eukaryotes. In most of these CBS domains containing proteins (CDCPs) studied so far, this domain is reported to regulate the activity of associated hetero-domain(s) in a polypeptide or another subunit in a protein complex (Baykov et al, 2011; Ereño-Orbea et al, 2013a). The importance of this domain is evident by the fact that mutation in this domain causes several physiological disorders in mammals (Ignoul and Eggermont, 2005; Giménez-Mascarell et al, 2019). As of today, the information available on CBS domains or the CDCPs in plants is very thin. Our present study on two CDCPs from rice, *OsCBSX3* {a polypeptide containing only the basic unit of CDCPs, i.e., only a pair of CBS domains (CBSX)} and *OsCBSCLC6* (a member of CLC anion transport protein), has gained knowledge on their gene expression, cellular localization, interaction, and plant growth and survival under stress conditions when overexpressed in the plant.

We found *OsCBSX3* to be highly expressed in the leaves and relatively less in the roots at both vegetative and reproductive stages. In addition, we observed the expression of *OsCBSX3* to be diurnally regulated in the shoots with a higher accumulation of its transcripts during the day. In the case of stress responses, its expression has previously been reported to be upregulated under drought stress (Wang et al, 2004). We also observed its induction in response to desiccation in shoots at the seedling stage in rice. Besides, we observed its induction by salinity, heat, and cold stress treatments as well, especially in the shoots of Pokkali (salt-tolerant) than in IR-64 (salt-sensitive). Our qRT-PCR-based expression profiling of *OsCBSX3* correlated well with the available Affymetrix GeneChip and mRNA-seq datasets retrieved from Genevestigator (Hruz et al, 2008). Therefore, based on our expression analysis, we propose that *OsCBSX3* mainly functions in shoot tissues and possibly has an important link with stress tolerance.

Two previous studies have reported two different subcellular localization of *OsCBSX3*, in the plasma membrane by Mou et al (2015), and in the nucleus by Huang et al (2018), while both studies were carried out in tobacco leaf epidermal cells. In contrast, we observed localization of *OsCBSX3* in the chloroplast, analyzed in rice protoplast as well as in tobacco leaf epidermis. Moreover, our result is in accordance with the prediction of its localization by TargetP 1.1 server. Furthermore, our yeast-two hybrid (Y2H) analysis indicated the interaction of *OsCBSX3* with a member of thioredoxin (Trx) protein, *OsTrx y*, which is also predicted to

be localized in the chloroplast. Such interaction of chloroplast localized CBSX homologs in *Arabidopsis* (AtCBSX1 and AtCBSX2) with the chloroplast localized Trxs members (AtTrx f, AtTrx m, AtTrx x, and AtTrx y) have also been reported earlier by Yoo et al (2011). Interestingly, the Trx that we found to interact with OsCBSX3 was identified to be an ortholog of AtTrx y, one of the members interacting with AtCBSX1 and AtCBSX2, hence we termed it as OsTrx y. However, the interaction of OsCBSX3 was found to be specific only to OsTrx y, as other putative chloroplast OsTrxs, as well as the putative members of mitochondrial and cytosolic OsTrxs, did not show interaction with OsCBSX3. In addition, the *Arabidopsis* AtTrx y and AtTrx m that are reported to interact with AtCBSX1 and AtCBSX2 (Yoo et al, 2011), and AtTrx o, that interacts with AtCBSX3 (Shin et al, 2020), also did not show interaction with the rice OsCBSX3, suggesting clear functional speciation among these CBSX proteins. Such specificity in interaction has also been observed in *Arabidopsis* AtCBSXs, as AtCBSX1 and AtCBSX2 interacted specifically with chloroplastic AtTrxs, while mitochondrial AtCBSX3 interacted only with Trx o2 among 10 different mitochondrial AtTrxs assayed (Yoo et al, 2011; Shin et al, 2020). Nevertheless, the interaction of OsCBSX3 with OsTrx y implies its function in redox-dependent pathways, as reported for AtCBSX1, AtCBSX2, and AtCBSX3 (Yoo et al, 2011; Shin et al, 2020). However, diverse Trx members that these CBSXs homologs interact with, may account for their functional divergence or involvement in distinct downstream pathways.

Besides finding OsTrx y as an interacting partner, another interesting feature observed during the Y2H assay is that when the putative N-terminal signal peptide residues were deleted from OsCBSX3 {referred to as OsCBSX3(-sp)}, the protein exhibited strong autoactivation activity. This indicates the possibility of OsCBSX3 having a role in transcriptional activation. Accordingly, Huang et al (2018) have reported the interaction of OsCBSX3 with OsDRAP1, an AP2/EREBP transcription factor in the nucleus. They reported an upregulation of *OsDRAP1* transcripts in response to various stresses, as we also observed for *OsCBSX3*. Therefore, in relation to the findings of Huang et al (2018), we consider that under stress conditions OsCBSX3 may interact with OsDRAP1 and co-localize into the nucleus where OsCBSX3 facilitates OsDRAP1 to mediate transcription of the downstream defensive genes. However, under a normal state, we suggest the interaction of OsCBSX3 with OsTrx y in the chloroplast, contributing to the maintenance of the basic developmental processes in the cell.

As the two Bateman modules (each consisting of a pair of CBS domains) are known to self-associate to form a CBS module (comprises four CBS domains), forming a homodimer in

the case of CDCP with a single pair of CBS domains (Baykov et al, 2011), the protein homo-oligomeric study based on Y2H assay suggests homodimerization of OsCBSX3. Notably, we observed the N-terminal amino acid residues preceding the first CBS domain (predicted as signal peptide by TargetP-1.1 server) to be essential for this homodimerization. Besides, these N-terminal residues were observed to be important for the interaction of OsCBSX3 with OsTrx y as well. The deletion of this region hindered the interaction between these two proteins. Whereas in the case of AtCBSX1 and AtCBSX2, the predicted N-terminal signal peptide residues are shown not to be essential for their homodimer structure formation (Yoo et al, 2011). However, the residues preceding CBS domains in SnRK $\beta\gamma$ -subunit have been identified as critical for the complex formation and activity of plant SnRK complex (Ramon et al., 2013). Therefore, we suggest the significance of N-terminal residues preceding the first CBS domain in OsCBSX3 in its homo-oligomerization (most preferably a homodimer) and the protein function.

We also attempted the study the properties of OsCBSX3 *in vitro* by expressing it in *E. coli*. However, despite repeated attempts and different approaches, we could not obtain soluble OsCBSX3. Protein aggregation or insolubility problems associated with CBS domains in CDCPs have been encountered by many labs as well (Jhee et al, 2000; Oyenarte et al, 2012; McCorvie et al, 2014) encouraging them to express these proteins with some modifications. Our endeavor to express modified OsCBSX3 with the deletion of predicted N-terminal signal peptide residues {referred to as OsCBSX3(-sp)} yielded some amount of protein in the soluble fraction, which we used further for oligomeric and ATP binding assays *in vitro*. In contrast to the Y2H result, the oligomeric analysis by gel permeation chromatography (GPC) using purified OsCBSX3(-sp) determined the protein in a monomeric state. However, we infer this *in vitro* result to be due to the deletion of the N-terminal region which we found to be critical for co-association of OsCBSX3 by Y2H assay. Likewise, *in vitro* ATP binding assay exhibited no binding of ATP to the purified OsCBSX3(-sp). Whereas the sequence analysis of OsCBSX3 CBS domains with the domain sequences from various CDCPs that are structurally characterized for ligand binding identified the presence of conserved residues in OsCBSX3 that would facilitate the binding of one nucleotide per its Bateman module. We presume the possible effects of N-terminal peptide deletion as well as the insolubility associated non-native folding of OsCBSX3(-sp) in our *in vitro* results, while we propose that OsCBSX3 is a homo-oligomeric protein and has a potential nucleotide-binding site.

To determine the physiological and growth functions of *OsCBSX3* *in planta*, we generated its overexpressing transgenic plants. The overexpression of two *Arabidopsis* CSBX homologs, *AtCBSX1* and *AtCBSX3* resulted in severe sterility and retarded plant growth, respectively in transgenic *Arabidopsis* (Yoo et al, 2011; Shin et al, 2020). In contrast, we observed normal fertilization, silique formation, and overall growth and development of transgenic *Arabidopsis* overexpressing *OsCBSX3*. Besides, the rice transgenic lines overexpressing *OsCBSX3* also exhibited normal growth and development. No growth difference in *OsCBSX3* overexpressing transgenic rice is reported by Mou et al (2015) as well. The overexpression of another member of *CBSX* from rice, *OsCBSX4*, in transgenic tobacco also exhibited no growth defect associated with transgene expression (Singh et al, 2012). Therefore, as suggested by Yoo et al (2011), plant CBSXs are likely to be involved in the Trx-mediated redox system, but each CBSX may have its distinct downstream functions.

As we observed that the expression of *OsCBSX3* is highly inducible to stresses, particularly to drought and salinity, we analyzed the performance of transgenic *Arabidopsis* and rice lines overexpressing *OsCBSX3* under these two stresses. The overexpression of two stress inducible CDCPs from rice, viz., *OsCBSX4* (*CBSX* member) and *OsCBSCBSPB4* {encodes a protein containing two pairs of CBS domains followed by a Phox/Bemp1 (PB1) domain}, have been previously reported to confer enhanced tolerance to transgenic tobacco plants under different abiotic stresses (Singh et al, 2012; Kumar et al, 2018). Our study also observed better tolerance and yield in both *Arabidopsis* and rice transgenic lines overexpressing *OsCBSX3* over WT under salinity and drought stress conditions. Apart from salinity and drought, the *OsCBSX3* overexpressing *Arabidopsis* lines also exhibited superior growth than the WT plants under oxidative stress conditions at the seedling level. These indicate the involvement of *OsCBSX3* in stress tolerance function.

In the rice genome, *OsCBSX3* completely overlaps with the 5' region of *OsCYP450* on the opposite strand in an antiparallel arrangement. In such gene arrangement, the expression of one gene is known to be regulated by the other through simple transcriptional interference to chromatin modulation or the formation of double-stranded RNA (dsRNA). The latter would result in RNA masking, RNA interference, or RNA editing (Latgé et al, 2018). The expression analysis carried out in the present study showed a similar expression pattern for *OsCBSX3* and *OsCYP450*, both being predominantly higher in shoots than in roots at different developmental stages. However, the transcripts of *OsCBSX3* in shoots were observed to be much higher than that of *OsCYP450*. Under stress conditions, the two genes showed reciprocal expression

patterns with higher induction of *OsCBSX3* correlating with the downregulation of *OsCYP450*. In *Arabidopsis*, such inverse gene expression regulation has been reported by Borsani et al (2005) in antisense overlapping gene pair of *SRO5* and *P5CDH* which upregulated and downregulated, respectively under salinity stress. They identified the downregulation of *P5CDH* to be resulting from its cleavage by siRNA produced from the processing of *SRO5-P5CDH* dsRNA transcripts, the regulation found to be important for salinity stress tolerance. Similar inverse transcript correlation has also been observed in another antisense overlapping gene pair, *KPL* and *ARI14*, from *Arabidopsis*, where *KPL-ARI14*-generated-siRNA resulted in depletion of *ARI14*, a critical factor for normal fertilization (Ron et al, 2010). In contrast, a positive expression regulation has also been reported for the overlapping gene pair, viz. *NFYA5* and *NERF*, with an increase in expression of one resulting in the increased expression of other. The siRNAs produced by them are found to repress the miR169 instead which targets and cleaves *NFYA5* (Gao et al, 2015). The expression pattern of *OsCBSX3* and *OsCYP450* suggests a possible interaction between their transcripts during stress conditions as observed by Borsani et al (2005) for *SRO5* and *P5CDH* in *Arabidopsis*. Surprisingly, in transgenic rice constitutively overexpressing *OsCBSX3*, no alteration in *OsCYP450* transcripts was observed in comparison to that in WT plants. Thus, we propose that if *OsCBSX3* transcripts regulate the transcripts of *OsCYP450*, the exonic region at least does not have any role in this regulation while the involvement of the intronic or UTR region in this regulation cannot be ignored.

Another important member of CDCPs characterized in this study (*OsCBSCLC6* from rice) belongs to the chloride channel (CLC) family that contains a CLC transmembrane domain followed by two CBS domains at the cytosolic C-terminal end. The CDCPs are present ubiquitously from prokaryotic to eukaryotic organisms. Notably, most of the prokaryotic CLCs lack cytosolic CBS domains. Distinct from the canonical Cl^- transport function, some plant CLC members are known to be involved in NO_3^- transport instead. Sequence analysis of *OsCBSCLC6* for the conserved residues in the selectivity filter (Bergsdorf et al, 2009; Zifarelli and Pusch, 2010), indicated its function in Cl^- transport and as a channel while not in NO_3^- transport and as an exchanger. Accordingly, the phylogenetic analysis also placed it closer to *AtCLC-g* from *Arabidopsis* which is suggested to function as Cl^- transporting channel protein (Nguyen et al, 2016). Based on the qRT-PCR analysis as well as the Affymetrix GeneChip and mRNA-seq expression datasets available from Genevestigator (Hruz et al, 2008), we observed the transcript of *OsCBSCLC6* to be predominantly higher in root tissues at both vegetative and

reproductive stages. Among the various stress treatments, *OsCBSCLC6* was found to be upregulated mainly by desiccation in the shoots of both Pokkali and IR64 seedlings.

In mammals, CLCs that function as channel proteins are found to be localized in the plasma membrane while others with exchanger functions are localized to the intracellular vesicles (Jentsch, 2015). The *Arabidopsis* genome possesses seven CLCs (*AtCLC-a* to *AtCLC-g*), and all of their encoded proteins are identified to be localized in the organellar membranes (Lv et al, 2009). In rice, two CLCs, viz. *OsCBSCLC1* and *OsCBSCLC2*, are shown to be localized in the tonoplast (Nakamura et al, 2006). Through the transient overexpression of *OsCBSCLC6*+GFP fusion protein in rice protoplasts, we found that *OsCBSCLC6* also localizes in the tonoplast.

In rice, the *Tos17* insertional mutant lines of two CLCs, *OsCBSCLC1* and *OsCBSCLC2*, have been reported to cause a reduction in plant growth and yield (Nakamura et al., 2006). Even in tobacco (*Nicotiana benthamiana*), the double knockout lines of *NbCLC-1a* and *NbCLC-1b* resulted in a reduced plant height and smaller flowers (Sun et al, 2018). We also observed a reduction in plant height, however in transgenic rice lines overexpressing *OsCBSCLC6*. This reduced height phenotype is observable at a late vegetative or reproductive stage. In *Arabidopsis*, Golgi localized *AtCLC-d* has been reported to maintain the acidification of the trans-Golgi network by transporting counter anion, while its mutant lines exhibited reduced root growth (von der Fecht-Bartenbach et al., 2007). Accordingly, we propose that *OsCBSCLC6* is critical for maintaining the cellular ionic homeostasis as its overexpression significantly affected plant growth. However, not affected by the reduction in plant height, the *OsCBSCLC6* overexpressing transgenic rice lines exhibited better physiological performance and higher grain yield than the WT plants under salinity and drought stress conditions. The overexpression of some members of CLCs from *Arabidopsis* and maize have also been reported to confer tolerance in plants to salinity and drought stresses (Wang et al, 2015; Nguyen et al, 2016; Hu et al, 2017).

For *in vitro* studies, we attempted different approaches to express the full length of *OsCBSCLC6* in *E. coli*; however, no expression was obtained in the bacterial system. Although the expression of its cytosolic fraction consisting of a pair of CBS domains was successful in *E. coli*, the expressed protein remained in the insoluble fraction and aggregated during its refolding by dialysis. Hence, we could not commence any *in vitro* study on *OsCBSCLC6*. Nevertheless, to analyze the homo-dimeric status of its cytosolic CBS domain pairs, we

performed a Y2H assay. However, we observed no interaction or association of its cytosolic fractions, implying no homo-oligomerization of its Bateman modules in contrast to the known homodimer formation by Bateman modules from various CLCs. However, as the affinity of a cytosolic region of human CLC-5 towards the plasma membrane has been reported (Mo et al, 2004), we presume such effect in our Y2H study of the cytosolic fragment of OsCBSCLC6.

6. Summary and Conclusions

Cystathionine β -Synthase (CBS) domain containing proteins (CDCPs) constitutes a large protein superfamily, having the CBS domain in common, that is prevalent across from prokaryotes to eukaryotes. Generally, the CBS domains are present in a pair or two pairs in these proteins. While some CDCPs are composed of only CBS domains, others harbor additional hetero-domain(s) as well. Nevertheless, in this divergent protein superfamily, the CBS domains have maintained their structural and functional conservation throughout the evolution. Structurally, the two CBS domains in a pair are known to tightly associate to form a Bateman module, and the two Bateman modules in turn have been identified to associate to form a CBS module. Functionally these domains are known to act as a regulatory unit modulating the activity of other functional domains in the same or different polypeptides through the binding of mainly the adenosine derived nucleotides. For e.g., the binding of S-adenosyl L-methionine (SAM / AdoMet) to a Bateman module results in allosteric activation of the catalytic pyridoxal-5'-phosphate (PLP) binding domain in Cystathionine β -Synthase enzyme, a protein from which the CBS domain acquired its name.

Notably, the significance of CDCPs, particularly of their CBS domains, has been realized since the identification of various physiological disorders linked to mutations in CBS domains in different systems. However, having such physiological importance, the studies on plant CDCPs are still very few. The genome of rice (*Oryza sativa* L. ssp. *japonica*) contains 37 genes encoding CDCPs, many of which have remained uncharacterized for their function. Therefore, considering these factors, the present study was formulated for the basic functional characterization of two important CDCP members from rice: (a) OsCBSX3 {a polypeptide containing only a pair of CBS domains (CBSX)} and (b) OsCBSCLC6 (a protein with CLC anion transport transmembrane domain and a pair of cytosolic CBS domains).

Starting with the OsCBSX3, we observed its gene to have higher expression by orders of magnitude in the shoots than in the roots of rice plants at both vegetative and reproductive stages. Its expression was found to be diurnally regulated in the shoots and showed induction under desiccation, salinity, heat, and cold stresses. Like *OsCBSX3*, the expression of *OsCYTOCHROME P450 (OsCYP450)*, which completely overlaps with *OsCBSX3* in the opposite strand, also exhibited higher expression in shoot than in the root tissues, however its transcript level was significantly lower than that of *OsCBSX3* in both the tissues. Importantly, the expression of these two genes was observed to be contrasting under some stress conditions;

the upregulation of *OsCBSX3* correlated with the downregulation of *OsCYP450*. However, the expression of *OsCYP450* remained unaffected in *OsCBSX3* overexpressing transgenic rice plants. Therefore, we suggest negative gene regulation interactions between these two genes, however not involving at least the exonic parts of *OsCBSX3*.

We identified OsCBSX3 to localize in the chloroplast in tobacco leaf epidermal cells as well as in the rice protoplast. The OsCBSX3 and its variant OsCBSX3(-sp) (deletion of the predicted signal peptide from N-terminus), expressed in *E. coli*, was observed to be highly insoluble. Although we were able to purify some amount of soluble OsCBSX3(-sp), the protein exhibited a high propensity towards aggregation. This purified protein behaved as a monomeric protein in gel permeation chromatography and exhibited no *in vitro* ATP binding in ATP binding assay. However, through yeast-two hybrid (Y2H) assay, we identified the oligomerization of OsCBSX3, and its sequence analysis indicated a potential binding of one adenosine nucleotide per each OsCBSX3 Bateman module. Through Y2H and BiFC assay, we also identified and confirmed its interaction specifically with OsTrx y (a putative thioredoxin y protein) among different Trx members screened. Besides, we found that the residues at the N-terminus of OsCBSX3 (a putative signal peptide residues) to be essential for OsCBSX3 homo-oligomerization as well as for its interaction with OsTrx y. Furthermore, the deletion of these N-terminal residues {OsCBSX3(-sp)} conferred auto-activation to the protein in the Y2H assay. Therefore, based on these findings, we suggest the existence of OsCBSX3 in the oligomeric state *in vivo* with a possible binding of one adenosine nucleotide per Bateman module. Our results also suggest the role of predicted N-terminal signal peptide residues in OsCBSX3 oligomerization and interaction with OsTrx y. And as proposed by Huang et al (2018), we hypothesize the possible function of OsCBSX3 as a transcription activator, however on the deletion of the N-terminal residues.

The physiological function of *OsCBSX3* in plants is also assessed by constitutively overexpressing it in transgenic plants. In contrast to sterility and growth-retarded phenotype reported in *Arabidopsis* overexpressing its CSBX homologs, *AtCBSX1* and *AtCBSX3*, respectively (Yoo et al, 2011; Shin et al, 2020), we observed normal growth of *Arabidopsis* transgenics overexpressing *OsCBSX3*. Likewise, no difference was observed in transgenic rice plants overexpressing *OsCBSX3*. Moreover, the overexpression of *OsCBSX3* resulted in better plant growth and yield under drought and salinity stresses in both *Arabidopsis* and rice, indicating their association with stress tolerance function.

In the case of OsCBSCLC6, a member of chloride channels (CLCs) in the CDCP superfamily, we observed its transcripts to be predominantly higher in the root tissues of a rice plant. In the shoot, we found its expression to be upregulated in response to desiccation at the seedling stage. When transiently overexpressed in rice protoplasts, OsCBSCLC6 exhibited localization in the tonoplast. The sequence-based analysis of the transmembrane domain of OsCBSCLC6 in relation to the sequences from well-characterized CLCs belonging to different organisms indicated OsCBSCLC6 to function in Cl^- transport as a channel, but not in NO_3^- transport and as an exchanger protein. Through the sequence analysis of its cytosolic CBS domains, we suggest the possible binding of one adenosine nucleotide per its Bateman module. The oligomeric analysis through Y2H assay indicated no association or oligomerization of their cytosolic CBS domains. However, as reported by Mo et al (2004), we consider the possible affinity of this cytosolic fragment toward the plasma membrane, thus interfering with the Y2H assay. In the *E. coli* system, the full-length OsCBSCLC6 failed to express, whereas the truncated cytosolic variants although expressed in *E. coli*, were retained in the insoluble fraction, and aggregated during refolding.

Growth assessment of the transgenic rice plant overexpressing *OsCBSCLC6* exhibited a reduction in plant height under control conditions. Such reduction in plant growth and yield has also been reported for some CLCs, however in their loss of function mutant lines (Nakamura et al., 2006; von der Fecht-Bartenbach et al., 2007; Sun et al, 2018). However, these transgenic plants exhibited healthier physiological performance and higher grain yield than the WT plants under drought and salinity stress conditions. Therefore, OsCBSCLC6 has a significant impact on plant growth and development as well as in stress tolerance.

Through all these findings, this study has revealed the basic and previously unknown functions of OsCBSX3 and OsCBSCLC6, the two distinct and important members of rice CDCPs. In addition, it has also opened up the many critical but interesting questions relating to: the mode of regulation between *OsCBSX3* and *OsCYP450*, the functional implications of OsCBSX3 and OsTrx y interactions and the physiological processes they are involved in, the probable function of OsCBSX3(-sp) as transcription activator, cellular ionic status in *OsCBSCLC6* overexpressing transgenics, and the physiological impact of OsCBSCLC6 at a cellular level associated with the reduction in plant height. The future work on these questions would deliver us complete knowledge on the functions of these two proteins in rice. Also, the knowledge acquired through this study would be useful in the functional characterization of many other plant CDCPs that have remained yet unexplored in plant sciences.

7. Bibliography

- Accardi A, Miller C. Secondary active transport mediated by a prokaryotic homologue of ClC Cl⁻ channels. *Nature*. 2004 Feb 26;427(6977):803-7. doi: 10.1038/nature02314. PMID: 14985752.
- Accardi A, Walden M, Nguitragool W, Jayaram H, Williams C, Miller C. Separate ion pathways in a Cl⁻/H⁺ exchanger. *J Gen Physiol*. 2005 Dec;126(6):563-70. doi: 10.1085/jgp.200509417. PMID: 16316975; PMCID: PMC2266597.
- Adkar-Purushothama CR, Iyer PS, Perreault JP. Potato spindle tuber viroid infection triggers degradation of chloride channel protein CLC-b-like and Ribosomal protein S3a-like mRNAs in tomato plants. *Sci Rep*. 2017 Aug 21;7(1):8341. doi: 10.1038/s41598-017-08823-z. PMID: 28827569; PMCID: PMC5566334.
- Aguado-Llera D, Oyenarte I, Martínez-Cruz LA, Neira JL. The CBS domain protein MJ0729 of *Methanocaldococcus jannaschii* binds DNA. *FEBS Lett*. 2010 Nov 5;584(21):4485-9. doi: 10.1016/j.febslet.2010.10.006. Epub 2010 Oct 8. PubMed PMID: 20934423.
- Akpinar BA, Aysar B, Lucas SJ, Budak H. Plant abiotic stress signaling. *Plant Signal Behav*. 2012 Nov;7(11):1450-5. doi: 10.4161/psb.21894. Epub 2012 Sep 18. Review. PubMed PMID: 22990453; PubMed Central PMCID: PMC3548870.
- Altamura C, Desaphy JF, Conte D, De Luca A, Imbrici P. Skeletal muscle ClC-1 chloride channels in health and diseases. *Pflugers Arch*. 2020 Jul;472(7):961-975. doi: 10.1007/s00424-020-02376-3. Epub 2020 May 2. PMID: 32361781.
- Amodeo GA, Rudolph MJ, Tong L. Crystal structure of the heterotrimer core of *Saccharomyces cerevisiae* AMPK homologue SNF1. *Nature*. 2007 Sep 27;449(7161):492-5. doi: 10.1038/nature06127. Epub 2007 Sep 12. PubMed PMID: 17851534.
- Anashkin VA, Baykov AA, Lahti R. Enzymes regulated via cystathionine β-synthase domains. *Biochemistry (Mosc)*. 2017 Oct;82(10):1079-1087. doi: 10.1134/S0006297917100017. Review. PubMed PMID: 29037129.
- Anashkin VA, Salminen A, Tuominen HK, Orlov VN, Lahti R, Baykov AA. Cystathionine β-Synthase (CBS) Domain-containing Pyrophosphatase as a Target for Diadenosine Polyphosphates in Bacteria. *J Biol Chem*. 2015 Nov 13;290(46):27594-603. doi: 10.1074/jbc.M115.680272. Epub 2015 Sep 23. PMID: 26400082; PMCID: PMC4646011.
- Armitano J, Redder P, Guimarães VA, Linder P. An essential factor for high Mg²⁺ tolerance of *Staphylococcus aureus*. *Front Microbiol*. 2016;7:1888. doi: 10.3389/fmicb.2016.01888. eCollection 2016. PubMed PMID: 27933050; PubMed Central PMCID: PMC5122736.
- Avila-Castañeda A, Gutiérrez-Granados N, Ruiz-Gayosso A, Sosa-Peinado A, Martínez-Barajas E, Coello P. Structural and functional basis for starch binding in the SnRK1 subunits AKINβ2 and AKINβγ. *Front Plant Sci*. 2014;5:199. doi: 10.3389/fpls.2014.00199. eCollection 2014. PubMed PMID: 24904601; PubMed Central PMCID: PMC4032982.
- Bak G, Lee EJ, Lee Y, Kato M, Segami S, Sze H, Maeshima M, Hwang JU, Lee Y. Rapid structural changes and acidification of guard cell vacuoles during stomatal closure require phosphatidylinositol 3,5-bisphosphate. *Plant Cell*. 2013 Jun;25(6):2202-16. doi: 10.1105/tpc.113.110411. Epub 2013 Jun 11. PMID: 23757398; PMCID: PMC3723621.

- Baker JL, Sudarsan N, Weinberg Z, Roth A, Stockbridge RB, Breaker RR. Widespread genetic switches and toxicity resistance proteins for fluoride. *Science*. 2012 Jan 13;335(6065):233-235. doi: 10.1126/science.1215063. Epub 2011 Dec 22. PMID: 22194412; PMCID: PMC4140402.
- Bateman A. The structure of a domain common to archaeobacteria and the homocystinuria disease protein. *Trends Biochem Sci*. 1997 Jan;22(1):12-3. doi: 10.1016/s0968-0004(96)30046-7. PubMed PMID: 9020585.
- Baykov AA, Tuominen HK, Lahti R. The CBS domain: a protein module with an emerging prominent role in regulation. *ACS Chem Biol*. 2011 Nov 18;6(11):1156-63. doi: 10.1021/cb200231c. Epub 2011 Oct 11. PMID: 21958115.
- Bergsdorf EY, Zdebik AA, Jentsch TJ. Residues important for nitrate/proton coupling in plant and mammalian CLC transporters. *J Biol Chem*. 2009 Apr 24;284(17):11184-93. doi: 10.1074/jbc.M901170200. Epub 2009 Mar 4. PMID: 19261613; PMCID: PMC2670123.
- Boi S, Solda G, Tenchini ML. Shedding light on the dark side of the genome: overlapping genes in higher eukaryotes. *Curr Genomics*. 2004;5(6):509-24. <https://doi.org/10.2174/1389202043349020>.
- Borsani G, Rugarli EI, Tagliatalata M, Wong C, Ballabio A. Characterization of a human and murine gene (CLCN3) sharing similarities to voltage-gated chloride channels and to a yeast integral membrane protein. *Genomics*. 1995 May 1;27(1):131-41. doi: 10.1006/geno.1995.1015. PMID: 7665160.
- Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell*. 2005 Dec 29;123(7):1279-91. doi: 10.1016/j.cell.2005.11.035. PubMed PMID: 16377568; PubMed Central PMCID: PMC3137516.
- Bose S, He H, Stauber T. Neurodegeneration upon dysfunction of endosomal/lysosomal CLC chloride transporters. *Front Cell Dev Biol*. 2021 Feb 23;9:639231. doi: 10.3389/fcell.2021.639231. PMID: 33708769; PMCID: PMC7940362.
- Brammer AE, Stockbridge RB, Miller C. F⁻/Cl⁻ selectivity in CLCF-type F⁻/H⁺ antiporters. *J Gen Physiol*. 2014 Aug;144(2):129-36. doi: 10.1085/jgp.201411225. PMID: 25070431; PMCID: PMC4113899.
- Brandt S, Jentsch TJ. ClC-6 and ClC-7 are two novel broadly expressed members of the CLC chloride channel family. *FEBS Lett*. 1995 Dec 11;377(1):15-20. doi: 10.1016/0014-5793(95)01298-2. PMID: 8543009.
- Buey RM, Fernández-Justel D, Marcos-Alcalde Í, Winter G, Gómez-Puertas P, de Pereda JM, Luis Revuelta J. A nucleotide-controlled conformational switch modulates the activity of eukaryotic IMP dehydrogenases. *Sci Rep*. 2017 Jun 1;7(1):2648. doi: 10.1038/s41598-017-02805-x. PubMed PMID: 28572600; PubMed Central PMCID: PMC5454003.
- Buey RM, Ledesma-Amaro R, Velázquez-Campoy A, Balsera M, Chagoyen M, de Pereda JM, Revuelta JL. Guanine nucleotide binding to the Bateman domain mediates the allosteric inhibition of eukaryotic IMP dehydrogenases. *Nat Commun*. 2015 Nov 12;6:8923. doi: 10.1038/ncomms9923. PubMed PMID: 26558346; PubMed Central PMCID: PMC4660370.
- Cañero DC, Roncero MIG. Influence of the chloride channel of *Fusarium oxysporum* on extracellular laccase activity and virulence on tomato plants. *Microbiology (Reading)*. 2008 May;154(Pt 5):1474-1481. doi: 10.1099/mic.0.2007/015388-0. PMID: 18451056.
- Cao Y, Schubert KR. Molecular cloning and characterization of a cDNA encoding soybean nodule IMP dehydrogenase. *Biochim Biophys Acta*. 2001 Sep 21;1520(3):242-6. doi: 10.1016/s0167-4781(01)00269-x. PubMed PMID: 11566360.

- Carpaneto A, Boccaccio A, Lagostena L, Di Zanni E, Scholz-Starke J. The signaling lipid phosphatidylinositol-3,5-bisphosphate targets plant CLC-a anion/H⁺ exchange activity. *EMBO Rep.* 2017 Jul;18(7):1100-1107. doi: 10.15252/embr.201643814. Epub 2017 May 23. PMID: 28536248; PMCID: PMC5494527.
- Chen YS, Kozlov G, Fakhri R, Yang M, Zhang Z, Kovrigin EL, Gehring K. Mg²⁺-ATP sensing in CNNM, a putative magnesium transporter. *Structure.* 2020 Mar 3;28(3):324-335.e4. doi: 10.1016/j.str.2019.11.016. Epub 2019 Dec 18. PubMed PMID: 31864811.
- Choudhury NR, Malik PS, Singh DK, Islam MN, Kaliappan K, Mukherjee SK. The oligomeric Rep protein of Mungbean yellow mosaic India virus (MYMIV) is a likely replicative helicase. *Nucleic Acids Res.* 2006;34(21):6362-77. doi: 10.1093/nar/gkl903. Epub 2006 Nov 16. PMID: 17142233; PMCID: PMC1669733.
- Citovsky V, Lee LY, Vyas S, Glick E, Chen MH, Vainstein A, Gafni Y, Gelvin SB, Tzfira T. Subcellular localization of interacting proteins by bimolecular fluorescence complementation *in planta*. *J Mol Biol.* 2006 Oct 6;362(5):1120-31. doi: 10.1016/j.jmb.2006.08.017. Epub 2006 Aug 11. PMID: 16949607.
- Collart FR, Osipiuk J, Trent J, Olsen GJ, Huberman E. Cloning and characterization of the gene encoding IMP dehydrogenase from *Arabidopsis thaliana*. *Gene.* 1996 Oct 3;174(2):217-20. doi: 10.1016/0378-1119(96)00045-5. PubMed PMID: 8890737.
- Cookson SJ, Williams LE, Miller AJ. Light-dark changes in cytosolic nitrate pools depend on nitrate reductase activity in *Arabidopsis* leaf cells. *Plant Physiol.* 2005 Jun;138(2):1097-105. doi: 10.1104/pp.105.062349. Epub 2005 May 20. PMID: 15908593; PMCID: PMC1150423.
- Cornuel JF, Moraillon A, Guéron M. Participation of yeast inosine 5'-monophosphate dehydrogenase in an *in vitro* complex with a fragment of the C-rich telomeric strand. *Biochimie.* 2002 Apr;84(4):279-89. doi: 10.1016/s0300-9084(02)01400-1. PubMed PMID: 12106905.
- Corral-Rodríguez MÁ, Stuver M, Abascal-Palacios G, Diercks T, Oyenarte I, Ereño-Orbea J, de Opakua AI, Blanco FJ, Encinar JA, Spiwok V, Terashima H, Accardi A, Müller D, Martínez-Cruz LA. Nucleotide binding triggers a conformational change of the CBS module of the magnesium transporter CNNM2 from a twisted towards a flat structure. *Biochem J.* 2014 Nov 15;464(1):23-34. doi: 10.1042/BJ20140409. PubMed PMID: 25184538; PubMed Central PMCID: PMC7318797.
- Crozet P, Margalha L, Confraria A, Rodrigues A, Martinho C, Adamo M, Elias CA, Baena-González E. Mechanisms of regulation of SNF1/AMPK/SnRK1 protein kinases. *Front Plant Sci.* 2014;5:190. doi: 10.3389/fpls.2014.00190. eCollection 2014. Review. PubMed PMID: 24904600; PubMed Central PMCID: PMC4033248.
- Dadinova LA, Soshinskaia EY, Jeffries CM, Svergun DI, Shtykova EV. Tetrameric structures of inorganic CBS-pyrophosphatases from various bacterial species revealed by small-angle x-ray scattering in solution. *Biomolecules.* 2020 Apr 7;10(4). doi: 10.3390/biom10040564. PubMed PMID: 32272694; PubMed Central PMCID: PMC7226116.
- Davies WP. An historical perspective from the Green Revolution to the gene revolution. *Nutr Rev.* 2003 Jun;61(6 Pt 2):S124-34. doi: 10.1301/nr.2003.jun.S124-S134. Review. PubMed PMID: 12908744.
- Davis-Kaplan SR, Askwith CC, Bengtzen AC, Radisky D, Kaplan J. Chloride is an allosteric effector of copper assembly for the yeast multicopper oxidase Fet3p: an unexpected role for intracellular chloride channels. *Proc Natl Acad Sci U S A.* 1998 Nov 10;95(23):13641-5. doi: 10.1073/pnas.95.23.13641. PMID: 9811853; PMCID: PMC24872.

- De Angeli A, Monachello D, Ephritikhine G, Frachisse JM, Thomine S, Gambale F, Barbier-Brygoo H. The nitrate/proton antiporter AtCLCa mediates nitrate accumulation in plant vacuoles. *Nature*. 2006 Aug 24;442(7105):939-42. doi: 10.1038/nature05013. Epub 2006 Jul 26. PMID: 16878138.
- De Angeli A, Moran O, Wege S, Filleur S, Ephritikhine G, Thomine S, Barbier-Brygoo H, Gambale F. ATP binding to the C terminus of the *Arabidopsis thaliana* nitrate/proton antiporter, AtCLCa, regulates nitrate transport into plant vacuoles. *J Biol Chem*. 2009 Sep 25;284(39):26526-32. doi: 10.1074/jbc.M109.005132. Epub 2009 Jul 27. PMID: 19636075; PMCID: PMC2785341.
- de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res*. 2006 Jul 1;34(Web Server issue): W362-5. doi: 10.1093/nar/gkl124. PubMed PMID: 16845026; PubMed Central PMCID: PMC1538847.
- Diédhiou CJ, Gollmack D. Salt-dependent regulation of chloride channel transcripts in rice. *Plant Sci*. 2006 Apr;170(4):793-800. doi: 10.1016/j.plantsci.2005.11.014.
- Doyle J. DNA Protocols for Plants. In: Hewitt G.M., Johnston A.W.B., Young J.P.W. (eds) *Molecular Techniques in Taxonomy*. NATO ASI Series (Series H: Cell Biology). 1991:57. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-83962-7_18.
- Dutzler R, Campbell EB, Cadene M, Chait BT, MacKinnon R. X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature*. 2002 Jan 17;415(6869):287-94. doi: 10.1038/415287a. PMID: 11796999.
- Dutzler R, Campbell EB, MacKinnon R. Gating the selectivity filter in ClC chloride channels. *Science*. 2003 Apr 4;300(5616):108-12. doi: 10.1126/science.1082708. Epub 2003 Mar 20. PMID: 12649487.
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*. 2000 Jul 21;300(4):1005-16. doi: 10.1006/jmbi.2000.3903. PubMed PMID: 10891285.
- Ereño-Orbea J, Majtan T, Oyenarte I, Kraus JP, Martínez-Cruz LA. Structural basis of regulation and oligomerization of human cystathionine β-synthase, the central enzyme of transsulfuration. *Proc Natl Acad Sci U S A*. 2013b Oct 1;110(40):E3790-9. doi: 10.1073/pnas.1313683110. Epub 2013 Sep 16. PubMed PMID: 24043838; PubMed Central PMCID: PMC3791738.
- Ereño-Orbea J, Oyenarte I, Martínez-Cruz LA. CBS domains: Ligand binding sites and conformational variability. *Arch Biochem Biophys*. 2013a Dec;540(1-2):70-81. doi: 10.1016/j.abb.2013.10.008. Epub 2013 Oct 23. Review. PubMed PMID: 24161944.
- Estévez R, Boettger T, Stein V, Birkenhäger R, Otto E, Hildebrandt F, Jentsch TJ. Barttin is a Cl⁻ channel beta-subunit crucial for renal Cl⁻ reabsorption and inner ear K⁺ secretion. *Nature*. 2001 Nov 29;414(6863):558-61. doi: 10.1038/35107099. PMID: 11734858.
- Felix G, Duran JD, Volko S, Boller T. Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant J*. 1999 May;18(3):265-76. doi: 10.1046/j.1365-3113.1999.00265.x. PMID: 10377992.
- Feng L, Campbell EB, Hsiung Y, MacKinnon R. Structure of a eukaryotic CLC transporter defines an intermediate state in the transport cycle. *Science*. 2010 Oct 29;330(6004):635-41. doi: 10.1126/science.1195230. Epub 2010 Sep 30. PMID: 20929736; PMCID: PMC3079386.

- Feng N, Qi C, Hou YJ, Zhang Y, Wang DC, Li DF. The C2'- and C3'-endo equilibrium for AMP molecules bound in the cystathionine-beta-synthase domain. *Biochem Biophys Res Commun.* 2018 Mar 4;497(2):646-651. doi: 10.1016/j.bbrc.2018.02.124. Epub 2018 Feb 15. PubMed PMID: 29453981.
- Fernández-Justel D, Peláez R, Revuelta JL, Buey RM. The Bateman domain of IMP dehydrogenase is a binding target for dinucleoside polyphosphates. *J Biol Chem.* 2019 Oct 4;294(40):14768-14775. doi: 10.1074/jbc.AC119.010055. Epub 2019 Aug 15. PubMed PMID: 31416831; PubMed Central PMCID: PMC6779442.
- Foster JW. *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nat Rev Microbiol.* 2004 Nov;2(11):898-907. doi: 10.1038/nrmicro1021. PMID: 15494746.
- Gao W, Liu W, Zhao M, Li WX. *NERF* encodes a RING E3 ligase important for drought resistance and enhances the expression of its antisense gene *NFYA5* in Arabidopsis. *Nucleic Acids Res.* 2015 Jan;43(1):607-17. doi: 10.1093/nar/gku1325. Epub 2014 Dec 16. PubMed PMID: 25514924; PubMed Central PMCID: PMC4288204.
- Gao XQ, Liu CZ, Li DD, Zhao TT, Li F, Jia XN, Zhao XY, Zhang XS. The Arabidopsis KIN β subunit of the SnRK1 complex regulates pollen hydration on the stigma by mediating the level of reactive oxygen species in pollen. *PLoS Genet.* 2016 Jul;12(7): e1006228. doi: 10.1371/journal.pgen.1006228. eCollection 2016 Jul. PubMed PMID: 27472382; PubMed Central PMCID: PMC4966946.
- Gaxiola RA, Yuan DS, Klausner RD, Fink GR. The yeast CLC chloride channel functions in cation homeostasis. *Proc Natl Acad Sci U S A.* 1998 Mar 31;95(7):4046-50. doi: 10.1073/pnas.95.7.4046. PMID: 9520490; PMCID: PMC19960.
- Geelen D, Lurin C, Bouchez D, Frachisse JM, Lelièvre F, Courtial B, Barbier-Brygoo H, Maurel C. Disruption of putative anion channel gene AtCLC-a in Arabidopsis suggests a role in the regulation of nitrate content. *Plant J.* 2000 Feb;21(3):259-67. doi: 10.1046/j.1365-313x.2000.00680. x. PMID: 10758477.
- Geiger D, Scherzer S, Mumm P, Stange A, Marten I, Bauer H, Ache P, Matschi S, Liese A, Al-Rasheid KA, Romeis T, Hedrich R. Activity of guard cell anion channel SLAC1 is controlled by drought-stress signaling kinase-phosphatase pair. *Proc Natl Acad Sci U S A.* 2009 Dec 15;106(50):21425-30. doi: 10.1073/pnas.0912021106. Epub 2009 Dec 2. PMID: 19955405; PMCID: PMC2795561.
- Giménez-Mascarell P, González-Recio I, Fernández-Rodríguez C, Oyenarte I, Müller D, Martínez-Chantar ML, Martínez-Cruz LA. Current structural knowledge on the CNNM family of magnesium transport mediators. *Int J Mol Sci.* 2019 Mar 6;20(5). doi: 10.3390/ijms20051135. Review. PubMed PMID: 30845649; PubMed Central PMCID: PMC6429129.
- Giménez-Mascarell P, Oyenarte I, Hardy S, Breiderhoff T, Stuijver M, Kostantin E, Diercks T, Pey AL, Ereño-Orbea J, Martínez-Chantar ML, Khalaf-Nazzal R, Claverie-Martin F, Müller D, Tremblay ML, Martínez-Cruz LA. Structural Basis of the Oncogenic Interaction of Phosphatase PRL-1 with the Magnesium Transporter CNNM2. *J Biol Chem.* 2017 Jan 20;292(3):786-801. doi: 10.1074/jbc.M116.759944. Epub 2016 Nov 29. PMID: 27899452; PMCID: PMC5247653.
- Gojon A, Krouk G, Perrine-Walker F, Laugier E. Nitrate transceptor(s) in plants. *J Exp Bot.* 2011 Apr;62(7):2299-308. doi: 10.1093/jxb/erq419. Epub 2011 Jan 14. PMID: 21239382.
- Gómez-García I, Oyenarte I, Martínez-Cruz LA. The crystal structure of protein MJ1225 from *Methanocaldococcus jannaschii* shows strong conservation of key structural features seen in the eukaryal gamma-AMPK. *J Mol Biol.* 2010 May 28;399(1):53-70. doi: 10.1016/j.jmb.2010.03.045. Epub 2010 Apr 9. PubMed PMID: 20382158.

- Graves AR, Curran PK, Smith CL, Mindell JA. The Cl⁻/H⁺ antiporter CIC-7 is the primary chloride permeation pathway in lysosomes. *Nature*. 2008 Jun 5;453(7196):788-92. doi: 10.1038/nature06907. Epub 2008 Apr 30. PMID: 18449189.
- Grieschat M, Guzman RE, Langschwager K, Fahlke C, Alekov AK. Metabolic energy sensing by mammalian CLC anion/proton exchangers. *EMBO Rep*. 2020 Jun 4;21(6): e47872. doi: 10.15252/embr.201947872. Epub 2020 May 10. PMID: 32390228; PMCID: PMC7271328.
- Gulerez I, Funato Y, Wu H, Yang M, Kozlov G, Miki H, Gehring K. Phosphocysteine in the PRL-CNNM pathway mediates magnesium homeostasis. *EMBO Rep*. 2016 Dec;17(12):1890-1900. doi: 10.15252/embr.201643393. Epub 2016 Nov 17. PubMed PMID: 27856537; PubMed Central PMCID: PMC5283600.
- Guo W, Zuo Z, Cheng X, Sun J, Li H, Li L, Qiu JL. The chloride channel family gene CLCd negatively regulates pathogen-associated molecular pattern (PAMP)-triggered immunity in Arabidopsis. *J Exp Bot*. 2014 Mar;65(4):1205-15. doi: 10.1093/jxb/ert484. Epub 2014 Jan 21. PMID: 24449384; PMCID: PMC3935575.
- Harada H, Kuromori T, Hirayama T, Shinozaki K, Leigh RA. Quantitative trait loci analysis of nitrate storage in Arabidopsis leading to an investigation of the contribution of the anion channel gene, *AtCLC-c*, to variation in nitrate levels. *J Exp Bot*. 2004 Sep;55(405):2005-14. doi: 10.1093/jxb/erh224. Epub 2004 Aug 13. PMID: 15310822.
- Hardy S, Uetani N, Wong N, Kostantin E, Labbé DP, Bégin LR, Mes-Masson A, Miranda-Saavedra D, Tremblay ML. The protein tyrosine phosphatase PRL-2 interacts with the magnesium transporter CNNM3 to promote oncogenesis. *Oncogene*. 2015 Feb 19;34(8):986-95. doi: 10.1038/onc.2014.33. Epub 2014 Mar 17. PubMed PMID: 24632616.
- Hattori M, Iwase N, Furuya N, Tanaka Y, Tsukazaki T, Ishitani R, Maguire ME, Ito K, Maturana A, Nureki O. Mg²⁺-dependent gating of bacterial MgtE channel underlies Mg²⁺ homeostasis. *EMBO J*. 2009 Nov 18;28(22):3602-12. doi: 10.1038/emboj.2009.288. Epub 2009 Oct 1. PubMed PMID: 19798051; PubMed Central PMCID: PMC2782099.
- Hattori M, Tanaka Y, Fukai S, Ishitani R, Nureki O. Crystal structure of the MgtE Mg²⁺ transporter. *Nature*. 2007 Aug 30;448(7157):1072-5. doi: 10.1038/nature06093. Epub 2007 Aug 15. PubMed PMID: 17700703.
- Hechenberger M, Schwappach B, Fischer WN, Frommer WB, Jentsch TJ, Steinmeyer K. A family of putative chloride channels from Arabidopsis and functional complementation of a yeast strain with a CLC gene disruption. *J Biol Chem*. 1996 Dec 27;271(52):33632-8. doi: 10.1074/jbc.271.52.33632. PMID: 8969232.
- Herdean A, Nziengui H, Zsiros O, Solymosi K, Garab G, Lundin B, Spetea C. The Arabidopsis thylakoid chloride channel AtCLCe functions in chloride homeostasis and regulation of photosynthetic electron transport. *Front Plant Sci*. 2016 Feb 9; 7:115. doi: 10.3389/fpls.2016.00115. PMID: 26904077; PMCID: PMC4746265.
- Hickey LT, N Hafeez A, Robinson H, Jackson SA, Leal-Bertioli SCM, Tester M, Gao C, Godwin ID, Hayes BJ, Wulff BBH. Breeding crops to feed 10 billion. *Nat Biotechnol*. 2019 Jul;37(7):744-754. doi: 10.1038/s41587-019-0152-9. Epub 2019 Jun 17. Review. PubMed PMID: 31209375.
- Hirata Y, Funato Y, Takano Y, Miki H. Mg²⁺-dependent interactions of ATP with the cystathionine-β-synthase (CBS) domains of a magnesium transporter. *J Biol Chem*. 2014 May 23;289(21):14731-9. doi:

- 10.1074/jbc.M114.551176. Epub 2014 Apr 6. PubMed PMID: 24706765; PubMed Central PMCID: PMC4031528.
- Hirayama T, Shinozaki K. Research on plant abiotic stress responses in the post-genome era: past, present and future. *Plant J.* 2010 Mar;61(6):1041-52. doi: 10.1111/j.1365-313X.2010.04124.x. Review. PubMed PMID: 20409277.
- Hoegg-Beiler MB, Sirisi S, Orozco IJ, Ferrer I, Hohensee S, Auberson M, Gödde K, Vilches C, de Heredia ML, Nunes V, Estévez R, Jentsch TJ. Disrupting MLC1 and GlialCAM and CIC-2 interactions in leukodystrophy entails glial chloride channel dysfunction. *Nat Commun.* 2014 Mar 19; 5:3475. doi: 10.1038/ncomms4475. PMID: 24647135.
- Hruz T, Laule O, Szabo G, Wessendorp F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics.* 2008;2008:420747. doi: 10.1155/2008/420747. Epub 2008 Jul 8. PubMed PMID: 19956698; PubMed Central PMCID: PMC2777001.
- Hu R, Zhu Y, Wei J, Chen J, Shi H, Shen G, Zhang H. Overexpression of *PP2A-C5* that encodes the catalytic subunit 5 of protein phosphatase 2A in *Arabidopsis* confers better root and shoot development under salt conditions. *Plant Cell Environ.* 2017 Jan;40(1):150-164. doi: 10.1111/pce.12837. Epub 2016 Oct 26. PMID: 27676158.
- Huang L, Wang Y, Wang W, Zhao X, Qin Q, Sun F, Hu F, Zhao Y, Li Z, Fu B, Li Z. Characterization of transcription factor gene *OsDRAP1* conferring drought tolerance in rice. *Front Plant Sci.* 2018; 9:94. doi: 10.3389/fpls.2018.00094. eCollection 2018. PubMed PMID: 29449862; PubMed Central PMCID: PMC5799227.
- Huang Y, Jin F, Funato Y, Xu Z, Zhu W, Wang J, Sun M, Zhao Y, Yu Y, Miki H, Hattori M. Structural basis for the Mg²⁺ recognition and regulation of the CorC Mg²⁺ transporter. *Sci Adv.* 2021 Feb;7(7). doi: 10.1126/sciadv.abe6140. Print 2021 Feb. PubMed PMID: 33568487; PubMed Central PMCID: PMC7875539.
- Huynh TN, Choi PH, Sureka K, Ledvina HE, Campillo J, Tong L, Woodward JJ. Cyclic di-AMP targets the cystathionine beta-synthase domain of the osmolyte transporter OpuC. *Mol Microbiol.* 2016 Oct;102(2):233-243. doi: 10.1111/mmi.13456. Epub 2016 Jul 26. PubMed PMID: 27378384; PubMed Central PMCID: PMC5118871.
- Ignoul S, Eggermont J. CBS domains: structure, function, and pathology in human proteins. *Am J Physiol Cell Physiol.* 2005 Dec;289(6):C1369-78. doi: 10.1152/ajpcell.00282.2005. Review. PubMed PMID: 16275737.
- Imes D, Mumm P, Böhm J, Al-Rasheid KA, Marten I, Geiger D, Hedrich R. Open stomata 1 (OST1) kinase controls R-type anion channel QUAC1 in *Arabidopsis* guard cells. *Plant J.* 2013 May;74(3):372-82. doi: 10.1111/tj.12133. Epub 2013 Mar 22. PMID: 23452338.
- Jämsen J, Baykov AA, Lahti R. Fast kinetics of nucleotide binding to *Clostridium perfringens* family II pyrophosphatase containing CBS and DRTGG domains. *Biochemistry (Mosc).* 2012 Feb;77(2):165-70. doi: 10.1134/S0006297912020071. PMID: 22348476.
- Jämsen J, Tuominen H, Salminen A, Belogurov GA, Magretova NN, Baykov AA, Lahti R. A CBS domain-containing pyrophosphatase of *Moorella thermoacetica* is regulated by adenine nucleotides. *Biochem J.* 2007 Dec 15;408(3):327-33. doi: 10.1042/BJ20071017. PubMed PMID: 17714078; PubMed Central PMCID: PMC2267367.

- Jentsch TJ, Pusch M. CLC chloride channels and transporters: structure, function, physiology, and disease. *Physiol Rev.* 2018 Jul 1;98(3):1493-1590. doi: 10.1152/physrev.00047.2017. PMID: 29845874.
- Jentsch TJ, Steinmeyer K, Schwarz G. Primary structure of *Torpedo marmorata* chloride channel isolated by expression cloning in *Xenopus* oocytes. *Nature.* 1990 Dec 6;348(6301):510-4. doi: 10.1038/348510a0. PMID: 2174129.
- Jentsch TJ. Discovery of CLC transport proteins: cloning, structure, function and pathophysiology. *J Physiol.* 2015 Sep 15;593(18):4091-109. doi: 10.1113/JP270043. Epub 2015 Aug 24. Review. PubMed PMID: 25590607; PubMed Central PMCID: PMC4594286.
- Jeong BC, Park SH, Yoo KS, Shin JS, Song HK. Change in single cystathionine β -synthase domain-containing protein from a bent to flat conformation upon adenosine monophosphate binding. *J Struct Biol.* 2013a Jul;183(1):40-6. doi: 10.1016/j.jsb.2013.04.013. Epub 2013 May 9. PubMed PMID: 23664870.
- Jeong BC, Park SH, Yoo KS, Shin JS, Song HK. Crystal structure of the single cystathionine β -synthase domain-containing protein CBSX1 from *Arabidopsis thaliana*. *Biochem Biophys Res Commun.* 2013b Jan 4;430(1):265-71. doi: 10.1016/j.bbrc.2012.10.139. Epub 2012 Nov 14. PubMed PMID: 23159611.
- Jeworutzki E, López-Hernández T, Capdevila-Nortes X, Sirisi S, Bengtsson L, Montolio M, Zifarelli G, Arnedo T, Müller CS, Schulte U, Nunes V, Martínez A, Jentsch TJ, Gasull X, Pusch M, Estévez R. GlialCAM, a protein defective in a leukodystrophy, serves as a ClC-2 Cl⁻ channel auxiliary subunit. *Neuron.* 2012 Mar 8;73(5):951-61. doi: 10.1016/j.neuron.2011.12.039. PMID: 22405205; PMCID: PMC3334819.
- Jhee KH, McPhie P, Miles EW. Domain architecture of the heme-independent yeast cystathionine beta-synthase provides insights into mechanisms of catalysis and regulation. *Biochemistry.* 2000 Aug 29;39(34):10548-56. doi: 10.1021/bi001020g. PubMed PMID: 10956046.
- Jossier M, Kroniewicz L, Dalmas F, Le Thiec D, Ephritikhine G, Thomine S, Barbier-Brygoo H, Vavasseur A, Filleur S, Leonhardt N. The *Arabidopsis* vacuolar anion transporter, AtCLC_c, is involved in the regulation of stomatal movements and contributes to salt tolerance. *Plant J.* 2010 Nov;64(4):563-76. doi: 10.1111/j.1365-313X.2010.04352.x. Epub 2010 Oct 5. PMID: 20822503.
- Kargas, G.; Londra, P.; Sgoubopoulou, A. Comparison of soil EC values from methods based on 1:1 and 1:5 soil to water ratios and EC_e from saturated paste extract-based method. *Water.* 2020 Apr;12(4):1010. <https://doi.org/10.3390/w12041010>.
- Kawasaki M, Uchida S, Monkawa T, Miyawaki A, Mikoshiba K, Marumo F, Sasaki S. Cloning and expression of a protein kinase C-regulated chloride channel abundantly expressed in rat brain neuronal cells. *Neuron.* 1994 Mar;12(3):597-604. doi: 10.1016/0896-6273(94)90215-1. PMID: 8155321.
- Kery V, Poneleit L, Kraus JP. Trypsin cleavage of human cystathionine beta-synthase into an evolutionarily conserved active core: structural and functional consequences. *Arch Biochem Biophys.* 1998 Jul 15;355(2):222-32. doi: 10.1006/abbi.1998.0723. PubMed PMID: 9675031.
- King NP, Lee TM, Sawaya MR, Cascio D, Yeates TO. Structures and functional implications of an AMP-binding cystathionine beta-synthase domain protein from a hyperthermophilic archaeon. *J Mol Biol.* 2008 Jun 27;380(1):181-92. doi: 10.1016/j.jmb.2008.04.073. Epub 2008 May 7. PubMed PMID: 18513746; PubMed Central PMCID: PMC2577872.
- Klemens CA, Chulkov EG, Wu J, Hye Khan MA, Levchenko V, Flister MJ, Imig JD, Kriegel AJ, Palygin O, Staruschenko A. Loss of Chloride Channel 6 (CLC-6) affects vascular smooth muscle contractility and arterial stiffness via alterations to golgi calcium stores. *Hypertension.* 2021 Feb;77(2):582-593. doi:

- 10.1161/HYPERTENSIONAHA.120.16589. Epub 2021 Jan 4. PMID: 33390052; PMCID: PMC7856014.
- Kobayashi K, Uchida S, Mizutani S, Sasaki S, Marumo F. Intrarenal and cellular localization of CLC-K2 protein in the mouse kidney. *J Am Soc Nephrol.* 2001 Jul;12(7):1327-1334. doi: 10.1681/ASN.V1271327. PMID: 11423561.
- Koetsier G and Cantor E. A practical guide to analyzing nucleic acid concentration and purity with microvolume spectrophotometers. *NEB Technical Note.* 2019:1-8.
- Kostantin E, Hardy S, Valinsky WC, Kompatscher A, de Baaij JH, Zolotarov Y, Landry M, Uetani N, Martínez-Cruz LA, Hoenderop JG, Shrier A, Tremblay ML. Inhibition of PRL-2·CNNM3 protein complex formation decreases breast cancer proliferation and tumor growth. *J Biol Chem.* 2016 May 13;291(20):10716-25. doi: 10.1074/jbc.M115.705863. Epub 2016 Mar 11. PubMed PMID: 26969161; PubMed Central PMCID: PMC4865918.
- Koutmos M, Kabil O, Smith JL, Banerjee R. Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine {beta}-synthase. *Proc Natl Acad Sci U S A.* 2010 Dec 7;107(49):20958-63. doi: 10.1073/pnas.1011448107. Epub 2010 Nov 16. PubMed PMID: 21081698; PubMed Central PMCID: PMC3000283.
- Kozhevnikova EN, van der Knaap JA, Pindyurin AV, Ozgur Z, van Ijcken WF, Moshkin YM, Verrijzer CP. Metabolic enzyme IMPDH is also a transcription factor regulated by cellular state. *Mol Cell.* 2012 Jul 13;47(1):133-9. doi: 10.1016/j.molcel.2012.04.030. Epub 2012 May 31. PubMed PMID: 22658723.
- Kumar R, Mustafiz A, Sahoo KK, Sharma V, Samanta S, Sopory SK, Pareek A, Singla-Pareek SL. Functional screening of cDNA library from a salt tolerant rice genotype Pokkali identifies mannose-1-phosphate guanyl transferase gene (*OsMPGI*) as a key member of salinity stress response. *Plant Mol Biol.* 2012 Aug;79(6):555-68. doi: 10.1007/s11103-012-9928-8. Epub 2012 May 29. PubMed PMID: 22644442.
- Kumar R, Subba A, Kaur C, Ariyadasa TU, Sharan A, Pareek A, Sopory SK, Singla-Pareek SL. OsCBSCBSPB4 is a two cystathionine- β -synthase domain-containing protein from rice that functions in abiotic stress tolerance. *Curr Genomics.* 2018 Jan;19(1):50-59. doi: 10.2174/1389202918666170228141706. PubMed PMID: 29491732; PubMed Central PMCID: PMC5817877.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol.* 2018 Jun 1;35(6):1547-1549. doi: 10.1093/molbev/msy096. PMID: 29722887; PMCID: PMC5967553.
- Kumari S, Sabharwal VP, Kushwaha HR, Sopory SK, Singla-Pareek SL, Pareek A. Transcriptome map for seedling stage specific salinity stress response indicates a specific set of genes as candidate for saline tolerance in *Oryza sativa* L. *Funct Integr Genomics.* 2009 Feb;9(1):109-23. doi: 10.1007/s10142-008-0088-5. Epub 2008 Jul 2. PubMed PMID: 18594887.
- Kushwaha HR, Singh AK, Sopory SK, Singla-Pareek SL, Pareek A. Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* (L.) Heynh and *Oryza sativa* L. reveals their developmental and stress regulation. *BMC Genomics.* 2009 Apr 28; 10:200. doi: 10.1186/1471-2164-10-200. PMID: 19400948; PMCID: PMC2694836.
- Labesse G, Alexandre T, Vaupré L, Salard-Arnaud I, Him JL, Raynal B, Bron P, Munier-Lehmann H. MgATP regulates allostery and fiber formation in IMPDHs. *Structure.* 2013 Jun 4;21(6):975-85. doi: 10.1016/j.str.2013.03.011. Epub 2013 May 2. PubMed PMID: 23643948.

- Lange PF, Wartosch L, Jentsch TJ, Fuhrmann JC. CIC-7 requires Ostm1 as a beta-subunit to support bone resorption and lysosomal function. *Nature*. 2006 Mar 9;440(7081):220-3. doi: 10.1038/nature04535. PMID: 16525474.
- Last NB, Stockbridge RB, Wilson AE, Shane T, Kolmakova-Partensky L, Koide A, Koide S, Miller C. A CLC-type F⁻/H⁺ antiporter in ion-swapped conformations. *Nat Struct Mol Biol*. 2018 Jul;25(7):601-606. doi: 10.1038/s41594-018-0082-0. Epub 2018 Jun 25. PMID: 29941917; PMCID: PMC6044475.
- Latgé G, Poulet C, Bours V, Josse C, Jerusalem G. Natural antisense transcripts: molecular mechanisms and implications in breast cancers. *Int J Mol Sci*. 2018 Jan 2;19(1). doi: 10.3390/ijms19010123. Review. PubMed PMID: 29301303; PubMed Central PMCID: PMC5796072.
- Leisle L, Ludwig CF, Wagner FA, Jentsch TJ, Stauber T. CIC-7 is a slowly voltage-gated 2Cl⁻/1H⁺-exchanger and requires Ostm1 for transport activity. *EMBO J*. 2011 Jun 1;30(11):2140-52. doi: 10.1038/emboj.2011.137. Epub 2011 Apr 28. PMID: 21527911; PMCID: PMC3117652.
- Leisle L, Xu Y, Fortea E, Lee S, Galpin JD, Vien M, Ahern CA, Accardi A, Bernèche S. Divergent Cl⁻ and H⁺ pathways underlie transport coupling and gating in CLC exchangers and channels. *Elife*. 2020 Apr 28;9:e51224. doi: 10.7554/eLife.51224. PMID: 32343228; PMCID: PMC7274781.
- Li M, Sun X, Di D, Zhang A, Qing L, Zhou T, Miao H, Fan Z. Maize AKINβγ proteins interact with P8 of rice black streaked dwarf virus and inhibit viral infection. *Viruses*. 2020 Dec 4;12(12). doi: 10.3390/v12121387. PubMed PMID: 33291518; PubMed Central PMCID: PMC7761811.
- Li WY, Wong FL, Tsai SN, Phang TH, Shao G, Lam HM. Tonoplast-located GmCLC1 and GmNHX1 from soybean enhance NaCl tolerance in transgenic bright yellow (BY)-2 cells. *Plant Cell Environ*. 2006 Jun;29(6):1122-37. doi: 10.1111/j.1365-3040.2005.01487.x. PMID: 17080938.
- Liao Q, Zhou T, Yao JY, Han QF, Song HX, Guan CY, Hua YP, Zhang ZH. Genome-scale characterization of the vacuole nitrate transporter Chloride Channel (CLC) genes and their transcriptional responses to diverse nutrient stresses in allotetraploid rapeseed. *PLoS One*. 2018 Dec 20;13(12): e0208648. doi: 10.1371/journal.pone.0208648. PMID: 30571734; PMCID: PMC6301700.
- Lin SC, Hardie DG. AMPK: sensing glucose as well as cellular energy status. *Cell Metab*. 2018 Feb 6;27(2):299-313. doi: 10.1016/j.cmet.2017.10.009. Epub 2017 Nov 16. Review. PubMed PMID: 29153408.
- Lísal J, Maduke M. The CIC-0 chloride channel is a 'broken' Cl⁻/H⁺ antiporter. *Nat Struct Mol Biol*. 2008 Aug;15(8):805-10. doi: 10.1038/nsmb.1466. Epub 2008 Jul 20. PMID: 18641661; PMCID: PMC2559860.
- López-Hernández T, Ridder MC, Montolio M, Capdevila-Nortes X, Polder E, Sirisi S, Duarri A, Schulte U, Fakler B, Nunes V, Scheper GC, Martínez A, Estévez R, van der Knaap MS. Mutant GlialCAM causes megalencephalic leukoencephalopathy with subcortical cysts, benign familial macrocephaly, and macrocephaly with retardation and autism. *Am J Hum Genet*. 2011a Apr 8;88(4):422-32. doi: 10.1016/j.ajhg.2011.02.009. Epub 2011 Mar 17. PMID: 21419380; PMCID: PMC3071909.
- López-Hernández T, Sirisi S, Capdevila-Nortes X, Montolio M, Fernández-Dueñas V, Scheper GC, van der Knaap MS, Casquero P, Ciruela F, Ferrer I, Nunes V, Estévez R. Molecular mechanisms of MLC1 and GLIALCAM mutations in megalencephalic leukoencephalopathy with subcortical cysts. *Hum Mol Genet*. 2011b Aug 15;20(16):3266-77. doi: 10.1093/hmg/ddr238. Epub 2011 May 30. PMID: 21624973.
- Lucas M, Encinar JA, Arribas EA, Oyenarte I, García IG, Kortazar D, Fernández JA, Mato JM, Martínez-Chantar ML, Martínez-Cruz LA. Binding of S-methyl-5'-thioadenosine and S-adenosyl-L-methionine to

- protein MJ0100 triggers an open-to-closed conformational change in its CBS motif pair. *J Mol Biol.* 2010 Feb 26;396(3):800-20. doi: 10.1016/j.jmb.2009.12.012. Epub 2009 Dec 21. PubMed PMID: 20026078.
- Lumbreras V, Alba MM, Kleinow T, Koncz C, Pagès M. Domain fusion between SNF1-related kinase subunits during plant evolution. *EMBO Rep.* 2001 Jan;2(1):55-60. doi: 10.1093/embo-reports/kve001. PubMed PMID: 11252725; PubMed Central PMCID: PMC1083798.
- Lurin C, Geelen D, Barbier-Brygoo H, Guern J, Maurel C. Cloning and functional expression of a plant voltage-dependent chloride channel. *Plant Cell.* 1996 Apr;8(4):701-11. doi: 10.1105/tpc.8.4.701. PMID: 8624442; PMCID: PMC161130.
- Lutts S, Kinet JM, Bouharmont J. NaCl-induced senescence in leaves of rice (*Oryza sativa* L.) cultivars differing in salinity resistance. *Ann. Bot.* 1996 Sept;78(3):389-98. <https://doi.org/10.1006/anbo.1996.0134>.
- Lv Q dan, Tang R jie, Liu H, Gao X, Li Y, Zheng H, Zhang H. Cloning and molecular analyses of the *Arabidopsis thaliana* chloride channel gene family. *Plant Sci.* 2009 May;176(5):650-61. doi: 10.1016/j.plantsci.2009.02.006.
- Mackintosh RW, Davies SP, Clarke PR, Weekes J, Gillespie JG, Gibb BJ, Hardie DG. Evidence for a protein kinase cascade in higher plants. 3-Hydroxy-3-methylglutaryl-CoA reductase kinase. *Eur J Biochem.* 1992 Nov 1;209(3):923-31. doi: 10.1111/j.1432-1033.1992.tb17364.x. PubMed PMID: 1358611.
- Mahmood NA, Biemans-Oldehinkel E, Poolman B. Engineering of ion sensing by the cystathionine beta-synthase module of the ABC transporter OpuA. *J Biol Chem.* 2009 May 22;284(21):14368-76. doi: 10.1074/jbc.M901238200. Epub 2009 Mar 27. PubMed PMID: 19329426; PubMed Central PMCID: PMC2682885.
- Marín MS, Casais R, Alonso JM, Parra F. ATP binding and ATPase activities associated with recombinant rabbit hemorrhagic disease virus 2C-like polypeptide. *J Virol.* 2000 Nov;74(22):10846-51. doi: 10.1128/jvi.74.22.10846-10851.2000. PMID: 11044135; PMCID: PMC110965.
- Markovic S, Dutzler R. The structure of the cytoplasmic domain of the chloride channel ClC-Ka reveals a conserved interaction interface. *Structure.* 2007 Jun;15(6):715-25. doi: 10.1016/j.str.2007.04.013. PMID: 17562318.
- Marmagne A, Vinauger-Douard M, Monachello D, de Longevialle AF, Charon C, Allot M, Rappaport F, Wollman FA, Barbier-Brygoo H, Ephritikhine G. Two members of the Arabidopsis CLC (chloride channel) family, AtCLCe and AtCLCf, are associated with thylakoid and Golgi membranes, respectively. *J Exp Bot.* 2007;58(12):3385-93. doi: 10.1093/jxb/erm187. Epub 2007 Sep 14. PMID: 17872921.
- Martínez-Cruz LA, Encinar JA, Kortazar D, Prieto J, Gómez J, Fernández-Millán P, Lucas M, Arribas EA, Fernández JA, Martínez-Chantar ML, Mato JM, Neira JL. The CBS domain protein MJ0729 of *Methanocaldococcus jannaschii* is a thermostable protein with a pH-dependent self-oligomerization. *Biochemistry.* 2009 Mar 31;48(12):2760-76. doi: 10.1021/bi801920r. PubMed PMID: 19267448.
- Matsumura Y, Uchida S, Kondo Y, Miyazaki H, Ko SB, Hayama A, Morimoto T, Liu W, Arisawa M, Sasaki S, Marumo F. Overt nephrogenic diabetes insipidus in mice lacking the CLC-K1 chloride channel. *Nat Genet.* 1999 Jan;21(1):95-8. doi: 10.1038/5036. PMID: 9916798.
- McCorvie TJ, Kopec J, Hyung SJ, Fitzpatrick F, Feng X, Termine D, Strain-Damerell C, Vollmar M, Fleming J, Janz JM, Bulawa C, Yue WW. Inter-domain communication of human cystathionine β -synthase: structural basis of S-adenosyl-L-methionine activation. *J Biol Chem.* 2014 Dec 26;289(52):36018-30.

- doi: 10.1074/jbc.M114.610782. Epub 2014 Oct 21. PubMed PMID: 25336647; PubMed Central PMCID: PMC4276868.
- McLean JE, Hamaguchi N, Belenky P, Mortimer SE, Stanton M, Hedstrom L. Inosine 5'-monophosphate dehydrogenase binds nucleic acids *in vitro* and *in vivo*. *Biochem J*. 2004 Apr 15;379(Pt 2):243-51. doi: 10.1042/BJ20031585. PubMed PMID: 14766016; PubMed Central PMCID: PMC1224093.
- Meyer S, Dutzler R. Crystal structure of the cytoplasmic domain of the chloride channel ClC-0. *Structure*. 2006 Feb;14(2):299-307. doi: 10.1016/j.str.2005.10.008. PMID: 16472749.
- Meyer S, Savaresi S, Forster IC, Dutzler R. Nucleotide recognition by the cytoplasmic domain of the human chloride transporter ClC-5. *Nat Struct Mol Biol*. 2007 Jan;14(1):60-7. doi: 10.1038/nsmb1188. Epub 2006 Dec 31. Erratum in: *Nat Struct Mol Biol*. 2007 Feb;14(2):172. PMID: 17195847.
- Mishra M, Wungrampha S, Kumar G, Singla-Pareek SL, Pareek A. How do rice seedlings of landrace Pokkali survive in saline fields after transplantation? Physiology, biochemistry, and photosynthesis. *Photosynth Res*. 2020 Jul 6;. doi: 10.1007/s11120-020-00771-6. [Epub ahead of print] PubMed PMID: 32632535.
- Mo L, Xiong W, Qian T, Sun H, Wills NK. Coexpression of complementary fragments of ClC-5 and restoration of chloride channel function in a Dent's disease mutation. *Am J Physiol Cell Physiol*. 2004 Jan;286(1):C79-89. doi: 10.1152/ajpcell.00009.2003. Epub 2003 Sep 17. PubMed PMID: 13679301.
- Monachello D, Allot M, Oliva S, Krapp A, Daniel-Vedele F, Barbier-Brygoo H, Ephritikhine G. Two anion transporters AtClCa and AtClCe fulfil interconnecting but not redundant roles in nitrate assimilation pathways. *New Phytol*. 2009;183(1):88-94. doi: 10.1111/j.1469-8137.2009.02837.x. Epub 2009 Apr 23. PMID: 19402883.
- Moomaw AS, Maguire ME. The unique nature of Mg²⁺ channels. *Physiology (Bethesda)*. 2008 Oct;23:275-85. doi: 10.1152/physiol.00019.2008. Review. PubMed PMID: 18927203; PubMed Central PMCID: PMC2711038.
- Mortimer SE, Xu D, McGrew D, Hamaguchi N, Lim HC, Bowne SJ, Daiger SP, Hedstrom L. IMP dehydrogenase type 1 associates with polyribosomes translating rhodopsin mRNA. *J Biol Chem*. 2008 Dec 26;283(52):36354-60. doi: 10.1074/jbc.M806143200. Epub 2008 Oct 30. PubMed PMID: 18974094; PubMed Central PMCID: PMC2605994.
- Mou S, Shi L, Lin W, Liu Y, Shen L, Guan D, He S. Over-expression of rice CBS domain containing protein, OsCBSX3, confers rice resistance to *Magnaporthe oryzae* inoculation. *Int J Mol Sci*. 2015 Jul 13;16(7):15903-17. doi: 10.3390/ijms160715903. PubMed PMID: 26184180; PubMed Central PMCID: PMC4519930.
- Mutte SK, Weijers D. Deep evolutionary history of the Phox and Bem1 (PB1) domain across eukaryotes. *Sci Rep*. 2020 Mar 2;10(1):3797. doi: 10.1038/s41598-020-60733-9. PubMed PMID: 32123237; PubMed Central PMCID: PMC7051960.
- Nakamura A, Fukuda A, Sakai S, Tanaka Y. Molecular cloning, functional expression and subcellular localization of two putative vacuolar voltage-gated chloride channels in rice (*Oryza sativa* L.). *Plant Cell Physiol*. 2006 Jan;47(1):32-42. doi: 10.1093/pcp/pci220. Epub 2005 Oct 25. PMID: 16249326.
- Nakashima K, Takasaki H, Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. NAC transcription factors in plant abiotic stress responses. *Biochim Biophys Acta*. 2012 Feb;1819(2):97-103. doi: 10.1016/j.bbagr.2011.10.005. Epub 2011 Oct 19. Review. PubMed PMID: 22037288.

- Neaogoe I, Stauber T, Fidzinski P, Bergsdorf EY, Jentsch TJ. The late endosomal CIC-6 mediates proton/chloride countertransport in heterologous plasma membrane expression. *J Biol Chem*. 2010 Jul 9;285(28):21689-97. doi: 10.1074/jbc.M110.125971. Epub 2010 May 13. PMID: 20466723; PMCID: PMC2898453.
- Nedelyaeva OI, Shuvalov AV, Karpichev IV, Beliaev DV, Myasoedov NA, Khalilova LA, Khramov DE, Popova LG, Balnokin YV. Molecular cloning and characterisation of SaCLCa1, a novel protein of the chloride channel (CLC) family from the halophyte *Suaeda altissima* (L.) Pall. *J Plant Physiol*. 2019 Sep;240:152995. doi: 10.1016/j.jplph.2019.152995. Epub 2019 Jun 13. PMID: 31252320.
- Nedelyaeva OI, Shuvalov AV, Mayorova OV, Yurchenko AA, Popova LG, Balnokin YV, Karpichev IV. Cloning and functional analysis of *SaCLCc1*, a gene belonging to the chloride channel family (CLC), from the halophyte *Suaeda altissima* (L.) Pall. *Dokl Biochem Biophys*. 2018 Jul;481(1):186-189. doi: 10.1134/S1607672918040026. Epub 2018 Aug 31. PMID: 30168055.
- Nguyen CT, Agorio A, Jossier M, Depré S, Thomine S, Filleur S. Characterization of the chloride channel-like, AtCLCg, involved in chloride tolerance in *Arabidopsis thaliana*. *Plant Cell Physiol*. 2016 Apr;57(4):764-75. doi: 10.1093/pcp/pcv169. Epub 2015 Nov 9. PMID: 26556649.
- Niemeyer MI, Cid LP, Zúñiga L, Catalán M, Sepúlveda FV. A conserved pore-lining glutamate as a voltage- and chloride-dependent gate in the CIC-2 chloride channel. *J Physiol*. 2003 Dec 15;553(Pt 3):873-9. doi: 10.1113/jphysiol.2003.055988. Epub 2003 Nov 14. PMID: 14617675; PMCID: PMC2343633.
- Ok SH, Yoo KS, Shin JS. CBSXs are sensor relay proteins sensing adenosine-containing ligands in *Arabidopsis*. *Plant Signal Behav*. 2012 Jun;7(6):664-7. doi: 10.4161/psb.19945. Epub 2012 May 14. PubMed PMID: 22580706; PubMed Central PMCID: PMC3442862.
- Oyenarte I, Majtan T, Ereño J, Corral-Rodríguez MA, Kraus JP, Martínez-Cruz LA. Purification, crystallization and preliminary crystallographic analysis of human cystathionine β -synthase. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 2012 Nov 1;68(Pt 11):1318-22. doi: 10.1107/S1744309112037219. Epub 2012 Oct 30. PubMed PMID: 23143240; PubMed Central PMCID: PMC3515372.
- Padmanaban S, Lin X, Perera I, Kawamura Y, Sze H. Differential expression of vacuolar H⁺-ATPase subunit c genes in tissues active in membrane trafficking and their roles in plant growth as revealed by RNAi. *Plant Physiol*. 2004 Apr;134(4):1514-26. doi: 10.1104/pp.103.034025. Epub 2004 Mar 29. PMID: 15051861; PMCID: PMC419827.
- Pareek A, Dhankher OP, Foyer CH. Mitigating the impact of climate change on plant productivity and ecosystem sustainability. *J Exp Bot*. 2020 Jan 7;71(2):451-456. doi: 10.1093/jxb/erz518. PubMed PMID: 31909813; PubMed Central PMCID: PMC6945998.
- Park E, Campbell EB, MacKinnon R. Structure of a CLC chloride ion channel by cryo-electron microscopy. *Nature*. 2017 Jan 26;541(7638):500-505. doi: 10.1038/nature20812. Epub 2016 Dec 21. PMID: 28002411; PMCID: PMC5576512.
- Park E, MacKinnon R. Structure of the CLC-1 chloride channel from *Homo sapiens*. *Elife*. 2018 May 29;7:e36629. doi: 10.7554/eLife.36629. PMID: 29809153; PMCID: PMC6019066.
- Pedersen TH, Riisager A, de Paoli FV, Chen TY, Nielsen OB. Role of physiological CIC-1 Cl⁻ ion channel regulation for the excitability and function of working skeletal muscle. *J Gen Physiol*. 2016 Apr;147(4):291-308. doi: 10.1085/jgp.201611582. PMID: 27022190; PMCID: PMC4810071.

- Pey AL, Majtan T, Sanchez-Ruiz JM, Kraus JP. Human cystathionine β -synthase (CBS) contains two classes of binding sites for S-adenosylmethionine (SAM): complex regulation of CBS activity and stability by SAM. *Biochem J*. 2013 Jan 1;449(1):109-21. doi: 10.1042/BJ20120731. PubMed PMID: 22985361.
- Pey AL, Martínez-Cruz LA, Kraus JP, Majtan T. Oligomeric status of human cystathionine beta-synthase modulates AdoMet binding. *FEBS Lett*. 2016 Dec;590(24):4461-4471. doi: 10.1002/1873-3468.12488. Epub 2016 Nov 23. PubMed PMID: 27861796.
- Phillips S, Brammer AE, Rodriguez L, Lim HH, Stary-Weinzinger A, Matulef K. Surprises from an unusual CLC homolog. *Biophys J*. 2012 Nov 7;103(9):L44-6. doi: 10.1016/j.bpj.2012.08.063. PMID: 23199933; PMCID: PMC3491686.
- Piccolo A, Malvezzi M, Houtman JC, Accardi A. Basis of substrate binding and conservation of selectivity in the CLC family of channels and transporters. *Nat Struct Mol Biol*. 2009 Dec;16(12):1294-301. doi: 10.1038/nsmb.1704. Epub 2009 Nov 8. PMID: 19898476; PMCID: PMC2920496.
- Pimkin M, Markham GD. The CBS subdomain of inosine 5'-monophosphate dehydrogenase regulates purine nucleotide turnover. *Mol Microbiol*. 2008 Apr;68(2):342-59. doi: 10.1111/j.1365-2958.2008.06153.x. Epub 2008 Feb 26. PubMed PMID: 18312263; PubMed Central PMCID: PMC2279236.
- Plana-Bonamaisó A, López-Begines S, Fernández-Justel D, Junza A, Soler-Tapia A, Andilla J, Loza-Alvarez P, Rosa JL, Miralles E, Casals I, Yanes O, de la Villa P, Buey RM, Méndez A. post-translational regulation of retinal IMPDH1 *in vivo* to adjust GTP synthesis to illumination conditions. *Elife*. 2020 Apr 7;9. doi: 10.7554/eLife.56418. PubMed PMID: 32254022; PubMed Central PMCID: PMC7176436.
- Poët M, Kornak U, Schweizer M, Zdebik AA, Scheel O, Hoelter S, Wurst W, Schmitt A, Fuhrmann JC, Planells-Cases R, Mole SE, Hübner CA, Jentsch TJ. Lysosomal storage disease upon disruption of the neuronal chloride transport protein CIC-6. *Proc Natl Acad Sci U S A*. 2006 Sep 12;103(37):13854-9. doi: 10.1073/pnas.0606137103. Epub 2006 Sep 1. PMID: 16950870; PMCID: PMC1564226.
- Polge C, Jossier M, Crozet P, Gissot L, Thomas M. Beta-subunits of the SnRK1 complexes share a common ancestral function together with expression and function specificities; physical interaction with nitrate reductase specifically occurs via AKINbeta1-subunit. *Plant Physiol*. 2008 Nov;148(3):1570-82. doi: 10.1104/pp.108.123026. Epub 2008 Sep 3. PubMed PMID: 18768910; PubMed Central PMCID: PMC2577271.
- Polge C, Thomas M. SNF1/AMPK/SnRK1 kinases, global regulators at the heart of energy control? *Trends Plant Sci*. 2007 Jan;12(1):20-8. doi: 10.1016/j.tplants.2006.11.005. Epub 2006 Dec 12. Review. PubMed PMID: 17166759.
- Ramon M, Ruelens P, Li Y, Sheen J, Geuten K, Rolland F. The hybrid four-CBS-domain KIN β subunit functions as the canonical γ subunit of the plant energy sensor SnRK1. *Plant J*. 2013 Jul;75(1):11-25. doi: 10.1111/tbj.12192. Epub 2013 May 15. PubMed PMID: 23551663; PubMed Central PMCID: PMC6599549.
- Rojas-Jiménez K, Sohlenkamp C, Geiger O, Martínez-Romero E, Werner D, Vinuesa P. A CIC chloride channel homolog and ornithine-containing membrane lipids of *Rhizobium tropici* CIAT899 are involved in symbiotic efficiency and acid tolerance. *Mol Plant Microbe Interact*. 2005 Nov;18(11):1175-85. doi: 10.1094/MPMI-18-1175. PMID: 16353552.
- Ron M, Alandete Saez M, Eshed Williams L, Fletcher JC, McCormick S. Proper regulation of a sperm-specific *cis*-nat-siRNA is essential for double fertilization in Arabidopsis. *Genes Dev*. 2010 May 15;24(10):1010-21. doi: 10.1101/gad.1882810. PubMed PMID: 20478994; PubMed Central PMCID: PMC2867206.

- Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A, Snuffin M, Rezantsev A, Popov D, Ryltsov A, Kostukovich E, Borisovsky I, Liu Z, Vinsavich A, Trush V, Quackenbush J. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*. 2003 Feb;34(2):374-8. doi: 10.2144/03342mt01. PubMed PMID: 12613259.
- Sahoo KK, Tripathi AK, Pareek A, Sopory SK, Singla-Pareek SL. An improved protocol for efficient transformation and regeneration of diverse indica rice cultivars. *Plant Methods*. 2011 Dec 30;7(1):49. doi: 10.1186/1746-4811-7-49. PubMed PMID: 22206592; PubMed Central PMCID: PMC3284416.
- Sasvari Z, Kovalev N, Nagy PD. The GEF1 proton-chloride exchanger affects tombusvirus replication via regulation of copper metabolism in yeast. *J Virol*. 2013 Feb;87(3):1800-10. doi: 10.1128/JVI.02003-12. Epub 2012 Nov 28. PMID: 23192874; PMCID: PMC3554144.
- Scheel O, Zdebek AA, Lourdel S, Jentsch TJ. Voltage-dependent electrogenic chloride/proton exchange by endosomal CLC proteins. *Nature*. 2005 Jul 21;436(7049):424-7. doi: 10.1038/nature03860. PMID: 16034422.
- Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc*. 2008;3(6):1101-8. doi: 10.1038/nprot.2008.73. PMID: 18546601.
- Scholl U, Hebeisen S, Janssen AG, Müller-Newen G, Alekov A, Fahlke C. Barttin modulates trafficking and function of Cl⁻-K channels. *Proc Natl Acad Sci U S A*. 2006 Jul 25;103(30):11411-6. doi: 10.1073/pnas.0601631103. Epub 2006 Jul 18. PMID: 16849430; PMCID: PMC1544099.
- Schrecker M, Korobenko J, Hite RK. Cryo-EM structure of the lysosomal chloride-proton exchanger CLC-7 in complex with OSTM1. *Elife*. 2020 Aug 4;9:e59555. doi: 10.7554/eLife.59555. PMID: 32749217; PMCID: PMC7440919.
- Schumacher K, Vafeados D, McCarthy M, Sze H, Wilkins T, Chory J. The Arabidopsis det3 mutant reveals a central role for the vacuolar H⁺-ATPase in plant growth and development. *Genes Dev*. 1999 Dec 15;13(24):3259-70. doi: 10.1101/gad.13.24.3259. PMID: 10617574; PMCID: PMC317205.
- Schuster CF, Bellows LE, Tosi T, Campeotto I, Corrigan RM, Freemont P, Gründling A. The second messenger c-di-AMP inhibits the osmolyte uptake system OpuC in *Staphylococcus aureus*. *Sci Signal*. 2016 Aug 16;9(441):ra81. doi: 10.1126/scisignal.aaf7279. PubMed PMID: 27531650; PubMed Central PMCID: PMC5248971.
- Shan X, Kruger WD. Correction of disease-causing CBS mutations in yeast. *Nat Genet*. 1998 May;19(1):91-3. doi: 10.1038/ng0598-91. PubMed PMID: 9590298.
- Shin JS, So WM, Kim SY, Noh M, Hyoung S, Yoo KS, Shin JS. CBSX3-Trxo-2 regulates ROS generation of mitochondrial complex II (succinate dehydrogenase) in Arabidopsis. *Plant Sci*. 2020 May;294:110458. doi: 10.1016/j.plantsci.2020.110458. Epub 2020 Feb 25. PubMed PMID: 32234226.
- Sikkema HR, van den Noort M, Rheinberger J, de Boer M, Krepel ST, Schuurman-Wolters GK, Paulino C, Poolman B. Gating by ionic strength and safety check by cyclic-di-AMP in the ABC transporter OpuA. *Sci Adv*. 2020 Nov;6(47). doi: 10.1126/sciadv.abd7697. Print 2020 Nov. PubMed PMID: 33208376; PubMed Central PMCID: PMC7673798.
- Simon DB, Bindra RS, Mansfield TA, Nelson-Williams C, Mendonca E, Stone R, Schurman S, Nayir A, Alpay H, Bakkaloglu A, Rodriguez-Soriano J, Morales JM, Sanjad SA, Taylor CM, Pilz D, Brem A, Trachtman H, Griswold W, Richard GA, John E, Lifton RP. Mutations in the chloride channel gene, *CLCNKB*, cause Bartter's syndrome type III. *Nat Genet*. 1997 Oct;17(2):171-8. doi: 10.1038/ng1097-171. PMID: 9326936.

- Singh AK, Kumar R, Pareek A, Sopory SK, Singla-Pareek SL. Overexpression of rice CBS domain containing protein improves salinity, oxidative, and heavy metal tolerance in transgenic tobacco. *Mol Biotechnol*. 2012 Nov;52(3):205-16. doi: 10.1007/s12033-011-9487-2. PubMed PMID: 22302312.
- Smith S, Boitz J, Chidambaram ES, Chatterjee A, Ait-Tihyaty M, Ullman B, Jardim A. The cystathionine- β -synthase domains on the guanosine 5'-monophosphate reductase and inosine 5'-monophosphate dehydrogenase enzymes from *Leishmania* regulate enzymatic activity in response to guanylate and adenylate nucleotide levels. *Mol Microbiol*. 2016 Jun;100(5):824-40. doi: 10.1111/mmi.13352. Epub 2016 Mar 10. PubMed PMID: 26853689; PubMed Central PMCID: PMC4879609.
- Sparkes IA, Runions J, Kearns A, Hawes C. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nat Protoc*. 2006;1(4):2019-25. doi: 10.1038/nprot.2006.286. PubMed PMID: 17487191.
- Steinmeyer K, Ortland C, Jentsch TJ. Primary structure and functional expression of a developmentally regulated skeletal muscle chloride channel. *Nature*. 1991 Nov 28;354(6351):301-4. doi: 10.1038/354301a0. PMID: 1659664.
- Steinmeyer K, Schwappach B, Bens M, Vandewalle A, Jentsch TJ. Cloning and functional expression of rat CLC-5, a chloride channel related to kidney disease. *J Biol Chem*. 1995 Dec 29;270(52):31172-7. doi: 10.1074/jbc.270.52.31172. PMID: 8537381.
- Stockbridge RB, Lim HH, Otten R, Williams C, Shane T, Weinberg Z, Miller C. Fluoride resistance and transport by riboswitch-controlled CLC antiporters. *Proc Natl Acad Sci U S A*. 2012 Sep 18;109(38):15289-94. doi: 10.1073/pnas.1210896109. Epub 2012 Sep 4. PMID: 22949689; PMCID: PMC3458365.
- Stöltzing G, Teodorescu G, Begemann B, Schubert J, Nabbout R, Toliat MR, Sander T, Nürnberg P, Lerche H, Fahlke C. Regulation of ClC-2 gating by intracellular ATP. *Pflugers Arch*. 2013 Oct;465(10):1423-37. doi: 10.1007/s00424-013-1286-0. Epub 2013 May 1. PMID: 23632988; PMCID: PMC3778897.
- Subba A, Tomar S, Pareek A, Singla-Pareek SL. The chloride channels: silently serving the plants. *Physiol Plant*. 2021 Apr;171(4):688-702. doi: 10.1111/ppl.13240. Epub 2020 Nov 5. PubMed PMID: 33034380.
- Sumimoto H, Kamakura S, Ito T. Structure and function of the PB1 domain, a protein interaction module conserved in animals, fungi, amoebas, and plants. *Sci STKE*. 2007 Aug 28;2007(401):re6. doi: 10.1126/stke.4012007re6. Review. PubMed PMID: 17726178.
- Sun H, Shen L, Qin Y, Liu X, Hao K, Li Y, Wang J, Yang J, Wang F. *CLC-Nt1* affects *Potato Virus Y* infection via regulation of endoplasmic reticulum luminal pH. *New Phytol*. 2018 Oct;220(2):539-552. doi: 10.1111/nph.15310. Epub 2018 Jul 19. PMID: 30022473.
- Sun W, Deng D, Liheng Y, Xiaojiao Z, Juan Y, Huixin P, Qiang Z. Overexpression of the chloride channel gene (*GmCLC1*) from soybean increases salt tolerance in transgenic *Populus deltoides* x *P. euramericana* 'Nanlin895'. *Plant Omics*. 2013 Sept;6(5):347-54.
- Suzuki T, Rai T, Hayama A, Sohara E, Suda S, Itoh T, Sasaki S, Uchida S. Intracellular localization of ClC chloride channels and their ability to form hetero-oligomers. *J Cell Physiol*. 2006 Mar;206(3):792-8. doi: 10.1002/jcp.20516. PMID: 16222710.
- Tampieri E, Baraldi E, Carnevali F, Frascaroli E, De Santis A. The activity of plant inner membrane anion channel (PIMAC) can be performed by a chloride channel (CLC) protein in mitochondria from seedlings of maize populations divergently selected for cold tolerance. *J Bioenerg Biomembr*. 2011 Dec;43(6):611-21. doi: 10.1007/s10863-011-9386-z. Epub 2011 Oct 12. PMID: 21989547.

- Taylor SC, Nadeau K, Abbasi M, Lachance C, Nguyen M, Fenrich J. The ultimate qPCR experiment: producing publication quality, reproducible data the first time. *Trends Biotechnol.* 2019 Jul;37(7):761-774. doi: 10.1016/j.tibtech.2018.12.002. Epub 2019 Jan 14. PMID: 30654913.
- Teardo E, Frare E, Segalla A, De Marco V, Giacometti GM, Szabò I. Localization of a putative ClC chloride channel in spinach chloroplasts. *FEBS Lett.* 2005 Sep 12;579(22):4991-6. doi: 10.1016/j.febslet.2005.08.005. PMID: 16115625.
- Tomita A, Zhang M, Jin F, Zhuang W, Takeda H, Maruyama T, Osawa M, Hashimoto KI, Kawasaki H, Ito K, Dohmae N, Ishitani R, Shimada I, Yan Z, Hattori M, Nureki O. ATP-dependent modulation of MgtE in Mg²⁺ homeostasis. *Nat Commun.* 2017 Jul 27;8(1):148. doi: 10.1038/s41467-017-00082-w. PubMed PMID: 28747715; PubMed Central PMCID: PMC5529423.
- Townley R, Shapiro L. Crystal structures of the adenylate sensor from fission yeast AMP-activated protein kinase. *Science.* 2007 Mar 23;315(5819):1726-9. doi: 10.1126/science.1137503. Epub 2007 Feb 8. PubMed PMID: 17289942.
- Trachsel E, Redder P, Linder P, Armitano J. Genetic screens reveal novel major and minor players in magnesium homeostasis of *Staphylococcus aureus*. *PLoS Genet.* 2019 Aug;15(8):e1008336. doi: 10.1371/journal.pgen.1008336. eCollection 2019 Aug. PubMed PMID: 31415562; PubMed Central PMCID: PMC6711546.
- Tuominen H, Salminen A, Oksanen E, Jämsen J, Heikkilä O, Lehtiö L, Magretova NN, Goldman A, Baykov AA, Lahti R. Crystal structures of the CBS and DRTGG domains of the regulatory region of *Clostridium perfringens* pyrophosphatase complexed with the inhibitor, AMP, and activator, diadenosine tetraphosphate. *J Mol Biol.* 2010 May 7;398(3):400-13. doi: 10.1016/j.jmb.2010.03.019. Epub 2010 Mar 19. PubMed PMID: 20303981.
- Uchida S, Sasaki S, Nitta K, Uchida K, Horita S, Nihei H, Marumo F. Localization and functional characterization of rat kidney-specific chloride channel, ClC-K1. *J Clin Invest.* 1995 Jan;95(1):104-13. doi: 10.1172/JCI117626. PMID: 7814604; PMCID: PMC295382.
- van Slegtenhorst MA, Bassi MT, Borsani G, Wapenaar MC, Ferrero GB, de Conciliis L, Rugarli EI, Grillo A, Franco B, Zoghbi HY, Ballabio A. A gene from the Xp22.3 region shares homology with voltage-gated chloride channels. *Hum Mol Genet.* 1994 Apr;3(4):547-52. doi: 10.1093/hmg/3.4.547. PMID: 8069296.
- Vandewalle A, Cluzeaud F, Peng KC, Bens M, Lüchow A, Günther W, Jentsch TJ. Tissue distribution and subcellular localization of the ClC-5 chloride channel in rat intestinal cells. *Am J Physiol Cell Physiol.* 2001 Feb;280(2):C373-81. doi: 10.1152/ajpcell.2001.280.2.C373. PMID: 11208533.
- von der Fecht-Bartenbach J, Bogner M, Dynowski M, Ludewig U. CLC-b-mediated NO₃⁻/H⁺ exchange across the tonoplast of Arabidopsis vacuoles. *Plant Cell Physiol.* 2010 Jun;51(6):960-8. doi: 10.1093/pcp/pcq062. Epub 2010 Apr 28. PMID: 20430762.
- von der Fecht-Bartenbach J, Bogner M, Krebs M, Stierhof YD, Schumacher K, Ludewig U. Function of the anion transporter AtCLC-d in the trans-Golgi network. *Plant J.* 2007 May;50(3):466-74. doi: 10.1111/j.1365-313X.2007.03061.x. Epub 2007 Mar 21. PMID: 17376158; PMCID: PMC1891005.
- Waldegger S, Jeck N, Barth P, Peters M, Vitzthum H, Wolf K, Kurtz A, Konrad M, Seyberth HW. Barttin increases surface expression and changes current properties of ClC-K channels. *Pflugers Arch.* 2002 Jun;444(3):411-8. doi: 10.1007/s00424-002-0819-8. Epub 2002 Apr 9. PMID: 12111250.
- Walia H, Wilson C, Condamine P, Liu X, Ismail AM, Zeng L, Wanamaker SI, Mandal J, Xu J, Cui X, Close TJ. Comparative transcriptional profiling of two contrasting rice genotypes under salinity stress during

- the vegetative growth stage. *Plant Physiol.* 2005 Oct;139(2):822-35. doi: 10.1104/pp.105.065961. Epub 2005 Sep 23. PubMed PMID: 16183841; PubMed Central PMCID: PMC1255998.
- Wang K, Preisler SS, Zhang L, Cui Y, Missel JW, Grønberg C, Gotfryd K, Lindahl E, Andersson M, Calloe K, Egea PF, Klaerke DA, Pusch M, Pedersen PA, Zhou ZH, Gourdon P. Structure of the human ClC-1 chloride channel. *PLoS Biol.* 2019 Apr 25;17(4):e3000218. doi: 10.1371/journal.pbio.3000218. PMID: 31022181; PMCID: PMC6483157.
- Wang S, Su SZ, Wu Y, Li SP, Shan XH, Liu HK, Wang S, Yuan YP. Overexpression of maize chloride channel gene *ZmCLC-d* in *Arabidopsis thaliana* improved its stress resistance. *Biol Plant.* 2015;59(1):55-64. doi: 10.1007/s10535-014-0468-8.
- Wang X, Ren X, Zhu L, He G. *OsBil*, a rice gene, encodes a novel protein with a CBS-like domain and its expression is induced in responses to herbivore feeding. *Plant Science.* 2004 June;166(6):1581-88. doi: 10.1016/j.plantsci.2004.02.011.
- Wege S, De Angeli A, Droillard MJ, Kroniewicz L, Merlot S, Cornu D, Gambale F, Martinoia E, Barbier-Brygoo H, Thomine S, Leonhardt N, Filleur S. Phosphorylation of the vacuolar anion exchanger AtCLCa is required for the stomatal response to abscisic acid. *Sci Signal.* 2014 Jul 8;7(333):ra65. doi: 10.1126/scisignal.2005140. PMID: 25005229.
- Wege S, Jossier M, Filleur S, Thomine S, Barbier-Brygoo H, Gambale F, De Angeli A. The proline 160 in the selectivity filter of the *Arabidopsis* NO₃⁻/H⁺ exchanger AtCLCa is essential for nitrate accumulation *in planta*. *Plant J.* 2010 Sep;63(5):861-9. doi: 10.1111/j.1365-313X.2010.04288.x. PMID: 20598093.
- Wei P, Che B, Shen L, Cui Y, Wu S, Cheng C, Liu F, Li MW, Yu B, Lam HM. Identification and functional characterization of the chloride channel gene, *GsCLC-c2* from wild soybean. *BMC Plant Biol.* 2019 Apr 1;19(1):121. doi: 10.1186/s12870-019-1732-z. PMID: 30935372; PMCID: PMC6444504.
- Wei P, Wang L, Liu A, Yu B, Lam HM. *GmCLC1* confers enhanced salt tolerance through regulating chloride accumulation in soybean. *Front Plant Sci.* 2016 Jul 25;7:1082. doi: 10.3389/fpls.2016.01082. PMID: 27504114; PMCID: PMC4959425.
- Wei Q, Liu Y, Zhou G, Li Q, Yang C, Peng S. Overexpression of *CsCLCc*, a chloride channel gene from *Poncirus trifoliata*, enhances salt tolerance in *Arabidopsis*. *Plant Mol Biol Rep.* 2013;31:1548-57. <https://doi.org/10.1007/s11105-013-0592-1>.
- Weinert S, Gimber N, Deuschel D, Stuhlmann T, Puchkov D, Farsi Z, Ludwig CF, Novarino G, López-Cayuqueo KI, Planells-Cases R, Jentsch TJ. Uncoupling endosomal CLC chloride/proton exchange causes severe neurodegeneration. *EMBO J.* 2020 May 4;39(9):e103358. doi: 10.15252/embj.2019103358. Epub 2020 Mar 2. PMID: 32118314; PMCID: PMC7196918.
- Wellhauser L, Luna-Chavez C, D'Antonio C, Tainer J, Bear CE. ATP induces conformational changes in the carboxyl-terminal region of ClC-5. *J Biol Chem.* 2011 Feb 25;286(8):6733-41. doi: 10.1074/jbc.M110.175877. Epub 2010 Dec 20. PMID: 21173145; PMCID: PMC3057859.
- White MM, Miller C. A voltage-gated anion channel from the electric organ of *Torpedo californica*. *J Biol Chem.* 1979 Oct 25;254(20):10161-6. PMID: 489590.
- Wilson D, Pethica R, Zhou Y, Talbot C, Vogel C, Madera M, Chothia C, Gough J. SUPERFAMILY--sophisticated comparative genomics, data mining, visualization and phylogeny. *Nucleic Acids Res.* 2009 Jan;37(Database issue):D380-6. doi: 10.1093/nar/gkn762. Epub 2008 Nov 26. PubMed PMID: 19036790; PubMed Central PMCID: PMC2686452.

- Wilson WA, Hawley SA, Hardie DG. Glucose repression/derepression in budding yeast: SNF1 protein kinase is activated by phosphorylation under derepressing conditions, and this correlates with a high AMP:ATP ratio. *Curr Biol*. 1996 Nov 1;6(11):1426-34. doi: 10.1016/s0960-9822(96)00747-6. PubMed PMID: 8939604.
- Wong TH, Li MW, Yao XQ, Lam HM. The GmCLC1 protein from soybean functions as a chloride ion transporter. *J Plant Physiol*. 2013 Jan 1;170(1):101-4. doi: 10.1016/j.jplph.2012.08.003. Epub 2012 Aug 24. PMID: 22921676.
- Wungrampha S, Joshi R, Rathore RS, Singla-Pareek SL, Govindjee, Pareek A. CO₂ uptake and chlorophyll a fluorescence of *Suaeda fruticosa* grown under diurnal rhythm and after transfer to continuous dark. *Photosynth Res*. 2019 Nov;142(2):211-227. doi: 10.1007/s11120-019-00659-0. Epub 2019 Jul 17. PMID: 31317383.
- Xiao B, Heath R, Saiu P, Leiper FC, Leone P, Jing C, Walker PA, Haire L, Eccleston JF, Davis CT, Martin SR, Carling D, Gamblin SJ. Structural basis for AMP binding to mammalian AMP-activated protein kinase. *Nature*. 2007 Sep 27;449(7161):496-500. doi: 10.1038/nature06161. Epub 2007 Sep 12. PubMed PMID: 17851531.
- Xing A, Ma Y, Wu Z, Nong S, Zhu J, Sun H, Tao J, Wen B, Zhu X, Fang W, Li X, Wang Y. Genome-wide identification and expression analysis of the CLC superfamily genes in tea plants (*Camellia sinensis*). *Funct Integr Genomics*. 2020 Jul;20(4):497-508. doi: 10.1007/s10142-019-00725-9. Epub 2020 Jan 3. PMID: 31897824.
- Yamasaki S, Dillenburg LR. Measurements of leaf relative water content in *Araucaria angustifolia*. *R Bras Fisiol Veg*. 1999;11(2):69-75.
- Yang G, Zou H, Wu Y, Liu H, Yuan Y. Identification and characterisation of candidate genes involved in chilling responses in maize (*Zea mays* L.). *Plant Cell Tiss Organ Cult*. 2011;106(1):127-41. doi: 10.1007/s11240-010-9900-8.
- Yoo KS, Ok SH, Jeong BC, Jung KW, Cui MH, Hyoung S, Lee MR, Song HK, Shin JS. Single cystathionine β -synthase domain-containing proteins modulate development by regulating the thioredoxin system in *Arabidopsis*. *Plant Cell*. 2011 Oct;23(10):3577-94. doi: 10.1105/tpc.111.089847. Epub 2011 Oct 21. PubMed PMID: 22021414; PubMed Central PMCID: PMC3229136.
- Yoo SD, Cho YH, Sheen J. *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat Protoc*. 2007;2(7):1565-72. doi: 10.1038/nprot.2007.199. PMID: 17585298.
- Zafar SA, Patil SB, Uzair M, Fang J, Zhao J, Guo T, Yuan S, Uzair M, Luo Q, Shi J, Schreiber L, Li X. DEGENERATED PANICLE AND PARTIAL STERILITY 1 (DPS1) encodes a cystathionine β -synthase domain containing protein required for anther cuticle and panicle development in rice. *New Phytol*. 2020 Jan;225(1):356-375. doi: 10.1111/nph.16133. Epub 2019 Sep 30. PubMed PMID: 31433495.
- Zafar SA, Uzair M, Khan MR, Patil SB, Fang J, Zhao J, Singla-Pareek SL, Pareek A and Li X. DPS1 regulates cuticle development and leaf senescence in rice. *Food and Energy Security*. 2021;10(1):e273. <https://doi.org/10.1002/fes3.273>.
- Zdebik AA, Zifarelli G, Bergsdorf EY, Soliani P, Scheel O, Jentsch TJ, Pusch M. Determinants of anion-proton coupling in mammalian endosomal CLC proteins. *J Biol Chem*. 2008 Feb 15;283(7):4219-27. doi: 10.1074/jbc.M708368200. Epub 2007 Dec 6. PMID: 18063579.

- Zhang H, Jin J, Jin L, Li Z, Xu G, Wang R, Zhang J, Zhai N, Chen Q, Liu P, Chen X, Zheng Q, Zhou H. Identification and analysis of the chloride channel gene family members in tobacco (*Nicotiana tabacum*). *Gene*. 2018 Nov 15;676:56-64. doi: 10.1016/j.gene.2018.06.073. Epub 2018 Jun 26. PMID: 29958955.
- Zhang H, Kozlov G, Li X, Wu H, Gulerez I, Gehring K. PRL3 phosphatase active site is required for binding the putative magnesium transporter CNNM3. *Sci Rep*. 2017 Mar 3;7(1):48. doi: 10.1038/s41598-017-00147-2. PubMed PMID: 28246390; PubMed Central PMCID: PMC5427921.
- Zhang X, Henriques R, Lin SS, Niu QW, Chua NH. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat Protoc*. 2006;1(2):641-6. doi: 10.1038/nprot.2006.97. Epub 2006 Jun 29. PMID: 17406292.
- Zhang Y, Su J, Duan S, Ao Y, Dai J, Liu J, Wang P, Li Y, Liu B, Feng D, Wang J, Wang H. A highly efficient rice green tissue protoplast system for transient gene expression and studying light/chloroplast-related processes. *Plant Methods*. 2011 Sep 30;7(1):30. doi: 10.1186/1746-4811-7-30. PMID: 21961694; PMCID: PMC3203094.
- Zhou GA, Qiu LJ. Identification and functional analysis on abiotic stress response of soybean Cl⁻ channel gene *GmCLCnt*. *Agric Sci China*. 2010;9(2):199-206. [https://doi.org/10.1016/S1671-2927\(09\)60084-5](https://doi.org/10.1016/S1671-2927(09)60084-5).
- Zifarelli G, Pusch M. Conversion of the 2 Cl⁻/1 H⁺ antiporter CIC-5 in a NO₃⁻/H⁺ antiporter by a single point mutation. *EMBO J*. 2009 Feb 4;28(3):175-82. doi: 10.1038/emboj.2008.284. Epub 2009 Jan 8. PMID: 19131966; PMCID: PMC2637338.
- Zivy M, Thiellement H, de Vienne D, Hofmann JP. Study on nuclear and cytoplasmic genome expression in wheat by two-dimensional gel electrophoresis : 1. First results on 18 alloplasmic lines. *Theor Appl Genet*. 1983 Jul;66(1):1-7. doi: 10.1007/BF00281838. PMID: 24263624.
- Zuhra K, Augsburg F, Majtan T, Szabo C. Cystathionine-β-Synthase: Molecular regulation and pharmacological inhibition. *Biomolecules*. 2020 Apr 30;10(5):697. doi: 10.3390/biom10050697. PMID: 32365821; PMCID: PMC7277093.

Appendix 1. Plant Medium Preparation

Table A1.1. Composition of Yoshida medium (Yoshida et al, 1976).

| Element | Reagent | Reagent (in g) required for 1 L of Stock solution in deionized water | Stock solution (in ml) required for 4 L of Working solution | Working conc. of an element (in ppm) | | | | |
|---------|---|--|---|--|------------------|-----|------|------|
| N | NH ₄ NO ₃ | 91.4 | 5 | 40 | | | | |
| P | NaH ₂ PO ₄ · 2H ₂ O | 40.3 | 5 | 10 | | | | |
| K | K ₂ SO ₄ | 71.4 | 5 | 40 | | | | |
| Ca | CaCl ₂ | 88.6 | 5 | 40 | | | | |
| Mg | MgSO ₄ · 7H ₂ O | 324 | 5 | 40 | | | | |
| Mn | MnCl ₂ · 4H ₂ O | 1.5 | 5 | 0.5 | | | | |
| Mo | (NH ₄) ₆ · MO ₇ O ₂₄ · 4H ₄ O | 0.074 | | Dissolved in 50 ml of concentrated H ₂ SO ₄ and made up the volume to 1 L in deionized water | 0.05 | | | |
| B | H ₃ BO ₃ | 0.934 | | | pH is set to 5.0 | 0.2 | | |
| Zn | ZnSO ₄ · 7H ₂ O | 0.035 | | | | - | 0.01 | |
| Cu | CuSO ₄ · 5H ₂ O | 0.031 | | | | | - | 0.01 |
| Fe | FeCl ₃ · 6H ₂ O | 7.7 | | | | | | - |
| | Citric acid monohydrate (C ₆ H ₈ O ₇ · H ₂ O) | 11.9 | - | | | | | |

The stock solutions are stored at 4°C. The working solution is prepared fresh as needed.

Yoshida medium is used to grow rice seedlings in hydroponics. This medium is also applied to tissue culture regenerated seedlings transferred onto vermiculite pots during hardening stage.

Table A1.2. Composition of Tobacco Leaf Infiltration Medium.

| Reagent | Stock solution | Stock required for 20 ml of working solution | Final concentration |
|--|--|--|---------------------|
| 2-(N-morpholino) ethanesulfonic acid / MES (C ₆ H ₁₃ NO ₄ S) | 100 mM, pH 5.6, 1.95 g in water up to 100 ml, filter sterilized and stored at 4 °C | 2 ml | 10 mM |
| Magnesium chloride hexahydrate (MgCl ₂ · 6H ₂ O) | 100 mM, 2.03 g in water up to 100 ml, filter sterilized and stored at 4 °C | 2 ml | 10 mM |
| 3',5'-dimethoxy-4'-hydroxy acetophenone / Acetosyringone (C ₁₀ H ₁₂ O ₄) | 500 mM, 20 mg in 200 µl DMSO, wrapped in Al foil and stored at -20 °C | 6 µl | 150 µM |
| Glucose | - | 100 mg | 0.5% |
| Sterile Milli-Q Water | - | 16 ml | - |

The working infiltration medium is always prepared fresh from the stock solutions on the day of infiltration.

Appendix 2. Solutions for Protoplast Isolation

Table A2.1. Composition of Protoplast Isolation Enzyme Solution.

| Reagent | Stock solution | Stock required for 15 ml of Working solution | Final concentration |
|---|--|--|---------------------|
| 2-(N-morpholino) ethanesulfonic acid / MES (C ₆ H ₁₃ NO ₄ S) | 0.2 M (pH 5.7), 1.17 g in water up to 30 ml, Filter sterilized | 0.75 ml | 10 mM |
| Mannitol (C ₆ H ₁₄ O ₆) | 0.8 M, 4.4 g in water up to 30 ml, Filter sterilized | 11.25 ml | 0.6 M |
| Calcium chloride dihydrate (CaCl ₂ · 2H ₂ O) | 1 M, 4.4 g in water up to 30 ml, Filter sterilized | 0.15 ml | 10 mM |
| Cellulase R-10 | - | 225 mg | 1.5% |
| Macerozyme R-10 | - | 112.5 mg | 0.75% |
| Bovine serum albumin / BSA | - | 15 mg | 0.1% |
| Sterile Milli-Q water | - | 2.85 ml | - |

The enzyme solution is prepared fresh. The solution (without CaCl₂ and BSA) is heated in at 50°C in a water-bath for 10 min to inactivate proteases and DNase, while enhancing the enzyme solubility. It is then cooled down to room temperature, and the required amount of CaCl₂ and BSA are added onto it. The solution is passed through 0.45 µm syringe filter and used for sample digestion.

Table A2.2. Composition of W5 Solution.

| Reagent | Stock Solution | Stock required for 30 ml of Working Solution | Final concentration |
|---|--|--|---------------------|
| 2-(N-morpholino) ethanesulfonic acid / MES (C ₆ H ₁₃ NO ₄ S) | 0.2 M (pH 5.7), 1.17 g in water up to 30 ml, Filter sterilized | 300 µl | 2 mM |
| Calcium chloride dihydrate (CaCl ₂ · 2H ₂ O) | 1 M, 4.4 g in water up to 30 ml, Filter sterilized | 3.75 ml | 125 mM |
| Sodium chloride (NaCl) | 5 M, 29.2 g in water up to 100 ml, Autoclaved | 924 µl | 154 mM |
| Potassium chloride (KCl) | 2 M, 14.9 g in water up to 100 ml, Autoclaved | 75 µl | 5 mM |
| Sterile Milli-Q water | - | 25 ml | - |

The prepared solution is passed through 0.45 µm syringe filter.

Table A2.3. Composition of MMG Solution.

| Reagent | Stock Solution | Stock required for 10 ml of Working Solution | Final Concentration |
|---|--|--|---------------------|
| 2-(N-morpholino) ethanesulfonic acid / MES (C ₆ H ₁₃ NO ₄ S) | 0.2 M (pH 5.7), 1.17 g in water up to 30 ml, Filter sterilized | 200 µl | 4 mM |
| Mannitol (C ₆ H ₁₄ O ₆) | 0.8 M, 4.4 g in water up to 30 ml, Filter sterilized | 5 ml | 0.4 M |
| Magnesium chloride hexahydrate (MgCl ₂ ·6H ₂ O) | 2 M, 12.2 g in water up to 30 ml, Filter sterilized | 75 µl | 15 mM |
| Sterile Milli-Q water | - | 4.725 ml | - |

The prepared solution is passed through 0.45 µm syringe filter.

Table A2.4. Composition of PEG Solution.

| Reagent | Stock Solution | Stock required for 10 ml of Working Solution | Final Concentration |
|---|--|--|---------------------|
| PEG 4000 | - | 4 g | 40 % |
| Mannitol (C ₆ H ₁₄ O ₆) | 0.8 M, 4.4 g in water up to 30 ml, Filter sterilized | 2.5 ml | 0.2 M |
| Calcium chloride dihydrate (CaCl ₂ ·2H ₂ O) | 1 M, 4.4 g in water up to 30 ml, Filter sterilized | 1 ml | 0.1 M |
| Sterile Milli-Q water | - | Maku up the volume to 10 ml | - |

Table A2.5. Composition of WI solution.

| Reagent | Stock Solution | Stock required for 5 ml of Working Solution | Final Concentration |
|---|--|---|---------------------|
| 2-(N-morpholino) ethanesulfonic acid / MES (C ₆ H ₁₃ NO ₄ S) | 0.2 M (pH 5.7), 1.17 g in water up to 30 ml, Filter sterilized | 100 µl | 4 mM |
| Mannitol (C ₆ H ₁₄ O ₆) | 0.8 M, 4.4 g in water up to 30 ml, Filter sterilized | 3.125 ml | 0.5 mM |
| Potassium chloride (KCl) | 2 M, 14.9 g in water up to 100 ml, Autoclaved | 50 µl | 20 mM |
| Sterile Milli-Q water | - | 1.725 | - |

Appendix A3. Buffers and Solutions for Yeast Two-Hybrid Assay

Table A3.1. Composition of 10X TE Buffer.

| Component | Amount required for 250 ml of 10X Stock Solution | 10 x Stock Concentration |
|-------------------------|--|--------------------------|
| Tris base | 3 g | 100 mM |
| Disodium EDTA dihydrate | 0.93 g | 10 mM |
| Milli-Q water | Up to 250 ml volume | - |

pH of the buffer is adjusted to ~7.5. The buffer is autoclaved and stored at room temperature.

Table A3.2. Composition of 10X LiAc Solution.

| Component | Amount required for 100 ml of 10X Stock Solution | 10 x Stock Concentration |
|---|--|--------------------------|
| Lithium acetate dihydrate or LiAc (C ₂ H ₃ LiO ₂ ·2H ₂ O) | 10.2 g | 1 M |
| Milli-Q water | Up to 100 ml volume | - |

pH of the buffer is adjusted to ~7.5 with diluted acetic acid. The buffer is autoclaved and stored at room temperature.

Table A3.3. Composition of 50% (w/v) PEG 3350 Solution.

| Component | Amount required for 100 ml Solution |
|---|-------------------------------------|
| Polyethylene glycol (PEG; Av. mol. wt.: 3350) | 50 g |
| Milli-Q water | Up to 100 ml volume |

The solution is autoclaved and stored at room temperature.

Table A3.4. Composition of TE-LiAc Solution.

| Component | Amount required for 10 ml of Solution |
|--------------------------|---------------------------------------|
| 10X TE buffer | 1 ml |
| 10X LiAc solution | 1 ml |
| Autoclaved Milli-Q water | 8 ml |

Table A3.5. Composition of TE-LiAc-PEG Solution.

| Components | Amount required for 10 ml of Solution |
|-------------------|---------------------------------------|
| 10X TE buffer | 1 ml |
| 10X LiAc solution | 1 ml |
| 50% PEG solution | 8 ml |

Appendix 4. Plant Protein Extraction Buffer

Table A4.1. Composition of Zivy's Buffer

| Components | Amount required for 50 ml of Solution | Final concentration |
|--------------------------------------|---------------------------------------|---------------------|
| Tris | 182 mg | 30mm |
| Ascorbic Acid | 8.8mg | 1mm |
| Na.EDTA | 18.5mg | 1mm |
| MgCl ₂ .6H ₂ O | 50.8mg | 5mm |
| DTT | 7.5mg | 1mm |

pH of the buffer is adjusted to ~ 8.5

Appendix A5. Buffers and Solutions for SDS-PAGE and Western Blot Assay

Table A5.1. Composition of 4X Laemmli Sample Buffer.

| Reagent | Amount required for 10 ml of 4X Laemmli Sample Buffer Stock | 4X Stock Concentration | 1X Concentration |
|------------------------|---|------------------------|------------------|
| 0.5 M Tris-HCl, pH 6.8 | 4 ml | 200 mM | 50 mM |
| Glycerol | 4 ml | 40% | 10% |
| SDS | 800 mg | 8% | 2% |
| 1% Bromophenol Blue | 0.8 ml | 0.08% | 0.02% |
| Sterile Milli-Q water | Make up the volume to 10 ml | - | - |

Prior to use in protein sample preparation, 8 μ l of β -mercaptoethanol is added to 100 μ l aliquot of 4X Laemmli sample buffer.

Table A5.2. Composition of SDS-PAGE Resolving Gel (for 20 ml volume).

| Component | 8% gel | 10% gel | 12% gel | 15% gel |
|---|-------------|-------------|-------------|-------------|
| Milli-Q water | 9.3 ml | 7.9 ml | 6.6 ml | 4.6 ml |
| 30% Acrylamide mix* | 5.3 ml | 6.7 ml | 8 ml | 10 ml |
| 1.5 M Tris-HCl, pH 8.8 | 5 ml | 5 ml | 5 ml | 5 ml |
| 10% SDS | 200 μ l | 200 μ l | 200 μ l | 200 μ l |
| 10% Ammonium persulfate (APS)** | 200 μ l | 200 μ l | 200 μ l | 200 μ l |
| N,N,N',N'-tetramethylethylene-diamine (TEMED)** | 12 μ l | 8 μ l | 8 μ l | 8 μ l |

* 30% Acrylamide mix is prepared by dissolving 29 g of acrylamide and 1 g of N,N'-methylene-bisacrylamide in Milli-Q water to a final volume of 100 ml. The solution is heated at 37°C to dissolve the chemicals. It is then passed through 0.45 μ m filter and stored in the dark at 4°C.

** After adding APS and TEMED to the solution, the gel starts to polymerize quickly, so these two reagents are added immediately before pouring the gel solution into the glass plates.

Table A5.3. Composition of SDS-PAGE Stacking Gel (for 6 ml volume).

| Component | 5% gel |
|---|-------------|
| Milli-Q water | 4.1 ml |
| 30% Acrylamide mix* | 1 ml |
| 1 M Tris-HCl, pH 6.8 | 750 μ l |
| 10% SDS | 60 μ l |
| 10% Ammonium persulfate (APS)** | 60 μ l |
| N,N,N',N'-tetramethylethylene-diamine (TEMED)** | 6 μ l |

* For preparation of 30% Acrylamide mix, see Appendix Table A5.2.

** APS and TEMED are added immediately before pouring the gel solution into the glass plates.

Table A5.4. Composition of 10X Tris-Glycine Electrophoresis Buffer.

| Reagent | Amount of reagent required for 1 L of 10X Stock Solution | 10 x Stock Concentration |
|---------------|--|--------------------------|
| Tris base | 30 g | 0.25 M |
| Glycine | 144 g | 1.92 M |
| SDS | 10 g | 1% |
| Milli-Q water | Up to 1 L volume | |

pH of the buffer should be ~8.3. The buffer is autoclaved and stored at room temperature.

Table A5.5. Composition of Coomassie staining and De-staining* Solutions.

| Component | Amount required for 1 L Solution |
|----------------------|----------------------------------|
| Coomassie Blue R-250 | 1.25 g |
| Methanol | 400 ml |
| Acetic Acid | 100 ml |
| Water | 500 ml |

* De-staining solution has the same reagent composition as that of Coomassie staining solution, except for Coomassie Blue R-250 which is absent in it.

Both staining and de-staining solutions are stored at room temperature.

Table A5.6. Composition of 10X Western Transfer Buffer*.

| Component | Amount required for 1 L of 10X Stock Solution | 10X Stock Concentration |
|---------------|---|-------------------------|
| Tris base | 30 g | 0.25 M |
| Glycine | 144 g | 1.92 M |
| SDS | 3 g | 0.3% |
| Milli-Q water | Up to 1 L volume | - |

pH of the buffer should be ~8.3. The buffer is autoclaved and stored at room temperature.

* To prepare 1X Western transfer buffer, 10X buffer is diluted in methanol and water at the ratio of 1:2:7.

Table A5.7. Composition of 10X Phosphate Buffered Saline (10X PBS).

| Components | Amount required for 1 L of 10X Stock Solution | 10X Stock Concentration |
|--|---|-------------------------|
| Sodium chloride (NaCl) | 80 g | 1.37 M |
| Potassium chloride (KCl) | 2 g | 27 mM |
| Sodium phosphate dibasic (Na ₂ HPO ₄) | 14.4 g | 100 mM |
| Potassium phosphate monobasic (KH ₂ PO ₄) | 2.4 g | 18 mM |
| Milli-Q water | Up to 1 L volume | - |

pH of the buffer is adjusted to ~7.4. The buffer is autoclaved and stored at room temperature.

Table A5.8. Composition of 10X Tris Buffered Saline (10X TBS).

| Components | Amount required for 100 ml of 10X Stock Solution | 10X Stock Concentration |
|------------------------|--|-------------------------|
| Tris base | 2.4 g | 0.2 M |
| Sodium chloride (NaCl) | 8.8 g | 1.5 M |
| Milli-Q water | Up to 100 ml volume | - |

pH of the buffer is adjusted to ~7.6. The buffer is autoclaved and stored at room temperature.

Appendix 6. Buffers for Agarose Gel Electrophoresis

A6.1 Preparation of Denaturing Agarose Gel (1.2%) for RNA

For 100 ml of gel solution, 1.2 g of agarose is melted in 85 ml of Milli-Q water by heating in a microwave oven. The temperature of the solution is allowed to fall down to about 60°C. Then after, 10 ml of 10X MOPS buffer and 5 ml of 37% formaldehyde are added, mixed by swirling and poured onto the gel casting tray.

Table A6.1. Composition of 10X MOPS Buffer for RNA Gel Electrophoresis.

| Reagent | Amount required for 1 L of 10X Stock solution | 10X Stock concentration |
|--|---|-------------------------|
| MOPS / 3-morpholinopropane-1-sulfonic acid | 41.8 g | 200 mM |
| Sodium acetate trihydrate | 6.8 g | 50 mM |
| Disodium EDTA dihydrate | 3.7 g | 10 mM |
| Deionized water | Up to 1 L volume | - |

pH of the buffer is adjusted to 7.0 with NaOH or glacial acetic acid. The buffer is Filter-sterilized (do not autoclave) and stored in the dark at room temperature.

Table A6.2. Composition of 50X TAE Buffer.

| Reagent | Amount required for 1 L of 50X Stock solution | 50X Stock concentration |
|---------------------|---|-------------------------|
| Tris base | 242 g | 2 M |
| Glacial acetic acid | 57.1 ml | 1 M |
| 0.5 M EDTA, pH 8.0 | 100 ml | 50 mM |
| Milli-Q water | Up to 1 L volume | - |

pH of the buffer is adjusted to ~8.5. The buffer is autoclaved and stored at room temperature.

Table A6.3. Composition of 5X TBE Buffer.

| Components | Amount required for 1 L of 5X Stock solution | 5X Stock concentration |
|--------------------|--|------------------------|
| Tris base | 54 g | 0.45 M |
| Boric acid | 27.5 g | 0.45 M |
| 0.5 M EDTA, pH 8.0 | 20 ml | 10 mM |
| Milli-Q water | Up to 1 L volume | - |

pH of the buffer is adjusted to ~8.3. The buffer is autoclaved and stored at room temperature.

Appendix 7. CTAB Buffer for Plant DNA Isolation

Table A7.1. Composition of CTAB Extraction Buffer.

| Component | Amount required for 10 ml of Working Solution | Concentration of Working Solution |
|---------------------------------|---|-----------------------------------|
| 1 M Tris (pH 8.0), autoclaved | 1 ml | 100 mM |
| 10% CTAB, autoclaved | 2 ml | 2% |
| 5 M NaCl, autoclaved | 2.8 ml | 1.4 M |
| 0.5 M EDTA (pH 8.0), autoclaved | 0.4 ml | 20 mM |
| β -mercaptoethanol* | 0.02 ml | 0.2% |
| Autoclaved Milli-Q water | 3.78 ml | |

* β -mercaptoethanol is added prior to use of the buffer.

After preparation, the working buffer solution is preheated at 65°C in a water bath for its subsequent use.

Appendix 8. Buffers and Solutions for Southern Blot Assay

A8.1 Preparation of X-ray film Developing Solution

Dissolve 7.5 g and 46.7 g of developer powder from Packet A and Packet B (Carestream, 4908216), respectively in deionized water making up the volume to 500 ml. The solution is prepared in amber bottle. The reagent is dissolved completely using magnetic stirrer, and the solution stored in the dark at room temperature.

A8.2 Preparation of X-ray film Fixing Solution

Dissolve 121.5 g of Fixer (Carestream, 4908232) in deionized water making up the volume to 500 ml. The solution is prepared in amber bottle. The reagent is dissolved completely using magnetic stirrer, and the solution stored in the dark at room temperature.

Table A8.1. Composition of Depurination Solution (0.25 N HCl Solution).

| Component | Amount required for 1 L Solution | Concentration |
|------------------------------|----------------------------------|---------------|
| Hydrochloric acid (HCl, 37%) | 20.7 ml | 0.25 N |
| Milli-Q water | 979.3 ml | - |

The solution is autoclaved and stored at room temperature.

Table A8.2. Composition of Denaturing Solution.

| Component | Amount required for 1 L Solution | Concentration |
|---------------|----------------------------------|---------------|
| NaOH | 20 g | 0.5 N |
| NaCl | 88 g | 1.5 M |
| Milli-Q water | Up to 1 L volume | - |

The solution is autoclaved and stored at room temperature.

Table A8.3. Composition of Neutralizing Solution.

| Component | Amount required for 1 L Solution | Concentration |
|---------------|----------------------------------|---------------|
| Tris base | 60.6 g | 0.5 M |
| NaCl | 88 g | 1.5 M |
| Milli-Q water | Up to 1 L volume | - |

pH of the solution is adjusted to ~7.5. The solution is autoclaved and stored at room temperature.

Table A8.4. Composition of 20X SSC Buffer.

| Component | Amount required for 2 L of 20X Stock Solution | 20X Stock Concentration |
|-----------------------------|---|-------------------------|
| Sodium chloride | 350 g | 3 M |
| Trisodium citrate dihydrate | 176 g | 0.3 M |
| Milli-Q water | Up to 2 L volume | - |

pH of the solution is adjusted to ~7.0. The solution is autoclaved and stored at room temperature.

Table A8.5. Composition of Wash Solution I.

| Component | Amount required for 300 ml Solution | Concentration |
|--------------------------|-------------------------------------|---------------|
| 20X SSC Solution | 30 ml | 2X SSC |
| 10% SDS Solution | 3 ml | 0.1% SDS |
| Autoclaved Milli-Q water | Up to 300 ml volume | - |

The solution is prepared fresh prior to use from the stock solutions.

Table A8.6. Composition of Wash Solution II.

| Component | Amount required for 300 ml Solution | Concentration |
|--------------------------|-------------------------------------|---------------|
| 20X SSC Solution | 7.5 ml | 0.5X SSC |
| 10% SDS Solution | 3 ml | 0.1% SDS |
| Autoclaved Milli-Q water | Up to 300 ml volume | - |

The solution is prepared fresh prior to use from the stock solutions.

Table A8.7. Composition of 5X Maleic Acid Buffer.

| Component | Amount required for 500 ml 5X Stock Solution | Concentration of 5X Stock Solution |
|---------------|--|------------------------------------|
| Maleic Acid | 29 g | 0.5 M |
| NaCl | 22 g | 0.75 M |
| Milli-Q water | Up to 500 ml volume | - |

pH of the solution is adjusted to ~7.5. The solution is autoclaved and stored at room temperature.

Table A8.8. Composition of Wash Buffer.

| Component | Amount required for 300 ml Solution | Concentration |
|-----------------------|-------------------------------------|-----------------------|
| 5X Maleic Acid Buffer | 60 ml | 1X Maleic Acid Buffer |
| Tween 20 | 900 μ l | 0.3% Tween 20 |
| Milli-Q water | 239 ml | - |

Table A8.9. Composition of 10X Blocking Solution.

| Component | Amount required for 100 ml of 10X Stock Solution | Concentration of 10X Stock Solution |
|--|--|-------------------------------------|
| Blocking reagent (DIG DNA Labeling and Detection Kit, Roche) | 10 g | 10% |
| 5X Maleic Acid Buffer | 20 ml | 1X Maleic Acid Buffer |
| Milli-Q water | Up to 100 ml volume | - |

The solution is heated in a microwave oven and stirred on a magnetic stirrer to dissolve the reagent. The solution is autoclaved and stored at 4°C.

1X Blocking Solution is prepared fresh prior to use by diluting 10X Blocking Solution in 1X Maleic Acid Buffer at 1:10 ratio.

Table A8.10. Composition of 10X Detection Buffer.

| Component | Amount required for 200 ml of 10X Stock | Concentration of 10X Stock |
|---------------|---|----------------------------|
| Tris base | 24.2 g | 1 M |
| NaCl | 11.6 g | 1 M |
| Milli-Q water | Up to 200 ml volume | - |

pH of the buffer is adjusted to ~9.5. The buffer is autoclaved and stored at room temperature.

Appendix 9. Primers sequences used in the study

Note that in cloning primers only the sequence corresponding to the gene has been listed in the table below. The required restriction site sequences based on the cloning sites in different vectors were added to these primers accordingly.

Table A9.1. List of Primers

| Primer Name | 5'<-----Sequence----->3' |
|---------------------------|------------------------------|
| OseEF-1 α _qRT_F | TTTCACTCTTGGTGTGAAGCAGAT |
| OseEF-1 α _qRT_R | GACTTCCTTCACGATTTCATCGTAA |
| OsCBSX3_qRT_F | CATCTGACAAAACAGTAATGGATG |
| OsCBSX3_qRT_R | AGCACGGGTAACATACTATC |
| OsCYP_qRT_F | CGTCGCTCTATTGAGGAGGATATG |
| OsCYP_qRT_R | CTTTGGGCAATGATGTAGGTTCC |
| OsTrx y_qRT_F | GGAGGTTGCAAAGTAGCATGTG |
| OsTrx y_qRT_R | TGTAAATGCAGATGCGATGGG |
| OsCBSCLC6_qRT_F | ATTGATACACGCACATACATTCTGG |
| OsCBSCLC6_qRT_R | CTCAATGACCAATCCAACGTGTATG |
| OsCBSCLC6_qRT_Transgene_F | TGACGAGGCACGACTTCATG |
| OsCBSCLC6_qRT_Transgene_R | AGGAGCGACCGGAATTTCTTG |
| AtUBQ10_qRT_F | GGCCTTGTATAATCCCTGATGAATAAG |
| AtUBQ10_qRT_R | AAAGAGATAACAGGAACGGAAACATAGT |
| VirC_F | AGCTCAAACCTGCTTCTG |
| VirC_R | ATCATTTGTAGCGACTTG |
| OsCBSX3_F | ATGGCCTGCATCAACACATTCC |
| OsCBSX3(-sp)_F | GCATCGACGTTTCATGGAC |
| OsCBSX3_R | AACCTCCAGCAGGGCTTCC |
| OsCBSX3_Screening_F | ATCGACGAGAACCCTGAAGC |
| OsCBSX3_Screening_R | TCTTTGACCCATGAGAAGCCC |
| OsTrx f_F | ATGGCGCTACGCCTCTCC |
| OsTrx f(-sp)_F | CAGGTGACGGAGGTGAACAAG |
| OsTrx f_R | TCAGCTTGACTTCACCGTCT |
| OsTrx m_F | ATGGCCTCCGCCCTCGCC |
| OsTrx m(-sp)_F | ACATCCATCCAAGTTCCAG |
| OsTrx m_R | TTACCTCTCGACAAATTTCTC |
| OsTrx y_F | ATGGCGGCCTTCACCTCCAC |
| OsTrx y_R | CTACTTTGCAACCTCCAAAGC |
| OsTrx o_F | ATGGCTCTGGCTCACCGACT |
| OsTrx o(-sp)_F | GGTGATTCCAGCATGGT |
| OsTrx o_R | TTACTTGTGAAGACTTTCCATG |
| OsTrx 1_F | ATGGGGTCTCTGTGTGGGA |
| OsTrx 1_R | TCAGGCGCTGTCAGCAAG |
| OsTrx 10_F | ATGGGCTCCTTCTTCTCGAC |
| OsTrx 10_R | TTAGGATGAGGAGGATGAAATG |

| | |
|-----------------------------|----------------------------|
| OsTrx 23_F | ATGGCCGCCGAGGAGGGAGT |
| OsTrx 23_R | TTAGGCAGAAGCAGATGCAG |
| Spermidine synthase_F | ATGGAGGGTGGAGGCGCAA |
| Spermidine synthase_R | CTAGGAAGCAGTCGGAATGGC |
| Initiation factor 2(-sp)_F | TGTTCCGTAGCACCAGATATTCT |
| Initiation factor 2_R | TCACTGCATGGTAGAGTACA |
| Glycosyl transferase(-sp)_F | TTCGACCGCTCTCAATCCG |
| Glycosyl transferase_R | TCAGGGAATTATGTTGCACTCA |
| Retrotransposon_F | ATGAGCAAAACAATTGGTAACTTCA |
| Retrotransposon_R | TCAGAGGTAATAGTAGGATCCTGCA |
| AtTrx m(-sp)_F | TGTGAAGCTCAGGAACTACTACC |
| AtTrx m_R | TCATGGCAAGAAGCTGTGCGAG |
| AtTrx o(-sp)_F | TCCACTCTCCGTTCCCTCGAC |
| AtTrx o_R | TCACTTGTAGAGCTGTTCCATGA |
| AtTrx y(-sp)_F | ATTGAAGCCAAGAAGCAGACA |
| AtTrx y_R | CTATGGCTTCACTTTTAGAGAATCCT |
| OsCBSCLC6_F | ATGGCGTGCACGGCTCCACG |
| OsCBSCLC6_(Δ1717)_F | CCGCACCTCGACGGCCAC |
| OsCBSCLC6_(Δ1753)_F | CAGCTCACCGTCGGCGAC |
| OsCBSCLC6_R | TCAGCCGCTGTTCGTTGGCTA |
| OsCBSCLC6_911b_Screening_F | CAAGGGTGGGCTGATCATGT |
| OsCBSCLC6_1239b_Screening_R | GAATCGGTTCAAGCAATGGC |
| OsCBSCLC6_1074b_F | GATCAACGAGAGGGGCAACGC |
| OsCBSCLC6_1641b_F | GCTGCTCATCTCCAAGACCG |
| OsCBSCLC6_1284b_R | GCTGGCCAAGTCGTTGTAGTG |

Publications:

1. Kumar R, **Subba A**, Kaur C, Ariyadasa TU, Sharan A, Pareek A, Sopory SK, and Singla-Pareek SL. OsCBSCBSP4 is a Two Cystathionine- β -Synthase Domain-containing Protein from Rice that Functions in Abiotic Stress Tolerance. *Curr Genomics*. **2018** Jan;19(1):50-59. doi: 10.2174/1389202918666170228141706. PMID: 29491732; PMCID: PMC5817877.
2. **Subba A**, Tomar S, Pareek A, and Singla-Pareek SL. The chloride channels: Silently serving the plants. *Physiol Plant*. **2021** Apr;171(4):688-702. doi: 10.1111/ppl.13240. Epub 2020 Nov 5. PMID: 33034380.

Posters:

1. **Subba A**, Tomar S, and Singla-Pareek SL. Deciphering the role of *OsCBSX3*, a CBS domain containing protein of rice. **ICGEB India Symposium #ISIN2019**. International Centre for Genetic Engineering and Biotechnology, New Delhi 110067, India. **2019**.
2. **Subba A**, Tomar S, Pareek A, and Singla-Pareek SL. Insights into the function of CBSX3 in rice. **Workshop on Plant Stress Biology and Food Security**. International Centre for Genetic Engineering and Biotechnology, New Delhi 110067, India. **2019**.
3. Tomar S, **Subba A**, Pareek A, and Singla-Pareek SL. Exploring CBS domain containing proteins – From stress tolerance to architecture in plants. **4th International Plant Physiology Congress**, CSIR-NBRI, Lucknow, **2018**.
4. **Subba A**, Tomar S, and Singla-Pareek SL. Functionally annotating “Unknowneome” of rice. **86th Annual Conference of Society of Biological Chemists (SBC-2017)**, JNU, New Delhi, **2017**.
5. **Subba A**, Ariyadasa TU, Kaur C, Pareek A, and Singla-Pareek SL. *OsCBSCBS3* is a stress-responsive ATP and/or AMP binding protein involved in abiotic stress adaptation in rice. **3rd International Plant Physiology Congress**, JNU, New Delhi, **2015**.

RESEARCH ARTICLE

OsCBSCBSPB4 is a Two Cystathionine- β -Synthase Domain-containing Protein from Rice that Functions in Abiotic Stress Tolerance

Ritesh Kumar¹, Ashish Subba¹, Charanpreet Kaur^{1,2}, Thilini U. Ariyadasa¹, Ashutosh Sharan², Ashwani Pareek², Sudhir K. Sopory¹ and Sneha L. Singla-Pareek^{1,*}

¹Plant Stress Biology Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, India;

²School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

Abstract: Cystathionine β -synthase (CBS) domains have been identified in a wide range of proteins of unrelated functions such as, metabolic enzymes, kinases and channels, and usually occur as tandem repeats, often in combination with other domains. In plants, CBS Domain-Containing Proteins (CDCPs) form a multi-gene family and only a few are so far been reported to have a role in development *via* regulation of thioredoxin system as well as in abiotic and biotic stress response. However, the function of majority of CDCPs still remains to be elucidated in plants. Here, we report the cloning, characterization and functional validation of a CBS domain containing protein, OsCBSCBSPB4 from rice, which possesses two CBS domains and one PB1 domain. We show that *OsCBSCBSPB4* encodes a nucleo-cytoplasmic protein whose expression is induced in response to various abiotic stress conditions in salt-sensitive IR64 and salt-tolerant Pokkali rice cultivars. Further, heterologous expression of *OsCBSCBSPB4* in *E. coli* and tobacco confers marked tolerance against various abiotic stresses. Transgenic tobacco seedlings over-expressing OsCBSCBSPB4 were found to exhibit better growth in terms of delayed leaf senescence, profuse root growth and increased biomass in contrast to the wild-type seedlings when subjected to salinity, dehydration, oxidative and extreme temperature treatments. Yeast-two hybrid studies revealed that OsCBSCBSPB4 interacts with various proteins. Of these, some are known to be involved in abiotic stress tolerance. Our results suggest that OsCBSCBSPB4 is involved in abiotic stress response and is a potential candidate for raising multiple abiotic stress tolerant plants.

ARTICLE HISTORY

Received: May 31, 2016
Revised: October 05, 2016
Accepted: October 20, 2016

DOI:
10.2174/1389202918666170228141706

Keywords: Cystathionine β -synthase (CBS) domain, Phox/Bemp1 (PB1) domain, Abiotic stress, Stress tolerance, Transgenic plants.

1. INTRODUCTION

Identified for the first time in archaeobacteria *Methanococcus janaschii* by Alexander Bateman [1], the Cystathionine β -Synthase (CBS) domain or the Bateman domain has been thereafter discerned in wide range of proteins in all kingdoms of life. This domain comprises of ~60 amino acids and is present in either pairs or quads in proteins, with each pair forming a tight association, referred to as CBS pair or Bateman module. They may exist either as a lone module (*e.g.* OsCBSX3) or fused to other diverse domains (*e.g.* OsCBSCCL6) in the protein [2, 3]. In humans, many hereditary diseases have been linked to mutations in the CBS domain of various proteins such as homocystinuria, caused by mutation in cystathionine beta synthase [4] and retinitis pigmentosa, caused by mutation in IMPDH [5], thereby emphasizing the significant role of this domain in the living systems. CBS domains are known to bind specific nucleotides (mostly AMP) and form energy sensing modules which

either activate or inhibit the other associated or interacting domains of various proteins [3, 6]. However, the precise role and regulation of proteins harboring this domain is still concealed, especially in plant systems.

The studies on CBS domain-containing proteins (CDCPs) have been initiated only recently in plants. In an effort to improve the stress tolerance in plants, Kumari *et al.* [7] analyzed the differential regulation of salinity stress-responsive genes among salt-tolerant (Pokkali) and salt-sensitive (IR64) genotypes of rice by subtractive cDNA approach wherein one of the potential genes was found to encode CBS domain-containing protein. The transcript accumulation of this hypothetical gene was differentially regulated in the contrasting genotypes of rice, indicating towards their probable role in salinity tolerance. This observation encouraged us to address the question ‘whether CBS Domain Containing Proteins (or CDCPs) are correlated to abiotic stress tolerance in plants?’

Previously we have identified 34 CDCPs in *Arabidopsis thaliana* and 59 in *Oryza sativa*. Our *in silico* expression analysis clearly indicated a potential role of some CDCPs in stress tolerance [2]. In this context, we have reported that

*Address correspondence to this author at the Plant Stress Biology Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, India; Tel: +91-11-26742357; Fax: +91-11-26742316; E-mail: sneh@icgeb.res.in

OsCBSX4, a single CBS domain containing protein from rice when over-expressed, imparts salinity, oxidative and heavy metal tolerance to transgenic tobacco plants [8]. Further in Arabidopsis, single CBS domain-containing protein, AtCBSX1 has also been reported to maintain cellular redox homeostasis via thioredoxin systems in response to changes in ATP:AMP ratio [9]. Recently, the role of CDCPs has also been indicated in resistance to *Magnaporthe oryzae* in rice [10] and tolerance to low nitrogen stress in soybean [11]. These reports indicate that CDCPs might play an important role in various cellular processes in plants.

In this study, we have characterized and functionally validated OsCBSCBSPB4, a CDCP containing two pairs of CBS domains and a Phox/Bemp1 (PB1) domain. *OsCBSCBSPB4* is specifically induced in response to salinity, oxidative and extreme temperature stresses in salt-sensitive IR64 and salt-tolerant Pokkali cultivars of rice. Our results show that over-expression of OsCBSCBSPB4 in tobacco enhances multiple abiotic stress tolerance, thereby suggesting an important role of this protein in plant stress response.

2. MATERIALS AND METHODS

2.1. Cloning and Sequence Analysis of OsCBSCBSPB4

The coding region of *OsCBSCBSPB4* (LOC_Os12g07190, RGAP 7 database) was amplified as 1,629 bp fragment from cDNA, prepared from salt-tolerant Pokkali rice. The amplicon was then cloned in TOPO-TA vector (Invitrogen) and sequenced (Macrogen, Korea).

ScanProsite tool [12] was used for analyzing the domain organization of OsCBSCBSPB4. For homology analysis, BLAST search was conducted using GenBank. Multiple sequence alignment was performed using Clustal W2 [13]. Neighbour joining method [14] was used to generate unrooted phylogenetic tree for different CBSCBSPB4 domain-containing proteins reported in various organisms using the MEGA7 software [15].

2.2. Stress Treatments

Seeds of IR64 and Pokkali rice were germinated hydroponically and grown at 28 \pm 1 $^{\circ}$ C. For salinity stress, 14-day-old rice seedlings of both the cultivars were transferred to 200 mM NaCl solution. For desiccation stress, seedlings were air dried on blotting paper, whereas, exposing the seedlings to 4 $^{\circ}$ C and 42 $^{\circ}$ C provided low and high temperature stresses, respectively. For oxidative stress, the seedlings were kept in 1 μ M methyl viologen (MV) solution. Seedlings grown in water for the same period were taken as control. The shoot tissues were harvested after 6 h of the stress treatments.

2.3. RNA Isolation and qRT-PCR

Total RNA was extracted using RaFlexTM (GeNei, India) from the two week old IR64 and Pokkali rice seedlings exposed to different abiotic stresses (cold, salinity, heat, MV and desiccation) as per the manufacturer's protocol. For preparation of cDNA, approximately 5 μ g of total RNA from the stressed and non-stressed samples was reverse transcribed using oligo (dT) primer and the first strand cDNA

Synthesis Kit (Fermentas, Life Sciences). Primers for real-time PCR were designed using Mac Vector 8.0 software. The reaction was performed using StepOneTM Real-Time PCR System (Applied Biosystems). eEF-1 α was used as the internal control. Three technical replicates were analyzed for each sample. The relative expression ratio of *OsCBSCBSPB4* was calculated using comparative Ct value method [16]. Heatmaps were generated using MeV software [17].

2.4. Heterologous Expression of OsCBSCBSPB4 Protein

For expression in *E. coli*, the coding region was amplified from the cloned cDNA for *OsCBSCBSPB4* in TOPO-TA vector using forward primer containing *EcoRI* site: 5' GGAATTCATGGTTCAAGGTAATTTAGAC 3' and reverse primer containing *XhoI* site: 5' CCGCTCGAGTCACTTTAGCTCGTTTCAG 3' and cloned into pET28a vector (Novagen) at *EcoRI* / *XhoI* sites to create pET28a-OsCBSCBSPB4 plasmid. The resulting plasmid was then transformed into BL21 (DE3) *E. coli* cells. Expression of OsCBSCBSPB4 protein was induced using 0.5 mM IPTG at 37 $^{\circ}$ C and cells were harvested after 4 h. To check the expression of OsCBSCBSPB4 protein, the cells were lysed and protein concentration was determined by Bradford's method [18]. The protein was then analyzed on 12% SDS PAGE and visualized by Coomassie staining.

2.5. Preparation of Polyclonal Antibodies Against OsCBSCBSPB4

For raising anti-OsCBSCBSPB4 antibodies, essentially same protocol was followed as described by Singh *et al.* [8]. In brief, the *OsCBSCBSPB4* cDNA was cloned in pET28a and overexpressed in *E. coli* (BL21-DE3) following which the recombinant protein was purified using Ni-NTA agarose as per the manufacturer's protocol (Qiagen, Germany) and used to raise polyclonal antibodies in rabbit.

2.6. Assessing Growth Pattern of OsCBSCBSPB4-transformed *E. coli* Under Various Abiotic Stresses

E. coli BL21 (DE3) cells transformed with either pET28a plasmid or pET28a-OsCBSCBSPB4 construct were grown in liquid LB medium at 37 $^{\circ}$ C (till OD₆₀₀ ~0.4) followed by induction with 0.5 mM IPTG and further grown for 13 h. At the time of IPTG induction, for salinity stress, NaCl was added to the growth medium to a final concentration of 200 mM. Likewise, for high and low temperature treatments, the bacterial cells were transferred to respective 42 $^{\circ}$ C and 15 $^{\circ}$ C, and for oxidative and dehydration treatments, 1 mM MV and 10% PEG 8000 were respectively added to the growth medium. Cell aliquots (1 mL) were taken after every 1 h and cell survival was estimated by measuring the absorbance at 600 nm. The O.D. represents the mean of three replicates of at least two independent recombinant bacterial cultures.

2.7. Preparation of OsCBSCBSPB4 Construct, Plant Transformation and Transgenic Screening

For plant transformation, *OsCBSCBSPB4* was initially cloned in pRT101 at *BamHI*/*EcoRI* sites. The 35S CaMV promoter-OsCBSCBSPB4 gene cassette was then excised from pRT101 and cloned in pGREEN0029 vector which carries an *NPT* (kanamycin) gene as the selectable marker.

Tobacco (*Nicotiana tabacum* cv. *petit havana*) leaf discs were transformed with *Agrobacterium tumefaciens* (LBA4404 strain) cells harboring pGREEN-OsCBSCBSPB4 construct following the procedure described previously [19]. Putative T₀ transgenic plants were regenerated from independent calli in the presence of kanamycin and further screened by PCR and Southern blotting. For PCR based screening method, genomic DNA from wild-type (WT) and transgenic lines was used as template. Forward primer was designed from 35S CaMV promoter sequence and *OsCBSCBSPB4*-specific sequence was used as reverse primer. For Southern hybridization, 15 µg of genomic DNA from PCR positive lines was digested with *EcoRI* and *BamHI*, blotted and probed using α^{32} P-dCTP labeled *OsCBSCBSPB4* gene, as described [8]. In addition, anti-OsCBSCBSPB4 antibodies were used to confirm OsCBSCBSPB4 protein expression in the transgenic plants through western blotting as described by Singla-Pareek *et al.* [20].

The seeds from the transgenic plants were germinated on kanamycin-containing medium to select transgenic T₁ seedlings. The positive transgenic plants were used for further growth analysis.

2.8. Leaf Disc Assay

Leaf discs of 1.0 cm diameter were excised from healthy and fully expanded tobacco leaves of similar age from the transgenic and WT plants (60 day old). The discs were floated in 6 mL solution of 200 mM NaCl (for stress) or water (for control) for 5 days. The treatment was carried out at 25°C and the experiment was repeated at least three times with different transgenic lines.

2.9. Stress-survival Assays of T₁ Transgenic Tobacco Plants

Seeds of two independent transgenic tobacco lines (line 3 and 10) over-expressing rice *OsCBSCBSPB4* gene were selected by germinating the seeds on kanamycin containing half-strength MS medium selection plates. After 7 days of germination, the seedlings were transferred to medium supplemented with 1 µM MV, 5% mannitol and 200 mM NaCl for oxidative, dehydration and salt stress, respectively. Seedlings were exposed to 4°C to provide low temperature stress and 42°C for high temperature stress. After 24 h of heat stress and 3 d of cold stress, plants were kept at 26±1°C for recovery.

2.10. Yeast Two-Hybrid Assay

The yeast two-hybrid screening was performed as described earlier with some modifications [21]. The *OsCBSCBSPB4* coding region was cloned in-frame at *EcoRI* site of pBD-GAL4 vector to make the bait plasmid pBD-OsCBSCBSPB4. Plasmid DNA of bait and prey (containing rice cDNA library) constructs were transformed into the *Saccharomyces cerevisiae* strain AH109. The transformants were grown on synthetic minimal medium plates lacking tryptophan, uracil, and leucine and adenine (4-DO). Colonies which were able to grow on 4-DO plates were then streaked on YPD medium plates and transferred to a filter paper to carry out the filter lift assay. After transfer, positive clones

were identified using X-gal based screening and plasmid was isolated from these positive clones, followed by sequencing to identify the putative interacting partners.

2.11. Sub-cellular Localization of OsCBSCBSPB4

OsCBSCBSPB4 was cloned in pMBP vector as translational fusion with GFP at the C-terminus under the control of 35S CaMV promoter at *BamHI* and *XbaI* site. Particle bombardment was used to introduce OsCBSCBSPB4:GFP fusion plasmid into the onion epidermal cells with a Biolistic PDS-1000/He system (BioRad) as described by Singh *et al.* [8]. The transformed onion epidermis was incubated at 28°C in dark for 16 h and GFP was detected in cells under a confocal microscope. For staining the nucleus, onion epidermal peels were incubated for 5 min with 100 nM of 4',6-diamidino-2-phenylindole (DAPI) (Invitrogen, Eugene, OR) prior to microscopic analysis. Cellular structure was visualized using bright-field optics.

3. RESULTS

3.1. Sequence Analysis of OsCBSCBSPB4

OsCBSCBSPB4 (LOC_Os12g07190) was PCR amplified as 1629 bp fragment from salt tolerant rice "Pokkali" cDNA and showed 100% sequence identity to *O. sativa* ssp. *japonica*. *OsCBSCBSPB4* is located on chromosome 12 and encodes a protein of 542 amino acids with a molecular mass of 58.3 kDa and a pI of 6.95. Domain analysis of the deduced amino acid sequence using ScanProsite database [12] confirmed the presence of two pairs of CBS domains, first pair lying in the region from 60-181 amino acids and the second pair being present within 227-354 amino acid residues (Fig. 1a). Along with CBS domains, a Phox/Bem1 (PB1) domain was also predicted in the OsCBSCBSPB4 protein at 407-493 amino acid residues.

Multiple sequence alignment revealed high amino acid sequence identity of OsCBSCBSPB4 with corresponding proteins from other species (Fig. 1b). OsCBSCBSPB4 shared 60% amino acid identity with AtCBSCBSPB4 from *Arabidopsis* and exhibited respectively 68%, 84%, 85% and 86% identity with corresponding proteins from *Nicotiana tabacum*, *Hordeum vulgare*, *Zea mays* and *Sorghum bicolor* (Fig. 1b). Furthermore, the phylogenetic tree revealed a closer relationship of OsCBSCBSPB4 with other grass family members than *Arabidopsis* (Fig. 1c).

3.2. OsCBSCBSPB4 is a Multiple-stress Inducible Gene

Quantitative Real time PCR (qRT-PCR) studies were carried out to determine the expression profile of *OsCBSCBSPB4* in the salt-sensitive IR64 and salt-tolerant Pokkali rice varieties in response to different abiotic stress conditions. For this, two-week old IR64 and Pokkali seedlings were subjected to different stress treatments *viz.* salinity, desiccation, oxidative, heat and cold stress for 6 h.

Under control (non-stressed) conditions, the expression of *OsCBSCBSPB4* was found to be slightly higher in Pokkali (1.26-fold) than in IR64. However, *OsCBSCBSPB4* expression increased several folds in both IR64 and Pokkali under stress conditions. In Pokkali rice, a 3.1-fold increase in *OsCBSCBSPB4* expression was observed after

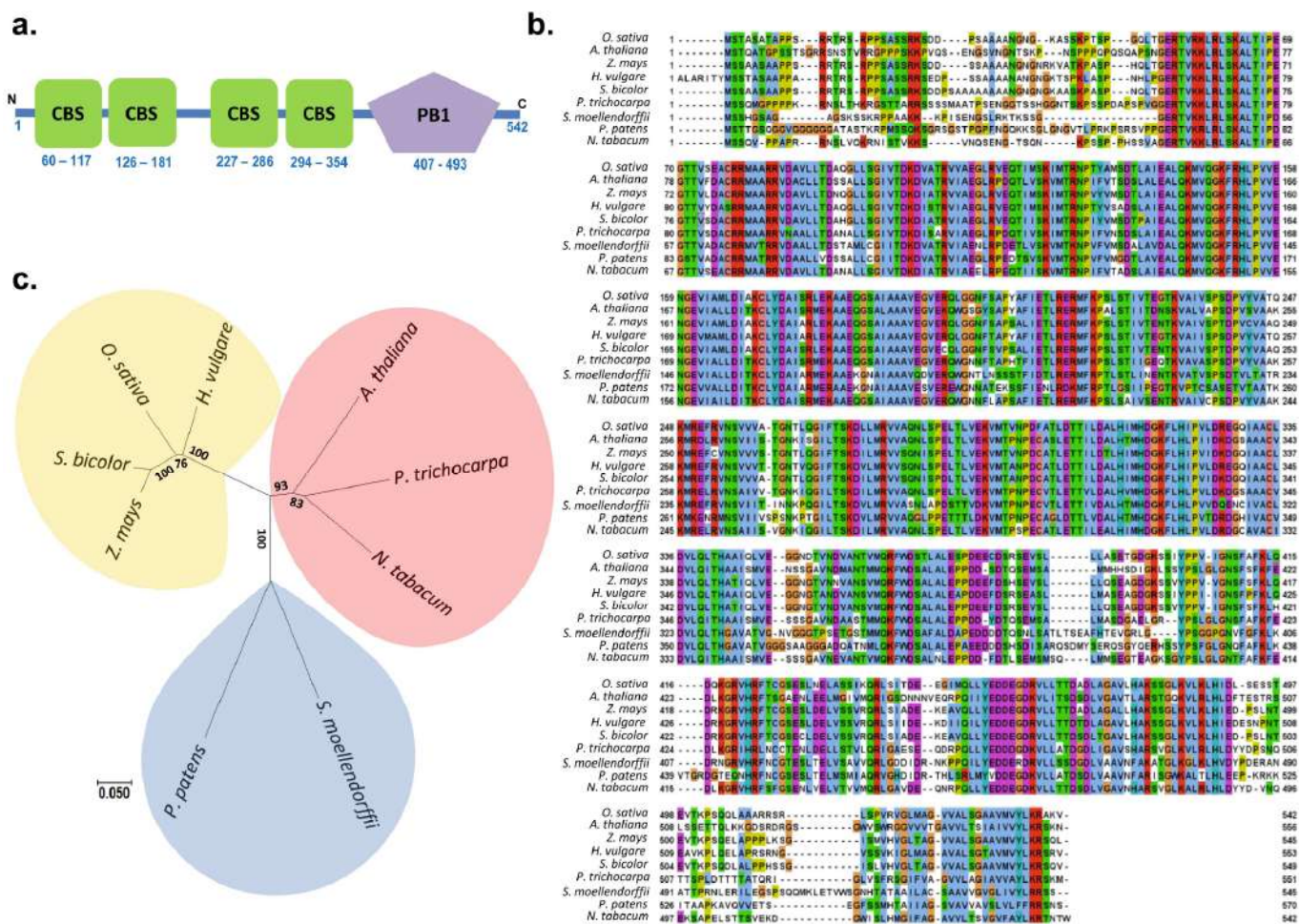


Fig. (1). Sequence analysis and comparison of *OsCBCSGBP4* with corresponding proteins from other species. (a) Schematic depiction of domains present in *OsCBCSGBP4* protein as determined by ScanProsite. Two pairs of CBS domains along with single PB1 domain have been indicated. (b) Multiple alignment of *OsCBCSGBP4* with other reported CBSGBP4 proteins. Clustal W2 was used for multiple alignment and Jalview was used to view the alignment. (c) The phylogenetic relation between CBSGBP4 domain containing proteins from different species. The evolutionary history was inferred using the Neighbor-Joining method. Evolutionary analysis was conducted in MEGA7.

6 h of NaCl (200 mM) treatment. Oxidative stress and low temperature treatment also induced *OsCBCSGBP4* expression to 3.7- and 3.3-fold, respectively. However, desiccation treatment led to only a marginal increase in *OsCBCSGBP4* transcript levels. Highest induction, up to 4.3-fold in transcript levels could be seen in response to heat treatment (Fig. 2 a,b).

Likewise in IR64 rice seedlings, *OsCBCSGBP4* expression was also induced in response to heat (7.7-fold), cold (1.6-fold), oxidative (1.8-fold) and salinity stress (2.6-fold) (Fig. 2 a,b). However, *OsCBCSGBP4* expression declined marginally after desiccation stress in IR64. Thus, *OsCBCSGBP4* was found to be highly stress-inducible gene being induced to different levels in IR64 and Pokkali rice genotypes, as can also be visualized through the heatmap (Fig. 2b).

3.3. Over-expression of *OsCBCSGBP4* in *E. coli* Enhances Multiple Abiotic Stress Tolerance

To get an insight into the role of *OsCBCSGBP4* in stress response, *OsCBCSGBP4* cDNA was cloned in pET28a vector to generate pET28a-*OsCBCSGBP4* construct followed

by transformation in *E. coli* (BL21-DE3). The expression of the recombinant *OsCBCSGBP4* protein in *E. coli* was confirmed by SDS-PAGE. A band corresponding to *OsCBCSGBP4* protein (~58 kDa) could be observed in the induced sample compared to the uninduced one (Fig. 3a). For stress treatments, *E. coli* cells were grown in liquid medium containing different stress-inducing agents after IPTG induction, such as 200 mM NaCl for salinity, 1 mM MV for oxidative and 10% PEG for dehydration stress. For high and low temperature treatments, bacterial cells were transferred to 42°C and 15°C, respectively immediately after IPTG induction. The survival rate of the cells (expressed as bacterial O.D.) upon treatment with different stresses is shown in (Fig. 3b). The bacterial cells transformed with pET28a plasmid (or empty vector) were used as control and these cells exhibited much slower growth rate under tested stress conditions. In contrast, *OsCBCSGBP4*-transformed bacterial cells possessed higher growth rates even after stress imposition and could continuously grow till 9 – 10 h before reaching a plateau (Fig. 3b). However, bacterial cells grown at 15°C exhibited slower growth when compared to that under other stresses but the growth rate was still more when compared to the empty vector-transformed bacterial cells. Thus, comparison of

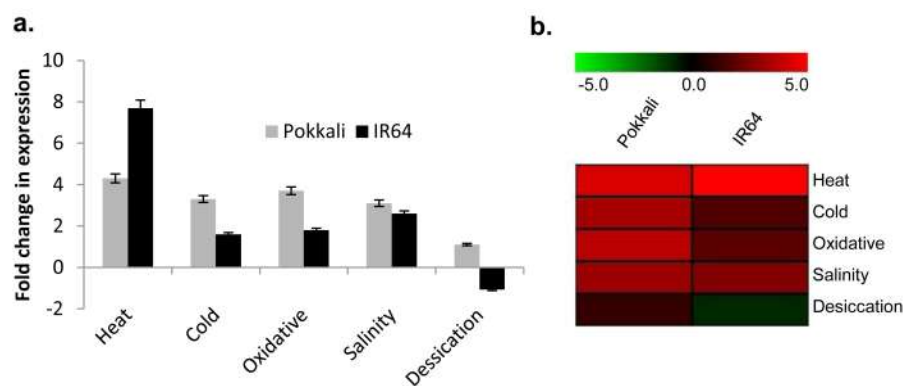


Fig. (2). Transcript profile of *OsCBSCBSPB4* in shoots of salt-sensitive (IR64) and salt-tolerant (Pokkali) rice cultivars. **(a)** Fold change in transcript levels (expression level relative to untreated control plants) of *OsCBSCBSPB4* in two-week old IR64 and Pokkali seedlings subjected to 6 h of salinity (NaCl 200 mM), oxidative (methyl viologen 1 μ M), desiccation (air drying) and high (42°C) and low (4°C) temperature stresses. The experiment was repeated twice with three replicates in each case. **(b)** Corresponding heat map of the *OsCBSCBSPB4* transcript in IR64 and Pokkali. Scale bar at the top represents low (green), intermediate (black) and high (red) expression levels.

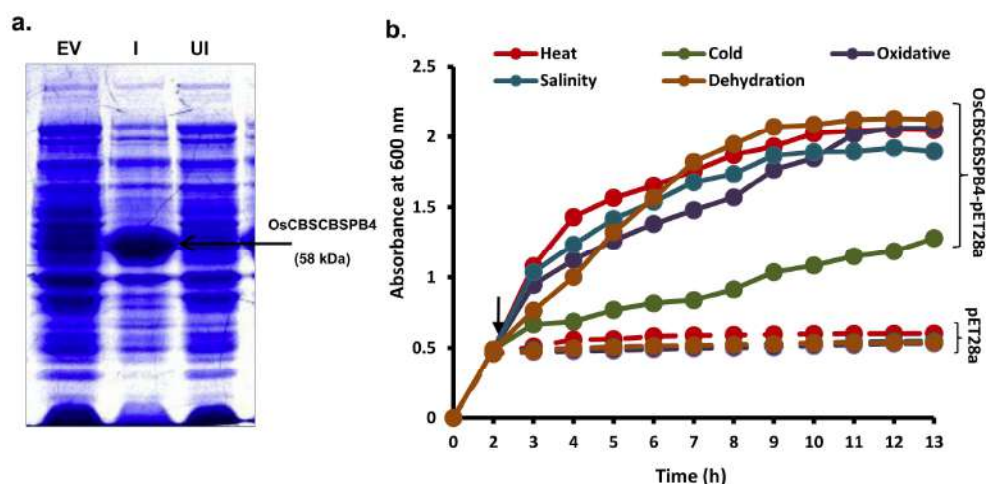


Fig. (3). Heterologous expression of *OsCBSCBSPB4* in *E. coli* and assessing the response to various abiotic stresses. **(a)** Coomassie stained SDS-PAGE showing crude protein extract from *E. coli* transformed with empty vector (EV), or *OsCBSCBSPB4* where 'UI' indicates uninduced protein and 'I' indicates protein induced with 0.5 mM IPTG. The induced band corresponding to *OsCBSCBSPB4* after 4 h of IPTG induction is visualized (marked by arrow). **(b)** Survival rate of *E. coli* BL21 (DE3) transformants in LB broth in response to various abiotic stress inducing agents such as, 200 mM NaCl for salinity, 10% PEG 8000 for dehydration, 1 mM MV for oxidative, 42°C for high temperature and 15°C for low temperature stress. pET28a-*OsCBSCBSPB4* (represented by solid lines) and the empty pET28a vector (taken as control; represented by dotted lines) were grown at 37°C for 2 h followed by 0.5 mM IPTG induction and stress treatments. Cell aliquots (1 mL) were taken after every 1 h. Cell survival was estimated by monitoring absorbance (O.D.) at 600 nm. O.D. represents the mean of three replicates of at least two independent recombinant bacterial cultures. Time point of addition of IPTG and stress imposition is marked by arrow.

growth pattern of *OsCBSCBSPB4*-transformed bacterial cells under stress conditions against pET28a-transformed cells revealed an important role of *OsCBSCBSPB4* in cellular response to abiotic stress.

3.4. The *OsCBSCBSPB4* protein is Localized in the Nucleus and Cytoplasm

To determine the sub-cellular localization of *OsCBSCBSPB4*, fusion construct of corresponding cDNA was fused in frame with GFP at C-terminus to give *OsCBSCBSPB4*:GFP fusion construct driven by the 35S CaMV promoter. The resulting chimeric protein (*OsCBSCBSPB4*:GFP) was transiently expressed in onion peel epidermal cells. Expression of *OsCBSCBSPB4*:GFP fusion protein was analyzed by confocal microscopy. As

shown in the (Fig. 4), *OsCBSCBSPB4*:GFP fusion protein was found to be localized majorly in the nucleus though some expression could be observed in the cytoplasm as well (Fig. 4a). On the other hand, free GFP expression could be seen in the whole cell (Fig. 4b). However, online nuclear localization prediction tools such as NLS mapper, did not predict any strong Nuclear Localization Signal (NLS) sequence in the *OsCBSCBSPB4* protein.

3.5. Over-expression of *OsCBSCBSPB4* in Transgenic Tobacco Confers Tolerance Towards Salinity Stress

To determine the functional significance of *OsCBSCBSPB4* in *planta*, full-length *OsCBSCBSPB4* cDNA was cloned in the plant transformation vector pGREEN via the shuttle vector pRT101 (Fig. 5a). The re-

sulting recombinant plasmid harboring *OsCBSCBSPB4* gene was transformed into tobacco *via Agrobacterium*-mediated transformation. The kanamycin resistant tobacco seedlings obtained after transformation were then screened for transgene by PCR (Fig. 5b). The morphological and growth characteristics of T₀ plants were found to be similar to the untransformed WT plants under control conditions. For subsequent analysis, ten T₀ transgenic lines were grown to maturity. A total of five independent PCR positive lines were confirmed by Southern hybridization for stable integration of the transgene. A band corresponding to 1.6 kb fragment could be observed in the transgenic plants but not in the wild-type (Fig. 5c). Further, the transgenic lines were checked for the expression of OsCBSCBSPB4 protein by western blot using anti-OsCBSCBSPB4 specific antibodies. The expression of OsCBSCBSPB4 protein in the transgenic lines could be confirmed through western blot (Fig. 5d).

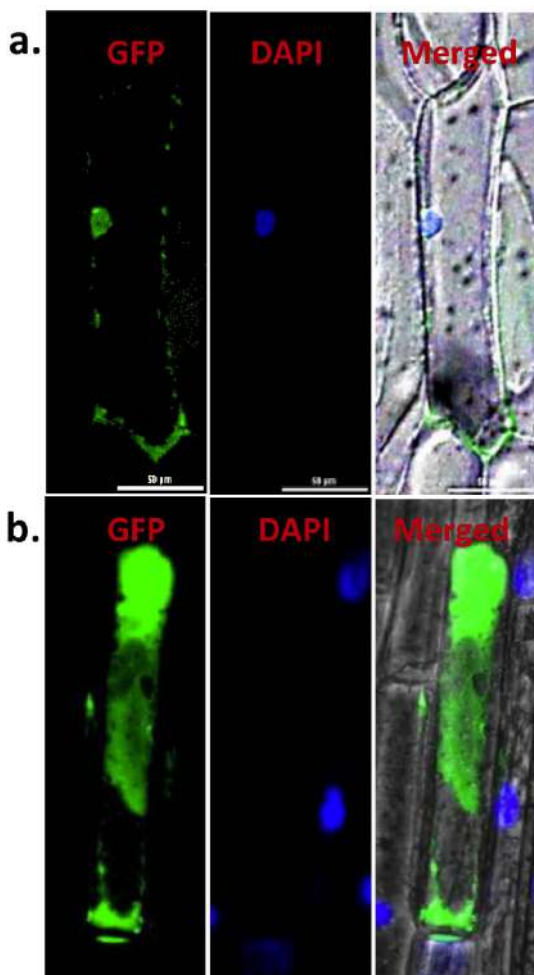


Fig. (4). OsCBSCBSPB4 localizes to both nucleus and cytoplasm. Onion peel epidermal cells were transiently transformed with constructs containing either (a) OsCBSCBSPB4:GFP or (b) GFP alone by particle bombardment method. Subcellular localization of OsCBSCBSPB4:GFP fusion protein or GFP alone was viewed under confocal microscope, 16 h after bombardment. DAPI was used to detect nuclei.

The ability of transgenic plants (T₀) to tolerate high levels of salinity stress was evaluated using leaf disc bioassay. The leaf discs from Southern positive transgenic lines,

namely, line 3 and 10 along with WT plants were floated onto 200 mM NaCl solution. Visual differences in the 'greenness' were observed after 5 days (Fig. 5e).

Bleaching of leaf tissues was also analyzed by measuring the remaining chlorophyll content of the leaf discs after the stress treatment (Fig. 5f). Under control conditions, no significant difference in chlorophyll content of the leaf discs could be noticed between WT and transgenic lines. However under salt stress, chlorophyll content drastically reduced by 90% in the leaf discs of WT plants. While in transgenic lines 3 and 10, a less significant decrease (by 40% for line 3 and 30% for line 10) was observed. Importantly, measured differences in the chlorophyll content of transgenic and WT lines were in conformity with the observed differences in 'greenness' in these plants.

3.6. OsCBSCBSPB4 Over-expressing Tobacco Plants Show Tolerance Towards Various Abiotic Stress Conditions

To further evaluate the role of *OsCBSCBSPB4* in stress tolerance, the T₁ transgenic seedlings were analyzed for stress response. In order to carry out systematic phenotypic comparison, WT and T₁ transgenic plants were germinated on half-strength MS plates and later transferred to medium supplemented with different stress inducers for salinity, dehydration, oxidative, heat and cold stress treatments. Under non-stress conditions, WT and transgenic plants exhibited no morphological differences in growth (Fig. 6a). However, in response to the applied stress conditions, the growth of WT plants was severely retarded (Fig. 6b-f). In contrast, OsCBSCBSPB4 over-expressing T₁ transgenic plants exhibited better growth under stress conditions (Fig. 6b-f).

Under salt stress, the root length was greatly reduced in the WT plants showing 70% decline, whereas both the transgenic lines exhibited only 20% reduction in root length under salinity conditions (Fig. 6g). Since *OsCBSCBSPB4* transcript levels were significantly increased under other stresses such as, high/low temperature, desiccation and oxidative stress, seedlings were also subjected to these stress treatments and the response was evaluated. The results indicated a response similar to that observed under salinity stress for both WT and T₁ transgenic lines under these stress conditions. Under all tested stress conditions, the reduction in shoot weight and root length was less in transgenic seedlings as compared to the WT seedlings. The root length was reduced by 23%, 25%, 28% and 24% under low temperature, high temperature, oxidative and dehydration stress, respectively in transgenic seedlings as compared to 70%, 68%, 71% and 67% reduction in WT plants (Fig. 6g). Similarly, the reduction in shoot weight in response to applied stress conditions was also less in the transgenic seedlings as compared to the WT seedlings (Fig. 6h).

3.7. Identification of Potential Interacting Partners of OsCBSCBSPB4

In order to identify the potential interacting partners of OsCBSCBSPB4, the yeast two-hybrid approach was used. We isolated and identified a total of 100 clones, out of which a few have been listed in (Table 1). These included transcription factors, such as RING-H2 finger protein and also the

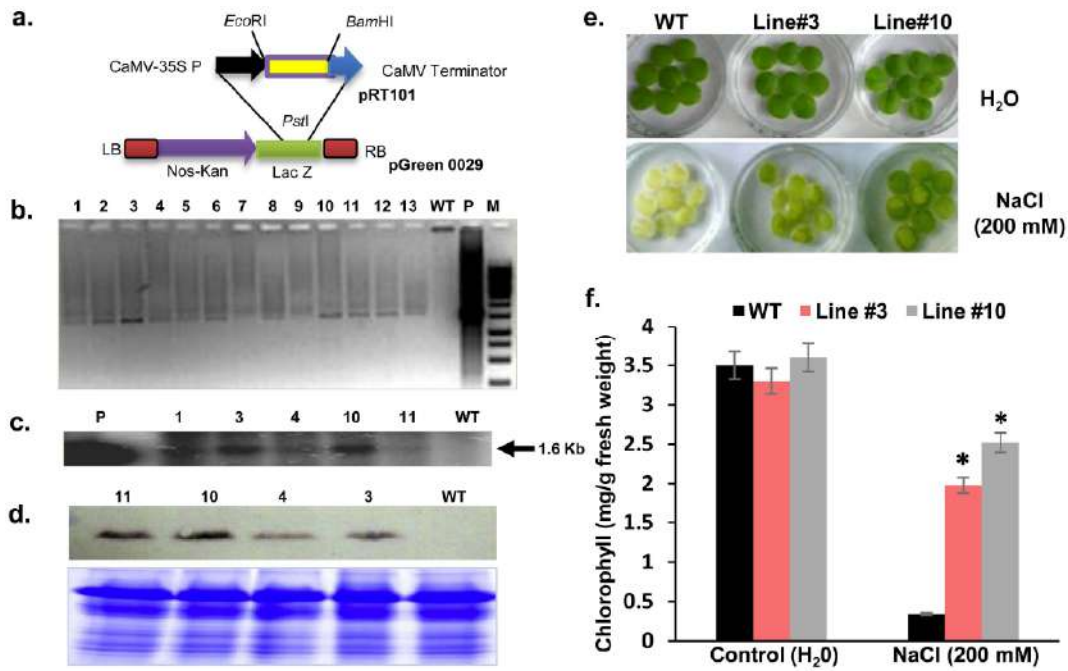


Fig. (5). Overexpression of *OsCBSCBSPB4* in transgenic tobacco confers salinity stress tolerance. (a) Schematic representation of *OsCBSCBSPB4*-pGREEN0029 construct used for tobacco transformation. (b) Confirmation of transgenic lines for the presence of *OsCBSCBSPB4* transgene by PCR (1-13, putative transgenic lines; WT, wild-type plant). (c) Southern blot analysis showing a band corresponding to 1.6 kb (marked by an arrow). Genomic DNA extracted from PCR lines (1, 3, 4, 10 and 11) was used for analysis along with WT. 'P' indicates positive control (taken as *OsCBSCBSPB4* amplicon). (d) Western blot analysis for detection of *OsCBSCBSPB4* protein in the transgenic tobacco lines using anti-*OsCBSCBSPB4* antibodies; corresponding coomassie-stained protein gel has been included below to show equal loading of protein. (e) Leaf disc senescence assay of wild-type (WT) and transgenic lines (line 3 and line 10) under 200 mM NaCl stress (5 days). Leaf discs floated on water served as control. (f) Chlorophyll content (mg/g fresh weight) of the corresponding leaf discs after 5 d of 200 mM NaCl stress. Experiment was repeated thrice and error bars indicate standard error. '*' indicates significant differences in comparison with the WT under respective conditions at p < 0.05 (Student's t-test).

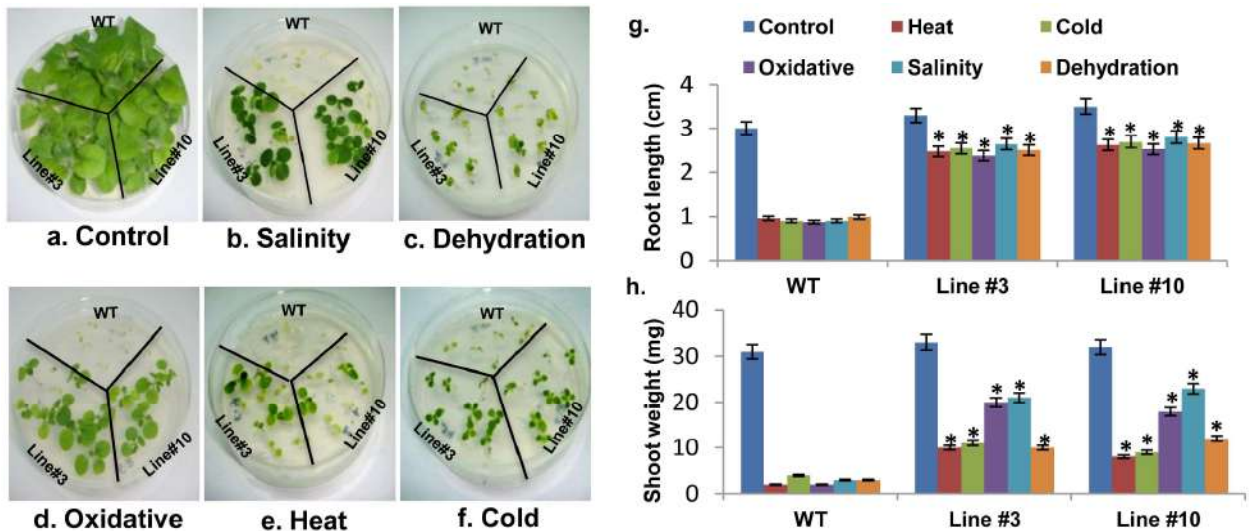


Fig. (6). Transgenic T₁ seedlings over-expressing *OsCBSCBSPB4* protein confers multiple abiotic stress tolerance. Seven day old seedlings of WT and transgenic lines (line 3 and line 10) were transferred to half strength MS medium for (a) control; or medium supplemented with either (b) NaCl (200 mM) for salinity or (c) Mannitol (5%) for dehydration or (d) MV (1 μM) for oxidative stress treatments. Seedlings were maintained under culture room conditions at 26±1°C and pictures taken after 15 d are shown. For high (e) and low (f) temperature stress, seedlings were exposed to 42°C for 24 h and 4°C for 72 h, respectively followed by recovery at 26±1°C. Graph showing comparative root length (g) and shoot weight (h) indicating severe inhibition in WT seedlings as compared to the transgenic seedlings under stress conditions. Experiment was repeated thrice and error bars indicate standard error. '*' indicates significant differences in comparison with the WT under respective conditions at p < 0.05 (Student's t-test).

Table 1. Potential interacting partners of OsCBSCBSPB4 identified through screening of rice library in yeast two-hybrid assay.

| Sl. No. | Putative Interacting Partners | Predicted Function |
|---------|---|---|
| 1 | Glutamine synthetase, catalytic domain containing protein | Glutamine biosynthetic process, nitrogen compound metabolic process |
| 2 | RING-H2 finger protein | Regulating growth/developmental processes, hormone signaling and responses to biotic and abiotic stresses in plants |
| 3 | Elongation factor protein | Protein biosynthesis |
| 4 | T-complex protein | Cytosolic protein folding machinery |
| 5 | LTPL9 - Protease inhibitor/seed storage/LTP family protein precursor | Protein of unknown function, induced in abiotic stress |
| 6 | Heat shock protein (Hsp 90) | For protein folding, assembly, translocation and degradation in many normal cellular processes, stabilize proteins and membranes, and can assist in protein refolding under stress conditions |
| 7 | Dihydrodipicolinate reductase | Dihydrodipicolinate reductase activity, catalyzes the second step of lysine biosynthesis |
| 8 | AMP-binding protein (or 4-coumarate-CoA ligase) | Biosynthesis of secondary metabolites |
| 9 | Amino acid permease | Regulates phloem amino acid composition; involved in the transport of amino acids into the cell |
| 10 | 60S ribosomal protein L31 | Structural constituent of ribosome |
| 11 | Magnesium-protoporphyrin IX monomethyl ester cyclase, chloroplast precursor | Involved in chlorophyll biosynthesis |
| 12 | Rieske domain containing protein | Involved in electron transfer chains of mitochondria and chloroplast |
| 13 | Histone H3 | Involved in the structure of chromatin |
| 14 | Acyl-desaturase, chloroplast precursor | Participates in polyunsaturated fatty acid biosynthesis in plastids |
| 15 | Lisencephaly type-1-like homology motif | Required for normal fertility and the first inter-node elongation |
| 16 | Photosystem II 10 kDa polypeptide, chloroplast precursor | Photosynthesis |
| 17 | DnaK family protein | Involved in protein folding, and help to protect cells from stress |
| 18 | Elongation factor Tu | Protein synthesis |
| 19 | Cysteine proteinase inhibitor precursor protein | Inhibition of proteases |
| 20 | Nitrogen regulatory protein P-II | Involved in the regulation of nitrogen metabolism |
| 21 | Phosphate carrier protein, mitochondrial precursor | Transporter activity |
| 22 | Metallothionein-like protein 3B | Metal binding proteins |
| 23 | Fructose-1,6-bisphosphatase | Carbohydrate metabolism |

proteins involved in stress response such as, heat shock proteins (Hsp90). Besides, we identified an AMP-binding domain containing protein as one of the potential interacting partner of OsCBSCBSPB4.

4. DISCUSSION

The mechanisms of stress tolerance exist in all living systems but these mechanisms may not be exactly same and may vary from species to species. Despite variations, the

format of cellular response to stress is generally believed to be conserved to large extent across prokaryotes as well as eukaryotes, including plants [22]. As abiotic stress ultimately affects the cellular gene-expression machinery, it is evident that a large number of genes are up or down regulated. Most of the genes involved in these pathways have not been characterized to date and have remained hypothetical. Therefore, it is essential to understand and execute functional approaches to establish the role of these genes with unknown or hypothetical functions, in order to improve our under-

standing of plant stress tolerance mechanisms [23]. Importantly, functionally analogous stress tolerant genes exist in both unicellular organisms and plants and hence, examining the functional significance of plant genes through their over-expression in simple organisms, such as yeast or bacteria seems to be promising. This approach has been tested by various groups [24-26].

Until now, only a few CDCPs have been characterized mainly from bacteria, humans and other animal systems. It is only recently that the identification and functional characterization of plant CDCPs is being undertaken [8, 10, 11]. In the present study, we have isolated an *OsCBSCBSPB4* gene from salt-tolerant rice cv. Pokkali and identified its novel function in multiple abiotic stress tolerance in bacteria as well as in model plant tobacco. *OsCBSCBSPB4* protein possesses two pairs of CBS domains which are known to be involved in a wide range of regulatory activities. For instance, single-CBS domain containing proteins, CBSX1 and CBSX2, from *Arabidopsis* have been identified as redox regulators that influence plant development *via* regulation of thioredoxin systems by sensing changes in adenosine-containing ligands [9]. Detailed studies on CBSX2 have revealed that it modulates the H₂O₂ status, which is linked to jasmonic acid response and in turn controls secondary wall thickening of the endothelial cells in anthers for dehiscence to occur [27]. Also, the four-CBS-domain (or two pairs of CBS domains) containing protein KINβγ, functions as the energy-sensing module of plant SnRK1 kinases through adenosine nucleotide binding [28].

In addition to CBS domains, *OsCBSCBSPB4* also possesses a PB1 domain, a protein interaction module that is conserved in animals, fungi, amoeba and plants. This domain is mainly involved in heterodimerization *via* interactions with PB1 domains of other proteins or with other protein domains. The canonical PB1 dimerization is required for the formation of complexes between proteins which are reported to be involved in activation of NADPH oxidase, polarity establishment in yeasts, cell polarization in animals and also in early cardiovascular development in mammals [29]. PB1 domain has also been observed in Auxin Response Factor (ARF) activators and Auxin/Indole 3-acetic acid inducible (Aux/IAA) repressors regulating the auxin responsive genes in plants [30]. Thus, presence of PB1 domain in *OsCBSCBSPB4* is indicative of the ability of this protein to interact with other proteins, thereby possibly transmitting cellular signals in the system.

CONCLUSION

Our results show that heterologous expression of *OsCBSCBSPB4* conferred tolerance to bacterial cells against multiple abiotic stresses. In agreement, we could observe an increase in the expression of *OsCBSCBSPB4* under these stresses in the two rice genotypes, IR64 and Pokkali. Thus, this salinity-responsive *OsCBSCBSPB4* gene may also be involved in conferring cold, heat and oxidative stress tolerance to plants as could be seen from the response of *OsCBSCBSPB4*-overexpressing transgenic lines to these stress conditions. In support of this, reports describing the ability of other CDCPs, such as *OsCBSX4*, in improving the tolerance of transgenic plants towards oxidative, salinity and heavy metal stress already exist [8].

In order to investigate the underlying mechanism behind observed stress tolerance conferred by *OsCBSCBSPB4*, yeast two hybrid studies were carried out which led to the identification of various nuclear and cytoplasmic proteins involved in stress response and tolerance as potential interacting partners of *OsCBSCBSPB4*. To name few, heat shock protein (Hsp90) [31], T-complex protein [32], LTPL9, a protease inhibitor/ seed storage/ LTP family protein precursor [33], metallothionein-like protein [34] and RING-H2 finger protein [35, 36] were identified as the potential interacting partners of *OsCBSCBSPB4*. Heat shock proteins and T-complex proteins are very well known molecular chaperones that assist protein folding and are induced under stress conditions [31, 32]. The transcription factor RING-H2 finger protein, is also associated with key roles in regulating growth/developmental processes, hormone signaling and responses to biotic and abiotic stresses in plants [35, 36]. In addition, we could also detect an AMP binding domain containing protein (or 4-coumarate-CoA ligase) as one of the interacting partners that is probably involved in the biosynthesis of plant secondary compounds [37]. It is induced by wounding and UV irradiation or upon fungal infection [38]. The observed interaction of *OsCBSCBSPB4* with an AMP binding domain containing protein indicates towards a role of *OsCBSCBSPB4* similar to that proposed for few CDCPs, where they bind adenosine nucleotides and transmit signals to the catalytic units. We believe that *OsCBSCBSPB4* can shuttle between nucleus and cytoplasm, as evident from its nucleo-cytoplasmic localization, and regulate the functions of various proteins involved in stress response which is further supported by the observed stress tolerance conferred by over-expression of *OsCBSCBSPB4* in bacteria as well as tobacco. Though no strong NLS sequence was predicted in the *OsCBSCBSPB4* protein but nuclear localization has been otherwise predicted for 4 CDCPs in soybean [11].

We hypothesize that *OsCBSCBSPB4*, like other CDCPs, regulates the activation of other proteins, specifically stress-proteins. This, we believe, may be the reason why its over-expression imparts tolerance against abiotic stress conditions. Taken together, *OsCBSCBSPB4* represents a potential candidate gene from rice, which may be playing an important role in stress tolerance mechanism in plants.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

This work was supported by internal grants of International Centre for Genetic Engineering and Biotechnology (ICGEB), and Department of Biotechnology (DBT), Government of India. TUA acknowledges ICGEB for providing a pre-doctoral fellowship.

REFERENCES

- [1] Bateman, A. The structure of a domain common to archaebacteria and the homocystinuria disease protein. *Trends Biochem. Sci.*, **1997**, *22*, 12-13.
- [2] Kushwaha, H.R.; Singh, A.K.; Sopory, S.K.; Singla-Pareek, S.L.; Pareek, A. Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* (L.) Heynh and *Oryza sativa* L. reveals their developmental and stress regulation. *BMC Genomics*, **2009**, *10*, 200-205.
- [3] Ereño-Orbea, J.; Oyente, I.; Martínez-Cruz, L.A. CBS domains: Ligand binding sites and conformational variability. *Arch. Biochem. Biophys.*, **2013**, *540*, 70-81.
- [4] Shan, X.; Dunbrack, R.L.; Christopher, S.A.; Kruger, W.D. Mutations in the regulatory domain of cystathionine beta synthase can functionally suppress patient-derived mutations in cis. *Hum. Mol. Genet.*, **2001**, *10*, 635-643.
- [5] Kennan, A.; Aherne, A.; Palfi, A.; Humphries, M.; McKee, A.; Stitt, A.; Simpson, D.A.; Demtroder, K.; Ortoft, T.; Ayuso, C.; Kenna, P.F.; Farrar, G.J.; Humphries, P. Identification of an *IMPDH1* mutation in autosomal dominant retinitis pigmentosa (RP10) revealed following comparative microarray analysis of transcripts derived from retinas of wild-type and Rho (-/-) mice. *Hum. Mol. Genet.*, **2002**, *11*, 547-557.
- [6] Baykov, A.A.; Tuominen, H.K.; Lahti, R. The CBS domain: A protein module with an emerging prominent role in regulation. *ACS Chem. Biol.*, **2011**, *6*, 1156-1163.
- [7] Kumari, S.; Sabharwal, V.P.; Kushwaha, H.R.; Sopory, S.K.; Singla-Pareek, S.L.; Pareek, A. Transcriptome map for seedling stage specific salinity stress response indicates a specific set of genes as candidate for saline tolerance in *Oryza sativa* L. *Funct. Integr. Genomics*, **2009**, *9*, 109-123.
- [8] Singh, A.K.; Kumar, R.; Pareek, A.; Sopory, S.K.; Singla-Pareek, S.L. Overexpression of rice CBS domain containing protein improves salinity, oxidative and heavy metal tolerance in transgenic tobacco. *Mol. Biotechnol.*, **2012**, *2*, 9847-9852.
- [9] Yoo, K.S.; Ok, S.H.; Jeong, B.C.; Jung, K.W.; Cui, M.H.; Hyoung, S.; Lee, M.R.; Song, H.K.; Shin, J.S. Single Cystathionine β-Synthase domain containing proteins modulate development by regulating the thioredoxin system in *Arabidopsis*. *Plant Cell*, **2011**, *23*, 3577-3594.
- [10] Mou, S.; Shi, L.; Lin, W.; Liu, Y.; Shen, L.; Guan, D.; He, H. Over-expression of rice CBS domain containing protein, *OsCBSX3*, confers rice resistance to *Magnaporthe oryzae* inoculation. *Int. J. Mol. Sci.*, **2015**, *16*, 15903-15917.
- [11] Hao, Q.; Shang, W.; Zhang, C.; Chen, H.; Chen, L.; Yuan, S.; Chen, S.; Zhang, X.; Zhou, X. Identification and comparative analysis of CBS domain-containing proteins in soybean (*Glycine max*) and the primary function of *GMCBS21* in enhanced tolerance to low nitrogen stress. *Int. J. Mol. Sci.*, **2016**, *17*, 620-637.
- [12] De Castro, E.; Sigrist, C.J.A.; Gattiker, A.; Bulliard, V.; Langendijk-Genevaux, P.S.; Gasteiger, E.; Bairoch, A.; Hulo, N. ScanProsite: Detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.*, **2006**, *34*, W362-W365.
- [13] Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; Thompson, J.D.; Gibson, T.J.; Higgins, D.G. Clustal W and Clustal X version 2.0. *Bioinformatics*, **2007**, *23*, 2947-2948.
- [14] Saitou, N.; Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **1987**, *4*, 406-425.
- [15] Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.*, **2016**, *33*(7), 1870-1874.
- [16] Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **2001**, *25*, 402-408.
- [17] Eisen, M.B.; Spellman, P.T.; Brown, P.O.; Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci., USA*, **1998**, *95*, 14863-14868.
- [18] Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **1976**, *72*, 248-254.
- [19] Mustafiz, A.; Sahoo, K.K.; Singla-Pareek, S.L.; Sopory, S.K. Metabolic engineering of glyoxalase pathway for enhancing stress tolerance in plants. *Methods Mol. Biol.*, **2010**, *639*, 95-118.
- [20] Singla-Pareek, S.L.; Reddy, M.K.; Sopory, S.K. Genetic engineering of the glyoxalase pathway in tobacco leads to enhanced salinity tolerance. *Proc. Natl. Acad. Sci., USA*, **2003**, *100*, 14672-14677.
- [21] Ito, T.; Chiba, T.; Ozawa, R.; Hattori, M.; Sakaki, Y. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci., U.S.A.*, **2001**, *98*, 4569-4574.
- [22] Kultz, D. Molecular and evolutionary basis of the cellular stress response. *Annu. Rev. Physiol.*, **2005**, *67*, 225-257.
- [23] Diédhiou, C.J.; Gollack, D. Salt-dependent regulation of chloride channel transcripts in rice. *Plant Sci.*, **2006**, *170*, 793-800.
- [24] Mundree, S.G.; Whittaker, A.; Thomson, J.A.; Farrant, J.M. An aldose reductase homolog from the resurrection plant *Xerophytaviscosa*. *Planta*, **2000**, *211*, 693-700.
- [25] Yamada, A.; Saitoh, T.; Mimura, T.; Ozeki, Y. Expression of mangrove allene oxide cyclase enhances salt tolerance in *Escherichia coli*, yeast and tobacco cells. *Plant Cell Physiol.*, **2002**, *4*, 903-910.
- [26] Yamada, A.; Tsutsumi, K.; Tanimoto, S.; Ozeki, Y. Plant RelA/Spot homolog confers salt tolerance in *Escherichia coli* and *Saccharomyces cerevisiae*. *Plant Cell Physiol.*, **2003**, *44*, 3-9.
- [27] Jung, K.W.; Kim, Y.Y.; Yoo, K.S.; Ok, S.H.; Cui, M.H.; Jeong, B.C.; Yoo, S.D.; Jeung, J.U.; Shin, J.S. A cystathionine-β-synthase domain-containing protein, CBSX2, regulates endotheic secondary cell wall thickening in anther development. *Plant Cell Physiol.*, **2013**, *54*, 195-208.
- [28] Ramon, M.; Ruelens, P.; Li, Y.; Sheen, J.; Geuten, K.; Rolland, F. The hybrid Four-CBS-Domain KINbc subunit functions as the canonical c subunit of the plant energy sensor SnRK1. *Plant J.*, **2013**, *75*, 11-25.
- [29] Sumimoto, H.; Kamakura, S.; Ito, T. Structure and function of the PB1 domain, a protein interaction module conserved in animals, fungi, amoebas, and plants. *Sci. STKE*, **2007**, *401*, re6.
- [30] Guilfoyle, T.J. The PB1 domain in Auxin response factor and Aux/IAA proteins: A versatile protein interaction module in the Auxin response. *Plant Cell*, **2015**, *27*, 33-43.
- [31] Wang, W.; Vinocur, B.; Shoseyov, O.; Altman, A. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.*, **2004**, *9*, 244-252.
- [32] Yokota, S.I.; Yanagi, H.; Yura, T.; Kubota, H. Upregulation of cytosolic chaperonin CCT subunits during recovery from chemical stress that causes accumulation of unfolded proteins. *Eur. J. Biochem.*, **2000**, *267*, 1658-1664.
- [33] Jung, H.W.; Kim, W.; Hwang, B.K. Three pathogen-inducible genes encoding lipid transfer protein from pepper are differentially activated by pathogens, abiotic, and environmental stresses. *Plant Cell Environ.*, **2003**, *26*, 915-928.
- [34] Kumar, G.; Kushwaha, H.R.; Panjabi-Sabharwal, V.; Kumari, S.; Joshi, R.; Karan, R.; Mittal, S.; Singla-Pareek, S.L.; Pareek, A. Clustered metallothionein genes are co-regulated in rice and ectopic expression of *OsMT1e-P* confers multiple abiotic stress tolerance in tobacco via ROS scavenging. *BMC Plant Biol.*, **2012**, *12*, 107.
- [35] Ko, J.H.; Yang, S.H.; Han, K.H. Upregulation of an *Arabidopsis* RING-H2 gene, *XERICCO*, confers drought tolerance through increased abscisic acid biosynthesis. *Plant J.*, **2006**, *47*, 343-355.
- [36] Liu, H.; Zhang, H.; Yang, Y.; Li, G.; Yang, Y.; Wang, X.; Basnayake, B.M.; Li, D.; Song, F. Functional analysis reveals pleiotropic effects of rice RING-H2 finger protein gene *OsBIRF1* on regulation of growth and defense responses against abiotic and biotic stresses. *Plant Mol. Biol.*, **2008**, *68*, 17-30.
- [37] Lindl, T.; Kreuzaler, F.; Hahlbrock, K. Synthesis of p-coumaroyl coenzyme A with a partially purified p-coumarate:CoA ligase from cell suspension cultures of soybean (*Glycine max*). *Biochim. Biophys. Acta.*, **1973**, *302*, 457-464.
- [38] Ehrling, J.; Büttner, D.; Wang, Q.; Douglas, C.J.; Somssich, I.E.; Kombrink, E. Three 4-coumarate: coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *Plant J.*, **1999**, *19*, 9-12.

SPECIAL ISSUE ARTICLE

The chloride channels: Silently serving the plants

Ashish Subba¹ | Surabhi Tomar¹ | Ashwani Pareek²  | Sneh L. Singla-Pareek¹ 

¹Plant Stress Biology, International Centre for Genetic Engineering and Biotechnology, New Delhi, India

²Stress Physiology and Molecular Biology Laboratory, School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

Correspondence

Sneh L. Singla-Pareek, Plant Stress Biology, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali, Marg, New Delhi 110067, India.
Email: sneh@icgeb.res.in

Funding information

International Centre for Genetic Engineering and Biotechnology, India.

Edited by: R. Deshmukh

Abstract

Chloride channels (CLCs), member of anion transporting proteins, are present ubiquitously in all life forms. Diverging from its name, the CLC family includes both channel and exchanger (proton-coupled) proteins; nevertheless, they share conserved structural organization. They are engaged in diverse indispensable functions such as acid and fluoride tolerance in prokaryotes to muscle stabilization, transepithelial transport, and neuronal development in mammals. Mutations in genes encoding CLCs lead to several physiological disorders in different organisms, including severe diseases in humans. Even in plants, loss of CLC protein function severely impairs various cellular processes critical for normal growth and development. These proteins sequester Cl^- into the vacuole, thus, making them an attractive target for improving salinity tolerance in plants caused by high abundance of salts, primarily NaCl. Besides, some CLCs are involved in NO_3^- transport and storage function in plants, thus, influencing their nitrogen use efficiency. However, despite their high significance, not many studies have been carried out in plants. Here, we have attempted to concisely highlight the basic structure of CLC proteins and critical residues essential for their function and classification. We also present the diverse functions of CLCs in plants from their first cloning back in 1996 to the knowledge acquired as of now. We stress the need for carrying out more in-depth studies on CLCs in plants, for they may have future applications towards crop improvement.

1 | INTRODUCTION

Membrane transport proteins are the chief biological components involved in maintaining the electrochemical potential and ionic balance by regulating the flux of various substrates across the membrane. These proteins are broadly divided into channels and transporters. In the case of channel proteins, the substrate diffuses

passively down their electrochemical gradients at rates of more than 10^6 molecules per seconds. On the other hand, transporters are slow ($1-1000$ molecules sec^{-1}) and allow the movement of the substrate against their concentration gradients by using energy either derived from hydrolysis of ATP or coupled to the movement of the secondary substrate, which is often a proton in plants (see Ashcroft *et al.*, 2009).

The chloride channel (CLC) protein family is one class of anion-conducting proteins present ubiquitously in all domains of life, including Archaea. The genesis of the CLC dates back to four decades when White and Miller (1979), while exploring for cation channels in membrane vesicles from the electric organ of *Torpedo californica* (electric ray fish), unexpectedly observed strong voltage-dependent anion conductance, selective to Cl^- . Later, Jentsch *et al.* (1990) isolated the gene encoding this anion channel from *Torpedo marmorata* and named it as CLC-0, the founder of all CLCs. The discovery of CLC-0 was promptly followed by the identification of CLCs in several organisms,

Abbreviations: ABA, Abscisic acid; AMP, Adenosine monophosphate; ATP, Adenosine triphosphate; CBS, Cystathionine- β -synthase; Ccc2, Cu+2 transporting P-type ATPase; CLC, Chloride channel; FET3, Ferrous transporter protein; Flg22, N-terminus derived Flagellin peptide; FLS2, Flagellin-sensing 2; NHX, Sodium/hydrogen exchanger; NRT, Nitrate transporter; OST1, Open stomata 1; PAMP, Pathogen-associated molecular pattern; PI(3,5)P2, Phosphatidylinositol-3,5-bisphosphate; PP2A-C5, C5 catalytic subunit of protein phosphatase 2A; PRR, Pattern recognition receptor; PSTVd, Potato spindle tuber viroid; PTI, Pathogen-associated molecular pattern triggered immunity; PVY, Potato virus Y; SnRK, Sucrose non-fermenting-related serine/threonine-protein kinase; TBSV, Tomato bushy stunt virus; V-ATPase, Vacuolar-type ATPase; vd-sRNA, Potato spindle tuber viroid derived small RNA; VIGS, Virus-induced gene silencing; VRC, Viral replicase complex.

and yet many more are to be identified. In plants, cloning of genes encoding CLCs was first reported in 1996 by Lurin *et al.* (1996) in tobacco (*NtCLC-1*) and Hechenberger *et al.* (1996) in *Arabidopsis* (*AtCLC-a - d*). While anion-selective conductance of *NtCLC-1* was confirmed electro-physiologically by expressing it in *Xenopus* oocytes (Lurin *et al.*, 1996), *AtCLC-d* was suggested to function in Cl^- transport as it functionally complemented the phenotype of yeast mutant carrying a mutation in its lone *CLC* gene (*Δ gef1*) (Hechenberger *et al.*, 1996).

CLCs were earlier considered to function as channel proteins, hence, named chloride channels (CLCs). In contrast, the identification of antiporter function (Cl^-/H^+) in *CLC-ec1* from *Escherichia coli* (Accardi and Miller, 2004) subsequently uncovered that the CLC protein family comprises of both anion channels and transporters. In mammals, five out of nine identified CLCs are transporters that localize in the intracellular vesicles, whereas CLC channels are present in the plasma membrane. However, all seven CLCs identified in *Arabidopsis thaliana* (*Arabidopsis*) show organellar membrane localization (Lv *et al.*, 2009), which raises the question if plant CLCs do not localize in the plasma membrane and whether they all serve as transporters. Besides, CLCs were initially anticipated for Cl^- transport activity; however, the identification of NO_3^- conduction by *Arabidopsis AtCLC-a* (De Angeli *et al.*, 2006) revealed that some CLCs are involved in NO_3^- transport, as NO_3^- being one of the most abundant anions in plants. Additionally, members of bacterial CLC-proteins (CLC-Fs), which form a phylogenetically distinct clade, have fluoride conductance instead. They are associated with fluoride riboswitches and impart fluoride tolerance by transporting F^- into liposomes (Baker *et al.*, 2012).

Some CLCs exhibit specific expression patterns, while others show extensive tissue coverage and accordingly, CLCs are involved in diverse cellular functions in all living systems. In mammals, the disruption of any isoform from nine identified CLCs leads to disease or growth defects, exhibiting functional non-redundancy (see Jentsch and Pusch, 2018). In *E. coli*, two CLC homologs are known to confer acid tolerance through the removal of excess intracellular protons while neutralizing the intracellular positive charged products with the Cl^- influx (Foster, 2004). In yeast, the lone existing CLC (*ScCLC* or *GEF1*) is involved in Cl^- influx into the Golgi vesicles, which facilitates Cu^{2+} transport by *Ccc2* (Cu^{2+} transporting P-type ATPase) and subsequent Cu^{2+} loading unto the iron transport multicopper oxidase (*FET3*), a ferrous transporter protein involved in high-affinity iron uptake (Davis-Kaplan *et al.*, 1998). Accordingly, yeast *Δ gef1* mutant exhibits an iron-limited growth defect and is sensitive to different salts and higher pH (Davis-Kaplan *et al.*, 1998; Gaxiola *et al.*, 1998). Thus, this mutant has been often used for complementation studies with plant CLCs. Despite the above-mentioned critical functions, CLCs studies in plant systems are still scarce. In this review, we shall briefly highlight the structure and conserved motifs obtained from *E. coli* and mammalian CLCs. Then, we shall focus on their current understanding in plant systems, gained mostly from studies on *Arabidopsis*.

2 | STRUCTURE AND CONSERVED RESIDUES GOVERNING THE FUNCTIONS OF CLCS

The ion-conducting transmembrane domain of CLC proteins is composed of 16–18 α -helices (referred in an alphabetical order in this review), which in eukaryotes is always followed by a cytoplasmic C-terminus CBS domain, but is present only in few prokaryotic CLCs. The crystal structure of CLCs from both *Salmonella enterica* and *E. coli* (Dutzler *et al.*, 2002) confirmed that CLCs form a homodimer protein, and each subunit forms its ion transporting pore, as initially proposed for *CLC-0* by Miller (1982). The structures from both these organisms showed high similarity, and each subunit consisted of 18 α -helices, which aligned tilted in the membrane and are of variable lengths, as shown for CLC from *E. coli* (Figure 1). The N-terminal half of the polypeptide is structurally related to the C-terminal half and these two halves orient in antiparallel architecture such that amino acid residues from four different regions are brought closer to form an ion conduction pathway (Figure 1). The residues forming this pore are highly conserved (called selectivity filter), comprising of sequences GSGIP (106–110 in helix D), G(K/R)EGP (146–150 in helix F), and GXFXP (355–359 in helix N), as well as Tyr 445 in helix R (shown as inset in Figure 1(B)). The Cl^- ion is coordinated by Ser 107, Ile 356, Phe 357 and Tyr 445 (Dutzler *et al.*, 2002).

The identification of *E. coli CLC-ec1* as transporter instead of channel protein (Accardi and Miller, 2004) discerned mammalian *CLC-3-7*, plant (*Arabidopsis*) *AtCLC-a* and *-b*, and red alga (*Cyanidioschyzon merolae*) *CmCLC* to function as a transporter. They all exhibited an exchange stoichiometry of 2 anions:1 proton (Scheel *et al.*, 2005; De Angeli *et al.*, 2006; Graves *et al.*, 2008; Feng *et al.*, 2010; Neagoe *et al.*, 2010; Guzman *et al.*, 2013). The structure of *C. merolae CmCLC* transporter superimposed very well with *CLC-ec1*, including the conserved residues of an ion conduction pathway (Figure 2). However, their sequences shared less identity (<25%; Feng *et al.*, 2010), thereby indicating the conservation of structure in CLCs. Also, the phylogenetically distinct F^- conducting bacterial CLC-Fs have been identified to function as transporters (Stockbridge *et al.*, 2012). The crystal structure of one of the CLC-Fs, “*CLC-eca*” from *Enterococcus casseliflavus*, complemented with the structure of other CLCs, but it possesses shorter helices (Last *et al.*, 2018). However, CLC-Fs exhibited 1:1 exchange stoichiometry, in exception to other CLC transporters (Stockbridge *et al.*, 2012).

In CLC transporters, two conserved glutamate residues have been identified as essential for imparting proton-coupled anion exchange activity. The first one is called the gating glutamate (E148 in *CLC-ec1*), which acts as a gate on the extracellular side, and when protonated, it is displaced away to permit Cl^- transport (Dutzler *et al.*, 2003). The other one is called proton glutamate (E203 in *CLC-ec1*), which resides on the intracellular surface of CLC transporters (Accardi *et al.*, 2005) and exchanges the proton with gating glutamate. Mutation in the gating glutamate abolished H^+ but not Cl^- transport in bacterial (*E. coli*),

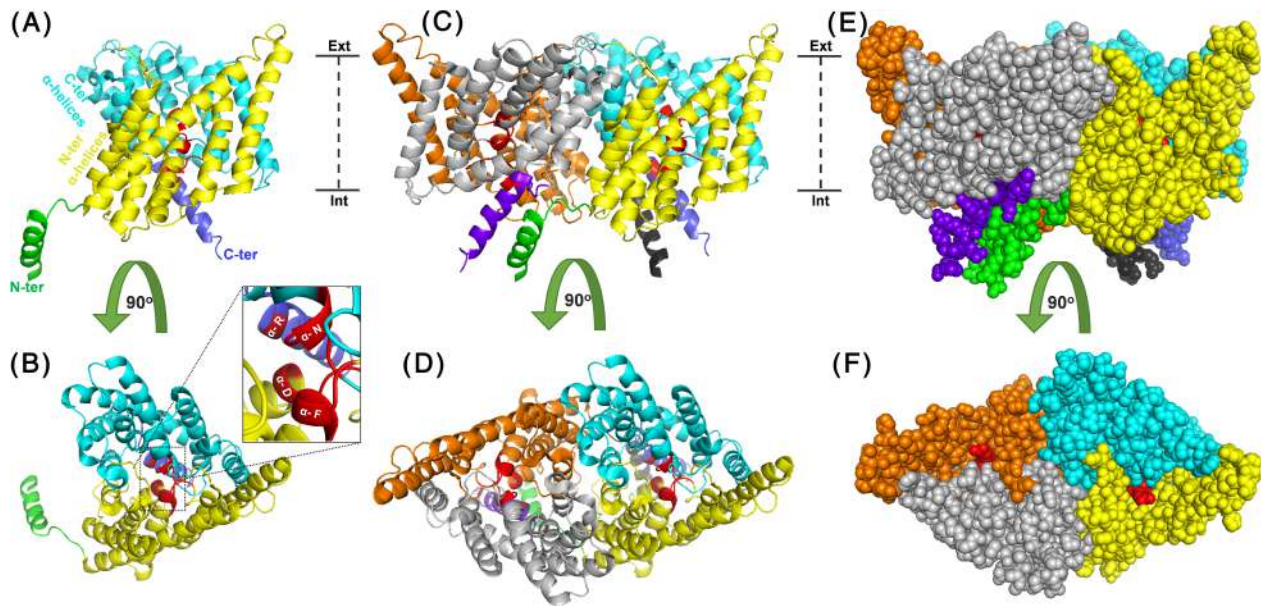


FIGURE 1 Structure depicting the monomeric and homodimeric CLC-ec1 protein (PDB ID: 1KPK) of *Escherichia coli*. Note the absence of CBS domain in this prokaryotic CLC protein. (A, B) Single subunit of CLC-ec1 homodimer, representing the structurally related N-terminal (colored yellow) and C-terminal (colored cyan) halves of a polypeptide, while the N-terminal end helix is colored green and the C-terminal end helix is colored blue. The helices forming the ion conduction pathway are highlighted in red and enlarged in the inset (B). Two subunits forming CLC-ec1 homodimer, with the other subunit colored in orange (N-terminal helices), gray (C-terminal helices), black (N-terminal end helix) and purple (C-terminal end helix), are depicted as cartoon (C, D) and spheres (E, F). The views in B, D, and F are 90° rotated images of A, C and E, respectively. The extracellular (Ext) and the intracellular (Int) sides are highlighted, with a dotted line representing the approximate transmembrane region. The structures have been modeled using PyMOL Molecular Graphics System, Version 2.3.3 Schrödinger, LLC

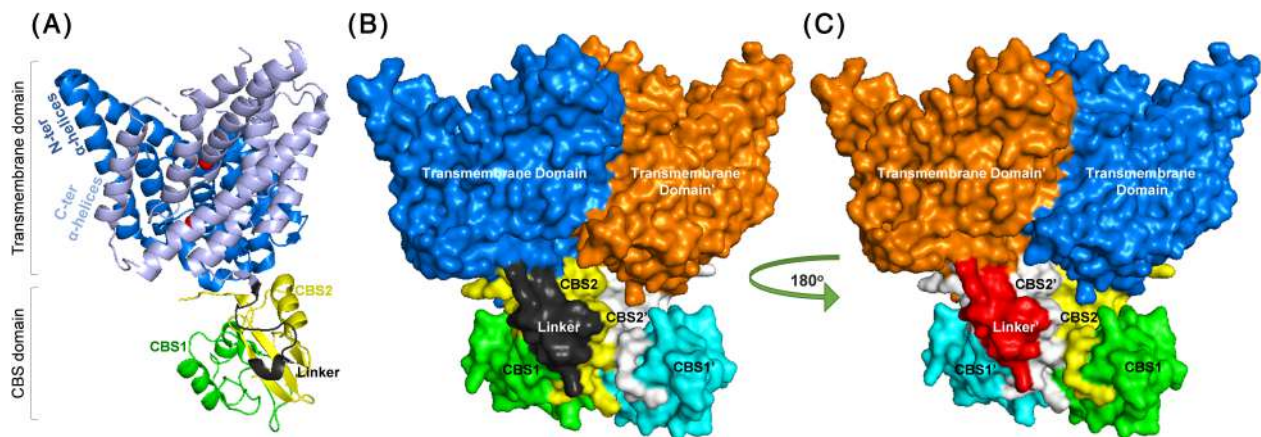


FIGURE 2 Structure depicting the monomeric and homodimeric CmCLC protein (PDB ID: 3ORG) of *Cyanidioschyzon merolae* (red alga). Note the presence of cytoplasmic C-ter CBS sub-domains connected to the transmembrane domain by a linker region in this eukaryotic protein. (A) Cartoon representation of CmCLC monomer with transmembrane domain (colored in blue and sky blue, representing structurally related N-ter and C-ter α -helices), cytoplasmic region (linker in black, CBS1 in green and CBS2 in yellow) and two bound Cl^- ions (in red). (B) Surface representation of CmCLC homodimer, with the other subunit [marked with (')] coloured orange for the transmembrane domain', cyan for CBS1' and white for CBS2'. (C) This view depicts the 180° rotated image of (B) and linker' is colored in red. The structure has been modeled using the PyMOL Molecular Graphics System, Version 2.3.3 Schrödinger, LLC

mammalian, as well as in plant CLC transporters (Accardi and Miller, 2004; Scheel *et al.*, 2005; Bergsdorf *et al.*, 2009; Leisle *et al.*, 2011). Similarly, the neutralization of proton glutamate eliminated H^+ transport with a reduced Cl^- conduction in CLC-ec1 (Accardi *et al.*, 2005). However, this mutation abolished both Cl^- and H^+ transport in mammalian and plant CLC transporters, while an

additional mutation of the gating glutamate restored Cl^- transport (Zdebek *et al.*, 2008; Bergsdorf *et al.*, 2009; Leisle *et al.*, 2011). The existence of these two conserved glutamates has been preliminarily used to assign either a transporter or channel function to the electrophysiologically uncharacterized CLCs. Nevertheless, exceptions may occur at the conserved proton glutamate, which is replaced by

threonine in CmCLC (Feng *et al.*, 2010), isoleucine in *Citrobacter koseri* CLC-ck2 (Phillips *et al.*, 2012), and valine (or rarely, isoleucine) in bacterial CLC-Fs (Stockbridge *et al.*, 2012), yet all three function as CLC transporters, raising concern to surmise CLC's function based upon proton glutamate.

In CLC channels, the corresponding position of the proton glutamate is occupied by valine (Accardi *et al.*, 2005), and its substitution with glutamate does not convert the CLC channels to exchangers (Zdebik *et al.*, 2008). However, the gating glutamate is present even in some CLC channel proteins, and its neutralization results in a constitutively-open type channel (Dutzler *et al.*, 2003; Niemeyer *et al.*, 2003). This gating glutamate in CLC channels has also been suggested to permit Cl⁻ transport by protonation of its side chain as in CLC transporters. Thus, CLC channels have also been proposed to act as 'broken' Cl⁻/H⁺ antiporters that have become leaky for Cl⁻ (Lísal and Maduke, 2008). Accordingly, both CLC channels and transporters are suggested to have evolutionary conserved and physically distinct routes for both Cl⁻ and H⁺ transport (Leisle *et al.*, 2020).

Interestingly, the structure of mammalian CLC channel proteins, CLC-K (Park *et al.*, 2017) and CLC-1 (Park and MacKinnon, 2018; Wang *et al.*, 2019) manifested high similarity to that of CLC transporters, with slight differences in the residues forming an ion-conducting pore that appeared critical for its channel activity. In CLC-K (from bovine), the loop connecting α -helices C and D is displaced, directing the serine residue (S121) of the ion-conducting pore towards the cytosol instead, which is suggested to remove the kinetic barrier in Cl⁻ conduction pathway and widen the pore (Park *et al.*, 2017). Similarly, in human CLC-1, the presence of amino acids with smaller hydrophilic side chains towards the Cl⁻ conducting pore, and shift in the position of ion binding residue Y578, is proposed to form wider pore size for channel function (Park and MacKinnon, 2018; Wang *et al.*, 2019). Accordingly, altering the residues in CLC-ec1 transporter, corresponding to the ones observed to widen the pore in CLC-K and CLC-1, led to a channel like activity of CLC-ec1 with remarkably increased Cl⁻ transport (Park and MacKinnon, 2018), further supporting the notion that CLC channels are a 'broken' form of Cl⁻/H⁺ antiporters.

The observation that AtCLC-a from *Arabidopsis* functions as NO₃⁻/H⁺ exchanger (De Angeli *et al.*, 2006) led to the identification of a 'Proline' in AtCLC-a in place of a conserved 'Serine' residue present in the signature sequence 'GSGIP' motif of Cl⁻ conducting CLCs (Zifarelli and Pusch, 2009). Accordingly, the substitution of proline to serine (P160S) in AtCLC-a results in an increased conductance of Cl⁻ over NO₃⁻, with both ions efficiently coupled to proton transport (Bergsdorf *et al.*, 2009; Wege *et al.*, 2010). Likewise, the substitution of serine to proline resulted in the transformation of *E. coli* CLC-ec1, mammalian CLC-5 and CLC-7 from Cl⁻/H⁺ to NO₃⁻/H⁺ exchanger (Picollo *et al.*, 2009; Zifarelli and Pusch, 2009; Leisle *et al.*, 2011). Besides, the serine to proline substitution also increased NO₃⁻ conductance in the CLC-0 channel (Bergsdorf *et al.*, 2009), signifying the role of these two residues in anion selectivity in both CLC transporters and channels. Sequence analysis of plant CLCs identified proline in the selective filter of at least one CLC protein in each seed-

producing plant examined. However, the same was not found in the case of CLCs from green algae, bryophyte, and lycophyte (Wege *et al.*, 2010). This has led to the conclusion that NO₃⁻-specific CLCs may have evolved in higher plants. Furthermore, this conserved 'Serine' residue is absent in CLC-Fs, and the equivalent position is occupied by a methionine or asparagine residue, which has been proposed to confer F⁻ selectivity in these proteins (Brammer *et al.*, 2014).

Another structural component of eukaryotic and few prokaryotic CLC proteins is the C-terminus cytoplasmic cystathionine- β -synthase (CBS) domain, composed of two sub-domains CBS1 and CBS2 (Figure 2). CBS domain-containing proteins, in general, are present in eukaryotes, prokaryotes, and archaea as well, existing either as a single domain or in association with other functional domains in a polypeptide (e.g., in CLCs). These are known to perform regulatory functions by binding to various ligands (Kushwaha *et al.*, 2009). Alike the transmembrane domain, the structural organization of the CBS domain also appears conserved in the CLC family. The CBS1 and CBS2 sub-domains form a pseudo-2-fold symmetry, tightly interacting via an interface formed by a pair of β -strands (Meyer and Dutzler, 2006; Markovic and Dutzler, 2007; Feng *et al.*, 2010). CBS domains from two CLC subunits form a homodimer, involving the major interactions between CBS2 sub-domains (Markovic and Dutzler, 2007; Meyer *et al.*, 2007). Besides, the CBS2 sub-domain and the polypeptide linker (connecting transmembrane domains with the CBS domain) make extensive contact with the transmembrane helices that extend into the ion conduction pathway (Feng *et al.*, 2010; Wang *et al.*, 2019). CBS domain also forms contact with the loop connecting α -helices, which forms the transmembrane dimer interface (Feng *et al.*, 2010). These indicate that CBS domains could influence ion conduction as well as the interactions between two transmembrane subunits (Feng *et al.*, 2010). Many CLCs mutations causing diseases have been found to occur in the CBS domain region, which further implies the functional significance of this domain (Feng *et al.*, 2010).

3 | CLCS IN PLANTS

As stated above, the pioneering work on CLCs in plants was initiated independently by Lurin *et al.* (1996) in tobacco and Hechenberger *et al.* (1996) in *Arabidopsis*. Lurin *et al.* (1996) isolated *NtCLC-1* that elicited anion-selective conductance when expressed in *Xenopus* oocytes. It exhibited the highest expression in the stem (Lurin *et al.*, 1996) and is localized in the endoplasmic reticulum (ER) (Sun *et al.*, 2018). Hechenberger *et al.* (1996) cloned four CLCs from *Arabidopsis* (*AtCLC-a - d*) of which *AtCLC-d* was able to rescue the iron-limited growth phenotype of Δ *gef1* mutant of yeast. However, this functional complementation was impeded when a mutation in CLCs linked to human diseases was incorporated in *AtCLC-d*, suggesting that *AtCLC-d* functions in Cl⁻ transport. Besides, *AtCLC-d* and GEF1 showed similar localization in the yeast cells (Hechenberger *et al.*, 1996), which was later confirmed to be the trans-Golgi network (von der Fecht-Bartenbach *et al.*, 2007). Subsequently, CLCs have been identified and studied in many plant species. They exhibit

TABLE 1 Chloride channels (CLCs) in plants: Localization, putative functions and conserved residues of their selectivity filter

| S. No. | Plant CLCs | Localization | CLC type | Gating glutamate | Proton glutamate | Pro/Ser filter | Role | References |
|--------|-------------------------|---------------------|---|------------------|------------------|----------------|---|--|
| 1 | AtCLC-a | Tonoplast | NO ₃ ⁻ /H ⁺ exchanger | + | + | Proline | NO ₃ ⁻ transport, stomatal regulation | De Angeli <i>et al.</i> (2006), Wege <i>et al.</i> (2010, 2014) |
| 2 | AtCLC-b | Tonoplast | NO ₃ ⁻ /H ⁺ exchanger | + | + | Proline | NO ₃ ⁻ efflux from vacuole | von der Fecht-Bartenbach <i>et al.</i> (2010) |
| 3 | AtCLC-c | Tonoplast | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | Cl ⁻ influx into the vacuole, stomatal regulation | Jossier <i>et al.</i> (2010), Hu <i>et al.</i> (2017) |
| 4 | AtCLC-d | Trans-Golgi Network | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | pH regulation of trans-Golgi network through anions flux, Biotic stress (negatively regulates PAMP) | Hechenberger <i>et al.</i> (1996), von der Fecht-Bartenbach <i>et al.</i> (2007), Guo <i>et al.</i> (2014) |
| 5 | AtCLC-e | Thylakoid | Cl ⁻ /H ⁺ exchanger (putative) | + | -(Serine) | -Aspartic acid | Cl ⁻ homeostasis in thylakoid, maintaining photosynthesis | Marmagne <i>et al.</i> (2007), Herdean <i>et al.</i> , (2016) |
| 6 | AtCLC-f | Golgi vesicles | Cl ⁻ /H ⁺ exchanger (putative) | + | -(Threonine) | Serine | - | Marmagne <i>et al.</i> (2007) |
| 7 | AtCLC-g | Tonoplast | Cl ⁻ channel (putative) | -(Alanine) | + | Serine | Cl ⁻ transport, stomatal regulation | Nguyen <i>et al.</i> (2016) |
| 8 | BnaCLC-a1/ BnaCLC-b1 | - | NO ₃ ⁻ /H ⁺ exchanger (putative) | + | + | Proline | - | Liao <i>et al.</i> (2018) |
| 9 | CsCLC-c | - | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | Cl ⁻ transport | Wei <i>et al.</i> (2013) |
| 10 | GsCLC-c2 | Tonoplast | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | Cl ⁻ transport into vacuole | Wei <i>et al.</i> (2019) |
| 11 | NtCLC-1 | ER | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | Cl ⁻ transport, ER pH regulation | Sun <i>et al.</i> (2018) |
| 12 | NbCLC-1a/ NbCLC-1b | - | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | - | Sun <i>et al.</i> (2018) |
| 13 | OsCLC-1/ OsCLC-2 | Tonoplast | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | Cl ⁻ transport | Nakamura <i>et al.</i> (2006) |
| 14 | SaCLC-a1 | - | NO ₃ ⁻ /H ⁺ exchanger (putative) | + | + | Proline | NO ₃ ⁻ transport | Nedelyaeva <i>et al.</i> (2019) |
| 15 | SaCLC-c1 | - | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | Cl ⁻ transport | Nedelyaeva <i>et al.</i> (2019) |
| 16 | ZmCLC-c | Mitochondria | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | Cl ⁻ transport | Tampieri <i>et al.</i> (2011) |
| 17 | ZmCLC-d | - | Cl ⁻ /H ⁺ exchanger (putative) | + | + | Serine | - | Yang <i>et al.</i> (2011) |

diverse expression patterns in various tissues, and are acknowledged to perform several functions ranging from nutrient uptake and assimilation to stomatal regulation and are associated with abiotic to biotic stresses.

The Arabidopsis genome possesses seven genes encoding CLCs, named as *AtCLC-a* to *-g*. These *AtCLCs* form two major phylogenetic sub-classes, *AtCLC-a* to *-d* and *-g*, form sub-class I, while *AtCLC-e* and *AtCLC-f* form sub-class II, which is phylogenetically closer to the prokaryotic CLCs (Lv *et al.*, 2009). *AtCLC-a* and

AtCLC-b are electro-physiologically characterized as NO₃⁻/H⁺ exchangers (De Angeli *et al.*, 2006; von der Fecht-Bartenbach *et al.*, 2010), that accordingly possess 'Proline' residue in the anion selectivity filter motif (as mentioned above in the structure section). The remaining *AtCLCs*, as well as the CLCs from other plants, have not been well-characterized electro-physiologically for channel or exchanger function and their selectivity. Hence, functional clues of these remaining plant CLCs for channel or transporter activity, and Cl⁻ or NO₃⁻ selectivity, have primarily been anticipated by

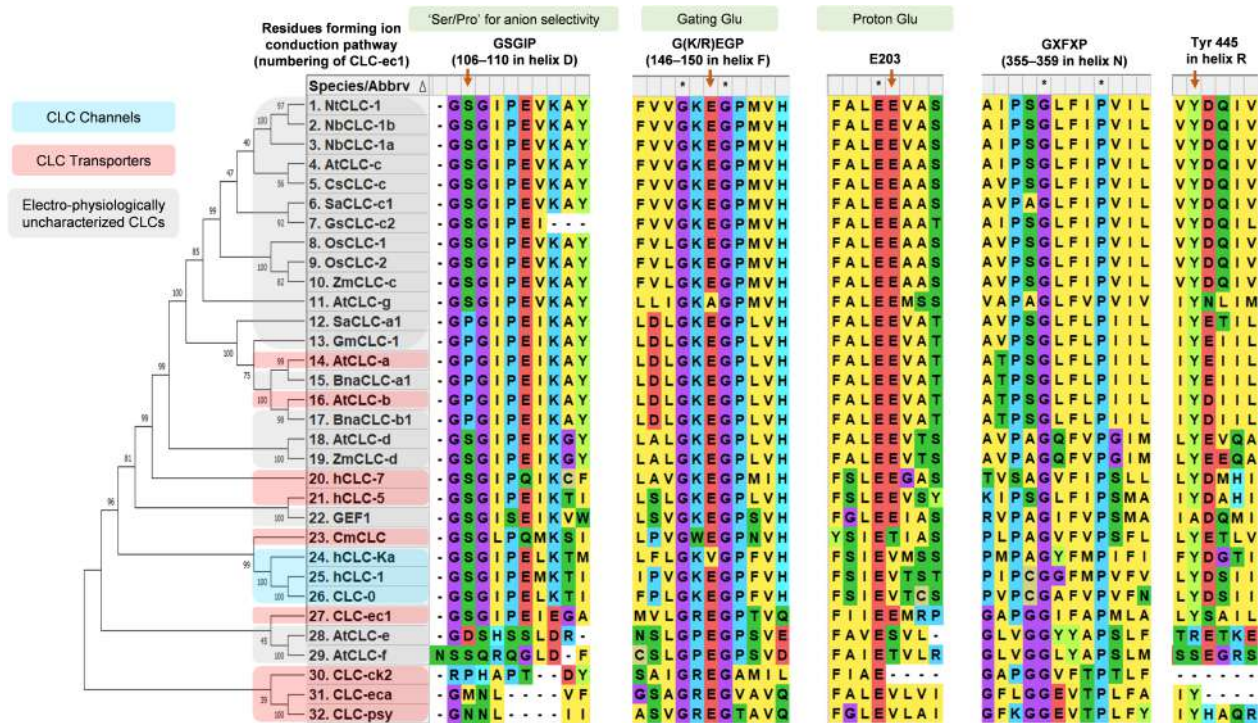


FIGURE 3 Image showing the phylogenetic evolution and the conservation of critical residues that form the ion conduction pore of CLCs reported from different organisms. The electro-physiologically characterized CLC channels are highlighted in blue, CLC transporters in pink while the uncharacterized CLCs are shown in gray. The sequence conservation in four conserved motifs that form the ion conduction pore and the proton glutamate is presented in the sequence alignment. Protein sequences have been taken from *Nicotiana tabacum* (NtCLC), *N. benthamiana* (NbCLC), *Arabidopsis thaliana* (AtCLC), *Poncirus trifoliata* (CsCLC), *Suaeda altissima* (SaCLC), *Glycine soja* (GsCLC), *Oryza sativa* (OsCLC), *Zea mays* (ZmCLC), *Glycine max* (GmCLC), *Brassica napus* (BnaCLC), *Homo sapiens* (hCLC), *Saccharomyces cerevisiae* (GEF1), *Cyanidioschyzon merolae* (CmCLC), *Torpedo marmorata* (CLC-0), *Escherichia coli* (CLC-ec1), *Citrobacter koseri* (CLC-ck2), *Enterococcus casseliflavus* (CLC-eca), *Pseudomonas syringae* (CLC-psy)

sequence analysis for the presence of Gating Glutamate and Proton Glutamate, and the occurrence of Serine or Proline residue in the conserved selectivity filter, respectively (as mentioned above in the structure section). Besides, knock-down studies have unveiled the probable Cl^- or NO_3^- specific functions of plant CLCs and their physiological significance in plants (Table 1). The sequence alignment of conserved residues in plant CLCs concerning the well-characterized CLCs from bacteria, yeast, alga, and humans has been depicted in Figure 3.

Based on the sequence homology, 10 genes encoding CLCs have been identified in *Oryza sativa* (Kushwaha *et al.*, 2009), 17 in *Nicotiana tabacum* (Zhang *et al.*, 2018), 8 in *Camellia sinensis* (Xing *et al.*, 2020), 11 CLCs in *Brassica rapa*, 10 in *B. oleracea*, and 22 in *B. napus* (Liao *et al.*, 2018), 7 in *Suaeda fruticosa* (Nedelyaeva *et al.*, 2018) and 8 in *Glycine max* (Wei *et al.*, 2019). Just as AtCLCs, most of these proteins form a major phylogenetic group (subclass I), while a few CLCs like AtCLC-e and AtCLC-f form a separate group (subclass II) and are closely related to bacterial CLCs. These two subclasses of plant CLC family are suggested to have originated from two distinct ancestral genes (Lv *et al.*, 2009). Through the sequence analysis of two CBS sub-domains at the C-terminus of AtCLCs, Lv *et al.* (2009) observed that CBS1 sub-domains have a higher sequence similarity within each other than the CBS2 sub-domains of two AtCLCs from subclass II,

while the opposite is observed for the AtCLCs in subclass I. The differences in gene expression in these two subclasses have also been reported, with a stronger expression of subclass I AtCLCs than two AtCLCs of subclass II (Lv *et al.*, 2009). Besides, expression of subclass II AtCLCs is observed to be restricted to shoots, while subclass I AtCLCs showed wide expression both in shoots and roots, with AtCLC-b being strongly expressed in roots than the aerial parts (Lv *et al.*, 2009; Jossier *et al.*, 2010). To have a better understanding of their diverse functions and significance, an attempt has been made to concisely outline their relevance in the plant system in the following sections.

3.1 | Plant CLCs for nitrate transport and storage

Nitrate (NO_3^-) is one of the primary forms of mineral nitrogen available to plants that is either stored in the vacuole or reduced to nitrite (NO_2^-) and then to ammonium (NH_4^+) for subsequent assimilation to N-containing organic molecules. Besides, NO_3^- also acts as a signaling molecule and regulates plant metabolism and development (Gojon *et al.*, 2011). In *Arabidopsis*, AtCLC-a and AtCLC-b share high sequence similarity, both are NO_3^-/H^+ antiporters and are localized in the tonoplast (De Angeli *et al.*, 2006; von der Fecht-Bartenbach

et al., 2010). AtCLC-a is the best characterized CLC in plants, critical for nitrate transport into the vacuole. Its expression has been found upregulated in response to nitrate application in both roots and shoots. Accordingly, its mutant, *clc-a*, showed a significant reduction of nitrate concentration in both roots and shoots under a high external nitrate level. Besides, the mutant lines exhibited more sensitivity to the herbicide sodium chlorate, an analog of nitrate (Geelen *et al.*, 2000). This was the first link between CLCs and NO_3^- , different from the non-plant CLC counterparts. The patch-clamp electrophysiological study of AtCLC-a on vacuole isolated from mesophyll cells confirmed AtCLC-a as a NO_3^-/H^+ (2:1) exchanger mediating NO_3^- influx to the vacuole (De Angeli *et al.*, 2006), consistent with the phenotype of reduced nitrate level in *clc-a* mutants. As binding of adenosine nucleotides to CBS domains modulates the activity of some mammalian CLCs (see Altamura *et al.*, 2020; Grieschat *et al.*, 2020), the presence of ATP resulted in a significant reduction in the activity of AtCLC-a, while AMP reduced the ATP-mediated AtCLC-a inhibition (De Angeli *et al.*, 2009). They suggested that AtCLC-a mediates nitrate transport depending upon cellular ATP:AMP ratio. This transport was anticipated to be correlated with the ATP:AMP-regulated nitrate assimilation in plants that takes place in the cytosol during the day, while nitrate is stored into the vacuole at night. Besides, substitution mutation at the CBS domain in AtCLC-a (D753A), corresponding to the ATP binding residue of hCLC-5 (Meyer *et al.*, 2007), eliminated the inhibitory effect of ATP on AtCLC-a, confirming the regulation of AtCLC-a by ATP in nitrate transport (De Angeli *et al.*, 2009).

Sequence comparison identified a proline (P160) in AtCLC-a to confer higher NO_3^- selectivity over Cl^- , which is replaced by serine in CLC isoforms with Cl^- selectivity (Zifarelli and Pusch, 2009). Likewise, substitution (P160S) makes AtCLC-a more conductive for Cl^- than NO_3^- , not altering the H^+ /anion coupling stoichiometry (Bergsdorf *et al.*, 2009; Wege *et al.*, 2010). AtCLC-a (P160S) failed to rescue the defect in nitrate accumulation of *clc-a* knockout Arabidopsis mutants (Wege *et al.*, 2010). However, AtCLC-a (P160S), but not wild type AtCLC-a, partially complemented the growth of *gef1* yeast mutant (as GEF1 also possesses serine in the selectivity filter). Whereas, AtCLC-a, but not AtCLC-a (P160S), rescued the growth of *gef1* on high nitrate condition (Wege *et al.*, 2010). These results support the alteration in anion selectivity of AtCLC-a by the substitution P160S.

AtCLC-a has strong expression in the guard cells, and its mutant (*clc-a*) exhibited a reduced stomatal aperture in response to light illumination and a defect in ABA-responsive stomatal closing (Wege *et al.*, 2014). While the light-responsive accumulation of anions into the vacuole during stomatal opening is associated with its known NO_3^- influx function, the defect in stomata closure indicates that AtCLC-a also executes NO_3^- efflux from the vacuole. An SnRK2 family kinase, AtOST1 (OPEN STOMATA 1 or SnRK2.6), activated by ABA in guard cells, is known to phosphorylate and subsequently activate several plasma membrane anion channels (Geiger *et al.*, 2009; Imes *et al.*, 2013). Wege *et al.* (2014) found that AtOST1 interacts with AtCLC-a in the tonoplast. They identified that OST1 phosphorylates Thr³⁸ at the N-terminal of AtCLC-a, which is found to be essential for anion efflux from the vacuole (Wege *et al.*, 2014). Accordingly, a

mutated AtCLC-a (T38A) overexpressed in *clc-a* mutant rescued the light-induced stomatal opening, but not the ABA-induced stomatal closing, implying that phosphorylation of AtCLC-a Thr³⁸ by OST1 is required for ABA-responsive NO_3^- efflux, leading to stomata closing (Wege *et al.*, 2014). Similarly, the C5 catalytic subunit of protein phosphatase 2A (PP2A-C5) has been found to interact with all tonoplast-localized AtCLCs. It is proposed that PP2A-C5 dephosphorylates these AtCLCs, thereby facilitating anions influx into the vacuole (Hu *et al.*, 2017). Hence, AtCLC-a may perform a reversible function through phosphorylation and dephosphorylation of its Thr³⁸ residue. Another evidence supporting the participation of AtCLC-a in ABA-induced stomata closing came recently from the study on phosphatidylinositol-3,5-bisphosphate (PI[3,5]P₂; Carpaneto *et al.*, 2017). PI(3,5)P₂ is a signaling lipid molecule known to induce vacuolar convulsion and acidification of the lumen, a process that is required for rapid ABA-induced stomatal closure of the guard cells (Bak *et al.*, 2013). In patch-clamp experiments on vacuoles, PI(3,5)P₂ inhibited AtCLC-a activity. Besides, PI(3,5)P₂ treatment decreased the vacuolar pH in WT but not in *clc-a* mutants (Carpaneto *et al.*, 2017). Hence, they suggested that inactivation of CLC-a by PI(3,5)P₂ inhibits the release of vacuolar H^+ for the uptake of NO_3^- , which indirectly leads to acidification of the vacuolar lumen and favors stomatal closing.

AtCLC-b, possessing proline residue in the selectivity filter, is another electro-physiologically characterized NO_3^-/H^+ exchanger. However, unlike *clc-a*, no reduction in nitrate content was observed in *clc-b* mutants (von der Fecht-Bartenbach *et al.*, 2010). Nitrate concentrations vary in plants with the diurnal activity of nitrate reductases that assimilate nitrate in the cytosol during the day, while storing it in vacuole at night (Cookson *et al.*, 2005). Likewise, von der Fecht-Bartenbach *et al.* (2010) observed diurnal variations in the expression of AtCLC-a and AtCLC-b from the gene expression database. AtCLC-b showed peak expression at the beginning of the light period, while AtCLC-a reached a maximum during the dark. Based on gene expression data and electrophysiological results, they suggested that AtCLC-b mediates the NO_3^- release from vacuoles during the day, which is then assimilated in the cytoplasm, while AtCLC-a stores NO_3^- into the vacuole at night.

As in Arabidopsis, *Brassica napus* *BnaCLC-a* members (*BnaCLC-a1-4*; sharing sequence similarity with AtCLC-a) showed strong induction of gene expression in both shoots and roots after NO_3^- replenishment following N starvation. *BnaCLC-b* was induced in the shoots by NO_3^- , while *BnaCLC-c* to -g members were either downregulated or showed no changes to NO_3^- treatment (Liao *et al.*, 2018). Even in *Suaeda altissima*, a euhalophyte plant, the expression of *SaCLC-a1* (that shares sequence similarity with AtCLC-a and AtCLC-b) increased in the roots in response to nitrate-deficiency (Nedelyaeva *et al.*, 2019). *SaCLC-a1* possesses both conserved glutamates (gating glutamate and proton glutamate) and proline in the selectivity filter, suggesting that it may have NO_3^-/H^+ activity as AtCLC-a and AtCLC-b. Although *SaCLC-a1* failed to complement Δ *gef1* mutant, substituting proline with serine (P188S) in the selectivity filter conferred partial growth restoration of Δ *gef1* cells (Nedelyaeva *et al.*, 2019). These studies imply that orthologs of

AtCLC-a and AtCLC-b in other plant species most probably possess a conserved function as NO_3^-/H^+ transporter.

3.2 | Plant CLCs with canonical chloride transport function

In *Arabidopsis*, the anion selectivity filter is occupied by a serine residue in AtCLC-c, AtCLC-d, and AtCLC-g. Besides, AtCLC-c and AtCLC-d have both gating and proton glutamates, indicating a function as Cl^-/H^+ transporter, whereas in AtCLC-g, the proton glutamate is present but the gating glutamate is replaced by alanine, indicating a function as Cl^- channel. AtCLC-c and AtCLC-g share high sequence identity, and both are localized in the tonoplast (Jossier *et al.*, 2010; Nguyen *et al.*, 2016), while AtCLC-d is localized in the trans-Golgi network (von der Fecht-Bartenbach *et al.*, 2007). Upon NaCl treatment, expression of AtCLC-d and AtCLC-g are reported to be induced in the shoots, while AtCLC-c is induced in both shoots and roots (Jossier *et al.*, 2010).

AtCLC-c is located in the major QTL for nitrate accumulation in *Arabidopsis* and its mutant, *clc-c*, accordingly, showed a reduction in nitrate as well as in chloride and citrate content (Harada *et al.*, 2004). However, Jossier *et al.* (2010) did not observe any difference in nitrate level in *clc-c* mutants and wild type (WT), contrasting to the report by Harada *et al.* (2004), which they reasoned to be due to different accessions used in these two studies. Also, the expression of AtCLC-c is down-regulated upon nitrate treatment (Harada *et al.*, 2004). Its expression is high in guard cells and pollens and increases furthermore by ABA treatment in guard cells, while it has weak expression in the roots (Jossier *et al.*, 2010). Accordingly, *clc-c* mutants, alike *clc-a*, exhibit impaired stomatal opening to light and insensitivity to ABA-induced stomatal closing. This defect in stomatal regulation is rescued by application of NO_3^- , but not Cl^- . Also, Cl^- content was significantly less in the guard cells of *clc-c* mutants (Jossier *et al.*, 2010). Thus, they proposed that AtCLC-c confers Cl^- influx into the vacuole and regulates the stomatal opening and that impaired Cl^- accumulation in the vacuole in mutants also affects its efflux from the vacuole, subsequently affecting the stomatal closing as well. Besides, *clc-c* mutants exhibited strong hypersensitivity to Cl^- with a reduction in the shoot as well as root weight (Jossier *et al.*, 2010), while the over-expression of AtCLC-c conferred salt tolerance to the plant (Hu *et al.*, 2017).

Similarly, AtCLC-g mutant (*clc-g*) also showed Cl^- -sensitive growth retardation (Nguyen *et al.*, 2016). The expression levels of both AtCLC-c and AtCLC-g are upregulated in response to NaCl treatment; however, AtCLC-g expression is higher in mesophyll cells but weak in guard cells where AtCLC-c expression is prominent (Jossier *et al.*, 2010; Nguyen *et al.*, 2016). Both *clc-c* and *clc-g* mutants are reported to accumulate higher Cl^- in the shoots on NaCl treatment (Nguyen *et al.*, 2016), while Jossier *et al.* (2010) reported no difference in Cl^- accumulation in the shoots of *clc-c* mutants and WT upon NaCl treatment. Double mutant plants (*clc-c:clc-g*) exhibited no difference from the single mutant of each gene to NaCl treatment (Nguyen

et al., 2016). Contrary to the observation by Jossier *et al.* (2010), the expression of AtCLC-g is repressed in the *clc-c* mutant background in controlled as well as NaCl-treated conditions. Similarly, the expression of AtCLC-c is also decreased in *clc-g* mutants but only in NaCl-treated conditions (Nguyen *et al.*, 2016). With these results, they suggested that AtCLC-g, functioning in mesophyll cells, is involved in Cl^- sequestration into the vacuole, and its gene expression coordinates with the expression of AtCLC-c in guard cells to regulate stomata. In mature leaves, AtCLC-g expression has also been observed in phloem cells and hydathodes and is hypothesized to prevent Cl^- over-accumulation in photosynthetically active leaves by partitioning the Cl^- into different organs via the phloem and by excreting it through hydathodes, respectively (Nguyen *et al.*, 2016).

In *Suaeda altissima*, the expression of SaCLC-c1 (homolog of AtCLC-c) increases in leaves with a rise in Cl^- concentrations (Nedelyaeva *et al.*, 2019). SaCLC-c1 possesses both gating and proton glutamates, and serine in the selectivity filter, indicating its role as Cl^-/H^+ antiporters (Nedelyaeva *et al.*, 2019). When expressed in *gef1* yeast mutant, SaCLC-c1 complemented the iron deficient, pH-sensitive, as well as salt-sensitive growth phenotype of the mutant, supporting SaCLC-c1 as a chloride transporter (Nedelyaeva *et al.*, 2018, 2019). However, AtCLC-c does not suppress the growth phenotype of the *gef1* yeast mutant (Hechenberger *et al.*, 1996). In rice, Nakamura *et al.* (2006) isolated two CLCs, OsCLC-1, and OsCLC-2, and the proteins encoded by both genes are localized in the tonoplast and share sequence similarity with AtCLC-c. These genes partially rescued the *gef1* yeast mutant to various salts and higher pH, indicating their possible role in chloride transport. Also, the expression of OsCLC-1 was found to be increased by NaCl treatment (Diédhiou and Gollmack, 2006; Nakamura *et al.*, 2006). The *Tos17* insertional mutant lines of both genes, *Osclc-1* and *Osclc-2*, exhibited a reduction in height at all plant stages (Nakamura *et al.*, 2006).

The homolog of AtCLC-c in tobacco, NtCLC-1, is proposed to act as a Cl^-/H^+ transporter based on conserved residues in the selectivity filter (Sun *et al.*, 2018). Two proteins present in *Nicotiana benthamiana*, NbCLC-1a, and NbCLC-1b, are highly homologous to NtCLC-1 and the knockout of genes encoding these two proteins slightly lowered the plant height and produced smaller flowers, which was fully rescued by the expression of either *NtCLC-1*, *NbCLC-1a*, or *NbCLC-1b*. However, the functional complementation was not successful by overexpressing *NtCLC-1* mutated at the conserved gating and proton glutamates, supporting its function as Cl^-/H^+ transporter (Sun *et al.*, 2018). NtCLC-1 is localized in the ER and, in *nbcl-1a:nbcl-1b* knockout lines, the pH of ER lumen was slightly lower (7.05 ± 0.13) than WT (7.37 ± 0.14), which was restored by the expression of *NtCLC-1* (Sun *et al.*, 2018). Additionally, among the 17 CLCs identified in tobacco, the expression of *NtCLC-2*, *NtCLC-3*, and *NtCLC-12* is induced in both root and shoot upon NaCl treatment, while the expression of *NtCLC-3*, *NtCLC-13*, and *NtCLC-15* is highly upregulated in shoot. In *NtCLC-2*-silenced plants, the Cl^- concentration was reduced, but it was increased in *NtCLC-13*-silenced plants (Zhang *et al.*, 2018).

In soybean (*Glycine max*) and its wild relative (*Glycine soja*), GmCLC-1 and GsCLC-c2, respectively, have been reported to be

involved in Cl^- transport (Li *et al.*, 2006; Wei *et al.*, 2019). GmCLC-1 is reported to accumulate Cl^- into the vacuole (Li *et al.*, 2006). Wong *et al.* (2013) observed a pH-dependent Cl^- conductance of GmCLC-1 (higher Cl^- current at pH 7.5 than at pH 5.5) by voltage-clamp experiment in *Xenopus* oocytes. Due to the presence of conserved gating and proton glutamate residues in GmCLC-1, they suggested its functions as a CLC antiporter. Also, its expression is strongly induced in roots upon Cl^- treatment (Wong *et al.*, 2013). The GmCLC-1 expression is reported to enhance the growth of Δgef1 yeast mutant in salt-containing medium (Wei *et al.*, 2016). However, it is contrasting to the fact that GmCLC-1 possesses a proline residue in the anion selectivity filter and clusters with AtCLC-a and AtCLC-b (NO_3^-/H^+ exchangers) in the phylogenetic branch (Figure 3). Moreover, GmCLC-1 showed equivalent NO_3^- conductance when expressed in *Xenopus* oocytes (Wei *et al.*, 2019), indicating it is more likely a NO_3^-/H^+ exchanger. GsCLC-c2 from *G. soja* clusters with AtCLC-c and AtCLC-g, possesses a serine in the selectivity filter and has both gating and proton glutamates, indicating it is a probable Cl^-/H^+ exchanger. Accordingly, its expression in *Xenopus* oocytes showed higher Cl^- conductance over NO_3^- , and it also rescued the growth of Δgef1 yeast mutant under Cl^- stress conditions (Wei *et al.*, 2019).

The trans-Golgi localized AtCLC-d exhibits a wide expression pattern, relatively strong in seedling roots (von der Fecht-Bartenbach *et al.*, 2007). It fully complemented the iron-sensitive growth phenotype of Δgef1 yeast mutant, indicating its function as a Cl^-/H^+ transporter (Hechenberger *et al.*, 1996). Its mutant, *clc-d*, showed a defect in root growth, while no difference in anion content was observed (von der Fecht-Bartenbach *et al.*, 2007). Similarly, partial loss of the trans-Golgi network-localized proton transporting V-ATPase activity has been associated with shorter root length and reduction in cell elongation (Schumacher *et al.*, 1999; Padmanaban *et al.*, 2004). Furthermore, treatment of plants with concanamycin A, a chemical that blocks V-ATPase, led to shorter hypocotyls in *clc-d* mutants as compared to WT, which was rescued by expressing AtCLC-d in the mutant (von der Fecht-Bartenbach *et al.*, 2007). It was suggested that AtCLC-d maintains V-ATPase-mediated acidification of the trans-Golgi network through an anion influx that neutralizes the positive charges (von der Fecht-Bartenbach *et al.*, 2007), as proposed for the vesicle-localized mammalian CLC transporters (See Jentsch, 2015).

3.3 | Plant CLCs involved in abiotic stress tolerance

Plant membrane transport proteins play a critical role in abiotic stresses, particularly in salinity stress, by either salt exclusion or vacuolar compartmentation (Gupta *et al.*, 2018; Nongpiur *et al.*, 2020). Studies on salt stress in plants have mostly been concerned with the effect of Na^+ , while the toxic effect of the associated Cl^- has been least emphasized (Joshi *et al.*, 2018). Plant CLCs, particularly Cl^- -specific CLCs, appear as an essential candidate to improve the salt tolerance by sequestering Cl^- in vacuoles. AtCLC-c overexpression is reported to enhance the tolerance of Arabidopsis to salt stress (Hu *et al.*, 2017),

while its disruption makes the plant hypersensitive to Cl^- (Jossier *et al.*, 2010). Hu *et al.* (2017) identified the interaction of AtCLC-c (cytoplasmic N-ter and C-ter) with a C5 catalytic subunit of the protein phosphatase 2A (PP2A-C5), overexpression of which also improved plant growth under salt condition, while its mutant showed a salt-sensitive phenotype. PP2A-C5 appears to function upstream of AtCLC-c since the overexpression of AtCLC-c in *pp2a-c5* mutant could not rescue its salt-sensitive phenotype. Besides, PP2A-C5 also showed interactions with the other three tonoplast-localized AtCLCs (AtCLC-a, -b, and -g; Hu *et al.*, 2017). They proposed that dephosphorylation by PP2A-C5 might activate the vacuolar CLCs transporting anions into the vacuoles, thus leading to increased salt tolerance.

In soybean, the wild species (*Glycine soja*) has been shown to better tolerate NaCl stress than the cultivated one (*Glycine max*) by restricting the Cl^- accumulation in the shoots (Zhang *et al.*, 2011). Genes encoding tonoplast-localized GsCLC-c2 isoforms are found to be differentially expressed and bear differences at the sequence level in these two species (Wei *et al.*, 2019). The overexpression of GsCLC-c2 conferred salt tolerance to plants by sequestering Cl^- into the vacuole in roots, thus, lowering the Cl^- level in the shoots (Wei *et al.*, 2019). Another gene in cultivated soybean, GmCLC-1, is induced by NaCl, polyethylene glycol (PEG), cold, and ABA treatments (Li *et al.*, 2006; Zhou and Qiu, 2010). Its overexpression in tobacco BY-2 cell lines conferred tolerance against NaCl treatment with higher Cl^- accumulation into the vacuole (Li *et al.*, 2006). Its gene expression as well as Cl^- conductance activity is found to be pH-dependant (Wong *et al.*, 2013). Furthermore, the overexpression of GmCLC1 conferred salt tolerance in soybean hairy roots as well as in Arabidopsis and poplar, by reducing the accumulation of Cl^- in shoots as a result of compartmentation of Cl^- into the vacuole of the roots (Zhou and Qiu, 2010; Sun *et al.*, 2013; Wei *et al.*, 2016). However, as mentioned in the previous section, the conserved residues in the selectivity filter and the phylogenetic clustering justifies for GmCLC1 being a NO_3^-/H^+ exchanger (Figure 3).

CsCLC-c (that shares high identity with AtCLC-c) from trifoliate orange (*Poncirus trifoliata*) also exhibited gene induction upon ABA, cold, and NaCl treatment. Its overexpression in Arabidopsis AtCLC-c mutant, *clc-c*, improved seedling growth, over the WT level, under salinity stress, with a reduction in Cl^- content in both roots and shoots in overexpressing lines (Wei *et al.*, 2013). In rice, the expression of OsCLC-1 in the salt-sensitive variety 'IR29' and the salt-tolerant variety 'Pokkali' was equivalent under control condition. Under NaCl treatment, its expression was transiently increased but reduced subsequently in the shoots of both varieties. However, this decrease was slow in Pokkali. In roots, its expression under NaCl stress decreased in IR29, whereas its level increased in Pokkali (Diédhiou and Gollmack, 2006). Accordingly, both varieties exhibited equivalent Cl^- accumulation under normal conditions. The NaCl treatment increased Cl^- accumulation in shoots and roots of both varieties; however, it was significantly less in Pokkali shoots (Diédhiou and Gollmack, 2006).

The maize CLCs, *ZmCLC-c*, and *ZmCLC-d* are induced by cold treatment, and the expression of *ZmCLC-c* is reported higher in the cold-tolerant population under control conditions (Tampieri

et al., 2011; Yang et al., 2011). ZmCLC-c is localized in the mitochondria where Cl^- and I^- fluxes were inhibited when treated with anti-ZmCLC-c antibody, suggesting that ZmCLC-c is a plant inner membrane anion channel (PIMAC; Tampieri et al., 2011). The induction of ZmCLC-d expression, in addition to cold, has also been observed in drought, heat, NaCl, ABA, and H_2O_2 treatments (Wang et al., 2015). Its overexpression in Arabidopsis is reported to enhance tolerance towards cold, drought, and salt stresses (Wang et al., 2015).

3.4 | Plant CLCs involved in plant defense responses

Many studies have documented the involvement of anion channels in plant defense responses (see De Angeli et al., 2007; Jeworutzki et al., 2010). Plant CLCs are also found to be involved in plant-pathogen interactions. Among AtCLC mutants (except *clc-f*), Guo et al. (2014) identified that *clc-d* mutants have an enhanced response to flg22, a conserved N-terminal region of flagellin acting as an elicitor of defense response like ROS production, callose deposition and expression of pathogen-associated marker genes (Felix et al., 1999). In contrast, AtCLC-d overexpressing lines were less responsive to flg22, indicating that AtCLC-d negatively regulates the pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI) (Guo et al., 2014). When inoculated with *Pseudomonas syringae* pv. tomato DC3000, *clc-d* mutants showed resistance, while AtCLC-d overexpressing lines were susceptible. Interestingly, the expression of AtCLC-d, as well as all other AtCLCs, was down-regulated by flg22 treatment (Guo et al., 2014). In *fls2* mutant plants (FLS2 is a plant cell surface pattern recognition receptor [PRR] for flg22 that triggers PTI), the expression of AtCLC-d remained unaffected by flg22. Besides, as compared to WT, the expression of AtCLC-d was higher in the *fls2* mutant and was reduced in FLS2 overexpressing line, indicating that the inhibition of AtCLC-d expression requires pre-recognition of flg22 by FLS2 (Guo et al., 2014). Correlating with the role of AtCLC-d in the acidification of trans-Golgi network (von der Fecht-Bartenbach et al., 2007), Guo et al. (2014) proposed that AtCLC-d probably affects the FLS2 endocytic transport and trafficking via the trans-Golgi network, critical for its PTI function.

Tobacco NtCLC-1 is found to interact and co-localize with the POTATO VIRUS Y (PVY) 6 K2 protein (Sun et al., 2018). Interaction of PVY 6 K2 with NtCLC-1 resulted in the alkalization of the ER, which was found to be necessary for viral intracellular replication and systemic infection (Sun et al., 2018). In tomato, CLC-b-like mRNA is identified as a target of Potato Spindle Tuber Viroid (PSTVd)-derived small RNA (vd-sRNA). The PSTVd infection led to the cleavage of CLC-b-like mRNA and a significant reduction in its transcript. Accordingly, the VIGS-mediated RNA silenced plants exhibited a stunting and leaf curling phenotype like viroid infection (Adkar-Purushothama et al., 2017).

Sasvari et al. (2013) observed impaired replication of Tomato Bushy Stunt Virus (TBSV) in *gef1* yeast mutant, which was found to be due to higher Cu^{2+} accumulation in the cytosol, an indirect effect of *gef1* mutant, that mainly affected the assembly of the viral replicase

complex (VRC). Similarly, chemical blocking of CLC in *Nicotiana benthamiana* also reduced TBSV RNA accumulation in both protoplasts and leaves of the whole plant (Sasvari et al., 2013), which was suggested to be due to cellular ionic imbalance caused by an alteration in Cl^- homeostasis. Even the human viruses are reported to induce chloride influx into the cells, and blocking of chloride channels or silencing of hCLCs has been shown to inhibit the viral replication (Zheng et al., 2014; Igloi et al., 2015). The CLCs of plant microbes, both fungi, and bacteria, also appear essential for infecting plants, as disruption of their CLCs are reported to cause a significant decrease in their virulence (Rojas-Jiménez et al., 2005; Cañero and Roncero, 2008).

4 | PHYLOGENETICALLY DISTINCT MEMBERS OF PLANT CLCS

Members of plant CLCs that are homologs of Arabidopsis AtCLC-e and AtCLC-f form a separate phylogenetic group and exhibit close identity with the prokaryotic CLCs. Both AtCLC-e and AtCLC-f possess conserved gating glutamate, but the proton glutamate in AtCLC-e and AtCLC-f is replaced by similar yet different residues, serine, and threonine, respectively. Proton glutamate in CmCLC, a Cl^-/H^+ exchanger from a thermophilic red alga (*C. merolae*), is also substituted with threonine (Feng et al., 2010), and the higher identity of these two AtCLCs with bacterial CLCs implies their exchanger function. Exceptionally, the anion selectivity filter is occupied by an aspartic acid in AtCLC-e, while serine is present in AtCLC-f, indicative of its function in Cl^- transport (Figure 3).

AtCLC-e is localized in the thylakoid membrane in chloroplasts. Illumination of dark-adapted leaves of its mutant, *clc-e*, exhibited an altered chlorophyll fluorescence, which is proposed to be due to altered ion homeostasis in the lumen of the thylakoid (Marmagne et al., 2007). Herdean et al. (2016) also observed lower photosynthetic performance of the *clc-e* mutants in the dark-adapted state. The mutants exhibited a mild change in proton motive force, and the photosynthetic electron transport of dark-adapted plants based on Chl *a* fluorescence induction (OJIP) was affected (Herdean et al., 2016). However, pre-treatment with KCl but not KNO_3 rescued the change in OJIP kinetics in mutants. Transmission electron microscopy analyses displayed a defect in the thylakoid network with a sizeable stromal region in the chloroplast of dark-adapted mutants, while no difference was observed in the chloroplast of mutants and WT plants adapted to the light. Thus, Herdean et al. (2016) proposed that AtCLC-e functions most probably in Cl^- efflux from the thylakoid to the stroma of the chloroplast after transition from light to dark. In contrast, Monachello et al. (2009) suggested the role of AtCLC-e in nitrate transport. They observed reduced nitrate accumulation in *clc-e* mutants, as observed by Geelen et al. (2000) in *clc-a* mutants. Both mutants, *clc-a* and *clc-e*, and their double mutant (*clc-a clc-e*), showed a reduction in net nitrate influx in roots, while the nitrite level was higher in these mutants as compared to WT. Accordingly, the expression of nitrate transporter genes, AtNRT2.1 and AtNRT1.1, was

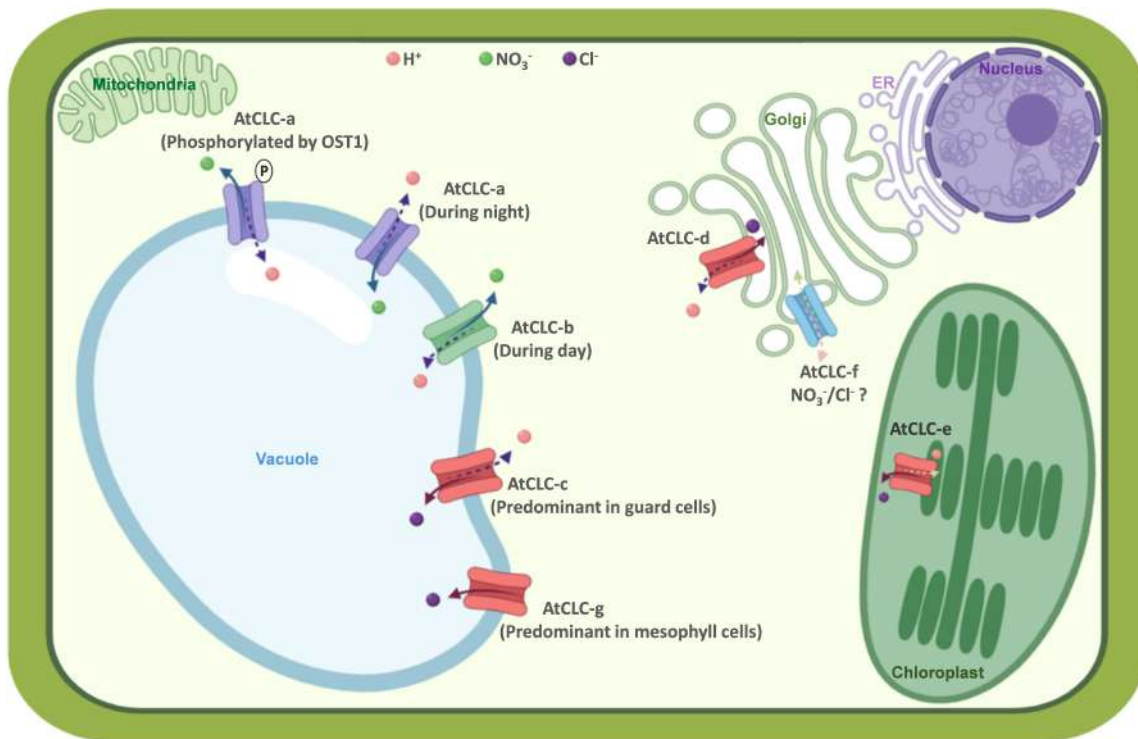


FIGURE 4 Schematic representation, showing localization and functions (established or putative) of all reported CLCs from Arabidopsis. Out of seven AtCLCs, four are localized in the vacuole, of which AtCLC-a and AtCLC-b have been reported to function as NO_3^- exchangers, while AtCLC-c and AtCLC-g are suggested to be involved in Cl^- transport. AtCLC-d and AtCLC-f are localized in the Golgi, of which AtCLC-d is known to regulate Golgi pH through Cl^- transport. AtCLC-e is localized in the thylakoid membrane and is reported to carry out Cl^- transport, important for efficient photosynthesis. The figure has been created with BioRender.com

strongly reduced in the roots of both *clc-a* and *clc-e* mutants, which was suggested to be due to an increasing nitrite level (Monachello *et al.*, 2009). They observed a significant reduction in the expression of *AtCLC-e* in *clc-a* mutants, while a slight increase in *AtCLC-a* transcripts was observed in *clc-e* mutants. With these results, Monachello *et al.* (2009) proposed that *AtCLC-e* functions in NO_3^- translocation into thylakoid lumen, disruption of which directly or indirectly increase nitrite level in the cytosol, affecting the cellular NO_3^- level, and that *AtCLC-e* and *AtCLC-a* may have an interconnecting function in nitrate assimilation. However, Herdean *et al.* (2016) inferred the study by Monachello *et al.* (2009) as an indirect effect due to possible alteration in the expression of NO_3^-/H^+ exchangers and instead emphasized *AtCLC-e*'s role in Cl^- transport.

The current knowledge on *AtCLC-f* is lacking, except that it is localized in Golgi vesicles and it can complement the iron as well as pH-sensitive growth of *Δgef1* yeast mutant (Marmagne *et al.*, 2007). However, its ortholog in spinach has been reported to localize in the outer membrane of the chloroplast (Teardo *et al.*, 2005). Consistent with the chloroplast localization, this protein is found to be expressed in shoots, but not roots. Besides, the treatment with p-chlorophenoxy-acetic acid (an inhibitor of CLC) impaired photosynthetic activity as visible by a reduction in oxygen evolution and altered F_v/F_m in isolated chloroplast and intact plant, respectively (Teardo *et al.*, 2005). Therefore, homologs of *AtCLC-e* and *-f* appear to be

localized in the chloroplast and maintain the ionic homeostasis suitable for photosynthetic reaction.

5 | CONCLUSION AND PERSPECTIVE

Plant CLCs carry out heterogeneous cellular functions. Intriguingly, all (as reported till date) show organellar localization but none to the plasma membrane. In case of animals, CLCs with exchanger functions are localized on intracellular vesicles, while CLC channels are localized on the plasma membrane. Besides, plant CLCs share more identity with the known mammalian CLC exchangers than mammalian CLC channels. These observations imply that plant CLCs may function as exchanger proteins and not channels. *AtCLC-a* and *AtCLC-b* are electro-physiologically characterized to function as NO_3^-/H^+ exchanger in Arabidopsis, while remaining CLCs are yet to be confirmed for their channel or exchanger functions. Although the analysis of conserved glutamate residues in the selectivity filter motifs suggests that plant CLCs include both channels and exchangers, the deviations from these conserved residues have also been reported in some CLCs.

The functions of plant CLCs appear conserved but non-redundant, throughout diverse plant species. Although some CLCs localize in the same organelle within the same cell, they appear to perform different functions based on the direction of anion transport, anion

selectivity, diurnal regulation, so on. (Figure 4). So far, plant salinity stress studies have remained focused on either sequestration or removal of cations and have earned less attention about the anions. Plant CLCs are shown to sequester Cl^- into the vacuole, and such CLCs, if expressed along with the proteins transporting cation into the vacuole (such as NHX exchangers), would facilitate efficient salt sequestration, while maintaining the vacuolar electroneutrality. Also, some members are suggested to excrete Cl^- by guttation, which may also enhance salt stress tolerance in plants. Lastly, the NO_3^- -specific plant CLCs could be interesting candidates for improving plant nitrogen use efficiency, which has become an economical as well as an environmental issue.

ACKNOWLEDGMENT

SLS-P acknowledge the facilities and support received from ICGER. AS and ST acknowledge the Department of Biotechnology, and University Grants Commission, Government of India, respectively, for providing research fellowship during their research work.

AUTHOR CONTRIBUTIONS

SLS-P conceived the study, AS and ST wrote the manuscript, and SLS-P and AP finalized the manuscript.

DATA AVAILABILITY STATEMENT

Data sharing does not apply to this article, as no new data were created or analyzed in this study.

ORCID

Ashwani Pareek  <https://orcid.org/0000-0002-2923-0681>

Sneh L. Singla-Pareek  <https://orcid.org/0000-0002-0521-2622>

REFERENCES

- Accardi, A. & Miller, C. (2004) Secondary active transport mediated by a prokaryotic homologue of CIC Cl^- channels. *Nature*, 427, 803–807.
- Accardi, A., Walden, M., Nguiragool, W., Jayaram, H., Williams, C. & Miller, C. (2005) Separate ion pathways in a Cl^-/H^+ exchanger. *The Journal of General Physiology*, 126, 563–570.
- Adkar-Purushothama, C.R., Iyer, P.S. & Perreault, J.P. (2017) Potato spindle tuber viroid infection triggers degradation of chloride channel protein CLC-b-like and ribosomal protein S3a-like mRNAs in tomato plants. *Scientific Reports*, 7, 8341.
- Altamura, C., Desaphy, J.F., Conte, D., De Luca, A. & Imbrici, P. (2020) Skeletal muscle CIC-1 chloride channels in health and diseases. *Pflügers Archiv*, 472, 961–975. <https://doi.org/10.1007/s00424-020-02376-3>.
- Ashcroft, F., Gadsby, D. & Miller, C. (2009) Introduction. The blurred boundary between channels and transporters. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364, 145–147.
- Bak, G., Lee, E.J., Lee, Y., Kato, M., Segami, S., Sze, H., et al. (2013) Rapid structural changes and acidification of guard cell vacuoles during stomatal closure require phosphatidylinositol 3,5-bisphosphate. *Plant Cell*, 25, 2202–2216.
- Baker, J.L., Sudarsan, N., Weinberg, Z., Roth, A., Stockbridge, R.B. & Breaker, R.R. (2012) Widespread genetic switches and toxicity resistance proteins for fluoride. *Science*, 335, 233–235.
- Bergsdorf, E.Y., Zdebik, A.A. & Jentsch, T.J. (2009) Residues important for nitrate/proton coupling in plant and mammalian CLC transporters. *The Journal of Biological Chemistry*, 284, 11184–11193.
- Brammer, A.E., Stockbridge, R.B. & Miller, C. (2014) F^-/Cl^- selectivity in CLCF-type F^-/H^+ antiporters. *The Journal of General Physiology*, 144, 129–136.
- Cañero, D.C. & Roncero, M.I.G. (2008) Influence of the chloride channel of *Fusarium oxysporum* on extracellular laccase activity and virulence on tomato plants. *Microbiology*, 154, 1474–1481.
- Carpaneto, A., Boccaccio, A., Lagostena, L., Di Zanni, E. & Scholz-Starke, J. (2017) The signaling lipid phosphatidylinositol-3,5-bisphosphate targets plant CLC-a anion/ H^+ exchange activity. *EMBO Reports*, 18, 1100–1107.
- Cookson, S.J., Williams, L.E. & Miller, A.J. (2005) Light-dark changes in cytosolic nitrate pools depend on nitrate reductase activity in *Arabidopsis* leaf cells. *Plant Physiology*, 138, 1097–1105.
- dan Lv, Q., jie, T.R., Liu, H., Gao, X., Li, Y., Zheng, H., et al. (2009) Cloning and molecular analyses of the *Arabidopsis thaliana* chloride channel gene family. *Plant Science*, 176, 650–661.
- Davis-Kaplan, S.R., Askwith, C.C., Bengtzen, A.C., Radisky, D. & Kaplan, J. (1998) Chloride is an allosteric effector of copper assembly for the yeast multicopper oxidase Fet3p: an unexpected role for intracellular chloride channels. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 13641–13645.
- De Angeli, A., Monachello, D., Ephritikhine, G., Frachisse, J.M., Thomine, S., Gambale, F., et al. (2006) The nitrate/proton antiporter AtCLCa mediates nitrate accumulation in plant vacuoles. *Nature*, 442, 939–942.
- De Angeli, A., Moran, O., Wege, S., Filleur, S., Ephritikhine, G., Thomine, S., et al. (2009) ATP binding to the C terminus of the *Arabidopsis thaliana* nitrate/proton antiporter, AtCLCa, regulates nitrate transport into plant vacuoles. *The Journal of Biological Chemistry*, 284, 26526–26532.
- De Angeli, A., Thomine, S., Frachisse, J.M., Ephritikhine, G., Gambale, F. & Barbier-Brygoo, H. (2007) Anion channels and transporters in plant cell membranes. *FEBS Letters*, 581, 2367–2374.
- Diédhiou, C.J. & Golldack, D. (2006) Salt-dependent regulation of chloride channel transcripts in rice. *Plant Science*, 170, 793–800.
- Dutzler, R., Campbell, E.B., Cadene, M., Chait, B.T. & MacKinnon, R. (2002) X-ray structure of a CIC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity. *Nature*, 415, 287–294.
- Dutzler, R., Campbell, E.B. & MacKinnon, R. (2003) Gating the selectivity filter in CLC chloride channels. *Science*, 300, 108–112.
- Felix, G., Duran, J.D., Volko, S. & Boller, T. (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant Journal*, 18, 265–276.
- Feng, L., Campbell, E.B., Hsiung, Y. & MacKinnon, R. (2010) Structure of a eukaryotic CLC transporter defines an intermediate state in the transport cycle. *Science*, 330, 635–641.
- Foster, J.W. (2004) *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nature Reviews Microbiology*, 2, 898–907.
- Gaxiola, R.A., Yuan, D.S., Klausner, R.D. & Fink, G.R. (1998) The yeast CLC chloride channel functions in cation homeostasis. *Proceedings of the National Academy of Sciences of the United States of America*, 95, 4046–4050.
- Geelen, D., Lurin, C., Bouchez, D., Frachisse, J.M., Lelièvre, F., Courtial, B., et al. (2000) Disruption of putative anion channel gene AtCLC-a in *Arabidopsis* suggests a role in the regulation of nitrate content. *The Plant Journal*, 21, 259–267.
- Geiger, D., Scherzer, S., Mumm, P., Stange, A., Marten, I., Bauer, H., et al. (2009) Activity of guard cell anion channel SLAC1 is controlled by drought-stress signaling kinase-phosphatase pair. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 21425–21430.

- Gojon, A., Krouk, G., Perrine-Walker, F. & Laugier, E. (2011) Nitrate transceptor(s) in plants. *Journal of Experimental Botany*, 62, 2299–2308.
- Graves, A.R., Curran, P.K., Smith, C.L. & Mindell, J.A. (2008) The Cl^-/H^+ antiporter CLC-7 is the primary chloride permeation pathway in lysosomes. *Nature*, 453, 788–792.
- Grieschat, M., Guzman, R.E., Langschwager, K., Fahlke, C. & Alekov, A.K. (2020) Metabolic energy sensing by mammalian CLC anion/proton exchangers. *EMBO Reports*, 21(6), e47872. <https://doi.org/10.15252/embr.201947872>.
- Guo, W., Zuo, Z., Cheng, X., Sun, J., Li, H., Li, L., et al. (2014) The chloride channel family gene CLCd negatively regulates pathogen-associated molecular pattern (PAMP)-triggered immunity in Arabidopsis. *Journal of Experimental Botany*, 65, 1205–1215.
- Gupta, B.K., Sahoo, K.K., Ghosh, A., Tripathi, A.K., Anwar, K., Das, P., et al. (2018) Manipulation of glyoxalase pathway confers tolerance to multiple stresses in rice. *Plant, Cell & Environment*, 41, 1186–1200.
- Guzman, R.E., Grieschat, M., Fahlke, C. & Alekov, A.K. (2013) CLC-3 is an intracellular chloride/proton exchanger with large voltage-dependent nonlinear capacitance. *ACS Chemical Neuroscience*, 4, 994–1003.
- Harada, H., Kuromori, T., Hirayama, T., Shinozaki, K. & Leigh, R.A. (2004) Quantitative trait loci analysis of nitrate storage in Arabidopsis leading to an investigation of the contribution of the anion channel gene, AtCLC-c to variation in nitrate levels. *Journal of Experimental Botany*, 55, 2005–2014.
- Hechenberger, M., Schwappach, B., Fischer, W.N., Frommer, W.B., Jentsch, T.J. & Steinmeyer, K. (1996) A family of putative chloride channels from Arabidopsis and functional complementation of a yeast strain with a CLC gene disruption. *The Journal of Biological Chemistry*, 271, 33632–33638.
- Herdean, A., Nziengui, H., Zsiros, O., Solymosi, K., Garab, G., Lundin, B., et al. (2016) The Arabidopsis thylakoid chloride channel AtCLCe functions in chloride homeostasis and regulation of photosynthetic electron transport. *Frontiers in Plant Science*, 7, 115.
- Hu, R., Zhu, Y., Wei, J., Chen, J., Shi, H., Shen, G., et al. (2017) Over-expression of PP2A-C5 that encodes the catalytic subunit 5 of protein phosphatase 2A in Arabidopsis confers better root and shoot development under salt conditions. *Plant, Cell & Environment*, 40, 150–164.
- Igloi, Z., Mohl, B.-P., Lippiat, J.D., Harris, M. & Mankouri, J. (2015) Requirement for chloride channel function during the hepatitis C virus life cycle. *Journal of Virology*, 89, 4023–4029.
- Imes, D., Mumm, P., Böhm, J., Al-Rasheid, K.A., Marten, I., Geiger, D., et al. (2013) Open stomata 1 (OST1) kinase controls R-type anion channel QUAC1 in Arabidopsis guard cells. *The Plant Journal*, 74, 372–382.
- Jentsch, T.J. (2015) Discovery of CLC transport proteins: cloning, structure, function and pathophysiology. *The Journal of Physiology*, 593, 4091–4109.
- Jentsch, T.J. & Pusch, M. (2018) CLC chloride channels and transporters: structure, function, physiology, and disease. *Physiological Reviews*, 98, 1493–1590.
- Jentsch, T.J., Steinmeyer, K. & Schwarz, G. (1990) Primary structure of *Torpedo marmorata* chloride channel isolated by expression cloning in *Xenopus* oocytes. *Nature*, 348, 510–514.
- Jeworutzki, E., Roelfsema, M.R., Anschütz, U., Krol, E., Elzenga, J.T., Felix, G., et al. (2010) Early signaling through the Arabidopsis pattern recognition receptors FLS2 and EFR involves Ca-associated opening of plasma membrane anion channels. *The Plant Journal*, 62, 367–378.
- Joshi, R., Sahoo, K.K., Tripathi, A.K., Kumar, R., Gupta, B.K., Pareek, A., et al. (2018) Knock-down of an inflorescence meristem-specific cytokinin oxidase–OsCKX2 in rice reduces yield penalty under salinity stress condition. *Plant, Cell & Environment*, 41, 936–946.
- Jossier, M., Kroniewicz, L., Dalmas, F., Le Thiec, D., Ephritikhine, G., Thomine, S., et al. (2010) The Arabidopsis vacuolar anion transporter, AtCLCc, is involved in the regulation of stomatal movements and contributes to salt tolerance. *The Plant Journal*, 64, 563–576.
- Kushwaha, H.R., Singh, A.K., Sopory, S.K., Singla-Pareek, S.L. & Pareek, A. (2009) Genome wide expression analysis of CBS domain containing proteins in *Arabidopsis thaliana* (L.) Heynh and *Oryza sativa* L. reveals their developmental and stress regulation. *BMC Genomics*, 10, 1–22.
- Last, N.B., Stockbridge, R.B., Wilson, A.E., Shane, T., Kolmakova-Partensky, L., Koide, A., et al. (2018) A CLC-type F^-/H^+ antiporter in ion-swapped conformations. *Nature Structural & Molecular Biology*, 25, 601–606.
- Leisle, L., Ludwig, C.F., Wagner, F.A., Jentsch, T.J. & Stauber, T. (2011) CLC-7 is a slowly voltage-gated $2\text{Cl}^-/1\text{H}^+$ -exchanger and requires Ostm1 for transport activity. *The EMBO Journal*, 30, 2140–2152.
- Leisle, L., Xu, Y., Fortea, E., Lee, S., Galpin, J.D., Vien, M., et al. (2020) Divergent Cl^- and H^+ pathways underlie transport coupling and gating in CLC exchangers and channels. *eLife*, 9, 51224.
- Li, W.Y., Wong, F.L., Tsai, S.N., Phang, T.H., Shao, G. & Lam, H.M. (2006) Tonoplast-located GmCLC1 and GmNHX1 from soybean enhance NaCl tolerance in transgenic bright yellow (BY)-2 cells. *Plant, Cell & Environment*, 29, 1122–1137.
- Liao, Q., Zhou, T., Yao, J.Y., Han, Q.F., Song, H.X., Guan, C.Y., et al. (2018) Genome-scale characterization of the vacuole nitrate transporter Chloride Channel (CLC) genes and their transcriptional responses to diverse nutrient stresses in allotetraploid rapeseed. *PLoS One*, 13, e0208648.
- Lísal, J. & Maduke, M. (2008) The CLC-0 chloride channel is a “broken” Cl^-/H^+ antiporter. *Nature Structural & Molecular Biology*, 15, 805–810.
- Lurin, C., Geelen, D., Barbier-Brygoo, H., Guern, J. & Maurel, C. (1996) Cloning and functional expression of a plant voltage-dependent chloride channel. *The Plant Cell*, 8, 701–711.
- Markovic, S. & Dutzler, R. (2007) The structure of the cytoplasmic domain of the chloride channel CLC-Ka reveals a conserved interaction interface. *Structure*, 15, 715–725.
- Marmagne, A., Vinauger-Douard, M., Monachello, D., de Longevialle, A.F., Charon, C., Allot, M., et al. (2007) Two members of the Arabidopsis CLC (chloride channel) family, AtCLCe and AtCLCf, are associated with thylakoid and golgi membranes, respectively. *Journal of Experimental Botany*, 58, 3385–3393.
- Meyer, S. & Dutzler, R. (2006) Crystal structure of the cytoplasmic domain of the chloride channel CLC-0. *Structure*, 14, 299–307.
- Meyer, S., Savaresi, S., Forster, I.C. & Dutzler, R. (2007) Nucleotide recognition by the cytoplasmic domain of the human chloride transporter CLC-5. *Nature Structural & Molecular Biology*, 14, 60–67.
- Miller, C. (1982) Open-state substructure of single chloride channels from *Torpedo* electroplax. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 299, 401–411.
- Monachello, D., Allot, M., Oliva, S., Krapp, A., Daniel-Vedele, F., Barbier-Brygoo, H., et al. (2009) Two anion transporters AtCLCa and AtCLCe fulfil interconnecting but not redundant roles in nitrate assimilation pathways. *The New Phytologist*, 183, 88–94.
- Nakamura, A., Fukuda, A., Sakai, S. & Tanaka, Y. (2006) Molecular cloning, functional expression and subcellular localization of two putative vacuolar voltage-gated chloride channels in rice (*Oryza sativa* L.). *Plant & Cell Physiology*, 47, 32–42.
- Neagoe, I., Stauber, T., Fidzinski, P., Bergsdorf, E.Y. & Jentsch, T.J. (2010) The late endosomal CLC-6 mediates proton/chloride counter transport in heterologous plasma membrane expression. *The Journal of Biological Chemistry*, 285, 21689–21697.
- Nedelyaeva, O.I., Shuvalov, A.V., Karpichev, I.V., Beliaev, D.V., Myasoedov, N.A., Khalilova, L.A., et al. (2019) Molecular cloning and characterisation of SaCLCa1, a novel protein of the chloride channel

- (CLC) family from the halophyte *Suaeda altissima* (L.) pall. *Journal of Plant Physiology*, 240, 152995.
- Nedelyaeva, O.I., Shuvalov, A.V., Mayorova, O.V., Yurchenko, A.A., Popova, L.G., Balnokin, Y.V., et al. (2018) Cloning and functional analysis of SaCLC1, a gene belonging to the chloride channel family (CLC), from the halophyte *Suaeda altissima* (L.) Pall. *Doklady. Biochemistry and Biophysics*, 481, 186–189.
- Nguyen, C.T., Agorio, A., Jossier, M., Depré, S., Thomine, S. & Filleur, S. (2016) Characterization of the Chloride Channel-like, AtCLCg, involved in chloride tolerance in *Arabidopsis thaliana*. *Plant & Cell Physiology*, 57, 764–775.
- Niemeyer, M.I., Cid, L.P., Zúñiga, L., Catalán, M. & Sepúlveda, F.V. (2003) A conserved pore-lining glutamate as a voltage- and chloride-dependent gate in the CIC-2 chloride channel. *The Journal of Physiology*, 553, 873–879.
- Nongpiur, R.C., Singla-Pareek, S.L. & Pareek, A. (2020) The quest for osmosensors in plants. *Journal of Experimental Botany*, 712, 595–607.
- Padmanaban, S., Lin, X., Perera, I., Kawamura, Y. & Sze, H. (2004) Differential expression of vacuolar H⁺-ATPase subunit c genes in tissues active in membrane trafficking and their roles in plant growth as revealed by RNAi. *Plant Physiology*, 134, 1514–1526.
- Park, E., Campbell, E.B. & MacKinnon, R. (2017) Structure of a CLC chloride ion channel by cryo-electron microscopy. *Nature*, 541, 500–505.
- Park, E. & MacKinnon, R. (2018) Structure of the CLC-1 chloride channel from *Homo sapiens*. *eLife*, 7, 36629.
- Phillips, S., Brammer, A.E., Rodriguez, L., Lim, H.H., Stary-Weinzinger, A. & Matulef, K. (2012) Surprises from an unusual CLC homolog. *Biophysical Journal*, 103, L44–L46.
- Piccolo, A., Malvezzi, M., Houtman, J.C.D. & Accardi, A. (2009) Basis of substrate binding and conservation of selectivity in the CLC family of channels and transporters. *Nature Structural & Molecular Biology*, 16, 1294–1301.
- Rojas-Jiménez, K., Sohlenkamp, C., Geiger, O., Martínez-Romero, E., Werner, D. & Vinuesa, P. (2005) A CLC chloride channel homolog and ornithine-containing membrane lipids of *Rhizobium tropici* CIAT899 are involved in symbiotic efficiency and acid tolerance. *Molecular Plant-Microbe Interactions*, 18, 1175–1185.
- Sasvari, Z., Kovalev, N. & Nagy, P.D. (2013) The GEF1 proton-chloride exchanger affects tombusvirus replication via regulation of copper metabolism in yeast. *Journal of Virology*, 87, 1800–1810.
- Scheel, O., Zdebik, A.A., Lourdel, S. & Jentsch, T.J. (2005) Voltage-dependent electrogenic chloride/proton exchange by endosomal CLC proteins. *Nature*, 436, 424–427.
- Schumacher, K., Vafeados, D., McCarthy, M., Sze, H., Wilkins, T. & Chory, J. (1999) The *Arabidopsis* det3 mutant reveals a central role for the vacuolar H⁺-ATPase in plant growth and development. *Genes & Development*, 13, 3259–3270.
- Stockbridge, R.B., Lim, H.H., Otten, R., Williams, C., Shane, T., Weinberg, Z., et al. (2012) Fluoride resistance and transport by riboswitch-controlled CLC antiporters. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 15289–15294.
- Sun, H., Shen, L., Qin, Y., Liu, X., Hao, K., Li, Y., et al. (2018) CLC-Nt1 affects potato virus Y infection via regulation of endoplasmic reticulum luminal pH. *The New Phytologist*, 220, 539–552.
- Sun, W., Deng, D., Liheng, Y., Xiaojiao, Z., Juan, Y., Huixin, P., et al. (2013) Overexpression of the chloride channel gene (GmCLC1) from soybean increases salt tolerance in transgenic *Populus deltoides* × *P. euramericana* 'Nanlin895' plant. *OMICS*, 6, 347–354.
- Tampieri, E., Baraldi, E., Carnevali, F., Frascaroli, E. & De Santis, A. (2011) The activity of plant inner membrane anion channel (PIMAC) can be performed by a chloride channel (CLC) protein in mitochondria from seedlings of maize populations divergently selected for cold tolerance. *Journal of Bioenergetics and Biomembranes*, 43, 611–621.
- Teardo, E., Frare, E., Segalla, A., De Marco, V., Giacometti, G.M. & Szabò, I. (2005) Localization of a putative CLC chloride channel in spinach chloroplasts. *FEBS Letters*, 579, 4991–4996.
- von der Fecht-Bartenbach, J., Bogner, M., Dynowski, M. & Ludewig, U. (2010) CLC-b-mediated NO₃⁻/H⁺ exchange across the tonoplast of *Arabidopsis* vacuoles. *Plant & Cell Physiology*, 51, 960–968.
- von der Fecht-Bartenbach, J., Bogner, M., Krebs, M., Stierhof, Y.D., Schumacher, K. & Ludewig, U. (2007) Function of the anion transporter AtCLC-d in the trans-Golgi network. *The Plant Journal*, 50, 466–474.
- Wang, K., Preisler, S.S., Zhang, L., Cui, Y., Missel, J.W., Grønberg, C., et al. (2019) Structure of the human CLC-1 chloride channel. *PLoS Biology*, 17, e3000218.
- Wang, S., Su, S.Z., Wu, Y., Li, S.P., Shan, X.H., Liu, H.K., et al. (2015) Overexpression of maize chloride channel gene ZmCLC-d in *Arabidopsis thaliana* improved its stress resistance. *Biologia Plantarum*, 59, 55–64.
- Wege, S., De Angeli, A., Droillard, M.J., Kroniewicz, L., Merlot, S., Cornu, D., et al. (2014) Phosphorylation of the vacuolar anion exchanger AtCLCa is required for the stomatal response to abscisic acid. *Science Signaling*, 7, ra65.
- Wege, S., Jossier, M., Filleur, S., Thomine, S., Barbier-Brygoo, H., Gambale, F., et al. (2010) The proline 160 in the selectivity filter of the *Arabidopsis* NO₃⁻/H⁺ exchanger AtCLCa is essential for nitrate accumulation in planta. *The Plant Journal*, 63, 861–869.
- Wei, P., Che, B., Shen, L., Cui, Y., Wu, S., Cheng, C., et al. (2019) Identification and functional characterization of the chloride channel gene, GsCLC-c2 from wild soybean. *BMC Plant Biology*, 19, 121.
- Wei, P., Wang, L., Liu, A., Yu, B. & Lam, H.M. (2016) GmCLC1 confers enhanced salt tolerance through regulating chloride accumulation in soybean. *Frontiers in Plant Science*, 7, 1082.
- Wei, Q., Liu, Y., Zhou, G., Li, Q., Yang, C. & Peng, S. (2013) Overexpression of CsCLCc, a chloride channel gene from *Poncirus trifoliata*, enhances salt tolerance in *Arabidopsis*. *Plant Molecular Biology Reporter*, 31, 1548–1557.
- White, M.M. & Miller, C. (1979) A voltage-gated anion channel from the electric organ of *Torpedo californica*. *The Journal of Biological Chemistry*, 254, 10161–10166.
- Wong, T.H., Li, M.W., Yao, X.Q. & Lam, H.M. (2013) The GmCLC1 protein from soybean functions as a chloride ion transporter. *Journal of Plant Physiology*, 170, 101–104.
- Xing, A., Ma, Y., Wu, Z., Nong, S., Zhu, J., Sun, H., et al. (2020) Genome-wide identification and expression analysis of the CLC superfamily genes in tea plants (*Camellia sinensis*). *Functional & Integrative Genomics*, 20, 497–508.
- Yang, G., Zou, H., Wu, Y., Liu, H. & Yuan, Y. (2011) Identification and characterisation of candidate genes involved in chilling responses in maize (*Zea mays* L.). *Plant Cell Tissue and Organ Culture*, 106, 127–141.
- Zdebik, A.A., Zifarelli, G., Bergsdorf, E.Y., Soliani, P., Scheel, O., Jentsch, T. J., et al. (2008) Determinants of anion-proton coupling in mammalian endosomal CLC proteins. *The Journal of Biological Chemistry*, 283, 4219–4227.
- Zhang, H., Jin, J., Jin, L., Li, Z., Xu, G., Wang, R., et al. (2018) Identification and analysis of the chloride channel gene family members in tobacco (*Nicotiana tabacum*). *Gene*, 676, 56–64.
- Zhang, X.K., Zhou, Q.H., Cao, J.H. & Yu, B.J. (2011) Differential Cl⁻/salt tolerance and NaCl-induced alternations of tissue and cellular ion fluxes in *Glycine max*, *Glycine soja* and their hybrid seedlings. *Journal of Agronomy and Crop Science*, 197, 329–339.
- Zheng, K., Chen, M., Xiang, Y., Ma, K., Jin, F., Wang, X., et al. (2014) Inhibition of herpes simplex virus type 1 entry by chloride channel inhibitors

- tamoxifen and NPPB. *Biochemical and Biophysical Research Communications*, 446, 990–996.
- Zhou, G.A. & Qiu, L.J. (2010) Identification and functional analysis on abiotic stress response of soybean Cl^- channel gene GmCLCnt. *Agricultural Sciences in China*, 9, 199–206.
- Zifarelli, G. & Pusch, M. (2009) Conversion of the 2 $\text{Cl}^-/1 \text{H}^+$ antiporter CIC-5 in a NO_3^-/H^+ antiporter by a single point mutation. *The EMBO Journal*, 28, 175–182.

How to cite this article: Subba A, Tomar S, Pareek A, Singla-Pareek SL. The chloride channels: Silently serving the plants. *Physiologia Plantarum*. 2020;1–15. <https://doi.org/10.1111/ppl.13240>