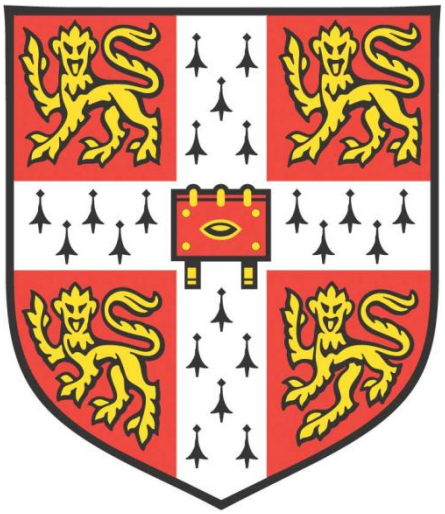
# MORPHOLOGICAL AND PHYSIOLOGICAL STUDIES OF THE CARBON CONCENTRATING MECHANISM IN *CHLAMYDOMONAS REINHARDTII*



**Kher Xing Chan**

**Fitzwilliam College**

**Department of Plant Sciences**

**University of Cambridge**

**This dissertation is submitted for the degree of Doctor of Philosophy**

**June 2018**

*Dedicated to God and Dr. Steven J Burgess*

*"Ask, and it will be given you;*

*search, and you will find;*

*knock, and the door will be opened for you.*

*For everyone who asks receives,*

*everyone who searches finds,*

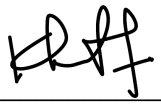
*and for everyone who knocks, the door will be opened."*

*(Matthew 7:7-8)*

# DECLARATION

This dissertation is the result of my own work and includes nothing, which is the outcome of work done in collaboration except where specifically indicated in the text. It has not been previously submitted, in part or whole, to any university of institution for any degree, diploma, or other qualification.

In accordance with the Degree Committee for Biology guidelines, this thesis is does not exceed 60,000 words.

Signed: 

Date: 25 MAY 2018

Kher Xing Chan

Cambridge

# ABSTRACT

*Chlamydomonas reinhardtii* possesses a single-cell-based $CO_2$-concentrating mechanism (CCM). The CCM is an important element of algal photosynthesis, metabolism, growth and biomass production, which works by increasing the concentration of inorganic carbon (Ci) in the pyrenoid, a dense RuBisCO-packed structure within the chloroplast. This suppresses RuBisCO oxygenase activity and associated photorespiration. The enhanced efficiency of $CO_2$ assimilation in the pyrenoid via CCM had been modelled theoretically as a requirement for successful CCM in higher plant systems. The ultimate aim of my research is to understand the biogenesis of the pyrenoid using a set of CCM mutants with pyrenoidal defects. Immunofluorescence methods and spot growth tests under different $CO_2$ concentrations were performed on mutants with CCM defects generated by an insertional mutagenesis screen. Morphological and physiological characterisation of these mutants revealed differences in the pyrenoid morphology, the ability for RuBisCO to aggregate into the pyrenoid and the formation of thylakoidal tubule network associated with the pyrenoid. The thylakoid tubule network may be linked to the transport of inorganic carbon into the pyrenoid as part of the CCM. Further characterisation of one of the mutants gave rise to the hypothesis that the gene of interest, Cre11.g467712 (*SAGA*), is a multi-functional anchor protein related to the structural formation of the pyrenoid and may be another essential component of the pyrenoid.

# ACKNOWLEDGEMENTS

Firstly, I am grateful to my supervisor, Professor Howard Griffiths, for the second chance to complete this dissertation and for your enthusiasm in CCM which inspires me from the first time we met. And special thanks to Professor Julian Hibberd, Professor Alison Smith and all the academics who had given me guidance throughout the three years in the Department of Plant Sciences. I am also comforted by the constant support of the legion of postdoctoral researchers in the department, especially Dr. Betty Chung, Dr. Francisco Navarro, Dr. Payam Mehrshahi, Dr. Katrin Gesler and Dr. Moritz Meyer for sharing their expertises with me. To Dr. Toomas Kivisild of the Division of Biological Anthropology, thank you for your kind advice when I faced challenges in my first year at Cambridge. Without you, I would not be where I am now.

I would like to thank Dr. Jeremy Skepper, Ms. Lyn Carter and everyone in the Cambridge Advanced Imaging Centre (CAIC) for their technical support on TEM imaging and kind understanding, and to Dr. Mike Deery, Dr. Marco Chiapello and everyone at the Cambridge Centre for Proteomics for the technical support for proteomics analysis. Special thanks to all the technicians and support staff of the Department of Plant Sciences for the assistance which made my life in the lab easier.

Without the collaboration, this dissertation would not have come into fruition. With this, thank you to Dr. Martin Jonikas, Dr. Alistair McCormick and their lab members who had supported my work through all the Skype calls we had over the past three years. Also a special acknowledgement to my Science cousin, Alan Itakura of Stanford University, for the time spent discussing about SAGA and life. To Professor James V. Moroney and Professor Hideya Fukuzawa for the inspiration and motivating discussions we had on CCM.

I also wish to acknowledge the Islamic Development Bank (IDB) and Cambridge Trust for the three-year funding through the Cambridge-IDB International Scholarship. Special acknowledgement to those who provided grants in my final year: University of Cambridge, Cambridge Philosophical Society, Cambridge Trust and Fitzwilliam College.

Thank you to Kayla Friedman and Malcolm Morgan of the Centre for Sustainable Development, University of Cambridge, UK for producing the Microsoft Word thesis template used to produce this document.

I would like to take this opportunity to thank specially those who had been in my cheerleading team and provided the mental health support I needed to complete this PhD marathon. To the Griffiths lab, for the push to move forward, especially to Jessica, Stephanie, Jamie, Luis, Yi and Myriam. To the Hibberd lab, especially Ivan, Susan, Marek, Jane and Greg, for being there, rain or shine (during tea times), accepting all of me, including those very rough patches in my journey. To my past college tutors, Dr. Helen Bettinson, Dr. John Cleaver, Prof. David Coomes and Dr. Susan Larsen, the Fitzwilliam College Chaplain, Rev. Helen Arnold, for all the advice and support which kept me going. To Fitzwilliam College and all the staff, for providing such a great network of support. To my Scripture meditation group, Matteo, Mel, Ewa, Imogen, Justin, Maria and Madda for the friendship, prayers and spiritual support. Not forgetting my writing up buddies on Twitter, Frauke, Mel, Kate, Sarah, Leah, and the rest of the #PhDgang, who kept me company and cheered me on during the lonely months of writing up.

Finally and most importantly, to my family and loved ones who believe in me, the Chan family, the Burgesses, Josephine Hiu and family, and Irene. A big thank you to my beloved one and only Dr. Steven James Burgess, for being there constantly, faithfully and lovingly, providing your endless support and love in this journey, as well as the intellectual stimulations and technical support.

To God, all glory and honour are Yours. Thank You for opening the door and providing me with angels to lead me along this journey.

# CONTENTS

**1 Introduction ..... 1**

1.1 PHOTOSYNTHESIS: ADDRESSING GLOBAL FOOD SECURITY IN THE MODERN SOCIETY ..... 1

*1.1.1 Rubisco and Photorespiration ..... 1*

*1.1.2 Inorganic Carbon Concentrating Mechanisms and Crop Yield Improvement 5*

1.1.2.1 CCM in Higher Plants ..... 5

1.1.2.2 CCM in Cyanobacteria and Algae ..... 7

*1.2* CARBON CONCENTRATING MECHANISM IN *CHLAMYDOMONAS REINHARDTII* ..... 14

*1.2.1 Pyrenoid, the Site of $CO_2$ Fixation ..... 15*

*1.2.2 Inorganic Carbon Transporters ..... 18*

*1.2.3 Carbonic Anhydrases ..... 19*

1.3 *CHLAMYDOMONAS* AS A CHASSIS FOR GENETIC ENGINEERING ..... 24

*1.3.1 Mutagenesis Screens ..... 24*

*1.3.2 Development of the Chlamydomonas Modular Cloning System ..... 25*

1.4 GENERAL AIMS AND HYPOTHESES OF THIS STUDY ..... 34

**2 Materials and Methods ..... 35**

2.1 PHYSIOLOGICAL ANALYSIS ..... 35

*2.1.1 Growth of Chlamydomonas reinhardtii strains ..... 35*

*2.1.2 Spot Growth Tests and Growth Experiments ..... 36*

*2.1.3 Oxygen Evolution ..... 36*

*2.1.4 Chlorophyll Extraction and Measurement ..... 37*

2.2 MICROSCOPY METHODS ..... 37

*2.2.1 Fixing and Embedding ..... 37*

*2.2.2 Immunogold Labelling ..... 38*

*2.2.3 Transmission Electron Microscopy for Image Acquisition ..... 38*

*2.2.4 Evaluation of Immunogold Labeling ..... 39*

*2.2.5 Indirect Immunofluorescence (IF) ..... 39*

2.2.5.1 Preparation of materials ..... 39

2.2.5.2 Method ..... 40
*2.2.6 Confocal Microscopy for Image Acquisition ..... 41*
2.3 PROTEOMICS METHODS ..... 41
*2.3.1 Protein Extraction for Western Blotting (except GFP-Trap samples) ..... 41*
*2.3.2 Protein Extraction for Proteomics Analysis ..... 41*
*2.3.3 Protein Acrylamide Gel Electrophoresis (PAGE) and Blotting of Protein Samples ..... 42*
*2.3.4 Co-immunoprecipitation using GFP-Trap_A (Chromotek, Germany) ..... 43*
*2.3.5 LC-MS/MS Analysis ..... 43*
*2.3.6 Native PAGE and Phosphoprotein Staining ..... 44*
2.3.6.1 Native PAGE and Western Blotting ..... 44
2.3.6.2 Phosphoprotein Staining ..... 44
2.4 MOLECULAR METHODS ..... 45
*2.4.1 Genomic DNA Extraction ..... 45*
*2.4.2 Genomic DNA PCR ..... 46*
*2.4.3 PCR Fragment Purification ..... 46*
*2.4.4 Plasmid Miniprep ..... 47*
*2.4.5 Golden Gate Cloning ..... 47*
2.4.5.1 List of Vector backbones and domesticated modules ..... 47
2.4.5.2 Primer Design ..... 50
2.4.5.3 Molecular biology reagents ..... 51
2.4.5.4 MoClo cloning protocol ..... 51
2.4.5.5 Restriction Enzyme Digestion ..... 52
2.4.5.6 Sanger Sequencing and Multiple Sequence Alignment ..... 52
2.4.5.7 Transformation ..... 53
2.4.5.8 Colony PCR ..... 53
2.5 BIOINFORMATICS ANALYSES ..... 54
*2.5.1 Databases ..... 54*
*2.5.2 Protein Prediction Programs ..... 54*
*2.5.3 Peptide Visualisation Software ..... 55*
*2.5.4 Multiple Sequence Alignment Web Tool ..... 55*
*2.5.5 Golden Gate Cloning ..... 55*
*2.5.6 Graphic Tools ..... 55*

**3 Pyrenoid Formation is Linked to Multiple Genes ..... 56**

3.1 INTRODUCTION ..... 56

*3.1.1 Engineering Photosynthesis to Address Global Food Demand ..... 56*

*3.1.2 Insertional Mutagenesis ..... 57*

*3.1.3 Screening for Pyrenoid Defects ..... 58*

3.2 RESULTS ..... 62

*3.2.1 Growth of CCM defective mutants ..... 62*

*3.2.2 Photosynthetic Rate of the CCM Mutants ..... 64*

*3.2.3 Quantification of RuBisCO aggregation and Pyrenoid Formation in Mutant Lines 66*

*3.2.4 Phenotypic Aberrations of CCM Defective Mutants ..... 74*

*3.2.5 Identification of Candidate Genes by Whole Genome Sequencing ..... 83*

3.3 DISCUSSION ..... 85

*3.3.1 Localisation of Rubisco in the Chloroplast of Chlamydomonas reinhardtii 85*

*3.3.2 The Biophysical Components of the Pyrenoid of Chlamydomonas reinhardtii ..... 87*

*3.3.3 Importance of the Biophysical Components in the Pyrenoid ..... 87*

*3.3.4 Multiple Genes Are Involved in the Biogenesis of the Pyrenoid ..... 90*

3.4 CONCLUSIONS AND FUTURE WORK ..... 93

**4 Modular Cloning in *Chlamydomonas reinhardtii* ..... 95**

4.1 INTRODUCTION ..... 95

*4.1.1 The Use of Chlamydomonas reinhardtii in Biotechnology ..... 95*

*4.1.2 Synthetic Biology in Chlamydomonas reinhardtii ..... 96*

*4.1.3 Development of Parts in Synthetic Biology ..... 98*

*4.1.4 Selection of Mutant with CCM Defect for Golden Gate Cloning ..... 98*

4.2 RESULTS ..... 99

*4.2.1 Morphological Observation of the Mutant ..... 99*

*4.2.2 Construction of Transformation Vector using Golden Gate Cloning ..... 104*

*4.2.3 Transformation and Validation of Transformants ..... 119*

4.3 DISCUSSION ..... 126

4.4 CONCLUSION AND FUTURE WORK ..... 129

**5 Structural Formation of the Pyrenoid ..... 131**

5.1 INTRODUCTION ..... 131

*5.1.1 The Structure of the Pyrenoid of Chlamydomonas reinhardtii ..... 131*

*5.1.2 Biogenesis of the Pyrenoid ..... 132*

*5.1.3 Mutants with Defective Pyrenoidal Tubule Network ..... 134*

5.2 RESULTS ..... 136

*5.2.1 Prediction of Structure and Domains of SAGA gene ..... 136*

*5.2.2 Localisation of SAGA in the Pyrenoid ..... 140*

*5.2.3 Candidate SAGA Protein Interactors ..... 155*

5.3 DISCUSSION ..... 171

*5.3.1 Bioinformatics of SAGA ..... 171*

*5.3.2 Localisation of SAGA in the Pyrenoid Aids the Carbon Concentrating Mechanism in Chlamydomonas reinhardtii ..... 173*

*5.3.3 SAGA Plays a Role in the Formation of Thylakoid Tubule Network in the Pyrenoid ..... 174*

5.4 CONCLUSION AND FUTURE WORK ..... 181

**6 Conclusions ..... 182**

6.1 THE PYRENOID IS A MULTI-COMPONENT SUBCOMPARTMENT OF THE CHLOROPLAST ..... 182

6.2 THE PYRENOID MATRIX IS ANCHORED TO A PYRENOID CORE ..... 185

6.3 SAGA IS A MULTI-FUNCTIONAL ANCHOR PROTEIN ..... 187

**7 References ..... 191**

**8 Appendices ..... 226**

# LIST OF TABLES

Table 1.1 The mechanism of CCMs in eukaryotic algae.. ..... 11
Table 1.2 Cyanobacteria and different phyla of eukaryotic algae which are known to possess carbon concentrating mechanisms ..... 12
Table 1.3 Comparison of the localisation of the algal CCM gene products in the *Chlamydomonas reinhardtii* (native), tobacco, *Arabidopsis* and *Xenopus* oocytes 13
Table 1.4 Summary of essential and putative components of CCM in *Chlamydomonas reinhardtii* ..... 21
Table 1.5 Summary of the commonly-used cloning techniques in genetic complementation assays ..... 28
Table 1.6 Nuclear and chloroplast promoters used for transgene expression in *Chlamydomonas reinhardtii* ..... 32
Table 1.7 Nuclear and chloroplast selectable markers used in *Chlamydomonas reinhardtii* ..... 33
Table 2.1 Vector backbones of Golden Gate MoClo Plant Tool Kit (Engler et al., 2014) and a domesticated Level 2 vector backbone (Dr. Francisco Navarro, University of Cambridge) used in this study. ..... 48
Table 2.2 Domesticated modules of Golden Gate cloning in *Chlamydomonas reinhardtii*. TU_4, pFJN43, pFJN45 and pFJN70 were kindly provided by Dr. Francisco Navarro, while pCCC3 and pCCN3 were constructed specifically for this study. . 49
Table 2.3 Primers used in assembly of the expression vector, pFPC2-1C4-HygR, including validation of levels 0, 1 and 2 modules. ..... 50
Table 3.1 List of the twelve mutants with pyrenoidal defects selected from the secondary CCM screens ..... 60
Table 3.2 Photosynthetic $K_{0.5}$ (Ci) of wild type and mutants indicating reduced affinity for inorganic carbon ($C_i$) ..... 65
Table 3.3 The number of cell profiles with and without pyrenoids in wild type *C. reinhardtii* and the CCM defective mutants ..... 69
Table 3.4 Pairwise comparisons between the cell sectional area of wild type and that of the mutants using Mann-Whitney-Wilcoxon test showed that all mutants except C4 had relatively similar cell area. ..... 70

Table 3.5 Pairwise comparisons between the pyrenoid section area of wild type and that of the mutants using Mann-Whitney-Wilcoxon test showed that all mutants except F8 and H10 had significantly different pyrenoid size compared to wild type. ..... 71
Table 3.6 Calculation of RuBisCO fraction in both the pyrenoid and stroma of wild type cells and the mutants acclimatised for 24 h at 0.04% $CO_2$ to induce CCM (n=20) 73
Table 3.7 Qualitative observations on 50 electron micrographs of the CCM defective mutants and the corresponding wild type ..... 77
Table 3.8 Summary of the top candidate gene for each CCM defective mutant used in this study ..... 84
Table 4.1 Whole genome sequencing showed potential locations of the insertional mutagenesis cassette in the genome of mutant C4. ..... 103
Table 4.2 $OD_{750}$ measurement for each colony grown on TP medium with 10 µg/ml hygromycin under CCM induced conditions. ..... 122
Table 5.1 Prediction of target and location of SAGA transit peptide using ChloroP, PredAlgo and TargetP suggested different locations and lengths. ..... 142
Table 5.2 Photosynthetic $K_{0.5}$ (Ci) of wild type, *saga* and two complemented lines, *saga::SAGA-Venus-3xFLAG A7* (SVA7) and *saga::SAGA-Venus-3xFLAG B2* (SVB2). ..... 144
Table 5.3 Summary list of potential protein interactors of SAGA identified by LC-MS/MS of the GFP-Trap_A coimmunoprecipitation assay. ..... 159
Table 5.4 List of phosphorylated proteins identified in the samples by LC-MS/MS of the GFP-Trap_A coimmunoprecipitation assay.. ..... 167

# LIST OF FIGURES

Figure 1.1 Types of carbon concentrating mechanisms to improve photosynthesis .....3
Figure 1.2 Schematic representation of the photosynthetic C3 pathway, with native photorespiratory C2 pathway and the Maier (magenta), Carvalho (green) and Kebeish (blue) bypasses .....6
Figure 1.3 A general illustration of CCM in cyanobacteria.....8
Figure 1.4 Carbon concentrating mechanism in *Chlamydomonas reinhardtii* .....17
Figure 3.1 A schematic diagram of CCM mutant selection.....58
Figure 3.2 Independent screens on the CCM defective mutant library.....59
Figure 3.3 Spot tests investigating the growth response of the wild type, cMJ030 (Zhang et al., 2014), *SSSO* which is a RbcS1 hybrid strain (Genkov et al., 2010) and the CCM defective mutants, including *epyc1* (Mackinder et al., 2016). .....63
Figure 3.4 Oxygen-evolving activity of WT and mutants in response to external dissolved inorganic carbon concentrations .....64
Figure 3.5 A schematic diagram of a pyrenoid in wild type *Chlamydomonas reinhardtii* .....66
Figure 3.6 Electron micrographs of the pyrenoid of the CCM defective mutants grown photoautotrophically at room temperature with illumination of about 50 µmol photons $m^{-2}s^{-1}$ under ambient level of $CO_2$ for 24 hours .....68
Figure 3.7 Boxplot showing the measurement of the cell section area of the wild type (WT) and the nine CCM defective mutants (n=30). .....69
Figure 3.8 Measurement of the pyrenoid sectional area of the CCM defective mutants displayed significant differences from the corresponding wild type (WT) with Kruskal-Wallis p-value of 1.7e-08 (n=30).....70
Figure 3.9 Fluorescence Western blot using anti-Rubisco antibody of wild type and CCM defective mutants grown at 5% (v/v) and 0.04% (v/v) $CO_2$ conditions .....71
Figure 3.10 Immunolocalisation of anti-RuBisCO antibody of *C. reinhardtii* wild type and mutant A8 cells grown on TP medium after 24 h CCM induction showed similar RuBisCO localisation pattern in the pyrenoid between wild type and the mutant .....72

Figure 3.11 Fluorescence Western blots of CCM defective mutants grown under 5% and 0.04% $CO_2$ conditions with antibodies of CCM-related proteins linked to the pyrenoid ..... 74
Figure 3.12 Electron micrograph of a wild type cell, cMJ030, at the median plane . .... 76
Figure 3.13 Electron micrographs of (A) wild type, (B) F8 and (C) C8 that showed presence of thylakoid tubule(s) in the pyrenoid ..... 77
Figure 3.14 Electron micrographs of (A) wild type and mutants (B) A5, (C) A7, (D) A8, (E) A11 and (F) C8 that showed multiple pyrenoids per cell ..... 79
Figure 3.15 Total pyrenoidal starch area of wild type and the mutants with p-value =6.5e-13, n=30. ..... 80
Figure 3.16 Total number of starch plates per pyrenoid of mutants differed significantly from wild type with p-value<2.2e-16, n=30 ..... 80
Figure 3.17 Examples of (A) wild type and mutants (B) F8 and (C) C4 with the presence of chloroplastic starch granules ..... 81
Figure 3.18 Summary of the qualitative assessments from the data of Table 3.7 using the four phenotypes observed on the electron micrographs of the CCM defective mutants. ..... 82
Figure 3.19 Current and hypothetical models of carbon concentrating mechanism of *C. reinhardtii* ..... 89
Figure 3.20 Indirect immunofluorescence with anti-Rubisco antibody on mutant F8 (Caspari, 2015) ..... 91
Figure 4.1 The basis of synthetic biology is formed by iterative cycles of a Design-Build-Test-Learn process ..... 97
Figure 4.2 Electron micrographs of wild type and mutant C4 whole cells and zoomed out images of the pyrenoids grown under different conditions to observe pyrenoid phenotypes ..... 101
Figure 4.3 Genomic structure of Cre02.g143387 showing the position of *AphVIII* insertion (red box) ..... 102
Figure 4.4 Secondary structure prediction of Cre02.g143387 using Phyre2 suggested that this protein was comprised of 50% disordered protein, 35% alpha helices and 8% of beta strands. ..... 102
Figure 4.5 NetPhos 3.1a predicted multiple phosphorylation sites in the protein encoded by the gene Cre02.g143387. ..... 104

Figure 4.6 The protein model for Cre02.g143387 was predicted by Phyre2 using *ab initio* modelling. ..... 104
Figure 4.7 Illustration showing the steps to transform the gene Cre02.g143387 into mutant C4 using Golden Gate Cloning system ..... 106
Figure 4.8 Cre02.g143387 genomic DNA was amplified from wild type *C. reinhardtii* cells using primer pair CC021 and CC022 . ..... 106
Figure 4.9 Level 0 module containing the coding sequence of gene Cre02.g143387 indicating the location of the cutting sites of *Bsa*I, *Blp*I and *Dpn*I in the module. ..... 107
Figure 4.10 Restriction enzyme digestion with *Bsa*I, *Blp*I and *Dpn*I in the level 0 module. ..... 108
Figure 4.11 Level 1 transcriptional unit containing Cre02.g143387 gene and a C-terminus YFP tag. ..... 110
Figure 4.12 PCR validation of the level 1 transcriptional unit using orientation-specific primer pairs ..... 111
Figure 4.13 The assembled level 2 plasmid contained the gene of interest with a YFP tag and hygromycin resistant gene (NPTII) . ..... 112
Figure 4.14 PCR validation was performed using primer pairs CC025+CC026 and CC029+CC030 to confirm that level 2 plasmids assembled using level 1 transcriptional unit contained the gene of interest in the correct orientation. ..... 112
Figure 4.15 Sequence alignment using Clustal Omega of the Sanger sequencing results to confirm the full sequence of the gene of interest in the assembled level 2 plasmid. The exons of the gene were underlined on the reference sequence, Cre02.g143387. Two mutations are found on the introns of the insert (highlighted in yellow): a nucleotide substitution (C>T) at position 420 and an insertion of a cytosine (C) at position 1190. The stop codon TAA was detected at the end of the insert (red letters, highlighted in yellow) ..... 113
Figure 4.16 Sequence alignment showed no difference between the gene product of Cre02.g143387 with the inferred amino acid sequence of the gene of interest in the level 2 plasmid. ..... 119
Figure 4.17 Seven-day old transformants plated on TAP agar plates with 10 µg/ml hygromycin grown under low light conditions at 25°C ..... 120

Figure 4.18 The selected colonies were grown in selective TAP medium containing 10 µg/ml hygromycin for three consecutive generations.....120
Figure 4.19 The selected colonies from previous generation were grown in this arrangement in TAP medium with 10 µg/ml hygromycin .....123
Figure 4.20 Colony PCR was performed using primer pair CC027+CC030 which amplified the region between CDS and 3'UTR in pFPC2-1C4-HygR plasmid ...123
Figure 4.21 (A) The transformants with positive amplification were cultured in TAP medium with 10 µg/ml hygromycin in a 24-well plate. (B) Confirmation that the transformants retained pFPC2-1C4-HygR using PCR amplification with primer pair CC027+CC030. (C) Growth spot test performed on the transformants to confirm the restoration of CCM phenotype . .....124
Figure 4.22 Validation of the transformants by observation of the subcellular localisation of the gene product of pFPC2-1C4-HygR in transformed mutant C4 cells compared to positive and negative controls .....125
Figure 5.1 A proposed model by Mackinder et al (2017) which shows the different protein-containing layers of the pyrenoid based on spatial interactome, microscopy and past literature. .....132
Figure 5.2 Electron micrographs of the pyrenoid of the wild type and *saga* mutants grown photoautotrophically at room temperature with illumination of about 50 µmol photons $m^{-2}s^{-1}$ under ambient level of $CO_2$ for 24 hours. Scale bar 500 nm. .....135
Figure 5.3 Genomic structure of *the SAGA* gene locus (ID: Cre11.g467712) showing the position of the *AphVIII* insertion (red box). .....136
Figure 5.4 Secondary structure prediction of *SAGA* using PHYRE2 suggested high degree of alpha helices and a starch-binding domain. .....136
Figure 5.5 (A) BLASTP of SAGA suggested a starch binding domain from CBM20 family and a SMC (Structural Maintenance of Chromosome) N terminal domain in the protein. NCBI Conserved Domain Database search showed that (B) SAGA (query) possessed all the highly conserved sites (#) of the CBM20 starch-binding feature 1 which acts as starch recognition site. (C) SAGA (query) aligned with the SBD feature 2 but did not possess the conserved active site of this feature. .....138
Figure 5.6 A dendrogram of pairwise relationships between SAGA and the suggested homologues in *Chlamydomonas reinhardtii*, *Volvox carteri*, *Ostreococcus*

*lucimarinus, Arabidopsis lyrata, Brassica rapa, Solanum lycopersicum* and *Selaginella moellendorffii*. ..... 139
Figure 5.7 PCOILS predicted high probability of coiled-coil secondary structure within the alpha helix region of SAGA. ..... 139
Figure 5.8 Top two predicted protein structures of SAGA based on alignment at different locations of the protein to crystal structures of ALIX/AIP1 and colicin e3. ..... 140
Figure 5.9 Schematic representation of pRAM118_SAGA plasmid constructed by Alan Itakura, Stanford University, using Gibson assembly ..... 141
Figure 5.10 Fluorescence Western blots of the wild type and the two complemented lines, SVA7 and SVB2, grown under 5% (v/v) $CO_2$ and 0.04% (v/v) $CO_2$ conditions with anti-GFP, anti-FLAG and anti-Rubisco antibodies ..... 142
Figure 5.11 Spot tests investigating the growth response of the wild type, cMJ030, *saga* and the complemented lines, SVA7 and SVB2 ..... 144
Figure 5.12 Oxygen-evolving activity of WT, *saga* and the complemented lines in response to external dissolved inorganic carbon concentrations ..... 145
Figure 5.13 The electron micrographs of the pyrenoid of wild type, *saga* and the two complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7)* and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) grown photoautotrophically at room temperature with illumination of about 50 µmol photons $m^{-2}s^{-1}$ under ambient level of $CO_2$ for 24 hours ..... 146
Figure 5.14 Electron micrographs of the wild type, *saga* and the two complemented lines grown photoautotrophically under 5% $CO_2$ (high $CO_2$), 1h and 24h after acclimation to ambient air (low $CO_2$) ..... 147
Figure 5.15 Measurement of the pyrenoid area of wild type, *saga* and the two complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) under 5% (v/v) $CO_2$ and upon 24h acclimation to 0.04% (v/v) $CO_2$ ..... 148
Figure 5.16 Measurement of the starch area per pyrenoid of wild type, *saga* and the two complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) under 5% (v/v) $CO_2$ and upon 24h acclimation to 0.04% (v/v) $CO_2$ ..... 149

Figure 5.17 Starch plates count per pyrenoid of wild type, *saga* and the two complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) under 5% (v/v) $CO_2$ and upon 24h acclimation to 0.04% (v/v) $CO_2$ ..... 150
Figure 5.18 SAGA-Venus-3xFLAG protein was localised within the pyrenoid of the complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) after 24h acclimation to ambient air (low $CO_2$) to induce CCM with wild type and *saga* as controls ..... 152
Figure 5.19 Subcellular localisation of Rubisco in wild type, *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) by indirect immunofluorescence assay with an anti-Rubisco antibody ..... 153
Figure 5.20 Subcellular localisation of SAGA in wild type, *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) by indirect immunofluorescence assay with anti-SAGA antibody ..... 154
Figure 5.21 (A) Native PAGE of total soluble protein of wild type, *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*). (B-D) Western blot using anti-GFP, anti-FLAG and anti-Rubisco antibodies..... 156
Figure 5.22 Example of Western blots using (A) anti-GFP and (B) anti-SAGA antibodies to confirm that SAGA-Venus-3xFLAG protein was pulled down only in the complemented lines during the co-immunoprecipitation assay performed with GFP-Trap_A (ChromoTek, Germany)..... 157
Figure 5.23 Venn diagram showed the number of proteins identified from the co-immunoprecipitation assay using GFP-Trap_A ..... 158
Figure 5.24 Phobius and PHYRE2 predicted that while SAGA was not a transmembrane protein ..... 164
Figure 5.25 NetPhos 3.1a predicted multiple phosphorylation sites in SAGA protein. 165
Figure 5.26 PAGE gel of total soluble proteins of (A) wild type, SVB2 (*saga::SAGA-Venus-3xFLAG B2*), (C) *saga* and SVA7 (*saga::SAGA-Venus-3xFLAG A7*) were treated first with Pro-Q® Diamond phosphoprotein gel stain to observe phosphorylated proteins and then with SYPRO® Ruby total protein stain for comparison to total soluble protein..... 166

Figure 5.27 Two examples of incomplete formation of thylakoid tubule (white arrow) penetrated into the pyrenoid matrix of a pyrenoid in the *saga* mutant although there were at least two pyrenoids per cell ..... 175
Figure 5.28 (A) The pattern of SAGA-mVenus fluorescence signal localisation. Scale bar 5 µm. Confocal sections of the localisation pattern of (B) LCI9 and (C) PSBP4 within the pyrenoid, extracted from Mackinder et al. (2017) ..... 176
Figure 5.29 Proposed position of SAGA in the pyrenoid layer based on the original image taken from Mackinder et al. (2017) ..... 177
Figure 5.30 A model of SAGA interaction in the pyrenoid. SAGA is proposed to act as an interface which anchors the pyrenoid matrix, the thylakoid tubule network and the starch sheath to form a functional pyrenoid. ..... 180
Figure 6.1 Proposed mode of action of SAGA with the cytoskeleton and/or microtubules in wild type *C. reinhardtii* ..... 184
Figure 6.2 (A to C) Models of the formation of an algal pyrenoid proposed by Meyer et al. (2017). (D) A revised model of pyrenoid biogenesis in *C. reinhardtii* based on the work of this doctoral research ..... 188

# LIST OF ABBREVIATIONS AND ACRONYMS

| | |
|---|---|
| µl | microliter |
| µM | micromolar |
| 2PG | phosphoglycolate |
| 3PGA | 3-phosphoglycerate |
| ATP | adenosine triphosphate |
| CAH or CA | carbonic anhydrases, e.g. CAH3 is carbonic anhydrase 3 |
| CAM | crassulacean acid metabolism |
| CCM | carbon or $CO_2$ concentrating mechanism |
| cDNA | complementary DNA |
| Ci | inorganic carbon |
| cMJ030 | wild type used in this study |
| $CO_2$ | carbon dioxide |
| $dH_2O$ | distilled water |
| DIC | differential interference microscopy |
| FBPA/SBPase | fructose-1,6-/sedoheptolose-1,7-bisphosphatase |
| gDNA | genomic DNA |
| GFP | green fluorescent protein |
| h | hour |
| $H^+$ | hydrogen ion |
| HC | high-$CO_2$ |
| $HCO_3^-$ | bicarbonate |
| HSTBSTT | high-salt TBS buffer with Tween 20 and Triton X-100 |
| IFT | intraflagellar transport |
| LC | low-$CO_2$ |
| LCI | low-$CO_2$-inducible, e.g. LCIA is low-$CO_2$-inducible protein A |

| | |
|---|---|
| LSU | large subunit |
| MeOH | methanol |
| min | minute |
| MoClo | modular cloning |
| $Na^+$ | sodium ion |
| NADPH | nicotinamide adenine dinucleotide phosphate |
| $O_2$ | oxygen |
| PAGE | protein acrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| rbcL | gene encoding Rubisco LSU |
| RbcS | gene encoding Rubisco SSU |
| Rubisco | ribulose-1,5-bisphosphate carboxylase/oxygenase |
| RuBP | ribulose-1,5-bisphosphate |
| s | second |
| SDS | sodium dodecyl sulphate |
| SSU | small subunit |
| TAP | tris-acetate-phosphate |
| TEM | transmission electron microscopy |
| TP | tris-phosphate |
| UV | ultraviolet |
| v/v | volume over volume |
| w/v | weight over volume |
| WT | wild type |

# LIST OF APPENDICES

Appendix 1 Alignment of the whole genome sequencing data of the CCM defective mutants used in this study..... 227

Appendix 2 The lack of visible thylakoid tubules network in the pyrenoids of mutant H10 cells. Scale bar 500 nm..... 237

Appendix 3 Pyrenoid morphology of wild type *Chlamydomonas reinhardtii* cells. Scale bar 500 nm..... 238

Appendix 4 The lack of visible thylakoid tubules network in the pyrenoids of *saga* cells. Scale bar 500 nm..... 239

Appendix 5 The presence of traversing thylakoid tubules in the pyrenoids of *saga* cells. Scale bar 500 nm..... 240

Appendix 6 The restoration of traversing thylakoid tubule network in the pyrenoids of *saga::SAGA-Venus-3xFLAG* A7 cells. Scale bar 500 nm..... 241

Appendix 7 The restoration of traversing thylakoid tubule network in the pyrenoids of *saga::SAGA-Venus-3xFLAG* B2 cells. Scale bar 500 nm..... 242

Appendix 8 Proteomic report displaying the protein identification probability of all the proteins detected by LC-MS/MS of the GFP-Trap_A coimmunoprecipitation assay..... 243

# 1 INTRODUCTION

## 1.1 Photosynthesis: Addressing Global Food Security in the Modern Society

### 1.1.1 Rubisco and Photorespiration

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco, EC 4.1.1.39) is the most abundant enzyme on earth and possesses both carboxylase and oxygenase activities (Spreitzer and Salvucci, 2002). Under the low oxygen, high carbon dioxide ($CO_2$) environment of Precambrian period, the earliest form of Rubisco was able to form 3-phosphoglycerate (3PGA) by fixing carbon dioxide to ribulose-1,5-bisphosphate (RuBP) (Andersson, 2008; Bauwe et al., 2010; Peterhänsel et al., 2013; Raven et al., 2012). Theemergence of oxygenic photosynthetic cyanobacteria ~2.4 Ga caused an increase in atmospheric oxygen concentrations, which coupled with the gradual decline of $CO_2$ levels due to glaciation (Canfield, 2005; Fletcher et al., 2008; Raven et al., 2012) led to an decrease in the ratio of carboxylation to oxygenation activity by Rubisco. Under current atmospheric gas concentrations, the higher cellular $O_2$ increases chances of the substitution of $O_2$ for $CO_2$ in the first reaction of the photosynthetic carbon fixation producing phosphoglycolate (2PG) despite the higher affinity of Rubisco for $CO_2$ (Bauwe et al., 2010; Peterhänsel et al., 2013; Raven et al., 2008; Raven et al., 2012). The photorespiratory $C_2$ cycle, where one molecule of 3PGA is recovered from two molecules of 2PG, allows photosynthesis under oxygenic conditions, and is suggested to

have co-evolved with oxygenic photosynthesis in cyanobacteria (Bauwe et al., 2010). However, this $C_2$ cycle increases the cost of carbon assimilation, whereby one carbon atom is lost in the process as photorespiratory $CO_2$ (Bauwe et al., 2010; Raven et al., 2008; Raven et al., 2012).

$C_3$ plants, such as beans, rice, wheat, potatoes and most of the temperate crops, fix $CO_2$ directly during photosynthesis in the day but lose about 20-30% of potential carbohydrate yield as photorespiratory $CO_2$ under moderate conditions and up to 50% in warm and dry conditions affecting crop yield, nitrogen and water usage (Andersson, 2008; Bauwe et al., 2010; Zhu et al., 2010) (Figure 1.1A, **Figure 1.2**). Despite significant increases in crop production during the Green Revolution period (1954-1980), the production per capita has started to decrease by 1981 due to environmental impacts, e.g. soil erosion, salinisation, waterlogging, urbanisation, nutrient depletion, overcultivation, overgrazing, acidification and soil compaction (Kendall and Pimentel, 1994, 2000). The improvement of crop yield for the past 10 years has been slow (Zhu et al., 2010), but it is crucial to maximise the yield (Covshoff and Hibberd, 2012) in order to cater for the projected increase of 70% in food demand by 2050 due to the exponential population growth (Freibauer et al., 2011). Peterhänsel et al. (2013) note that improving photosynthesis by reducing RuBP oxygenation to improve crop yield can be performed by two approaches: optimisation of enzymatic properties of Rubisco to reduce oxygenation potential (Ellis, 2010) or reduction of oxygenation activity by concentrating $CO_2$ around Rubisco (Hibberd et al., 2008).

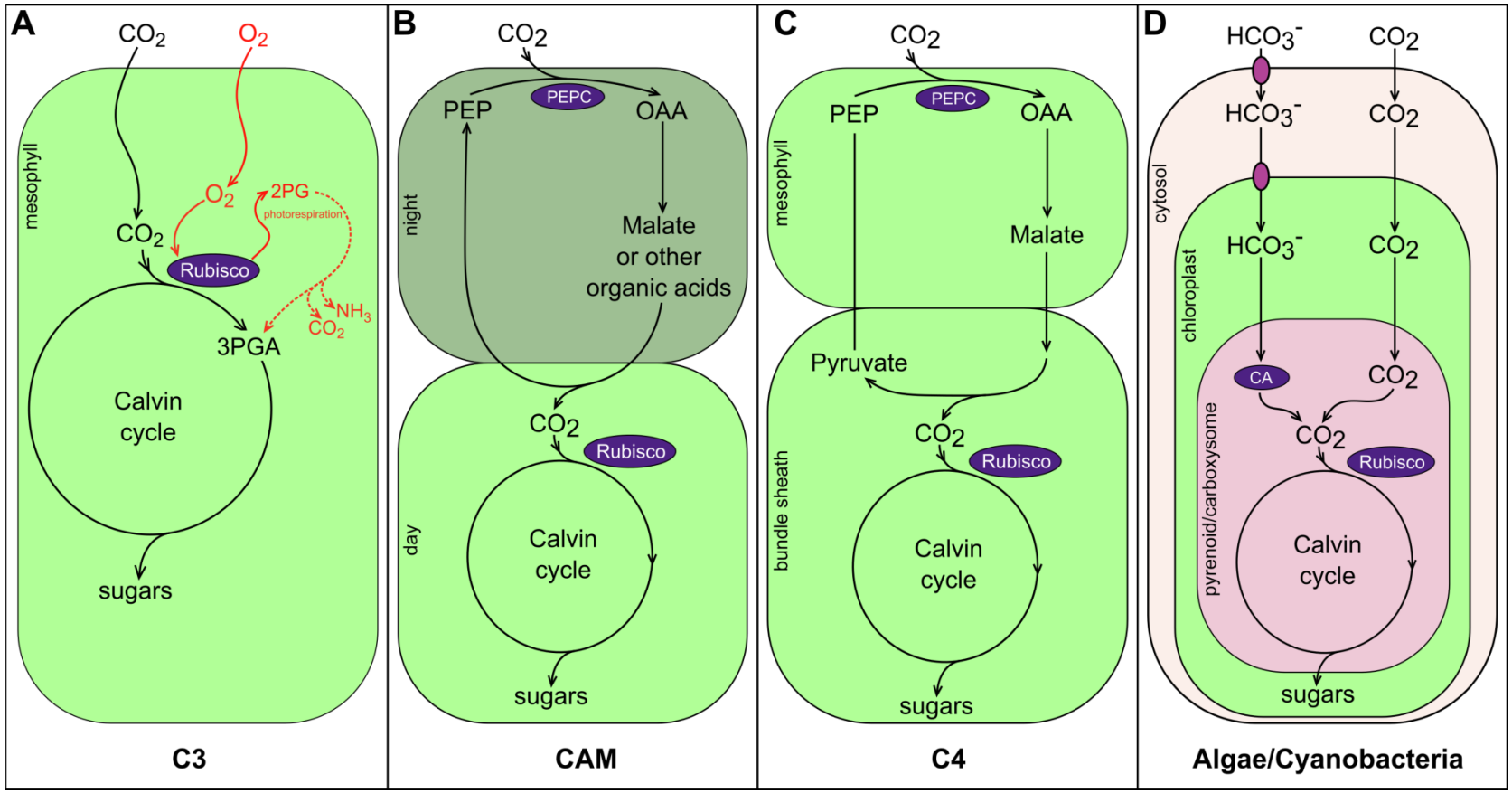



**Figure 1.1 Types of carbon concentrating mechanisms to improve photosynthesis: A) C3 photosynthesis, B) CAM photosynthesis, C) C4 photosynthesis, D) Algal/cyanobacterial photosynthesis. Each type of CCM involves a compartment where Rubisco is accumulated, either a day/night division in CAM plants or a biophysical compartment, e.g. the bundle sheath in $C_4$ plants, the carboxysomes in cyanobacteria and the pyrenoids in some algae, including *Chlamydomonas reinhardtii*.**

As generally known, in green algae and other oxygenic photosynthetic microorganisms, a functional Rubisco holoenzyme is formed by the assembly of eight nuclear-encoded small subunits and eight chloroplast-encoded large subunits (Spreitzer and Salvucci, 2002). The $CO_2$ active site in Rubisco holoenzyme is at the large subunit (LSU), but the small subunit (SSU) contains amino acid sequences which influence the carboxylation catalytic efficiency and the $CO_2/O_2$ specificity of the enzyme (Genkov et al., 2010). The $CO_2/O_2$ specificity factor is defined by the ratio of the catalytic efficiencies of carboxylation to oxygenation (Lee et al., 1993). Based on the current hypothesis, crop yield could be improved by decreasing the non-essential oxygenation or increasing the rate of carboxylation of Rubisco (Ellis, 2010). The work of Genkov et al. (2010) and Meyer et al. (2012) demonstrated that the $CO_2/O_2$ specificity factor of *Chlamydomonas* Rubisco could be altered by switching the Rubisco SSU of *Chlamydomonas* to SSU of another species. The latest review by Sharwood (2017) covered the topic of Rubisco engineering into crop plants, where he pointed out some individual examples of engineering Rubisco in higher plants: Ishikawa et al. (2011) engineered *Sorghum* RbcS into rice plants while Lin et al. (2014) expressed β-carboxysomes with active cyanobacterial Rubisco in the chloroplast of tobacco plants, but more work needs to be done as there is no example of improved photosynthetic rate yet.

The photorespiration pathway has always been regarded as an additional cost of carbon assimilation pathway and is often considered a wasteful process (Betti et al., 2016; Busch et al., 2017). This is due to the energy requirement for the photorespiratory cycle to recycle the 2-phosphoglycolate (2PG) which is a product of oxygenase reaction of Rubisco (Busch et al., 2017). The key hypothesis proposed by researchers is that engineering photorespiratory bypasses will improve recycling of 2PG which will then minimise $CO_2$ and ATP losses (Betti et al., 2016). So far, three successful photorespiratory bypasses have been demonstrated. The "Maier" bypass (**Figure 1.2**, magenta) improves photosynthesis and biomass production by the co-expression of glycolate oxidase, malate synthase and catalase of *E. coli* in the chloroplast of *A. thaliana* (Maier et al., 2012). The "Carvalho" bypass (**Figure 1.2**, green) which introduces *E. coli* glyoxylate carboligase and hydroxypyruvate isomerase into tobacco which would potentially reduce the cost associated with 2PG recycling (Carvalho et al., 2011). However, Carvalho et al. (2011) noted the presence of hydroxypyruvate isomerase in the trangenic plants was not detected, thus no conclusion could be drawn to the effect of this enzyme to improving photosynthesis (Betti et al., 2016). The

"Kebeish" bypass (**Figure 1.2**, blue) which introduces the *Escherichia coli* glycolate catabolic pathway directly into *Arabidopsis thaliana* to bypass photorespiration is shown to improve photosynthesis and biomass production (Kebeish et al., 2007). This bypass, introduced into the oilseed crop *Camelina sativa* by Dalal et al. (2015), was shown to reduce photorespiration and improve photosynthesis in the plant. The authors observed an improvement of seed yield while retaining the seed quality, higher vegetative biomass and faster development only under ambient or low $CO_2$ conditions (Dalal et al., 2015). Nölke et al. (2014) also showed that a partial "Kebeish" bypass in potato (*Solanum tuberosum*) with the expression of *E. coli* glycolate dehydrogenase polyprotein also exhibited increase in shoot biomass and tuber yield.

## 1.1.2 Inorganic Carbon Concentrating Mechanisms and Crop Yield Improvement

### 1.1.2.1 CCM in Higher Plants

Inorganic carbon concentrating mechanisms (CCMs) are present in almost all oxygenic photosynthetic organisms, they increase the $CO_2$ concentration and the $CO_2/O_2$ ratio at the site of Rubisco activity in order to suppress oxygenase activity under the low atmospheric concentrations of $CO_2$ (Raven et al., 2008). A small percentage of terrestrial plants possess a CCM in the form of either $C_4$ photosynthesis (8,100 species, 3% of terrestrial plants) (Kellogg, 2013; Sage, 2016) or crassulacean acid metabolism (CAM) (16,000 species, 7% of all terrestrial plants) (Dodd et al., 2002). The CAM pathway concentrates $CO_2$ around Rubisco separated, first by phosphoenolpyruvate carboxylase (PEPCase) in the night when stomata are opened to fix $CO_2$ into malate which is then decarboxylated in the day for fixation by Rubisco when stomata are closed (Keeley and Rundel, 2003) (Figure 1.1B). To date, there is a lack of example that demonstrated success in engineering CAM into $C_3$ plants, but a system dynamic modelling of CAM developed by Owen and Griffiths (2013) provides a multidirectional modelling framework that incorporates the gene regulatory networks, the metabolic network fluxes and the signalling pathways (Borland and Yang, 2013; DePaoli et al., 2014). However, as discussed in the reviews by DePaoli et al. (2014) and Borland et al. (2014), there is a need to further understand the multiple elements, including the complex gene regulation in the diurnal decarboxylation module, the water use efficiency, the stomatal regulation, and their correlations with one another in CAM photosynthesis before researchers could proceed to engineer CAM photosynthetic elements into $C_3$ crops.

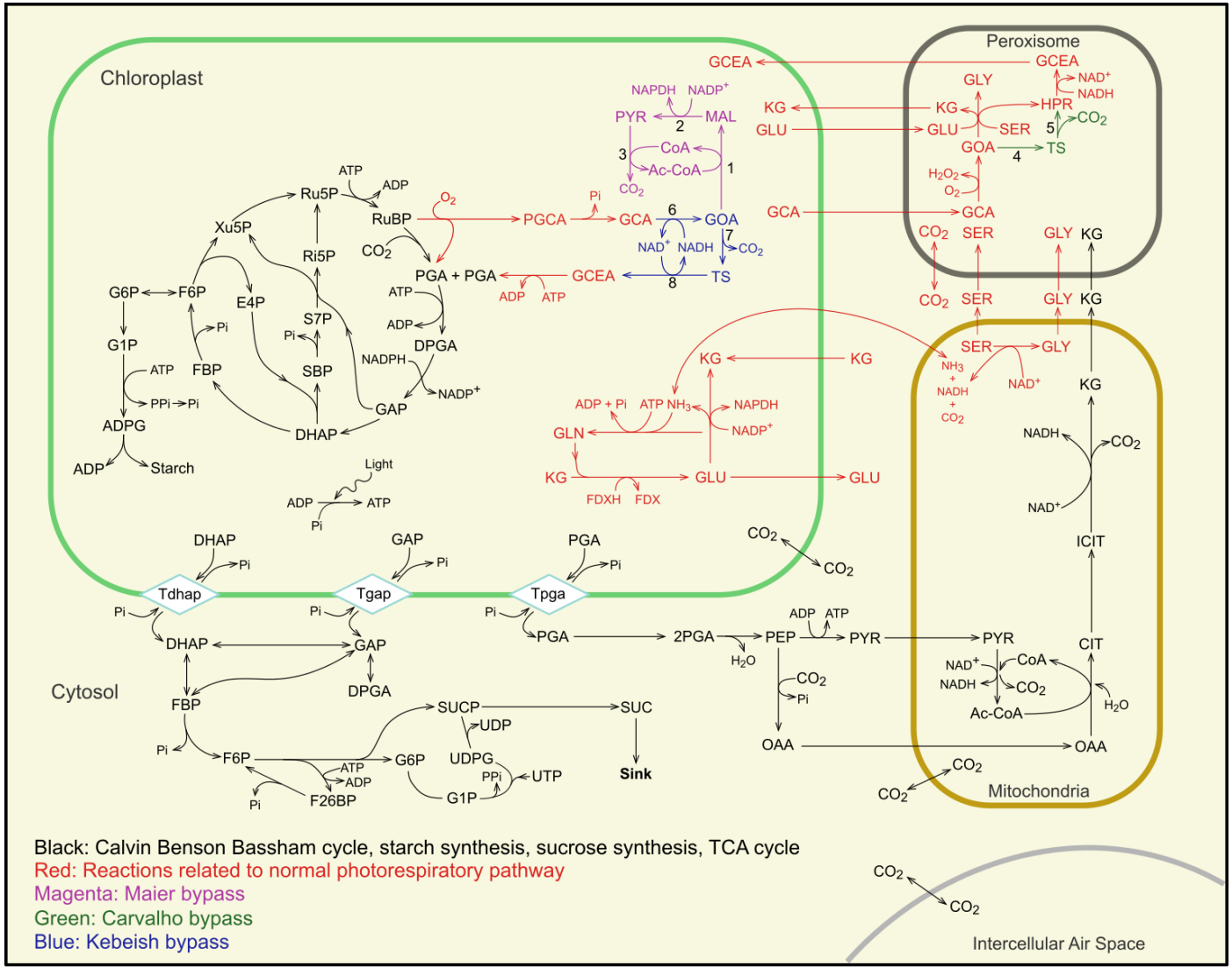



**Figure 1.2 Schematic representation of the photosynthetic C3 pathway, with native photorespiratory C2 pathway and the Maier (magenta), Carvalho (green) and Kebeish (blue) bypasses. 1, malate synthase (EC 2.3.3.9); 2, malate dehydrogenase (EC 1.1.1.40); 3, pyruvate dehydrogenase E1 component (EC 1.2.4.1); 4, gyoxylate carboligase (EC 4.1.1.47P); 5, hydroxypyruvate isomerase (EC 5.3.1.22); 6, glycolate dehydrogenase (EC 1.1.99.14); 7, glyoxylate carboligase (EC 4.1.1.47Chl); 8, 2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60).**

On the other hand, $C_4$ plants possess double carboxylation processes which are separated spatially in different tissues of the leaf, where malate is transported from mesophyll cells to bundle sheath cells to be decarboxylated and concentration of $CO_2$ occurs around Rubisco (Keeley and Rundel, 2003) (Figure 1.1C). With the aim of crop yield improvement, a few international projects have been launched to introduce $C_4$ traits into $C_3$ crop plants, especially in rice (Covshoff and Hibberd, 2012; Hibberd et al., 2008; Jungnick et al., 2014; Peterhänsel et al., 2013). However, as much as introducing the complicated $C_4$ pathway into $C_3$ plants is desirable (Covshoff and Hibberd, 2012; Borland et al., 2014) and a well-described $C_4$ biochemistry pathway is available, there is still a huge knowledge gap and multiple challenges to engineer $C_4$ photosynthetic elements into $C_3$ crops (Schuler et al., 2016). Some of the challenges faced by researchers working to engineer $C_4$ photosynthetic elements into $C_3$ plants include the poor understanding on the formation of Kranz anatomy, the lack of a $C_4$ model monocot species which is as established as *A. thaliana* with a simple and efficient genetic model system, limited knowledge on regulatory sequences and mechanistic understanding of these sequences, the need to expand the *cis*-regulatory elements for bundle sheath or mesophyll cell-specific expression and the identification of a core set of factors in plants which determines evolution of $C_4$ photosynthesis (Schuler et al., 2016). The latest progress made by Burgess et al. (2017) to understand the regulatory architecture of $C_4$ photosynthesis using *Sorghum bicolor*, *Zea mays* and *Setaria italica* in comparison to that of $C_3$ grass, *Brachypodium distachyon* provided an insight to the regulatory sequences in $C_4$ pathway.

#### 1.1.2.2 CCM in Cyanobacteria and Algae

At the same time, the potentially simpler CCM systems in cyanobacteria (Price et al., 2013), algae (Raven et al., 2012) and some terrestrial plants such as hornworts (Smith and Griffiths, 1996) have drawn the attention of researchers to consider introducing CCM components from aquatic organisms into higher plants (Peterhänsel et al., 2013; Price et al., 2013; Jungnick et al., 2014). These $CO_2$ fixation mechanisms involves active transport of $CO_2$, $HCO_3^-$ and/or $H^+$ across one or more membranes (Peterhänsel et al., 2013; Raven et al., 2008) (Figure 1.1D). Engineering aquatic CCM components into crop plants is thought to lead to the increase of $CO_2$ concentration around Rubisco thus improving the ability of the plants to tolerate water stress, reduce wasteful photorespiration (Jungnick et al., 2014). Cyanobacteria and green algae possess highly similar single-cell-based CCMs which are well-characterised, with multiple active

inorganic carbon ($C_i$) uptake systems, but they are known to be significantly different from carbon concentrating mechanism in higher plants (Moroney et al., 2013; Jungnick et al., 2014; Wang et al., 2015).

The cyanobacterial CCM is best studied and researchers have attempted to introduce the system into higher plants in order to raise the stromal dissolved inorganic carbon (DIC) concentration (Price et al., 2013; Jungnick et al., 2014; Rae et al., 2017). Most species of cyanobacteria have a constitutive CCM with basic components of carboxysomes and the associated components despite the variability in the components of plasma membrane and thylakoid membranes (Raven et al., 2017). The minimal requirements for CCM in cyanobacteria (Figure 1.3) are the presence of an energised transporter which accumulates $HCO_3^-$ in the cytosol, the absence of carbonic anhydrase (CA) from the cytosol and the anion-permeable carboxysomes that are packed with all the intracellular CA activity and Rubisco (Price et al., 2008; Rae et al., 2017; Raven et al., 2017) .

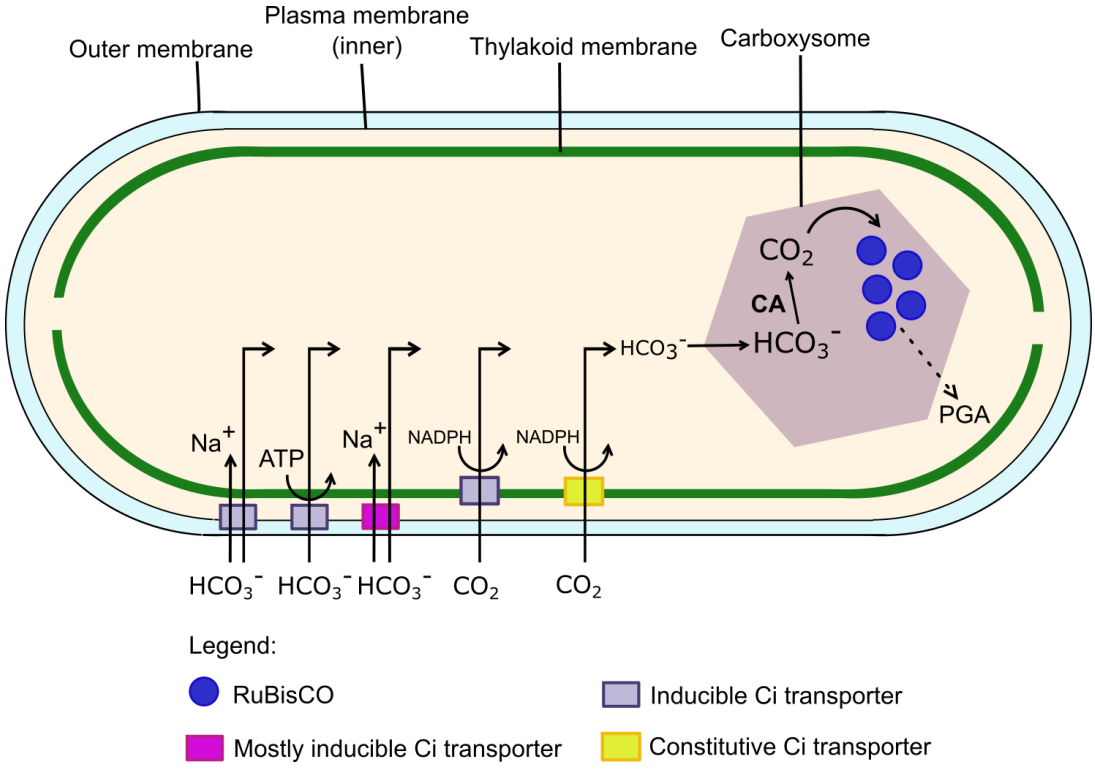



**Figure 1.3 A general illustration of CCM in cyanobacteria. The inducible, mostly inducible and constitutive inorganic carbon transporters are located in the inner plasma membrane and thylakoid membrane, while carbonic anhydrase is absent in the cytosol. Both carbonic anhydrase and Rubisco are located within the carboxysome. Redrawn after Rae et al. (2017)**

The product of the cyanobacterial CCM-related gene, *ictB*, was initially shown to be involved in the $HCO_3^-$ accumulation in the cells as a bicarbonate transporter (Bonfil et al., 1998) but later this role was disproved based on subsequent work with knockouts of four and five DIC uptake systems (Shibata et al., 2002; Xu et al., 2008; Price et al., 2013). Despite not knowing the exact function of *ictB* in *Synechococcus* sp, Lieman-Hurwitz et al. (2003) first demonstrated that expression of *ictB* in both transgenic *A. thaliana* and *Nicotiana tabacum* increases photosynthetic rate under limiting $CO_2$ conditions. The authors also pointed out that Rubisco activity of the transgenic *Arabidopsis* and tobacco plants under relative high humidity was not affected but when grown under limited humidity, the plants exhibited higher enzyme activity and accelerated photosynthetic rate (Lieman-Hurwitz et al., 2003). Subsequent successful expression of *ictB* in $C_3$ crop plants, such as in transgenic rice (Gong et al., 2015) and transgenic soy bean (Hay et al., 2017), further confirmed that *ictB* boosts photosynthetic rate when expressed in higher plants. In the field trials using transgenic soy bean plants expressing *ictB*, there were significant increment in the total mass per plant and a higher photosynthetic rates without changes in stomatal conductance in comparison to wild type (Hay et al., 2017).

Though not a CCM-related gene, the engineering of the cyanobacterial gene encoding the fructose-1,6-/sedoheptulose-1,7-bisphosphatase (FBPA/SBPase) in tobacco was the first demonstration of the improvement of $C_3$ photosynthesis using a gene from cyanobacteria (Miyagawa et al., 2001). This improvement was then affirmed by the overexpression of Arabidopsis cDNA encoding SBPase into tobacco plants (Lefebvre et al., 2005; Rosenthal et al., 2011). Intriguingly, the combined expression of two $C_3$ cycle enzymes, SBPase and FBPA and a cyanobacterial *ictB* into the same tobacco plant had resulted in a significant improvement in photosynthetic rate and biomass yield of the plants grown under greenhouse conditions as close to natural as possible (Simkin et al., 2015). This provides impactful evidence that multi-gene manipulation of photosynthesis could boost crop yield in future.

Unlike the cyanobacterial CCM, where the indicator for the existence of a CCM is unanimously based on the presence of carboxysomes in the cells, the CCM in eukaryotic algae have diverse inorganic carbon transport mechanisms (Raven et al., 2008). Most microalgae have some form of CCM but some green algae, e.g. *Coccomyxa* sp. Have low affinity for $CO_2$ (Colman et al., 2002) suggesting an absence of CCM. Due to the highly variable biochemical properties of algal Rubisco (Giordano et al.,

2005) and the different types of inorganic carbon acquisition systems (Table 1.1) used by different algal species, algal CCMs are suggested to be polyphyletic though very little is known about the underlying mechanism and molecular genetics of each phylum (Colman et al., 2002; Raven et al., 2008). Most evidence supports the idea that movement of inorganic carbon species and/or $H^+$ across the membrane is by active transport (Moroney and Ynalvez, 2007; John A. Raven et al., 2008; Meyer and Griffiths, 2013) forming the basis of algal CCM (Table 1.2).

In terms of the progress made in understand the algal CCM with the aim of crop yield improvement, most work is focussed on the CCM of *Chlamydomonas reinhardtii* (Mackinder, 2017; Meyer and Griffiths, 2013; Rae et al., 2017; Raven et al., 2008; Raven et al., 2017). The work of Spreitzer Lab to engineer Rubisco small subunit of higher plants, namely from spinach (*Spinacia oleracea*), sunflower (*Helianthus annuus* cv Sunspot) and *Arabidopsis* (*A. thaliana* cv. Columbia), into *C. reinhardtii*, was then used to demonstrate the important role of the *Chlamydomonas* Rubisco SSU in the formation of the pyrenoid in *C. reinhardtii* (Genkov et al., 2010; Meyer et al., 2012). However, these hybrid mutants do not possess a pyrenoid and had reduced photosynthetic rate compared to wild type (Genkov et al., 2010; Meyer et al., 2012).

Atkinson et al. (2016) demonstrated transient expression of a list of key algal CCM genes in tobacco leaves and stable transformation into Arabidopsis plants (Table 1.3). Out of the ten genes transformed, GFP-tagged HLA3 and LCIA proteins were targeted to the expected location in transgenic *Arabidopsis* but neither conferred positive effect on the photosynthetic rate of the plants (Atkinson et al., 2016). The effect of *Chlamydomonas* Rubisco SSU was followed up by the authors by the success in complementing *Arabidopsis* Rubisco SSU mutants with a modified *Arabidopsis* SSU and with *Chlamydomonas* SSU (Atkinson et al., 2017). The fact that the modified Arabidopsis SSU was able to assemble into a Rubisco holoenzyme complex without any catalytic defects (Atkinson et al., 2017) indicates it may be possibile to engineer an algal CCM into $C_3$ plants in order to accumulate $CO_2$ at a saturating level around Rubisco (Sharwood, 2017).

**Table 1.1 The mechanism of CCMs in eukaryotic algae. Extracted from Raven et al. (2008).**

| Mechanism | Location within organism | Phylogenetic distribution |
|---|---|---|
| Energised entry of $CO_2$ | Plasma membrane,<br>plastid envelope,<br>Rubisco within stroma or pyrenoid | Many algae,<br>some hornworts with CCMs,<br>some aquatic vascular plants with CCMs |
| Energised entry of $HCO_3^-$ | Plasma membrane,<br>plastid envelope,<br>Rubisco in stroma or pyrenoid | Many algae,<br>hornworts with CCMs,<br>some aquatic vascular plants with CCMs |
| Energised flux of $H^+$ to cell wall, conversion of $HCO_3^-$ to $CO_2$ | Plasma membrane,<br>$CO_2$ flux to Rubisco in stroma | Some algae including characean green algae,<br>hornworts with CCMs,<br>some aquatic vascular plants with CCMs |

**Table 1.2 Cyanobacteria and different phyla of eukaryotic algae which are known to possess carbon concentrating mechanisms. Extracted from (Raven et al., 2012).**

| Taxon | Occurrence of CCM |
|---|---|
| Cyanobacteria | Present in all known species |
| Chlorophyta: Prasinophyceae | Present in all known species |
| Chlorophyta: Chlorophyceae | Present in almost all known species |
| Chlorophyta: Trebouxiophyceae | CCM present or absent known species |
| Chlorophyta: Ulvophyceae | Present in most,<br>Absent in some,<br>$C_4$ in one |
| Streptophyta: Charophyceae | Present in almost all |
| Streptophyta: Embryophytes | Usually absent,<br>pyrenoid-based CCM present in some anthocerophytes,<br>$C_4$ or CAM in some freshwater tracheophytes,<br>CCMs not involving $C_4$ or CAM in some freshwater and all marine tracheophytes |
| Rhodophyta: Bangiophyceae | Present in almost all |
| Rhodophyta: Florideophyceae | Present in many,<br>Absent from some marine and many freshwater |
| Ochrista: Bacillariophyceae | Present in most |
| Ochrista: Fucophyceae | Present in most |
| Ochrista: Tribophyceae | Present in most |
| Ochrista: Chrysophyceae and Synurophyceae | Absent in all |
| Haptophyta: Coccolithaceae | Present |

**Table 1.3 Comparison of the localisation of the algal CCM gene products in the *Chlamydomonas reinhardtii* (native), tobacco, *Arabidopsis* and *Xenopus* oocytes, as demonstrated by Atkinson et al. (2016). The genes were transiently expressed in tobacco or transformed into *Arabidopsis* and *Xenopus* oocytes to identify the location of the algal gene products tagged with fluorescence proteins. *The authors also showed that the gene products of algal *CCP1* and *CCP2* were localised to the mitochondria in *C. reinhardtii* instead of the previously proposed location of the chloroplast membrane.**

| Gene | Putative function in native algal CCM | Location in *C. reinhardtii* | Location in Tobacco (T), *Arabidopsis* (A) & *Xenopus* oocyte (X) |
|---|---|---|---|
| *HLA3* | Ci uptake into cytosol | Plasma membrane | Plasma membrane (T)<br>Plasma membrane (A)<br>Plasma membrane (X) |
| *LCI1* | Ci uptake into cytosol | Plasma membrane | Plasma membrane (T) |
| *LCIA* | Ci transport from cytosol to stroma | Chloroplast membrane | Chloroplast envelope (T)<br>Chloroplast envelope (A)<br>Plasma membrane (X) |
| *CCP1** | Ci transport from cytosol to stroma | Mitochondria | Mitochondria (T) |
| *CCP2** | Ci transport from cytosol to stroma | Mitochondria | Mitochondria (T) |
| *CAH1* | $CO_2$ and $HCO_3^-$ at cell surface | Periplasmic space | Cell periphery (T) |
| *CAH6* | Recapture $CO_2$ leaking from pyrenoid | Stroma/ Flagella | Cytosol (T) |
| *CAH3* | Terminal dehydration of $HCO_3^-$ to $CO_2$ to saturate Rubisco in the pyrenoid | Thylakoid lumen | Cytosol & chloroplast (T) |
| *LCIB* | $CO_2$ uptake or trapping of stromal $CO_2$, pyrenoid localised | Stroma | Stroma (T) |
| *LCIC* | $CO_2$ uptake or trapping of stromal $CO_2$, pyrenoid localised | Stroma | Stroma (T) |

## *1.2* Carbon Concentrating Mechanism in *Chlamydomonas reinhardtii*

Almost 50% of the world's photosynthesis is performed by aquatic microorganisms (Moroney and Ynalvez, 2007; Mackinder et al., 2016). However, as the concentration of dissolved $CO_2$ is less than the $K_m(CO_2)$ of Rubisco, and $O_2$ competes with $CO_2$ for the Rubisco active site, the enzyme functions at only about 25% of its carboxylation capacity, which is highly inefficient. Moreover, $CO_2$ diffuses 10,000 times slower in an aqueous solution than in air and the availability of dissolved $CO_2$ varies with pH, so that $CO_2$ is the major form of inorganic carbon ($C_i$) at acidic pH while only a small fraction of $C_i$ is made up of $CO_2$ under alkaline conditions (Badger et al., 1980; Moroney and Ynalvez, 2007). Due to these challenges, many aquatic photosynthetic organisms have evolved carbon concentrating mechanisms to improve photosynthesis.

The observations that unicellular algae have high affinity for inorganic carbon (Berry et al., 1976) and a lack measurable photorespiration (Lloyd et al., 1977) lead to the suggestion that unicellular green algae do not rely solely on the diffusive entry of $CO_2$ for photosynthesis. The first experimental evidence of the presence of CCM in green alga, *C. reinhardtii*, was presented by Badger et al. in 1980 and since then, the CCM in this species has been studied extensively (for reviews of recent advances, see: Moroney and Ynalvez 2007; Wang et al. 2011, 2015, Meyer and Griffiths 2013, 2015; Jungnick et al. 2014; Mackinder 2017; Meyer et al. 2017). The CCM in a wild-type *C. reinhardtii* cell (Figure 1.4A and B) is inducible under low-$CO_2$ (LC) conditions, and involves three key components: 1) a pyrenoid in the chloroplast, which is the site of Ci accumulation and Rubisco aggregation, 2) inorganic carbon transporters which actively transport Ci into the cell, and 3) the conversion of accumulated Ci to $CO_2$ by carbonic anhydrase (CA) activity (Moroney and Ynalvez, 2007; Jungnick et al., 2014; Wang et al., 2015; Mackinder, 2017) (Figure 1.4C). When the CCM is fully induced, the cell is able to concentrate Ci against a free energy gradient increasing the $CO_2/O_2$ ratio at the Rubisco fixation site which in turn will increase photosynthetic rate and reduce photorespiration. This provides high capacity for net organic carbon production at low external Ci level (Renberg et al., 2010). All the essential and putative components of CCM in *C. reinhardtii* were summarised in Table 1.4 at the end of this section.

### 1.2.1 Pyrenoid, the Site of $CO_2$ Fixation

The ability of green algae and cyanobacteria to fix carbon under an environment with a low concentration of dissolved $CO_2$ is highly dependent on concentrating $CO_2$ around Rubisco (Moroney and Ynalvez, 2007). In all photosynthetic cyanobacteria, Rubisco is concentrated in carboxysomes, which are an essential component of the cell (Figure 1.1D & Figure 1.3) (Giordano et al., 2005; Rae et al., 2017). On the other hand, not all algae possess a pyrenoid. In algae which have pyrenoids, for example, in *Chlamydomonas reinhardtii*, Rubisco is transported and accumulated into this subcompartment which is inducible by both light and $CO_2$ (Mitchell et al., 2014; Rae et al., 2017).

The pyrenoid is defined as an electron-dense proteinaceous microcompartment within the chloroplast (Kuchitsu et al., 1988; Morita et al., 1997; Borkhsenious et al., 1998) which is known to respond to changes of $CO_2$ (Miyachi et al., 1986; Tsuzuki et al., 1986). One of the first descriptions of the presence of a pyrenoid in *Chlamydomonas* was published in Jahrbücher für Wissenschaftliche Botanik by Dill in 1895 which was then compiled in the Journal of Royal Microscopical Society in 1896. The morphological and taxonomical characteristics of the pyrenoid were later described by Griffiths (1980). In 1971, Holdsworth showed Rubisco is present in the pyrenoid of the green alga *Eremosphaera viridis*. By 1988, Kuchitsu et al. demonstrated Rubisco is present in the pyrenoid of *Chlamydomonas reinhardtii*. Subsequent immunolocalisation experiments with an anti-Rubisco antibody determined that almost 99% of Rubisco was aggregated into the pyrenoid under LC conditions (Lacoste-Royal and Gibbs, 1987; Süss et al., 1995; Morita et al., 1997; Borkhsenious et al., 1998; He et al., 2003). Aside from the large and small subunits of Rubisco, other components of the pyrenoid include Rubisco activase (Holdsworth, 1971; Süss et al., 1995; He et al., 2003; Pollock et al., 2003), phosphoribulokinase (PRK) (Holdsworth, 1971; McKay and Gibbs, 1991), nitrate reductase (NR) (Lopez-Ruiz et al., 1985), a putative linker protein, EPYC1/LCI5 (Mackinder et al., 2016), and a small portion of water-soluble photosynthetic enzymes, including ribose-5-phosphate isomerase (RPI), ferredoxin-$NADP^+$ reductase (FNR) and ATP synthase (Süss et al., 1995).

Other CCM components in and around the pyrenoid include a network of pyrenoidal tubules formed by thylakoid membranes traversing the pyrenoid (Borkhsenious et al., 1998; Engel et al., 2015);CAH3 which localises within the tubules and is proposed to function in the supply of $CO_2$ to Rubisco by dehydrating $HCO_3^-$ (Moroney and Ynalvez,

2007; Sinetova et al., 2012; Wang et al., 2015) and LCIB and LCIC, which are homologous LC-induced proteins which form a heteromultimeric complex (Wang et al., 2015).

LCIB is localised to the stroma, itis proposed to catalyse the unidirectional hydration of $CO_2$ to $HCO_3^-$ (Duanmu et al., 2009) and is crucial for cell survival under LC conditions (Wang et al., 2015). Yamano et al. (2010) identified a complex of LCIB and LCIC that accumulates in the peri-pyrenoidal space and proposed that LCIB might either play a role in recapturing $CO_2$ which leaks out from the pyrenoid, or function as a structural barrier which traps $CO_2$ leaking from the pyrenoid. Despite the fact that LCIB has a similar structurr to β-carbonic anhydrases, no carbonic anhydrase activity has been observed (Jin et al., 2016).

Although the exact function of the LCIB/LCIC complex has yet to be verified (Yamano et al., 2010; Wang et al., 2015) it has been shown to interact with LCI11 (Cre16.g663450), while LCIC interacts with Cre16.g662600. Both LCI11 and Cre16.g662600 are suggested to be putative bestrophins, which are known to be chloride ion channels permeable to $HCO_3^-$ in mammalian cells (Qu and Hartzell, 2008). Mackinder et al. (2017) also demonstrated a direct interaction between the two putative bestrophins and another bestrophin-like protein (Cre16.g663400).

Apart from the components found within the pyrenoid and the traversing thylakoid tubule network, a starch sheath which surrounds the pyrenoid in *C. reinhardtii* is usually observed when CCM is induced (Griffiths, 1970; Kuchitsu et al., 1988b; Ramazanov et al., 1994; Moroney and Ynalvez, 2007; Mackinder et al., 2017). Starch granules are observed in the chloroplast and around a proto-pyrenoid when cells are grown at high $CO_2$ while a starch sheath surrounding an enlarged pyrenoid is usually observed when CCM is induced (Kazuyuki Kuchitsu et al., 1988; Villarejo et al., 1996; Moroney and Ynalvez, 2007; Meyer et al., 2017). Controversially, Villarejo et al. (1996) showed that the absence of a pyrenoidal starch sheath does not affect the photosynthetic ability of the cells, suggesting that the starch sheath is not directly involved in the CCM. However, the dramatic changes in the arrangement of the pyrenoidal starch sheath when cells are grown under low $CO_2$ led to the speculation that this arrangement would help to increase the density of Rubisco packing by providing a rigid structure encasing the enzymes (Engel et al., 2015). Another hypothesis is that this pyrenoidal starch sheath functions to prevent leakage of $CO_2$ which is accumulated in the pyrenoid (Meyer and Griffiths, 2015).

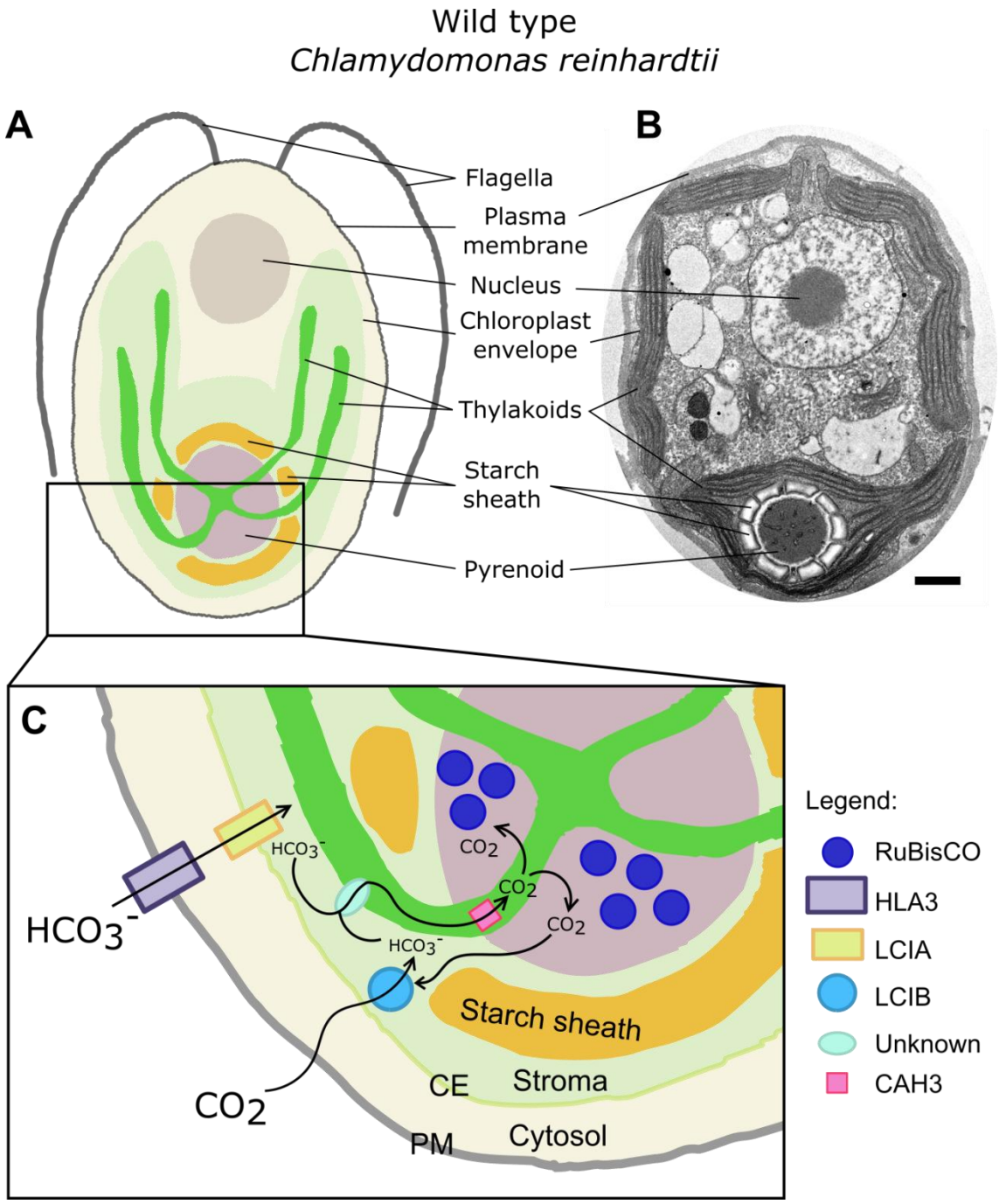



**Figure 1.4 Carbon concentrating mechanism in *Chlamydomonas reinhardtii*. (A) A labelled illustration of a wild type cell of *Chlamydomonas reinhardtii*. (B) Transmission electron micrograph of a wild type *C. reinhardtii* cell. Scale bar 500 nm. (C) The summary of the carbon concentrating mechanism of *C. reinhardtii* under low $CO_2$ conditions with key basic components.**

### 1.2.2 Inorganic Carbon Transporters

The membrane proteins LCI1, LCIA and HLA3 are associated with transport of Ci into the cells at LC levels (Meyer and Griffiths, 2013; Jungnick et al., 2014; Wang et al., 2014, 2015). Ohnishi et al. (2010) showed that the plasma membrane-bound LCI1 is involved in the transport of $HCO_3^-$ during CCM induction. In separate research, LCIA, a member of the formate/nitrite transporter (FNT) family, was shown to be induced under LC conditions leading to uptake of $NO_2^-$ and $HCO_3^-$ (Miura et al., 2004). Another plasma membrane-bound protein involved in the uptake of Ci is HLA3, which is a member of the ATP-binding cassette (ABC) transporter superfamily (Im and Grossman, 2002). Duanmu et al. (2009) confirmed the role of Ci transport by HLA3 and LCIA based on a reduced growth rate, photosynthetic Ci affinity and Ci accumulation in the HLA3/LCIA double knockdown mutant. Recently, LCI1 and HLA3 were shown to form a complex with ACA4 (autoinhibited $Ca^{2+}$-ATPase 4, Cre10.459200) that plays a role in $HCO_3^-$ uptake at the plasma membrane (Mackinder et al., 2017).

The master regulator of the CCM of *C. reinhardtii* is a zinc-finger type transcription regulator, CCM1 (also known as CIA5) (Fang et al., 2012; Wang et al., 2014). Although the expression of CCM1 (CIA5) is not dependent on the levels of $CO_2$ in the cell, the regulation of most of the LC-inducible genes including the induction of the active Ci transporters, LCI1, LCIA and HLA3 under LC conditions requires posttranslational activation of this protein (Moroney et al., 1989; Fukuzawa et al., 2001; Xiang et al., 2001; Miura et al., 2004; Fang et al., 2012). Activation of CCM1 also affects the induction of LCR1, a MYB DNA-binding transcriptional regulator, which in turn activates a periplasmic carbonic anhydrase, CAH1 and LCI6 (of unknown function) by transmission of the amplified LC signal from CCM1 (Yoshioka et al., 2004).

A recent study by Wang et al. (2016) showed the presence of a new CCM regulator, which is a $Ca^{2+}$-binding protein (CAS, Cre12.g497300) that localises in the thylakoid membrane. When the CCM is repressed, CAS is dispersed throughout the chloroplast. In response to light and $CO_2$, CAS moves to the thylakoid tubules in the pyrenoid and is responsible for maintenance of 13 low-$CO_2$-inducible genes, including inorganic carbon transporters, HLA3 and LCIA (Wang et al., 2016). This suggests the existence of a regulatory pathway involved in chloroplast-retrograde signalling from CAS localised in the chloroplast to the nuclear genes in CCM (Wang et al., 2016). Interestingly, accumulation of CCM1 is independent of CAS indicating the possibility that CAS and CCM1 function in parallel to each other in CCM regulation.

The latest interactome study carried out Mackinder et al. (2017) showed that HLA3 and LCI1 formed a complex with each other and with ACA4 (autoinhibited $Ca^{2+}$-ATPase 4, Cre10.g459200), which the authors suggested to be maintaining $H^+$ gradient to help $HCO_3^-$ uptake by HLA3/LCI1 complex. HLA3 is also shown to interact with an EF-hand-containing $Ca^{2+}$/calmodulin-dependent protein kinase (Cre13.g571700) which might be involved in regulating post-translational modification of HLA3, and an adenylate/guanylate cyclase (CYG63, Cre05.g236650) that has a potential role in sensing availability of inorganic carbon at plasma membrane (Mackinder et al., 2017).

The recent reviews on CCM in *C. reinhardtii* proposed the existence of a thylakoid-membrane localised $HCO_3^-$ transporter which would pump $HCO_3^-$ into the thylakoid tubule to be converted into $CO_2$ by a carbonic anhydrase, CAH3 (Meyer and Griffiths, 2015; Mackinder, 2017). The CIA8 protein encoded by *Cia8* (Cre09.g395700) is a putative chloroplast $HCO_3^-$ transporter predicted to be membrane-bound transporter from the SBF/BASS subfamily (Machingura et al., 2017). Due to the low level expression of *Cia8* under high $CO_2$ condition with down-regulation of other transporters, the authors proposed that CIA8 possesses regulatory function aside from direct $HCO_3^-$ uptake from chloroplast cytosol into the thylakoid stroma for $CO_2$ fixation (Machingura et al., 2017).

### 1.2.3 Carbonic Anhydrases

Another key feature of CCM in *C. reinhardtii* is the presence of carbonic anhydrases (CAs) which convert the accumulated $HCO_3^-$ in the cell to $CO_2$ prior to fixation by Rubisco in the pyrenoid. The genome sequence of *C. reinhardtii* (Sabeeha S. Merchant et al., 2007) reveals a large number of encoded CAs but only 10 putative CAs which are involved in the CCM have been identified (Spalding, 2008; Jungnick et al., 2014). CAH1, a periplasmic CA, was one of the first LC-induced genes identified in *C. reinhardtii* while the isozyme, CAH2, is the only one induced under elevated $CO_2$ conditions (>1% (v/v) $CO_2$) (Van and Spalding, 1999; Spalding, 2008). CAH3 is a key CA which is associated with photosystem II and localised within the thylakoid lumen which functions to dehydrate $HCO_3^-$ to $CO_2$ in the pyrenoid for fixation by Rubisco (Karlsson et al., 1998; Spalding, 2008; Blanco-Rivero et al., 2012; Sinetova et al., 2012; Mackinder et al., 2017). When the CCM is induced, CAH3 is phosphorylated and relocalised to the intrapyrenoidal thylakoid lumen (Blanco-Rivero et al., 2012; Sinetova et al., 2012). CAH4 and CAH5 are LC-induced mitochondrial CA isoforms that are

proposed to be involved in pH buffering and anaplerotic reactions (Spalding, 2008). Other recently identified CA isoforms include CAH6, CAH7, CAH8, CAH9 and CAG3 (GLP1) (Spalding, 2008; Wang et al., 2015). Functions of other isoforms, except CAH6, have yet to be determined so far. CAH6 was initially proposed to play a role in recapturing $CO_2$ by conversion to $HCO_3^-$ in the alkaline stroma (Mitra et al., 2004). However, the recent localisation study of CAH6 using fluorescent protein tagging suggested that the protein is localised in the flagella and no signal was detected in the stroma (Mackinder et al., 2017). Interestingly, the authors also showed that CAH6 was detected in the flagella proteome (Pazour et al., 2005) and is an abundant intraflagellar transport (IFT) cargo (Engel et al., 2012). With the evidence suggesting that *Chlamydomnas* shows chemotaxis (Choi et al., 2016) and the possible involvement of carbonic anhydrases in $CO_2$-sensing in plants (Hu et al., 2010), it is possible that CAH6 facilitates the role of $CO_2$-sensing in *Chlamydomonas*.

**Table 1.4 Summary of essential and putative components of CCM in *Chlamydomonas reinhardtii* (Wang et al. 2015; Mackinder et al. 2016; Machingura et al. 2017; Mackinder 2017, the Jonikas lab (personal communication), denoted with *, and this study, denoted with **)**

| Proteins | Cellular location | Putative function | Protein family | Associated $CO_2$ acclimation states |
|---|---|---|---|---|
| Carbonic anhydrases | | | | |
| CAH1 | Periplasm | Periplasmic CA | Carbonic anhydrase, alpha-type | Low or very-low |
| CAH2 | Periplasm | Periplasmic CA | Carbonic anhydrase, alpha-type | High |
| CAH3 | Thylakoid lumen | Key CA | Carbonic anhydrase, alpha-type | Low and very low |
| CAH4 | Mitochondria | Mitochondrial CA | Carbonic anhydrase, beta-type | Low or very-low |
| CAH5 | Mitochondria | Mitochondrial CA | Carbonic anhydrase, beta-type | Low or very-low |
| CAH6 | Flagella | Chloroplast stromal CA | Carbonic anhydrase, beta-type | Unknown |
| Ci transporters | | | | |
| HLA3 | Plasma membrane | $HCO_3^-$ transport | ABC transporter | Very-low |
| LCI1 | Plasma membrane | Ci transport | Unknown | Low or very-low |
| LCIA | Chloroplast envelope | $HCO_3^-$ transport | Formate-nitrite transporter (FNT) | Very-low |
| CCP1 | Mitochondria | Ci transport? | Mitochondrial carrier protein | Very-low |

| Proteins | Cellular location | Putative function | Protein family | Associated $CO_2$ acclimation states |
|---|---|---|---|---|
| CCP2 | Mitochondria | Ci transport? | Mitochondrial carrier protein | Very-low |
| CemA | Chloroplast envelope | Proton extrusion | Cyanobacterial PcxA | Unknown |
| RHP1 | Plasma membrane | $CO_2$ channel | Rh (Rhesus) protein | High |
| RHP2 | Plasma membrane | $CO_2$ channel | Rh (Rhesus) protein) | High |
| CIA8 | Chloroplast envelope | $HCO_3^-$ transport | SBF/BASS subfamily membrane-bound transporter | Low |
| $CO_2$ uptake | | | | |
| LCIB | Pyrenoid periphery | $CO_2$ uptake | Unknown | Low and very-low |
| LCIC | Pyrenoid periphery | $CO_2$ uptake | Unknown | Low and very-low |
| LCID | Unknown | $CO_2$ uptake? | Unknown | Unknown |
| LCIE | Unknown | $CO_2$ uptake? | Unknown | Unknown |
| Gene Expression | | | | |
| CIA5/CCM1 | Nucleus | Transcription regulator | Zinc finger protein | Low and very-low |
| LCR1 | Nucleus | Transcription factor | Myb-type transcription factor | Low and very low |

| Proteins | Cellular location | Putative function | Protein family | Associated $CO_2$ acclimation states |
|---|---|---|---|---|
| CAS | Thylakoid | Chloroplast retrograde signalling | $Ca^{2+}$-binding protein | Low |
| Pyrenoid function | | | | |
| CIA6 | Unknown | Pyrenoid formation | SET domain methyltransferase | Low or very-low |
| EPYC1/LCI5 | Pyrenoid matrix | Pyrenoid matrix formation | Unknown | Low |
| SAGA** | Pyrenoid tubule? | Pyrenoid tubule formation | CBM20 starch binding domain | Low |
| LCI9 | Pyrenoid mesh | Pyrenoid starch formation | CBM20 starch binding domain | Low |
| SAGA-like* | Unknown | Pyrenoid starch formation | CBM20 starch binding domain | Low |
| RCA1 | Pyrenoid matrix | Pyrenoid matrix formation | Unknown | All |
| Cre06.g259100 | Pyrenoid matrix | Pyrenoid matrix formation | Unknown | ? |
| PSBP4 | Pyrenoid periphery | Pyrenoid puncta formation | Unknown | ? |
| SBE3 | Starch sheath | Starch formation | ? | ? |
| STA2 | Starch sheath | Starch formation | ? | ? |

## 1.3 *Chlamydomonas* as a Chassis for Genetic Engineering

### 1.3.1 Mutagenesis Screens

The first mutant screen of *C. reinhardtii* performed by Sager (1954) isolated streptomycin resistant cells, *sr-1* and *sr-2*, from ultraviolet-irradiated cell suspension. By 1960, Levine developed screening method to isolate photosynthetic mutants by testing the ability of the cells to incorporate radioactive $CO_2$. Using this method, he characterised a photosynthetic mutant, *ac-21*, which lost the ability of photosynthetic phosphorylation but retained carboxylation activity of Rubisco (Levine, 1960). A few years later, Bennoun and Levine (1967) developed a chlorophyll fluorescence screen for photosynthetic mutants which is still used extensively. Molecular studies of *C. reinhardtii* were initiated in 1960s with the characterisation of the nuclear and chloroplast DNA (Rochaix, 2002).

The first successful nuclear genome transformation using a yeast gene as a selectable marker by Rochaix and van Dillewijn (1982) led to subsequent development of *Chlamydomonas* selectable markers to be incorporated into transformation vectors for high frequency insertional mutagenesis (Kindle, 1990; Tam and Lefebvre, 1993; Rochaix, 2002). Paromomycin resistance is widely used as a marker for nuclear transformation and promoter tagging. The *hsp70A::5'rbcS2::InI::aphVIII::3'rbcS2* hybrid cassette is one of the first stable enzymatic selection markers for nuclear genome manipulation in *C. reinhardtii* (Sizova et al., 2001). Since homologous recombination in *C. reinhardtii* occurs at a very low frequency and DNA can integrate randomly in the nuclear genome, Tam and Lefebvre (1993) showed nuclear transformation can be used for insertional mutagenesis as well as for complementation assay.

With the release of the first draft of *C. reinhardtii* genome by Grossman et al. (2003) and subsequently a highly annotated version by Merchant et al. (2007), insertional mutagenesis in forward genetics research is useful to identify genes involved in given pathways (Dent et al., 2005). The “Spot Test” was first designed by Spreitzer and Mets (1981) to recover photoautotrophic mutants generated by insertional mutagenesis and has since been used as a standard screening method for high-$CO_2$-requiring mutants (Spalding et al., 1983; Moroney et al., 1989; Van et al., 2001; Dent et al., 2005; Genkov et al., 2010; Meyer et al., 2012; Zhang et al., 2014). Methods of identification of the flanking region of the cassette in the insertional mutants in *C. reinhardtii* include plasmid rescue (Tam and Lefebvre, 1993), thermal asymmetric interlaced PCR (Dent et

al., 2005), restriction enzyme site-direction amplification PCR (González-Ballester et al., 2005), 3'-rapid amplification of cDNA ends (Meslet-Cladière and Vallon, 2012), SiteFinding PCR (Li et al., 2012), and ChlaMmeSeq (Zhang et al., 2014). Whole genome sequencing (Merchant et al., 2007) is an alternative method to resolve the insertional mutants which could not be identified using the methods mentioned above.

The development of both insertional mutagenesis and screening methods led to discovery of various CCM mutants, including a mutant line of the CCM master regulator, *cia5*, isolated by Moroney et al. (1989) from UV-irradiated cell suspension and later characterised by Fukuzawa et al. (2001) and Xiang et al. (2001) with mutation at a zinc-finger type transcription regulator and known to possess defective pyrenoids. Other CCM mutants with pyrenoid defects which were isolated include a Rubisco activase mutant, *Rca*, which does not have functional Rubisco due to the lack of Rubisco activase and possesses defective pyrenoids (Pollock et al., 2003), the *18-7G* mutant with a missense mutation in the gene for the Rubisco large subunit and 'empty' pyrenoids (Spreitzer et al., 1985; Rawat et al., 1996), the *cia6* mutant, with insertion in the *CIA6* gene, contains mislocalised Rubisco and non-wild-type-like pyrenoid structure (Ma et al., 2011) and the *abl* mutants which have mislocalised LCIB protein with significantly smaller pyrenoid (Yamano et al., 2014).

The insertional mutagenesis screen conducted by the Jonikas Lab (Zhang et al., 2014) resulted in the discovery of *epyc1*, a knockout mutant of the Rubisco linker protein EPYC1 (also known as LCI5) which affects the formation of the pyrenoid matrix causing defective CCM (Mackinder et al., 2016). The mutants from the same CCM screen were subjected to two independent secondary screens by the Griffiths group (University of Cambridge) using immunofluorescence with anti-Rubisco antibody (Caspari, 2015) and differential interference contrast (DIC) microscopy (Dr. Moritz Meyer, personal communication) to select CCM mutants with pyrenoidal defects. Further work on this set of CCM mutants will be discussed in Chapter 3 of this thesis.

### 1.3.2 Development of the *Chlamydomonas* Modular Cloning System

In order to understand the roles played by the putative genes involved in *Chlamydomonas* CCM, genomic complementation assay (Rochaix, 1995) is carried out to rescue the phenotype loss in the CCM defective mutants. The main discoveries of CCM protein functions which utilised the complementation assay including the mode of action of the "master switch" of CCM, CCM1/CIA5 (Fukuzawa et al., 2001; Xiang et al.,

2001; Chen et al., 2017), the Myb transcription factor, LCR1 which regulates the periplasmic carbonic anhydrase, CAH1 (Yoshioka et al., 2004), a thylakoid lumen carbonic anhydrase 3 (CAH3) (Karlsson et al., 1998), the CIA6 which plays a role in formation of the pyrenoid (Ma et al., 2011), an inorganic carbon carbon acquisition gene CIA7 (Ynalvez and Moroney, 2008), the function of Rubisco small subunit in pyrenoid formation (Genkov and Spreitzer, 2009; Meyer et al., 2012), the function of Rubisco linker protein, EPYC1/LCI5 (Mackinder et al., 2016) and most recently, the function of the chloroplast retrograde signalling $Ca^{2+}$-binding protein (CAS) (Wang et al., 2016).

As stated in the methods of the studies carried out to understand the roles of CCM-related genes, the first step to run a genetic complementation assay to rescue the phenotype loss is to construct the plasmid which contains the gene that had been knocked-out during mutagenesis (Karlsson et al., 1998; Fukuzawa et al., 2001; Xiang et al., 2001; Yoshioka et al., 2004; Ynalvez and Moroney, 2008; Ma et al., 2011; Meyer et al., 2012; Mackinder et al., 2016; Wang et al., 2016; Shen et al., 2017). Upon transformation, the second part of genetic complementation assay is the screening, selection and characterisation of positive transformants. The growth spot test (Spreitzer and Mets, 1981) and oxygen evolution are widely used in the algal CCM community as methods to observe restoration of phenotypes (Fukuzawa et al., 2001; Xiang et al., 2001; Yoshioka et al., 2004; Ynalvez and Moroney, 2008; Ma et al., 2011; Meyer et al., 2012; Mackinder et al., 2016; Wang et al., 2016; Shen et al., 2017) .

With the improvement of the annotation of the *Chlamydomonas* genome project since the first genome of *C. reinhardtii* was published (Merchant et al., 2007), the process of understanding the gene prior to designing the complementation assay becomes easier with the availability of different datasets available through JGI v5.5 (Blaby et al., 2014). Multiple cloning methods have been used for genetic complementation assays (Table 1.5), with Gibson assembly being the preferred in the cited papers. However, the flexibility of modular cloning in the ability to assemble a multi-gene construct (Chapter 4) suggests the possibility of replacing the more time-consuming and rigid Gibson assembly with a modular cloning system with interchangeable genetic parts.

In plant science research, the Plant MoClo Tool Kit (Engler et al., 2014) which provides 96 different basic modules which can be domesticated is pioneering the use of a low-cost, interchangeable system. In terms of *Chlamydomonas*-specific research, multiple promoters (Table 1.6) and antibiotic selection genes (Table 1.7) had been studied and published (Barrera and Mayfield, 2013). With this knowledge and the existing Plant

MoClo Tool Kit, the *Chlamydomonas* MoClo Kit is a project in progress led by multiple algal research groups in UK and Europe (Dr. Francisco Navarro and Dr. Payam Mehrshahi, University of Cambridge, personal communication) which could benefit future *Chlamydomonas* research in understanding the functions of the putative genes in not just CCM but in other areas as well.

**Table 1.5 Summary of the commonly-used cloning techniques in genetic complementation assays extracted from Binder et al. (2014) and Addgene (2017).**

| Method | Description | Advantages | Disadvantages |
|---|---|---|---|
| Restriction enzyme cloning | Restriction enzymes are used to cut double-stranded DNA into fragments with precise 5' or 3' single-strand overhangs (sticky ends) or blunt ends.<br>A subsequent ligation reaction will fuse two pieces of DNA with complementary overhangs. | Availability of hundreds of enzymes with specific recognition sites and predictable results.<br>Relatively cheap.<br>Easy to subclone DNA fragments with the restriction sites into plasmids with the same sites. | Multiple steps of digestion-cloning-transformation-selection to assemble a single plasmid. |
| Gateway® Recombination Cloning | Recombination-based cloning method which allows easy movement of a DNA fragment with specific recombination sites from one plasmid to another in a single recombination reaction.<br>When the DNA fragment is cloned into a donor plasmid (entry clone), the DNA could be easily shuttled into any compatible Gateway® Destination vector. | Simple process.<br>Less time-consuming comparing to restriction enzyme-based cloning method. | Reagents are very expensive.<br>Limited to assembly of maximum four elements in one step.<br>The assembled construct contain predefined linker sequences. |

| | | | |
|---|---|---|---|
| TOPO® Cloning | Also known as TA cloning.<br>Amplification with Taq polymerase leaves a single adenosine (A) overhang at the 3'end of the product which could be fused to cloning-ready TOPO® vectors with DNA topoisomerase I. | Easy way to clone fresh PCR product into a plasmid which could be sequenced to ensure the product is mutation-free prior to subcloning to another vector.<br>TOPO® ready Gateway® Entry plasmids are available. | Very few plasmid backbones which are TOPO® ready but it is not be feasible to create a TOPO® vector in the lab.<br>Efficiency is dependent of the polymerase used, while ligation efficiency depends on the single A overhangs which could degrade over time. |
| Isothermal Assembly Reaction (Gibson Assembly) | Gibson assembly utilises the properties of 5' exonuclease, polymerase and ligase for successful assembly of construct.<br>The 3' single-stranded overhangs are generated using 5' exonuclease which digests the 5' end of the double stranded DNA fragments.<br>When performed in a reaction mixture which contains two or more DNA fragments with 20-40bp homology at their ends, the DNA fragments with | Most commonly used method.<br>Allows simple assembly of multiple fragments of DNA in a chosen orientation with no unwanted sequence at the junctions.<br>Long overlapping region between fragments ensures correct order of assembled fragments. | Time-consuming and expensive to design and synthesise long primers.<br>Parts could not be reused, interchanged nor shared due to fragment-specific long primers. |

| | | | |
|---|---|---|---|
| | complementary sticky ends will anneal.<br>The polymerase would fill the remaining region of ssDNA and the ligase would fuse the nicks generating a single DNA fragment. | | |
| Type IIS Assembly (Golden Gate/MoClo) | Modular cloning system that uses the unique properties of type IIS restriction enzymes to cut dsDNA at a specific distance from the recognition sites, allowing customised overhangs.<br>The customised overhangs will allow multiple DNA parts (promoters, genes, terminators, tags, etc.) to be assembled correctly in a single reaction. | It is a one-step, one-pot reaction where digestion and ligation are carried out in a single reaction using a single restriction enzyme.<br>This preserves directionality of the cloning reaction.<br>No scar is retained in the final product as the restriction sites both in the insert and plasmid would be removed.<br>DNA parts can be reused and shared with other labs and species-specific domestication could be done.<br>Enables multi-gene construct to be assembled. | Type IIS enzyme recognition sites in the fragment need to be removed prior to assembly by creating silent point mutation in the internal sites of the fragments.<br>Flanking sequences which differ only by one base would result in unintended ligation products. |
| Sequence and | This method creates 5' overhangs on both | Lower cost comparing to Gibson | Requires sequence validation at every |

| | | | |
|---|---|---|---|
| Ligation Independent Cloning | the vector and the insert using the 3'-5' exonuclease activity of T4 DNA polymerase.<br>The T4 polymerase will continue to function as an exonuclease in the presence of a single free dNTP until a base is exposed on the single strand overhang complementary to the free nucleotide.<br>Then T4 will switch to its polymerase activity to add the free base.<br>Sequence assembly occurs in *E. coli* utilising the DNA repair power of the bacteria. | assembly as T4 polymerase is cheaper. | step due to error-prone PCR amplification of inserts and assembly into linearised backbone vectors.<br>Limited by the availability of DNA sequence termini without secondary structures. |

**Table 1.6 Nuclear and chloroplast promoters used for transgene expression in *Chlamydomonas reinhardtii*. Extracted from Barrera and Mayfield (2013)**

| Promoter | Description | Genome | Source | Host |
|---|---|---|---|---|
| *hsp70A* | Heat shock protein 70A | Nuclear | *C. reinhardtii* | *C. reinhardtii* |
| *cyc6* | Cytochrome c6 | Nuclear | *C. reinhardtii* | *C. reinhardtii* |
| *psaD* | Photosystem I complex protein | Nuclear | *C. reinhardtii* | *C. reinhardtii* |
| *nia1* | Nitrate reductase | Nuclear | *C. reinhardtii* | *C. reinhardtii* |
| *cop* | Chlamyopsin | Nuclear | *C. reinhardtii* | *C. reinhardtii* |
| *rbcs2* | Small subunit ribulose bisphosphate carboxylase | Nuclear | *C. reinhardtii* | *C. reinhardtii* |
| *CaMV* | Cauliflower mosaic virus 35S | Nuclear | Cauliflower mosaic virus | *C. reinhardtii* |
| *Nos* | Nopaline synthase | Nuclear | *Agrobacterium tumefaciens* | *C. reinhardtii* |
| *β-2-tub* | β-2-tubulin | Nuclear | *C. reinhardtii* | *C. reinhardtii* |
| *psbA* | Photosystem II protein D1 | Chloroplast | *C. reinhardtii* | *C. reinhardtii* |
| *psbD* | Photosystem II protein D2 | Chloroplast | *C. reinhardtii* | *C. reinhardtii* |
| *atpA* | ATPase alpha subunit | Chloroplast | *C. reinhardtii* | *C. reinhardtii* |
| *rbcL* | Ribulose bisphosphate carboxylase large subunit | Chloroplast | *C. reinhardtii* | *C. reinhardtii* |
| *cabII-1* | Chlorophyll-*ab* binding protein | Chloroplast | *C. reinhardtii* | *C. reinhardtii* |

**Table 1.7 Nuclear and chloroplast selectable markers used in *Chlamydomonas reinhardtii*. Extracted from Barrera and Mayfield (2013).**

| Marker | Selection |
|---|---|
| *ARG9* | Arginine prototrophy |
| *aph7* | Hygromycin B resistance (aminoglycoside phosphotransferase) |
| *ALS* | Sulfometuron methyl resistance (mutated endogenous acetalolactase gene) |
| *aphVIII* | paromomycin/kanamycin resistance (aminoglycoside phosphotransferase) |
| *act-2* | Cycloheximide resistance (mutated endogenous ribosomal L41) |
| *aphA-6* | Kanamycin/amikacin resistance (aminoglycoside phosphotransferase) |
| *PPX1* | Porphyric herbicide resistance (mutated endogenous protoporphyrinogen oxidase) |
| *ble* | Zeocin resistance (stoichiometrically neutralises phleomycins) |
| *NIC7* | Nicotinamide prototrophy |
| *THI-10* | Thiamine prototrophy |
| *cat* | Chloramphenicol resistance (acetyltransferase) |
| *CRY1-1* | Cryptopleurine/emetine resistance (mutated endogenous ribosomal S14) |
| *nptII* | Neomycin resistance (phosphotransferase) |
| *aadA* | Spectinomycin/Streptomycin resistance (adenylyltransferase) |
| *oee-1* | Oxygen-evolving enhancer protein (restores photosynthesis) |
| *ARG7* | Arginine prototrophy |
| *NIT1 (NIA1)* | Nitrate prototrophy |

## 1.4 General Aims and Hypotheses of this Study

Since the first written record on the existence of the pyrenoid in *Chlamydomonas* by Dill (1895), the understanding the role of the pyrenoid in CCM has significantly improved. There is an increasing knowledge on key CCM components, especially with the publication of the spatial interactome study which reveals the protein-protein interaction between and among key CCM proteins (Mackinder et al., 2017). However, there remains a void on how the structure of the pyrenoid is formed, the minimal components required to recreate a pyrenoid structure in the chloroplast and the function of the thylakoid tubule network in the pyrenoid.

As part of the Combining Algal and Plant Photosynthesis (CAPP) project, the work completed for this doctoral thesis focused on the main vision of CAPP project to assemble a pyrenoid structure in higher plants. The strains used in this thesis were selected mutants and the corresponding wild type from the insertional mutagenesis screen conducted by the Jonikas Lab as part of the CAPP project. The main objectives of this thesis are:

1. To identify the key components of the biogenesis of the pyrenoid (Chapter 3). The approaches taken to characterise the nine mutants selected from the insertional mutagenesis screen include physiological and detailed morphological observation of each of the mutants, as well as whole genome sequencing to confirm the identity of the genes related to the defects exhibited by the mutants.
2. To propose function of a gene identified in the knockout mutants from the previous chapter which is functionally unannotated (Chapter 4). The experimental design included the use of Golden Gate cloning technique on the mutant to attempt genetic complementation. A new level zero part containing the coding sequence of the gene of interest was generated and assembled with parts from the *Chlamydomonas* MoClo kit.
3. To understand the role of the thylakoid tubule network (the pyrenoid core) in the biogenesis of the pyrenoid (Chapter 5). Complemented lines of the mutant with defects in the formation of the thylakoid tubule network in the pyrenoid were used to understand the role played by the gene knocked out in the mutant. Co-immunoprecipitation and LC-MS/MS was used as an approach to identify potential protein interactors of the protein which is proven crucial to form a structurally functional pyrenoid.

# 2 MATERIALS AND METHODS

## 2.1 Physiological Analysis

### 2.1.1 Growth of *Chlamydomonas reinhardtii* strains

All *Chlamydomonas reinhardtii* mutant strains (A5, A7, A8, A11, C4, C8, D4, F8 and H10) and the corresponding wild type (cMJ030) were kindly provided by Dr. Martin Jonikas, University of Princeton, NJ (Previously, Carnegie Institute for Science, Stanford, CA) (Zhang et al., 2014) as part of the CAPP (Combining Algal and Plant Photosynthesis) collaborative project. Complemented lines, *saga*::*SAGA-Venus-3xFLAG* A7 (SVA7) and *saga*::*SAGA-Venus-3xFLAG* B2 (SVB2) were provided by Alan Itakura (Stanford University) under the same CAPP collaboration. The RbcS1 hybrid, (SSSO with a spinach SSU and *Chlamydomonas* LSU) was originally generated in the Spreitzer Lab (Genkov et al., 2010), and was initially thought to lack a pyrenoid, although subsequent studies have revealed that the defining knotted tubules are present in this mutant (Caspari et al., 2017; Meyer et al., 2017).

*Chlamydomonas reinhardtii* mutant strains (walled) were maintained on solid Tris acetate phosphate (TAP) medium (Spreitzer and Mets, 1981) with Hutner's trace elements (Hutner et al., 1950) and 20 mg/l of paromomycin in the dark, while cMJ030 and SSSO were maintained in solid TAP medium in the dark. For experiments which required photoautotrophic growth and CCM induction, starter cultures from fresh plates were first inoculated in liquid TAP medium. About 5 x $10^4$ cells were pelleted at 3000 x

g at room temperature and transferred to Tris-phosphate (TP) medium and bubbled with 5% (v/v) carbon dioxide ($CO_2$) until mid-log phase. Unless stated, all liquid cultures were grown in an Innova 42 incubator (New Brunswick Scientific, Enfield, CT, USA) at 25°C, 125 rpm shaking with light intensity of 50-100 µmol photons $m^{-2}$ $s^{-1}$. Induction of CCM was performed by switching gas supply of 5% (v/v) $CO_2$ to ambient air supply (0.04% (v/v) $CO_2$) for up to 24 h during analyses. In all, data described in this thesis as high $CO_2$ or HC, indicates that cultures were supplied with 5% (v/v) $CO_2$, while low $CO_2$ or LC indicates 0.04% (v/v) $CO_2$ in the ambient air supply.

TAP medium was prepared by adding 10 ml of 100X phosphate stock (14.34 g $K_2HPO_4$ and 7.26 g $KH_2PO_4$ made up to 1L with $dH_2O$), 10 ml of 2M Tris stock (242.2 g Tris base and 149 ml concentrated HCl made up to 1L with $dH_2O$), 10 ml of 1M acetate stock (24.2 g sodium acetate trihydrate made up to 200 ml with $dH_2O$), 50 ml of Beijerinck's Solution (8 g $NH_4Cl$, 1 g $CaCl_2.2H_2O$ and 2 g $MgSO_4.7H_2O$ made up to 1L with $dH_2O$), 1 ml of Hutner's trace elements (Hutner et al., 1950). TP medium was prepared using the same recipe minus the addition of the acetate stock.

### 2.1.2 Spot Growth Tests and Growth Experiments

Actively growing cells were first suspended to similar cell densities of about $OD_{750}$=0.2 in TP medium and diluted by factors of 1, 10, 100 and 1000, and spotted (10 µl) onto solid TP plates and grown in different $CO_2$ concentrations (5% and 0.04%) for 7 days with a control set grown heterotrophically on solid TAP in the dark.

### 2.1.3 Oxygen Evolution

*Chlamydomonas* cells grown in Tris-phosphate medium were harvested by centrifugation at 3,000 x g for 3 min at 20°C and resuspended in 25 mM HEPES-KOH (pH 7.3) to the concentration of ~200 µg chlorophyll $ml^{-1}$. Aliquots of cells (1 ml) were added to an OXYV1 Hansatech Oxyview System (Hansatech Instruments, King's Lynn, UK) maintained at 23°C using a circulating water bath. The chamber was sealed and illuminated with 200-300 µmol photons $m^{-2}$ $s^{-1}$ until the cells had depleted the internal inorganic carbon storage. When net oxygen evolution ceased, 10-µl aliquots of sodium bicarbonate were added to the cells at 30-s intervals. The cumulative concentrations of $HCO_3^-$ after each addition were as follows: 2.5, 5, 10, 25, 50, 100, 250, 500, 1000 and 2000 µM. The rate of oxygen evolution was recorded per second using the PicoLog 1216 Data Logger (Pico Technologies, St. Neots, UK). $K_{0.5}$ values was calculated using

the Excel spreadsheet created for the PhD work of Dr. Madeline Mitchell (Mitchell, 2014), where the equation describing Michaelis-Menten kinetics was fitted to curves of external inorganic carbon versus photosynthetic rate and solved for $K_{0.5}$ ($K_m$) and the maximum rate ($V_{max}$) using the least squares method.

### 2.1.4 Chlorophyll Extraction and Measurement

Cells were harvested by centrifugation at 5,000 x g for 5 min at 20°C then resuspended in 1 ml of methanol before centrifugation at the same speed for 5 min to pellet the cell debris. The absorbance of the supernatant was measured at 652, 665 and 750 nm. Total chlorophyll concentration was calculated using the equations of Porra et al. (1989).

## 2.2 Microscopy Methods

### 2.2.1 Fixing and Embedding

*Chlamydomonas* cells at mid-log phase were harvested by centrifugation at 5,000 x g and fixed for 1 h at room temperature with fixation buffer (2% (v/v) glutaraldehyde, 0.01% (v/v) hydrogen peroxide in TP medium). Samples were then fixed and osmicated for 1 h at room temperature in an osmium mix (1% (v/v) $OsO_4$, 1.5% (w/v) $K_3[Fe(CN)_6]$, 2 mM $CaCl_2$). Bulk staining was performed with 2% (w/v) uranyl acetate for 1 h at room temperature. At the interval of each of the subsequent step, three 5 min washes with distilled water were performed. Samples were dehydrated progressively in 70% (v/v) and 95% (v/v) ethanol, followed by two washes in 100% ethanol and 100% acetonitrile.

The samples were then embedded in epoxy resin (34% (w/v) Quetol 651, 44% (w/v) nonenyl succinic anhydride, 20% (w/v) methyl-5-norbonene-2,3-dicarboxylic anhydride, and 2% (w/v) catalyst dimethylbenzylamine (Agar Scientific, Essex, UK)). The fixed cells were first mixed with mixture of acetonitrile and epoxy resin at a ratio of 1:1 and left to settle overnight. Over the next two days, the cells were refreshed with 100% epoxy resin. Lastly, the samples were cured at 60°C for at least 24 h.

Sections of 50 nm thickness were prepared with a Leica Ultracut UCT (Leica Microsystems, Milton Keynes, UK), mounted on 300 mesh copper grids (for straight imaging) or 300 mesh nickel grids (for immunogold labeling), and counterstained with uranyl acetate followed by lead citrate (Ms. Lyn Carter, Cambridge Advanced Imaging Centre, UK).

The uranyl acetate solution was prepared by saturating 100 cm of 50% (v/v) methanol with uranyl acetate and leaving to stand for 2 days. The lead citrate stain was prepared by mixing 1.33 g of lead nitrate and 1.76 g of trisodium citrate in 30 ml of boiled and cooled double distilled water. The solution was mixed by shaking for 30 min and added with 0.8 ml of 10M NaOH, then diluted to 50 ml. The lead citrate staining solution was left to stand overnight before use. For staining, a small aliquot of each of the solutions were added to individual 1.5 ml tubes and centrifuged at maximum speed for 3 min just before use to remove any precipitates. The clear supernatant of each solution would be used for staining. The copper grids were first stained in the uranyl acetate solution. The stain was jet-washed off with 50% (v/v) methanol and then blotted dry. Then, the grids were transferred to lead citrate stain and after that the stain was jet-washed off using distilled water and blotted dry.

### 2.2.2 Immunogold Labelling

Samples mounted on 300 mesh nickel grids were treated with 4% (w/v) Na-meta-periodate for 15 min and 1% (w/v) periodic acid for 5 min to remove superficial osmium and unmask the epitopes. Then, the samples were blocked for 5 min in blocking buffer of 1% BSA (w/v) in high-salt Tris-buffered saline (HSTBSTT) (0.05% (v/v) Tween 20, 0.05% (v/v) Triton X-100 and 500 mM NaCl). The sections were then incubated overnight at room temperature with diluted antibody (1:1000 dilutions of rabbit anti-Rubisco antibody in blocking buffer) followed by two 5-min washes in HSTBSTT and two 5-min washes in $dH_2O$. To detect bound antibody, the sections were incubated with a secondary antibody (1:200 dilutions of goat anti-rabbit 15-nm gold conjugates (BBI Solutions, Cardiff, UK) in blocking buffer) for 1 h at room temperature followed by two 5-min washes in HSTBSTT and two 5-min washes in $dH_2O$. The sections were dried and kept for further observation.

### 2.2.3 Transmission Electron Microscopy for Image Acquisition

All sections were examined using Tecnai G2 80-200kv transmission electron microscope (FEI Company, Hillboro, OR, USA) and imaged with AMT Image Capture Engine software (Advanced Microscopy Techniques, Woburn, MA, USA). All image analyses were performed using ImageJ2 (Fiji) (Schindelin et al., 2012; Schneider et al., 2012; Rueden et al., 2017).

## 2.2.4 Evaluation of Immunogold Labeling

Sections were examined using TEM and evaluation of immunogold labeling was made by using the cell counter and measurement functions of ImageJ2 (Fiji). The percentage of aggregation of gold particles in pyrenoid and chloroplast stromal area was calculated using these equations:

$$d_i = \frac{n_i}{A_i}$$

$$N_i = (d_i - d_k) \times n_i$$

$$\%N_i = \frac{N_i}{N_i + N_j} \times 100$$

$where,$
$d = density,$
$n = number\ of\ particles\ counted,$
$N = normalised\ number\ of\ particles,$
$A = area,$
$i = pyrenoid,$
$j = chloroplast\ stroma,$
$k = background$

## 2.2.5 Indirect Immunofluorescence (IF)

### 2.2.5.1 Preparation of materials

#### *2.2.5.1.1 Slides*

After washing with acid alcohol, all the slides used for immunofluorescence were coated with 0.1% (v/v) poly-L-lysine by adding 10 µl of the solution and smearing across the slide with a cover slip. The slides were then allowed to dry and could be stored for up to seven days. A humid chamber for each sample was prepared by wetting two layers of filter paper placed in a petri dish with distilled water.

#### *2.2.5.1.2 10X PBS Solution*

A total of 80.1 g of NaCl, 26.8 g of $Na_2HPO_4.7H_2O$, 2.7 g of $KH_2PO_4$ and 2 g of KCl were solubilised and make up to 1L with $dH_2O$. This 10X PBS was autoclaved prior to use. 1X PBS (137 mM NaCl, 10 mM $Na_2HPO_4$, 2 mM $KH_2PO_4$, 2.7 mM KCl) solution was prepared fresh with one part of 10X PBS and 9 parts of $dH_2O$ and filter sterilised before use.

*2.2.5.1.3 Cold Methanol*

The volume of 100% methanol calculated based on the usage of the experiment was pre-cooled at -20°C prior to the experiment.

*2.2.5.1.4 Permeating Solution*

This solution was prepared by adding 2% (v/v) of Triton X-100 to 1X PBS.

*2.2.5.1.5 PBS-Mg Solution*

This solution was prepared by adding 5 mM $MgCl_2$ to 1X PBS.

*2.2.5.1.6 Blocking Solution*

This solution was prepared with 1% BSA, 1% cold water fish gelatin, 0.05% Triton X-100 and 0.05% Tween20 in 1X PBS.

*2.2.5.1.7 Antibody solution*

All antibodies used would be diluted using Blocking Solution. Primary antibodies used for indirect immunofluorescence were anti-Rubisco (1:2,000) and anti-SAGA antibodies (1:2,000 dilution) while the secondary antibody used was the Alexa Fluor 488 Goat anti-Rabbit IgG (H+L) antibody (Invitrogen; 1:1,000).

2.2.5.2 Method

About 1 x $10^6$ cells were concentrated to a volume of 500 µl and fixed on the poly-L-lysine-coated slides for 10 min and the excess media was removed from the slides by blot-drying. The slides coated with cells were placed in a coplin jar and about 75 ml of cold methanol was added to the jar. The slides were incubated at -20°C for about 10 min. After then, the dehydrated cells were rehydrated by incubating the slides in 1X PBS for 5 min. This step was repeated twice. After that, Permeating Solution was added to the jar and incubated for 10 min at room temperature followed by two washes in PBS-Mg solution for 10 min each. The slides with treated cells were removed from the coplin jar and air-dried. About 80 µl of Blocking Solution was pipetted onto each cover slip and the slides were inverted onto the cover slips followed by 30 min incubation at room temperature in the humid chambers. After that, about 80 µl of primary antibody was added to the cells as described previously. The slides were then placed in the respective humid chambers and incubated overnight at 4°C.

The next day, the slides were placed into the coplin jar and washed thrice with 1X PBS for 5 min at room temperature. Secondary antibody was added to the slides as described previously and incubated at room temperature for 1 h. After the three washes with 1X

PBS for 5 min each, the slides were then air-dried and mounted. First, about 25 µl of ProLong® Gold Antifade Reagent (Life Technologies) was added onto the slide and a cover slip was then inverted over the droplet. The mounted slides were left to dry overnight in the dark at 4°C.

### 2.2.6 Confocal Microscopy for Image Acquisition

*Chlamydomonas* cells were examined using Leica TCS SP5 confocal microscope (Leica Microsystems) and imaged with the Leica Application Suite Advanced Fluorescence (LAS AF) software (Leica Microsystems). All image analyses were performed using ImageJ2 (Fiji) (Schindelin et al., 2012; Schneider et al., 2012; Rueden et al., 2017).

The excitation wavelength used for indirect immunofluorescence was at 500 nm while for mVenus and YFP was at 515 nm. Chlorophyll autofluorescence excitation wavelength was set at 680 nm.

## 2.3 Proteomics Methods

### 2.3.1 Protein Extraction for Western Blotting (except GFP-Trap samples)

*Chlamydomonas* cells at mid-log phase were harvested by centrifugation (5 min, 4°C, 5000 x g) and resuspended in 1.5 ml of ice-cold extraction buffer (10 mM $MgCl_2$, 50 mM Bicine, 10 mM $NaHCO_3$, 1 mM dithiothreitol treated with cOmplete™ Mini EDTA-free protease inhibitors (Roche, Basel Switzerland). Cells were sonicated for 10 min at 4°C using a Diagenode Bioruptor Standard (Diagenode, Liège, Belgium) on high power in pulsed mode of 45 sec pulses with a 10 sec interval in between pulses. Unbroken cells and cell debris were separated from the supernatant (soluble protein fraction) by centrifugation at 17,000 x g for 15 min at 4°C and stored at -80°C until analysis. Bradford assay was performed by adding 50 µl of total soluble protein to 1.5 ml of Bradford reagent (Sigma-Aldrich; Bradford 1976).

### 2.3.2 Protein Extraction for Proteomics Analysis

The method for protein extraction was according to Section 2.3.1 but the Lysis Buffer contained 10 mM Tris-Cl, pH 7.5, 150 mM NaCl and 0.5 mM EDTA, pH 8.0, treated with 1 mM phosphatase inhibitors, sodium fluoride, 0.3 mM sodium orthovanadate, and cOmplete™ Mini EDTA-free protease inhibitors (Roche, Basel, Switzerland).

### 2.3.3 Protein Acrylamide Gel Electrophoresis (PAGE) and Blotting of Protein Samples

Five microgram of total soluble protein samples (unless stated) were denatured in 2x Laemmli buffer (Bio-Rad Laboratories) with 0.05% β-mercaptoethanol at 95°C for 4 min and electrophoresed at 200 V for 35 min using 1x Running Buffer (2.5 mM Tris, 19.2 mM glycine, 0.01% SDS, pH 8.3) on 15-well 12% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad Laboratories). Gels were run in duplicate in parallel in the same electrophoretic tank.

One gel was stained with GelCode® Blue Stain Reagent (Thermo Scientific, US) while the other was electroblotted to a polyvinylidene fluoride (PVDF) membrane (Amersham™ Hybond® P, Sigma-Aldrich, US) using a Mini Trans-Blot Electrophoretic Transfer Cell system (Bio-Rad Laboratories) with Buffer A (25 mM Tris, pH 8.3, 192 mM glycine, 20% MeOH (v/v), 0.025% SDS (w/v) at 100 V for 1 h or 30 V overnight at 4°C. After two 5-min washes with TBST, the PVDF membrane was then incubated with blocking buffer of 5% (w/v) skimmed milk in 1x TBST buffer (137 mM NaCl, 2.7 mM KCl, 19 mM Tris-base, pH 7.4, Cold Spring Harbor Protocols) for 1 h at room temperature followed by 3 x 5 min washes in TBST. The blot was then incubated for 1 h in primary antibody diluted in TBST containing 2.5% (w/v) skimmed milk at room temperature, washed again 3 x 5 min in TBST. The membrane was next incubated with goat anti-rabbit IgG (H+L) 800 CW (1:5000 dilution, LI-COR) in TBST containing 1% (w/v) skimmed milk for 1 h at room temperature prior to 2 x 5 min washes in TBST and one 5 min wash in TBS.

The antibodies used for Western blotting included anti-Rubisco antibody (1:50,000 dilution) which was a gift from Prof. John Gray (University of Cambridge), anti-LCIB and anti-LCIC antibodies were from Asst. Prof. Takashi Yamano (University of Kyoto; 1:5,000 and 1:10,000 dilutions, respectively), anti-EPYC1 and anti-SAGA antibodies (1:2,000 and 1:3,000 dilutions, respectively) were from Dr. Martin Jonikas (Princeton University), anti-CAH3 antibody (Agrisera, Sweden; 1:2000 dilution), anti-GFP antibody (abcam, Cambridge; 1,5000 dilution), anti-Histone H3 antibody (abcam, Cambridge, 1:10,000 dilution) and monoclonal anti-FLAG antibody (Sigma-Aldrich, Germany; 1:2,000 dilution). All antibodies were polyclonal antibodies produced in rabbit except anti-FLAG antibody which was produced in mouse.

Visualisation and quantification was carried out with the LI-COR Odyssey® scanner and software (LI-COR Biosciences). Quantification of fluorescence signal was performed on single channel with Image Studio Lite (LI-COR Biosciences).

### 2.3.4 Co-immunoprecipitation using GFP-Trap_A (Chromotek, Germany)

About 1 mg of total soluble protein for each strain (wild type, *saga*, SVA7 and SVB2) acclimated to low-$CO_2$ for 24 h was extracted using method described in 2.3.2. Unless stated, all steps were performed at 4°C and all solutions were treated with 1 mM phosphatase inhibitors, sodium fluoride, 0.3 mM sodium orthovanadate, and cOmplete™ Mini EDTA-free protease inhibitors (Roche, Basel, Switzerland).

First, the GFP-Trap_A beads were vortexed and 25 µl of bead slurry was transferred into 500 µl of ice-cold Wash Buffer (10 mM Tris-Cl, pH 7.5, 150 mM NaCl, 0.5 mM EDTA, pH 8.0, 0.01% NP-40). The mix was centrifuged at 2,500 x g for 2 min and supernatant was discarded. This bead-washing step was repeated twice. About 1-2 mg of total soluble protein extracted using method 2.3.2 was diluted to 2.5 ml with Dilution Buffer (10 mM Tris-Cl, pH 7.5, 150 mM NaCl, 0.5 mM EDTA, pH 8.0) and then added to the GFP-Trap_A beads. The mixture was tumbled end-over-end for 1 h. After that, the tubes were centrifuged and the supernatant was transferred to another tube and labelled "Supernatant". The beads were then washed with 500 µl of ice-cold Wash Buffer and the supernatant was transferred to a new tube labelled "Wash 1". The washing step was repeated using 500 µl of Dilution Buffer and the supernatant was labelled "Wash 2". Next, the GFP-Trap_A beads were resuspended in 30 µl of 2X Laemmli buffer (Bio-Rad Laboratories) with 0.05% β-mercaptoethanol and boiled for 10 min at 95°C to dissociate the immunocomplexes from the beads.

One microliter of the final elute was used to run SDS-PAGE and blotted with anti-GFP and anti-SAGA antibodies to confirm that SAGA-Venus-3xFLAG protein was pulled down in the complemented lines only before submitting the samples for LC-MS/MS.

### 2.3.5 LC-MS/MS Analysis

The GFP-Trap_A final elute for each sample (WT, *saga*, SVA7 and SVB2) was run for 1 cm into a 12% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad Laboratories) and then stained with GelCode® Blue Stain Reagent (Thermo Scientific, US). The samples were processed by the Cambridge Centre for Proteomics for peptide identification.

The gel-based samples were processed using trypsin digestion and the peptides were separated by high-resolution nanoscale chromatography and the high mass accuracy MS/MS spectra were acquired automatically. The peptide data were searched against UniProt and JGI Phytozome *Chlamydomonas reinhardtii* v5.5 databases using the Mascot search algorithm. The search results were converted to Scaffold format which could be opened using Scaffold 4 (Proteome Software, Inc., Oregon, USA) to visualise dataset. Support for data interpretation was provided by Dr. Marco Chiapello, the Research Data Technician of CCP.

### 2.3.6 Native PAGE and Phosphoprotein Staining

#### 2.3.6.1 Native PAGE and Western Blotting

Total soluble protein for each strain (wild type, *saga*, SVA7 and SVB2) grown under high $CO_2$ or acclimated to low-$CO_2$ for 24 h was extracted using method described in 2.3.2.

Samples containing 10 µg of total soluble protein were added with equal volume of 2X Sample Buffer (62.5 mM Tris-HCl, pH 6.8, 25% glycerol, 1% Bromophenol Blue) and electrophoresed at 150 V for 50 min using 1x Running Buffer (25 mM Tris, 192 mM glycine) on 10-well 4-20% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad Laboratories). Gels were run in duplicate in parallel in the same electrophoretic tank.

One gel was stained with GelCode® Blue Stain Reagent (Thermo Scientific, US) while the other gel was electroblotted as described in method 2.3.3 using anti-GFP, anti-Rubisco, anti-EPYC and anti-SAGA antibodies.

#### 2.3.6.2 Phosphoprotein Staining

Upon running SDS-PAGE on the total soluble protein, phosphoprotein staining was performed using a modified version of the Pro-Q® Diamond stain protocol (Invitrogen) to suit the container size. First, the gel was immersed in about 25 ml of fix solution (50% methanol, 10% acetic acid) and incubated at room temperature with gentle agitation for at least 30 min. This step was repeated once to ensure all SDS was washed out of the gel. After that, the gel was incubated thrice, each time in about 50 ml of ultrapure water with gentle agitation for 10 min to remove all the methanol and acetic acid from the gel. Next, the gel was incubated in about 30 ml of Pro-Q® Diamond phosphoprotein gel stain with gentle agitation in the dark for 60 min. The gel was then destained with 30 ml of Pro-Q® Diamond phosphoprotein gel destain solution with

gentle agitation for 30 min at room temperature, protected from light. This procedure was repeated twice, allowing optimal destaining time of about 1.5 h. The gel was washed twice with ultrapure water at room temperature for 5 min per wash. For visualisation of phosphoproteins, the gel was scanned using Typhoon 9210 (Amersham Biosciences) with excitation source of 532 nm laser and emission filter of 560 nm long-pass.

After staining with Pro-Q® Diamond stain and imaged with Typhoon 9210, the gel was stained with SYPRO® Ruby protein gel in order to help with determining the relative phosphorylation state of a given protein. First, the gel was incubated directly in about 30 ml of SYPRO® Ruby gel stain solution overnight on an orbital shaker. The next day, the gel was transferred to a clean container and washed with 30 ml of wash solution (10% methanol, 7% acetic acid) for 30 min. Before imaging, the gel was rinsed twice in ultrapure water for 5 min each time to prevent possible corrosive damage to the imager. Gel visualisation was done on Typhoon 9210 (Amersham Biosciences) with excitation source of 450 nm and emission filter of 610 nm band-pass.

## 2.4 Molecular Methods

### 2.4.1 Genomic DNA Extraction

The genomic DNA extraction method was modified from Lim (2012). A loopful of *C. reinhardtii* cells were resuspended in 500 µl of 1X TEN buffer (150 mM NaCl, 10 mM EDTA, 10 mM Tris-HCl, pH 8.0) and then centrifuged at maximum speed for 10 min. The supernatant was removed and the pellet was resuspended in 750 µl of CTAB buffer (2% (w/v) CTAB, 100 mM Tris-HCl, pH 8.0, 1.4 M NaCl, 20 mM EDTA, 2% β-mercaptoethanol). The cell suspension in CTAB was then incubated at 65°C for 1h and centrifuged at maximum speed for 5 min. The cell lysate was added to an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma-Aldrich) and mixed well before centrifuging at maximum speed for 5 min. The aqueous layer was extracted again with an equal volume of chloroform-isoamyl alcohol (24:1). The DNA in the clear aqueous phase was precipitated with an equal volume of isopropanol at -20°C overnight, washed twice with 70% (v/v) ethanol by centrifuging at maximum speed each time. The DNA pellet was dissolved with 40 µl of ultrapure water and the concentration was measured using Nano Drop™ 2000 Spectrophotometer (Thermo Scientific) according to manufacturer's instruction.

### 2.4.2 Genomic DNA PCR

Routine PCR to amplify the coding region of Cre02.g143387 was performed using Phusion Hot Start II High-Fidelity DNA polymerase (ThermoScientific) using the following reaction mix: 1X Phusion GC Reaction Buffer (ThermoScientific), 2 mM $MgCl_2$, 3% DMSO, 0.5 µM primer CC021, 0.5 µM primer CC022, 0.2 mM dNTP mix, 0.02 U/µl of Phusion HSII HF DNA polymerase and 50 ng of genomic DNA isolated from wild type *C. reinhardtii*. Total volume used was 50 µl. The PCR condition used consisted of the following: 1) Initial denaturation at 98°C for 30 s, 2) 5 cycles of denaturation at 98°C for 10 s, annealing at 66.4°C for 30 s, extension at 72°C for 1 min, 3) 25 cycles of denaturation at 98°C for 10 s and extension at 72°C for 1 min, 4) final extension at 72°C for 10 min. The optimal annealing temperature was achieved after optimisation using gradient PCR with annealing temperatures ranging from 60 to 67°C.

### 2.4.3 PCR Fragment Purification

The PCR product from method 2.4.2 was electrophoresed on a 1% (w/v) agarose gel in 1X TAE buffer. The DNA fragment was excised from the agarose gel with a clean, sharp scalpel and purified using QIAquick Gel Extraction Kit (QIAGEN). The excised gel was weighed in a colourless tube and 3 volumes of Buffer QG were added to 1 volume of gel. The mix was incubated at 50°C for 10 min (or until the gel slice had completely dissolved). The tube was vortexed every 2 min during incubation to help dissolve gel. After that, the mixture was checked to ensure the colour was yellow. A QIAquick spin column was placed in a provided 2 ml collection tube. The sample was applied to the QIAquick column and centrifuged at 17,000 x g for 1 min to bind the DNA to the column. The flow-through was discarded and the QIAquick column was placed back in the same collection tube. 500 µl of Buffer QG was added to the column and centrifuged for 1 min to remove all traces of agarose. Next, 0.75 ml of Buffer PE was added to the QIAquick column and centrifuged for 1 min. The flow-through was discarded and the column was centrifuged for an additional 1 min to remove residual ethanol from the buffer. The QIAquick column was placed into a clean 1.5 ml tube. To elute DNA, 30 µl of water was added to the center of the QIAquick membrane and the column was left standing for 1 min before centrifuging for 1 min. The DNA concentration was measured using Nano Drop™ 2000 Spectrophotometer (Thermo Scientific) according to manufacturer's instruction.

### 2.4.4 Plasmid Miniprep

Plasmid isolation after each of *E. coli* transformation step was performed according to the instruction manual of QIAprep® Miniprep Kit (QIAGEN). Single colonies were inoculated in 5 ml of LB medium containing an appropriate selective antibiotic and incubated for 12-16 h at 37°C with vigorous shaking. The bacterial cells were harvested by centrifugation at 17,000 x g for 3 min at 20°C and the supernatant was decanted. The cell pellet was resuspended in 250 µl of Buffer P1 (added with RNase A and LyseBlue reagent) and transferred to a 2-ml microfuge tube. 250 µl of Buffer P2 was added to the cells and the tube was inverted gently for 4 to 6 times to mix thoroughly. The cell suspension would turn blue after the addition of Buffer P2. After that, 350 µl of Buffer N3 was added and mixed by inverting the tube immediately and thoroughly, where the solution would turn cloudy and colourless. The mix was centrifuged at 17,000 x g for 10 min. The QIAprep 2.0 spin column was placed on the collector tube, and 800 µl of supernatant was applied. The column was centrifuged for 1 min and the flow-through was discarded. The spin column was then washed by adding 0.5 ml of Buffer PB and centrifuged for 1 min with the flow-through discarded. Next, the QIAprep 2.0 spin column was added with 0.75 ml of Buffer PE and centrifuged for 1 min. After discarding the flow-through, the column was centrifuged for addition 1 min to remove residual wash buffer. Finally, the plasmid DNA was eluted by adding 50 µl of water to the QIAprep 2.0 column and left standing for 1 min before final centrifugation for 1 min. The plasmid DNA concentration was measured using Nano Drop™ 2000 Spectrophotometer (Thermo Scientific) according to manufacturer's instruction.

### 2.4.5 Golden Gate Cloning

#### 2.4.5.1 List of Vector backbones and domesticated modules

The vector backbones of the domesticated modules were part of the Golden Gate Modular Cloning (MoClo) Plant Parts Kit (Engler et al., 2014). The domesticated modules kindly provided by Dr. Francisco Navarro (University of Cambridge) would be part of the Chlamy MoClo Kit (publication in preparation) that had been codon-optimised for *Chlamydomonas reinhardtii* (Table 2.1 and Table 2.2).

**Table 2.1 Vector backbones of Golden Gate MoClo Plant Tool Kit (Engler et al., 2014) and a domesticated Level 2 vector backbone (Dr. Francisco Navarro, University of Cambridge) used in this study.**

| Label | Vector Name | Resistance | Description | RE |
|---|---|---|---|---|
| D1 | pAGM1276 | Spectinomycin | Level zero acceptor for NT1 modules | *Bpi*I |
| E2 | pAGM1251 | Spectinomycin | Level zero acceptor for Pro+5U(f) modules | *Bpi*I |
| F2 | pAGM1287 | Spectinomycin | Level zero acceptor for CDS1 no stop modules | *Bpi*I |
| H2 | pICH41308 | Spectinomycin | Level zero acceptor for CDS1 modules | *Bpi*I |
| B4 | pICH47811 | Carbenicillin | Level 1 acceptor. Position 2. Reverse orientation | *Bsa*I |
| C5 | pICH41744 | Spectinomycin | End-link 2 for assembling 2 level one part into a level 2 acceptor | *Bpi*I |
| pFPC2 | pFPC2 | Kanamycin | Level 2 acceptor | *Bpi*I |

**Table 2.2 Domesticated modules of Golden Gate cloning in *Chlamydomonas reinhardtii*. TU_4, pFJN43, pFJN45 and pFJN70 were kindly provided by Dr. Francisco Navarro, while pCCC3 and pCCN3 were constructed specifically for this study.**

| Vector Name | Resistance | Module name + 4bp overhangs | Module description | RE |
|---|---|---|---|---|
| TU_4 | Carbenicillin | GGAG_AphVIII(hygromycin)_CGCT | Level 1 plasmid. Position 1. Reverse orientation | *Bpi*I |
| TU_1C4 | Carbenicillin | GGAG_Cre02.g143387_CGCT | Level 1 plasmid. Position 2. Reverse orientation. | *Bpi*I |
| pFJN43 | Spectinomycin | GGAG_Pro-PSAD_5U-PSAD_AATG | Promoter + 5'UTR (Chlamy PSAD) | *Bsa*I |
| pFJN45 | Spectinomycin | GCTT_3U-PSAD_Ter-PSAD_CGCT | 3'UTR + Terminator (Chlamy PSAD) | *Bsa*I |
| pFJN70 | Spectinomycin | TTCG_CT-YFP_GCTT | C-terminus tag (YFP) | *Bsa*I |
| pCCC3 | Spectinomycin | AATG_CDS1(C3)_GCTT | CDS1_1C4 | *Bsa*I |
| pCCN3 | Spectinomycin | AATG_CDS1ns(N3)_TTCG | CDS1ns_1C4 | *Bsa*I |

2.4.5.2 Primer Design

All PCR primers (Table 2.3) used in this project were designed using Primer3web (version 4.1.0) (Koressaar and Remm, 2007; Untergasser et al., 2012).

**Table 2.3 Primers used in assembly of the expression vector, pFPC2-1C4-HygR, including validation of levels 0, 1 and 2 modules.**

| Primer ID | Sample | Location | Primer Direction | Primer Sequence |
|---|---|---|---|---|
| CC021 | gDNA | CDSns | Forward | TT-GAAGAC-AT-AATG-CGCCCGAGGCTGACAGGCG |
| CC022 | gDNA | CDS1ns | Reverse | TT-GAAGAC-AT-CGAA-TTAGGTGCCTTGGGTGTGC |
| CC025 | Plasmid | 5U-CDS | Forward | TGAAGGCAGGCAAATGCTCATG |
| CC026 | Plasmid | | Reverse | GCGAGCTAGCCAAATGTGTGTT |
| CC029 | Plasmid | CDS-YFP-3U | Forward | TCTACACTTAAGCCCCCAAGCC |
| CC030 | Plasmid | | Reverse | CCACCATGTTGCTGCTGTGATT |
| C4-F | Plasmid | CDS | Forward | GCATTCGCGCCTCCCTTCC |
| C4-R1 | Plasmid | CDS | Reverse | GCACTCGTGAAGCGCACAC |
| C4-R2 | Plasmid | CDS | Reverse | GCATCTACGCAGGTGTAGACGT |

Primer pair CC021 and CC022 used to amplify genomic DNA region of Cre02.g143387 was designed by incorporating 1) extra bases for enzyme binding (TT), 2) *Bpi*I recognition site (GAAGAC), 3) spacer (AT) as *Bpi*I cuts two nucleotides at 3', 4) specific cutting site for the module, in this case, the CDSns module, and 5) the primer specific for the region (Cre02.g143387) in preparation for level 0 module assembly. In the forward primer, the anneal sequence started after the Met codon of the gene as Met was incorporated in the overhang.

Primer pairs CC025+CC026 and CC029+CC030 were designed for validation of plasmids and inserts. Primers C4-F, C4-R1 and C4-R2 were designed to sequence the full length of Cre02.g143387 in the plasmid.

#### 2.4.5.3 Molecular biology reagents

The restriction enzymes, *Bsa*I and *Bpi*I, were purchased from Thermo Scientific and T4 DNA ligase was purchased from New England Biolabs (NEB).

#### 2.4.5.4 MoClo cloning protocol

Level 0 CDSns module, pCCN3, was assembled by the restriction-ligation reaction in one tube which contained 75 ng of acceptor plasmid (pAGM1827), 150 ng of insert DNA, 5U of *Bpi*I (Thermo Scientific), 200U of T4 DNA ligase (NEB), 1X T4 Ligation Buffer (NEB) and 1X BSA in a final reaction volume of 20 µl. The reaction was incubated in a thermocycler using the assembly protocol of a cycle of 37°C for 20 s, followed by 26 cycles of 37°C for 3 min, 16°C for 4 min, a cycle of 50°C for 5 min and finally a cycle of 80°C for 5 min. Five microliter of ligated plasmid or 75 ng of backbone vector was added to 50 µl of competent DH5α *E. coli* cells that were thawed on ice. The reaction mix was incubated on ice for 20 min and heat-shocked at 42°C for 45 s and immediately incubated on ice for 2 min. After that, 450 µl of SOC medium was added to the cells and incubated at 37°C for 1 h. During the incubation period, the selective LB plates containing 50 µg/ml of spectinomycin were prepared by spreading 40 µl of 100 mM IPTG (final concentration of 0.5 mM) and 120 µl of 20 mg/µl X-gal (final concentration of 40 µg/ml) onto each of the selective LB agar plate. After the 1-h incubation, 50 µl of transformed cells were spread on the LB plates and incubated at 37°C overnight. The next day, about 4 white colonies were selected and grown in 5 ml of LB broth plus spectinomycin overnight for plasmid miniprep (method 2.4.4) and other downstream experiments. Restriction enzyme digestions using *Bsa*I, *Blp*I and *Dpn*I (2.4.5.5) determined that the gene of interest was inserted properly into the acceptor plasmid while Sanger sequencing (2.4.5.6) was used to validate the gene sequence.

For assembly of level 1 plasmid, the reaction mix contained 75 ng of acceptor plasmid (pICH47811), 150 ng each of pFJN43, pFJN45, pFJN70 and pCCN3, 5U of *Bsa*I (ThermoFisher), 200U of T4 DNA ligase (NEB), 1X T4 Ligation Buffer (NEB), 1X BSA in a 20-µl final volume. The assembly protocol and transformation method were as described for construction of level 0 module. The transformed DH5α cells were spread on selective LB plates containing 10 µg/ml carbenicillin, 0.5 mM IPTG and 40 µg/ml X-gal. Miniprep (2.4.4) was performed on 3 colonies and colony PCR (2.4.5.8) was performed for validation of the assembled transcriptional unit, TU_1C4.

Level 2 multigene construct consisted of a transcription unit carrying the gene of interest (TU_1C4) and a transcription unit with hygromycin resistance gene (TU_4). The reaction mix contained 75 ng of acceptor plasmid (pFPC2), 75 ng of end-linker 2 (C5 of the Plant MoClo Kit), 75 ng of TU_4, 150 ng of TU_1C4, 5U of *Bpi*I (ThermoFisher), 200 U of T4 DNA ligase (NEB), 1X T4 ligation buffer (NEB) and 1X BSA in a final reaction volume of 20 µl. The reaction was incubated for 45 cycles of 37°C for 2 min, 16°C for 5 min, followed by a cycle of 50°C for 5 min and finally at 80°C for 10 min. The transformation was as described above and the transformed cells were spread on selective LB plates containing 50 µg/ml of kanamycin. All three white colonies grown after overnight incubation at 37°C were selected and the sequence validated by colony PCR (2.4.5.8).

#### 2.4.5.5 Restriction Enzyme Digestion

In order to confirm that the gene of interest was inserted correctly into the level 0 acceptor plasmid, restriction enzyme digestions were performed using *Bsa*I (ThermoFisher), *Blp*I (NEB) and *Dpn*I (NEB).

The 20-µl reaction mix for *Bsa*I digestion consisted of 1U/µl of enzyme *Bsa*I, 1X Buffer G (ThermoFisher), and 1 µl of pCCN3 which was incubated at 37°C for 1h followed by enzyme deactivation at 65°C for 20 min. The product was electrophoresed in a 1% agarose gel in 1X TAE buffer.

As for *Blp*I reaction, 1 U/µl of *Blp*I, 1X Buffer 2 (NEB) and 1 µl of pCCN3 were combined with DNase- and RNase-free water to make up the final volume of 20 µl. The incubation time was 1 h and the digestion product was run directly on a 1% agarose gel in 1X TAE as this enzyme could not be deactivated by heat.

The reaction mix for *Dpn*I was similar to that of *Blp*I except the enzyme was deactivated by increasing the temperature to 80°C for 20 min after the hour-long incubation at 37°C. Agarose electrophoresis condition was as described above.

#### 2.4.5.6 Sanger Sequencing and Multiple Sequence Alignment

The level zero CDSns module, pCCN3, assembled was sent for Sanger sequencing by the Source Bioscience Sequencing Team (www.sourcebioscience.com) using forward primer, AG122 (5'-AATACGCAAACCGCCTCTC-3') and reverse primer, AG123 (AAATAGGCGTATCACGAGG) provided by Aleix Gorchs-Rovira (University of Cambridge).

The sequence data was then aligned to the genomic sequence of Cre02.g143387 downloaded from the *Chlamydomonas reinhardtii* v5.5 genome deposited in JGI Phytozome portal (https://phytozome.jgi.doe.gov/pz/portal.html) using the Clustal Omega Multiple Sequence Alignment web tool from the EMBL-EBI website (https://www.ebi.ac.uk/Tools/msa/clustalo) (Sievers et al., 2011; McWilliam et al., 2013; Li et al., 2015).

#### 2.4.5.7 Transformation

All cells used for transformation were inoculated in Tris acetate phosphate (TAP) medium and supplemented with Hutner's trace elements (Hutner et al., 1950). The growth conditions were constant illumination at 50 µmol photons $m^{-2}$ $s^{-1}$ at 25°C without aeration. When the cells reached the density of about 1 to 2 million cells $ml^{-1}$ at mid-log phase, the cells were harvested at 3,000 x g for 3 min at 20°C and washed once with fresh TAP medium. After that, the cell pellet was resuspended gently in 10 ml GeneArt Max Efficiency Transformation Reagent (Invitrogen) and centrifuged at 1,000 x g for 10 min. This step was repeated and the pellet was resuspended in 250 µl TAP supplemented with 40 mM sucrose. About 1.5 µg of linearised DNA was added to the cell suspension and incubated at 4°C for 5 min. The cells were transferred to an ice-cold GenePulser Electroporation cuvette (Bio-Rad Laboratories) and then electroporated using Gene Pulser® II Electroporation System (Bio-Rad Laboratories) at a capacitance of 25 µF, a voltage of 800 V, with no shunt resistance. The elapsed time was in the range between 9.5 to 15 ms. Following electroporation, the cells were transferred into a 15 ml tube containing 10 ml of TAP medium plus 40 mM sucrose and placed horizontally in the Innova 42 incubator (New Brunswick Scientific, Enfield, CT, USA) (120 rpm) under dim light of about 2 to 3 µmol photons $m^{-2}$ $s^{-1}$ at 25°C for 14 to 16 h. After incubation, the cells were harvested by centrifugation at 1,000 x g for 15 min and resuspended in 300 µl of TAP medium. For selection of transformed colonies, 100 µl of cells were spread on TAP agar plates supplemented with 10 µg/ml of hygromycin and incubated at 25°C under constant light of 50 µmol photons $m^{-2}$ $s^{-1}$ for 5 to 7 days. Colonies which grow on the selective TAP agar were selected for further analysis.

#### 2.4.5.8 Colony PCR

Validation of level 1 and level 2 modules were performed using the Phire Plant Direct PCR Kit (Thermo Scientific) with primer pairs CC025+CC026 and CC029+CC030, respectively.

DNA extraction was performed by first adding 2 µl of cells grown in TAP medium containing 10 µg/ml hygromycin to 18 µl of Phire dilution buffer and then incubated at -20°C for 20 min. After that, the cell suspension was centrifuged at 3,500 x g for 10 min. For PCR, 3 µl of the supernatant was used.

The PCR reaction contained 0.2 µl of Phire Hot Start II DNA polymerase, 1X Phire Plant PCR buffer which included dNTPs and $MgCl_2$, 0.5 µM of forward and reverse primers, 3 µl of algal DNA sample to make up the final volume of 20 µl. The amplification protocol started with initial denaturation at 98°C for 3 min followed by 30 cycles of denaturation at 98°C for 15 s, annealing at 65°C for 15 s and extension at 72°C for 1 min. The final extension was set at 72°C for 7 min. The PCR products were then electrophoresed on 1% agarose gel in 1X TAE and stained with ethidium bromide for visualisation under UV light.

## 2.5 Bioinformatics Analyses

Software, web-based programs and databases used in this project for protein identification, sequence alignments, and Golden Gate cloning are listed below:

### 2.5.1 Databases

Phytozome v12 (https://phytozome.jgi.doe.gov/pz/portal.html)

NCBI (https://www.ncbi.nlm.nih.gov/)

UniProt (http://www.uniprot.org/)

### 2.5.2 Protein Prediction Programs

PHYRE2 Protein Fold Recognition Server (Kelley et al., 2015) (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index)

PredAlgo (Tardif et al., 2012) (https://giavap-genomes.ibpc.fr/cgi-bin/predalgodb.perl?page=main)

PCOILS (Gruber et al., 2006; Alva et al., 2016) (https://toolkit.tuebingen.mpg.de/#/tools/pcoils)

TargetP (Emanuelsson et al., 2000) (http://www.cbs.dtu.dk/services/TargetP/)

ChloroP (Emanuelsson et al., 1999) (http://www.cbs.dtu.dk/services/ChloroP/)

DeepLoc (Almagro Armenteros et al., 2017) (http://www.cbs.dtu.dk/services/DeepLoc/)

Web BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi)

### 2.5.3 Peptide Visualisation Software

Scaffold 4 (http://www.proteomesoftware.com/products/scaffold/)

### 2.5.4 Multiple Sequence Alignment Web Tool

Clustal Omega (Sievers et al., 2011) (https://www.ebi.ac.uk/Tools/msa/clustalo)

### 2.5.5 Golden Gate Cloning

Primer3 (Koressaar and Remm, 2007; Untergasser et al., 2012) (http://primer3.ut.ee/)

The Benchling Platform (https://benchling.com/)

Serial Cloner 2.6 (http://serialbasics.free.fr/Serial_Cloner.html)

ApE Plasmid Editor (http://biologylabs.utah.edu/jorgensen/wayned/ape/)

SnapGene® Viewer (http://www.snapgene.com/products/snapgene_viewer/)

### 2.5.6 Graphic Tools

ImageJ2 (FIJI) (Schindelin et al., 2012; Schneider et al., 2012; Rueden et al., 2017)

RStudio (https://www.rstudio.com/)

Inkscape (https://inkscape.org/en/)

# 3 PYRENOID FORMATION IS LINKED TO MULTIPLE GENES

## 3.1 Introduction

### 3.1.1 Engineering Photosynthesis to Address Global Food Demand

In order to address the issue of global food demand by 2050, researchers are looking for methods to improve the existing crops by engineering photosynthesis and yield potential (Long et al., 2015). This is due to the stagnation of the Green Revolution (Kendall and Pimentel, 1994, 2000; Long et al., 2015), for example, the average yields per hectare of rice in China, India and Indonesia during 1970-1980 increased by the average of 36%, but the yield increase had dwindled to about 7% between 2000 and 2010 (Long, 2014; Long et al., 2015). Therefore, it is crucial to identify ways to engineer photosynthesis and yield potential in order to meet the desired 70% increase in food demand by the next 30 years (Long, 2014).

Researchers have been attempting to introduce potentially simpler CCM systems from cyanobacteria (Price et al., 2013), algae (Raven et al., 2012) and some terrestrial plants such as hornworts (Smith and Griffiths, 1996), into higher plants (Jungnick et al., 2014; Peterhänsel et al., 2013; Price et al., 2013). Atkinson et al. (2016) demonstrated

transient expression of a number of the key components of CCM of *Chlamydomonas reinhardtii* in *Arabidopsis*, and transgenic plants grew as well as wild type. The potential for improved growth when crop plant is engineered with component(s) of simple CCM system was recently demonstrated by Hay et al. (2017) where a cyanobacterial gene, *ictB* was engineered into soybean that yielded transgenic plants with improved photosynthesis, biomass production and seed yield. Thus identifying additional components comprising the algal CCM remains a key goal, currently.

### 3.1.2 Insertional Mutagenesis

The *Chlamydomonas* mutants used in the current study were generated by Dr. Leif Pallesen and the Jonikas Lab using an insertional mutagenesis high throughput screen (Zhang et al., 2014) (Figure 3.1A). The transformation cassette was prepared by digestion of plasmid pMJ016c, purified and used for electroporation transformation (Shimogawara et al., 1998). The wild type, cMJ030, used in this study was a progeny from $4A^{-}/D66^{+}$ crosses, with mating type *minus* provided by Dr. Martin Jonikas, Princeton University (Zhang et al., 2014). The wild type cMJ030 cell has minus mating type which displays these properties: both photoautotrophic and heterotrophic growth, greening in the dark, normal screening ability, high transformation efficiency, efficient recovery from cryogenic storage in liquid nitrogen and *cw15*-like cell wall morphology (Zhang et al., 2014). The paromomycin-resistant transformants were selected and transferred for screening of CCM defects based on differential sensitivity to elevated/ambient $CO_2$. Mutants that grew at elevated $CO_2$ level but were unable to grow under ambient $CO_2$ conditions were classified as CCM defective (Figure 3.1B).

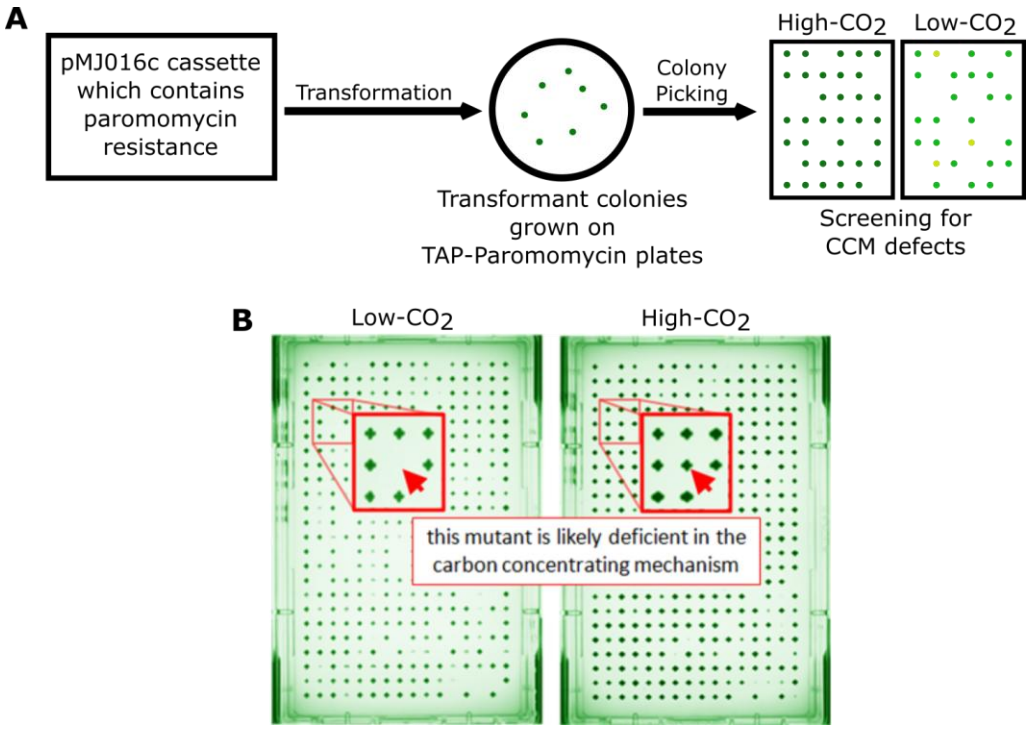



**Figure 3.1 A schematic diagram of CCM mutant selection. (A) CCM defective mutant selection. Paromomycin cassettes were introduced into cMJ030 cells by electroporation. Colonies were picked individually and arrayed at 384 colonies per plate on fresh TAP plates containing 20 µg/ml paromomycin and cultured under high $CO_2$ or low $CO_2$ conditions. (B) Mutants with CCM deficiencies were selected based on the inability of the mutants to grow under low $CO_2$ conditions but were rescued under high $CO_2$ conditions.**

### 3.1.3 Screening for Pyrenoid Defects

*Chlamydomonas reinhardtii* cells are able to activate a biophysical carbon concentrating mechanism (CCM) to accumulate inorganic carbon in the pyrenoid where the $CO_2$-fixing Rubisco is aggregated that will increase photosynthetic rate under low $CO_2$ conditions (Badger et al., 1980; Kuchitsu et al., 1988a; Engel et al., 2015; Mackinder, 2017). The pyrenoid of a wild type cell is a proteinaceous microcompartment in the chloroplast (Kuchitsu et al., 1988a; Morita et al., 1997; Borkhsenious et al., 1998) known to respond to the changes in $CO_2$ concentrations (Tsuzuki et al., 1986). The three dimensional reconstruction of *C. reinhardtii* cell components by Engel et al. (2015) also suggested that Rubisco complexes are clustered in the pyrenoid. This was also demonstrated previously using immunogold labelling of anti-Rubisco antibody by Lacoste-Royal & Gibbs (1987), Süss et al. (1995), Morita et al. (1997), Borkhsenious et al. (1998) and He et al. (2003), where 90% of Rubisco aggregated into the pyrenoid

under low $CO_2$ conditions. Although the pyrenoid is constantly present under various conditions (Borkhsenious et al., 1998), the size of the pyrenoid increases when CCM is induced (Mackinder et al., 2016).

Based on this knowledge, the CCM defective mutants from the previous screen by the Jonikas Lab were subjected to two independent screens for pyrenoid defects using immunofluorescence with an anti-Rubisco antibody (Caspari, 2015) and differential interference contrast (DIC) microscopy (Dr. Moritz Meyer, Princeton University) (Figure 3.2) based on the following characteristics: 1) multiple pyrenoids, 2) reduced pyrenoid size and 3) mislocalised pyrenoid. A total of twelve mutants with pyrenoid defects were selected from these screens for further characterisation (Table 3.1), of which ten were used for spot test and oxygen evolution, but nine were selected for subsequent microscopy analyses. The other three mutants were characterised by summer students and Dr. Moritz Meyer.

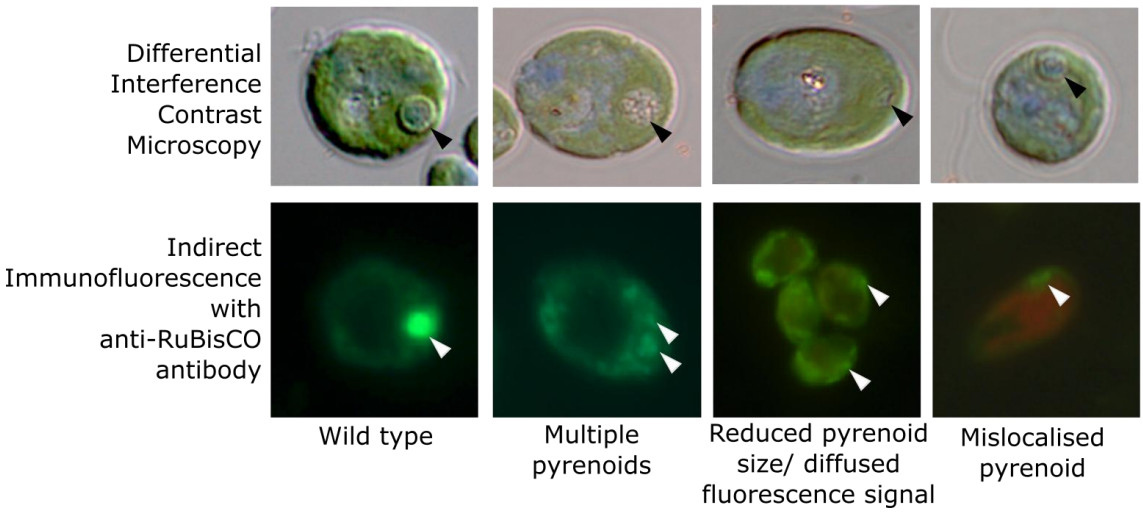



**Figure 3.2 Independent screens on the CCM defective mutant library using differential interference contrast microscopy (top panel, courtesy of Dr. Moritz Meyer) and indirect immunofluorescence with anti-Rubisco antibody (bottom panel, courtesy of Dr. Oliver Caspari) revealed mutants with abnormal pyrenoid phenotypes: Multiple pyrenoids, reduced pyrenoid size and mislocalised pyrenoids. Arrows indicate the position(s) of pyrenoid in the cells.**

**Table 3.1 List of the twelve mutants with pyrenoidal defects selected from the secondary CCM screens. In this study, ten mutants were initially used for spot tests and oxygen evolution but subsequently only nine mutants were subjected to microscopy analysis. All mutants were sent for whole genome sequencing but only the results of the initial 10 mutants used for spot tests were included in this chapter. *Work completed by other analysts. **Data from WGS are not included in this thesis. MM, Dr. Moritz Meyer; CC, Cindy Chan (author); CW, Charlie Whittaker.**

| Mutant Name | Spot Tests & Oxygen Evolution | Transmission Electron Microscopy | Whole Genome Sequencing | Analysts |
|---|---|---|---|---|
| A4 | No | No | Yes** | MM |
| A5 | Yes | Yes | Yes | CC |
| A7 | Yes | Yes | Yes | CC |
| A8 | Yes | Yes | Yes | CC |
| A11 | Yes | Yes | Yes | CC |
| C4 | Yes | Yes | Yes | CC |
| C8 | Yes | Yes | Yes | CC |
| D4 | Yes | Yes | Yes | CC |
| F8 | Yes | Yes | Yes | CC |
| G6 | Yes | Yes* | Yes | MM/CC |
| H9 | Yes* | Yes* | Yes** | MM/CW |
| H10 | Yes | Yes | Yes | CC |

The main focus of this chapter is the morphological and physiological characterisation of the mutants with pyrenoid defects based on these hypotheses surrounding the lack of a fully functional CCM under low $CO_2$ conditions:

1. The Rubisco trafficking mechanism of the mutants is affected since almost all Rubisco is trafficked into the pyrenoid when CCM is induced (Borkhsenious et al., 1998).
2. The size of the pyrenoid in the mutants is smaller than that of the wild type due to the lack of all components required to form a fully-functional pyrenoid as demonstrated in *epyc1* mutant (Mackinder et al., 2016).
3. The CCM impairment of the mutants is caused by the lack of the pyrenoidal tubules that are suggested to play the roles of delivering $CO_2$, and exchange of metabolites for carbon fixation in the pyrenoid (Engel et al., 2015).
4. The defects shown in the pyrenoid of the mutants are caused by knockout of known CCM genes controlling the formation of the pyrenoid (Fukuzawa et al., 2001; Ohnishi et al., 2010; Sinetova et al., 2012; Yamano et al., 2014).

## 3.2 Results

### 3.2.1 Growth of CCM defective mutants

Spot tests were used to confirm the compromised growth phenotype of the putative CCM defective mutants (Figure 3.3). The controls included in this experiment include the corresponding wild type, cMJ030 (Zhang et al., 2014) as positive control. The negative controls were a pyrenoid-less *RbcS1* hybrid mutant, *SSSO* (Genkov et al., 2010), and an essential pyrenoid-localised linker protein mutant, *epyc1* (Mackinder et al., 2016). The *epyc1* mutant contains the same insertional mutagenesis cassette as the mutants in this study.

A control was set up under heterotrophic conditions with cells grown on TAP (tris acetate phosphate) agar in the dark using acetate as a carbon source, to test whether the strains used in this experiment possessed growth defects unrelated to $CO_2$ requirement (Figure 3.3A, D & G). With the exception of A5 (Figure 3.3A), which showed a yellowing phenotype, all strains grew normally. Additionally, the mutants could grow on TAP agar with paramomycin (data not shown), as confirmation that the cells contained at least a copy of the paromomycin resistance gene included in the original insertional mutagenesis cassette.

Cells were also grown on TP (tris phosphate) agar at 5% (v/v) $CO_2$ (high $CO_2$) under light intensity of 50 µmol photons $m^{-2}s^{-1}$ to assess light-adapted growth phenotypes. The CCM would be repressed under this condition, thus the growth of the mutants were indistinguishable from the corresponding wild type (Figure 3.3B, E & H).

In order to assess the presence of a functional CCM, cells were grown on TP agar at 0.04% (v/v) $CO_2$ (low $CO_2$) under similar light intensity. The corresponding wild type showed normal growth, indicating the presence of a functional CCM while both negative controls, *SSSO* (Figure 3.3C & F) and *epyc1* (Figure 3.3C, F & I), grew poorly, which is consistent with previous reports demonstrating that pyrenoid formation is essential for CCM induction under low $CO_2$ conditions (Genkov et al., 2010; Mackinder et al., 2016). The rest of the mutants, except F8 and G6, in this study displayed a slower growth when compared to wild type suggesting the cells possessed CCM defects (Figure 3.3C & F). Mutants F8 and G6 exhibited growth indistinguishable from wild type cells.

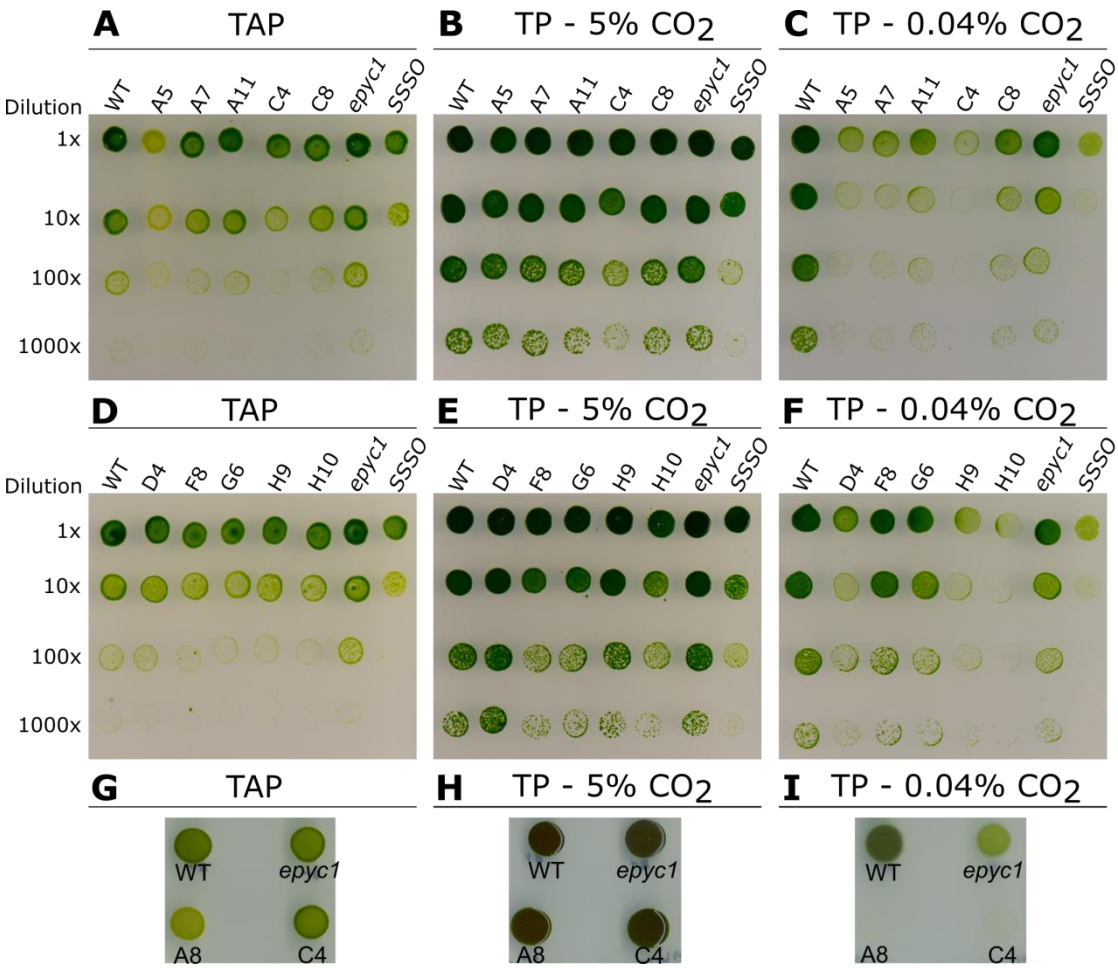



**Figure 3.3 Spot tests investigating the growth response of the wild type, cMJ030 (Zhang et al., 2014), *SSSO* which is a RbcS1 hybrid strain (Genkov et al., 2010) and the CCM defective mutants, including *epyc1* (Mackinder et al., 2016). (A, D & G) Control tris-acetate-phosphate (TAP) plate was kept in the dark at room temperature, while tris-phosphate (TP) plates were kept at about 50 µmol photons $m^{-2}$ $s^{-1}$ at room temperature either at (B, E & H) 5% (v/v) $CO_2$ (high $CO_2$) or at (C, F & I) air level of $CO_2$ for 7 days. Mutants used for further characterisation were A5, A7, A8, A11, C4, C8, D4, F8 and H10.**

### 3.2.2 Photosynthetic Rate of the CCM Mutants

The rate of photosynthetic oxygen evolution under different concentrations of inorganic carbon was used to determine the extent of any inducible carbon concentrating mechanism. The photosynthetic $K_{0.5}$ (Ci) value of the corresponding wild type under low $CO_2$ conditions was on average at 48.2 µM Ci which was in the range of photosynthetic responses of cells expressing CCM of 10-100 µM Ci (Wang et al., 2014). Under low $CO_2$ conditions, most mutants had lower half maximum photosynthetic rates compared to wild type, except C8, F8 and G6 (Figure 3.4 and Table 3.2) suggesting the mutants had reduced affinity for Ci. On the other hand, mutants F8 and G6 showed no significant different in the photosynthetic $K_{0.5}$ (Ci) values, of 53.9 and 86.5 µM Ci respectively, indicating that these mutants possessed a functional CCM similar to that of WT. Under high $CO_2$ (HC) conditions, all mutants except G6 had higher photosynthetic $K_{0.5}$ (Ci) than the corresponding wild type. LC-acclimated WT and mutant F8 cells achieved half maximum photosynthetic rates at below 250 µM Ci while the other mutants required higher concentrations of $C_i$ (Figure 3.4 and Table 3.2). The reduced affinity for $C_i$ of the mutants was consistent with the spot tests (Figure 3.3), indicating that these mutants have knockouts in genes related to CCM induction in *C. reinhardtii*.

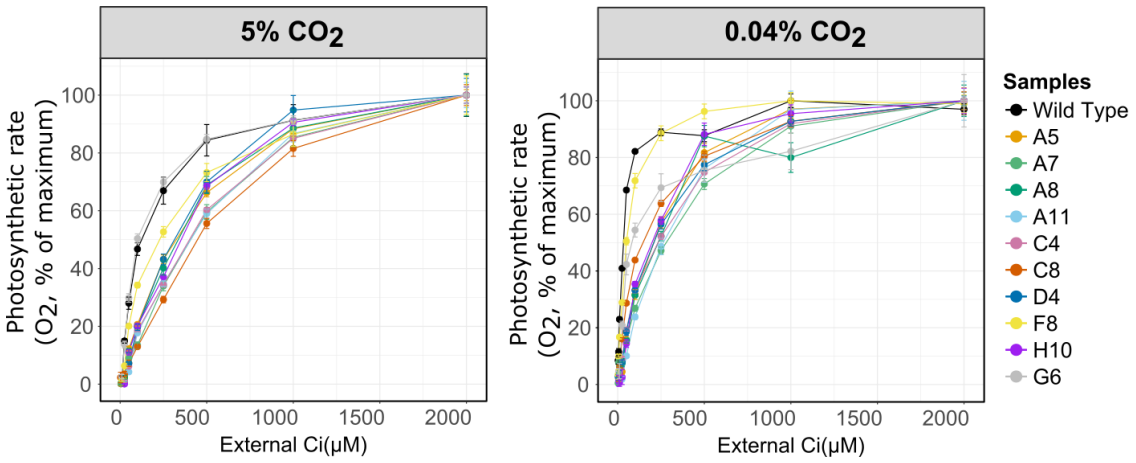



**Figure 3.4 Oxygen-evolving activity of WT and mutants in response to external dissolved inorganic carbon concentrations. WT and mutants were grown in 5% $CO_2$ and 0.04% $CO_2$ conditions. Photosynthetic oxygen evolution was measured using a Clark-type oxygen electrode in the presence of different concentrations of $NaHCO_3$ at the light intensity of about 300 µmol photons $m^{-2}s^{-1}$.**

**Table 3.2 Photosynthetic $K_{0.5}$ (Ci) of wild type and mutants indicating reduced affinity for inorganic carbon ($C_i$). Mutants were grown under high $CO_2$ (5% (v/v) $CO_2$) or acclimated at low $CO_2$ (0.04% (v/v) $CO_2$) for 24 h prior to the oxygen evolution measurements. Data are shown ± standard error, which were obtained from at least three independent experiments. Ci, inorganic carbon; HC, high $CO_2$; LC, low $CO_2$, ****, p<0.0001; ***, 0.0001<p<0.001; **, 0.001<p<0.01; *, 0.01<p<0.1; ns, p>0.1.**

| Sample | Condition | $K_{0.5}$ (Ci) (µM) | p-value | Significance |
|---|---|---|---|---|
| WT | HC | 148.4 ± 24.2 | - | - |
| | LC | 48.2 ± 17.1 | - | - |
| A5 | HC | 483.5 ± 38.1 | 0.00743 | ** |
| | LC | 273.9 ± 10.9 | 0.00571 | ** |
| A7 | HC | 771.3 ± 139.8 | $1.70 \times 10^{-5}$ | **** |
| | LC | 356.9 ± 54.9 | 0.00038 | *** |
| A8 | HC | 453.4 ± 151.6 | 0.01356 | * |
| | LC | 258.5 ± 113.6 | 0.00924 | ** |
| A11 | HC | 711.2 ± 105.0 | $5.90 \times 10^{-5}$ | **** |
| | LC | 364.3 ± 79.1 | 0.0003 | *** |
| C4 | HC | 677.7 ± 29.4 | 0.00012 | *** |
| | LC | 273.2 ± 22.3 | 0.00584 | ** |
| C8 | HC | 883.6 ± 18.2 | $1.70 \times 10^{-6}$ | **** |
| | LC | 139.4 ± 18.2 | 0.22871 | ns |
| D4 | HC | 506.2 ± 77.9 | 0.00467 | ** |
| | LC | 257.6 ± 29.9 | 0.00949 | ** |
| F8 | HC | 263.9 ± 57.9 | 0.32058 | ns |
| | LC | 53.9 ± 8.9 | 0.93857 | ns |
| G6 | HC | 134.3 ± 14.3 | 0.90276 | ns |
| | LC | 86.5 ± 28.1 | 0.60855 | ns |
| H10 | HC | 545.1 ± 64.0 | 0.00207 | ** |
| | LC | 255.5 ± 68.7 | 0.01013 | * |

### 3.2.3 Quantification of RuBisCO aggregation and Pyrenoid Formation in Mutant Lines

In order to investigate the proposed pyrenoid defects of the mutants, cells were grown in liquid TP medium and acclimatised for 24 h at ambient $CO_2$ to assess the extent that RuBisCO content, or the aggregation of RuBisCO into the pyrenoid in the mutants differs from the wild type cells.

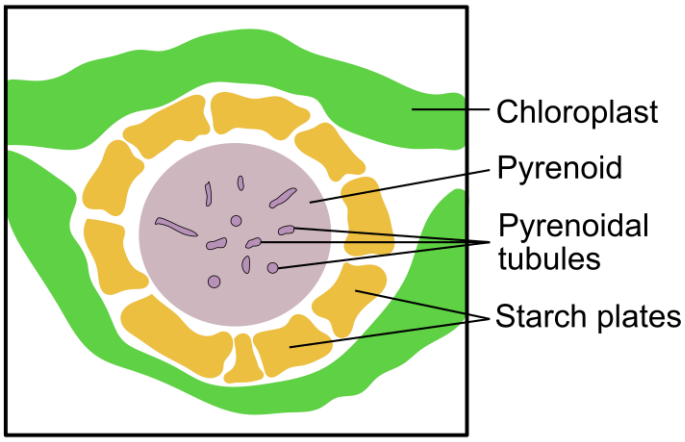



**Figure 3.5 A schematic diagram of a pyrenoid in wild type *Chlamydomonas reinhardtii*. The electron-dense pyrenoid is surrounded by starch plates with fenestrations where the thylakoid tubules penetrate into the pyrenoid to form a network of pyrenoidal tubules.**

The pyrenoid of a wild type cell (Figure 3.5 and Figure 3.6A) is an electron-densed subcompartment which is surrounded by starch plates with fenestrations where the thylakoid tubules penetrate into the pyrenoid and form a knotted network of tubules. Although the transmission electron micrographs showed that the mutants retained the ability to form pyrenoids, these were aberrant. In general, mutant lines lacked a visible thylakoid tubule network within the pyrenoid, and the starch sheath surrounding the pyrenoid was abnormal in shape and arrangement (Figure 3.6 B-J). Scoring of cells for the presence of a pyrenoid using TEM images at low magnification, showed that 45.9% of wild type cells possessed a pyrenoid, which is similar to that reported by Genkov et al. (2010), while the mutants displayed frequencies between 20-31.7% (Table 3.3). Not every cell in an image is expected to contain a pyrenoid due to the plane of the section, an aspect affected by the ratio of pyrenoid to the cell size (Rawat et al., 1996; Genkov et al., 2010).

Measurement of cell section area of wild type and mutants were performed using 30 high magnification TEM images (Figure 3.7). The average cell area of wild type cells

was 24.9 $\mu m^2$ while the average cell area of the mutants ranged from 15.48 $\mu m^2$ to 24.61 $\mu m^2$. Global comparison using the Kruskal-Wallis one-way analysis of variance test on the cell area of wild type and mutants yielded p-value of 0.046. Pairwise comparison between wild type and each of the mutants showed that the cell area of the mutant C4 was significantly different from wild type (Table 3.4). Cell areas of other mutants were similar to that of the wild type suggesting the lower observable frequency of pyrenoid-positive cells could be dependent on reduced pyrenoid size.

As there was no significant difference in cell size apart from mutant C4, the pyrenoid area of the wild type and mutants was measured (Figure 3.8) since the frequency of observing a pyrenoid is dependent on the ratio of pyrenoid size to cell size (Rawat et al., 1996). Wild type cells had an average pyrenoid cross-section area of 1.25 $\mu m^2$ compared to the significantly different pyrenoid area of the mutants, except F8 and H10, which ranged from 0.39 $\mu m^2$ to 0.82 $\mu m^2$ whereas the average pyrenoid area of mutants F8 and H10 were 0.90 $\mu m^2$ and 1.17 $\mu m^2$, respectively. A global comparison generated a p-value of $1.7 \times 10^{-8}$, indicating significant differences among the samples. Pairwise comparison indicated significant difference between the sectional area of the pyrenoid of the wild type and the mutants except mutants F8 and H10 (Table 3.5).

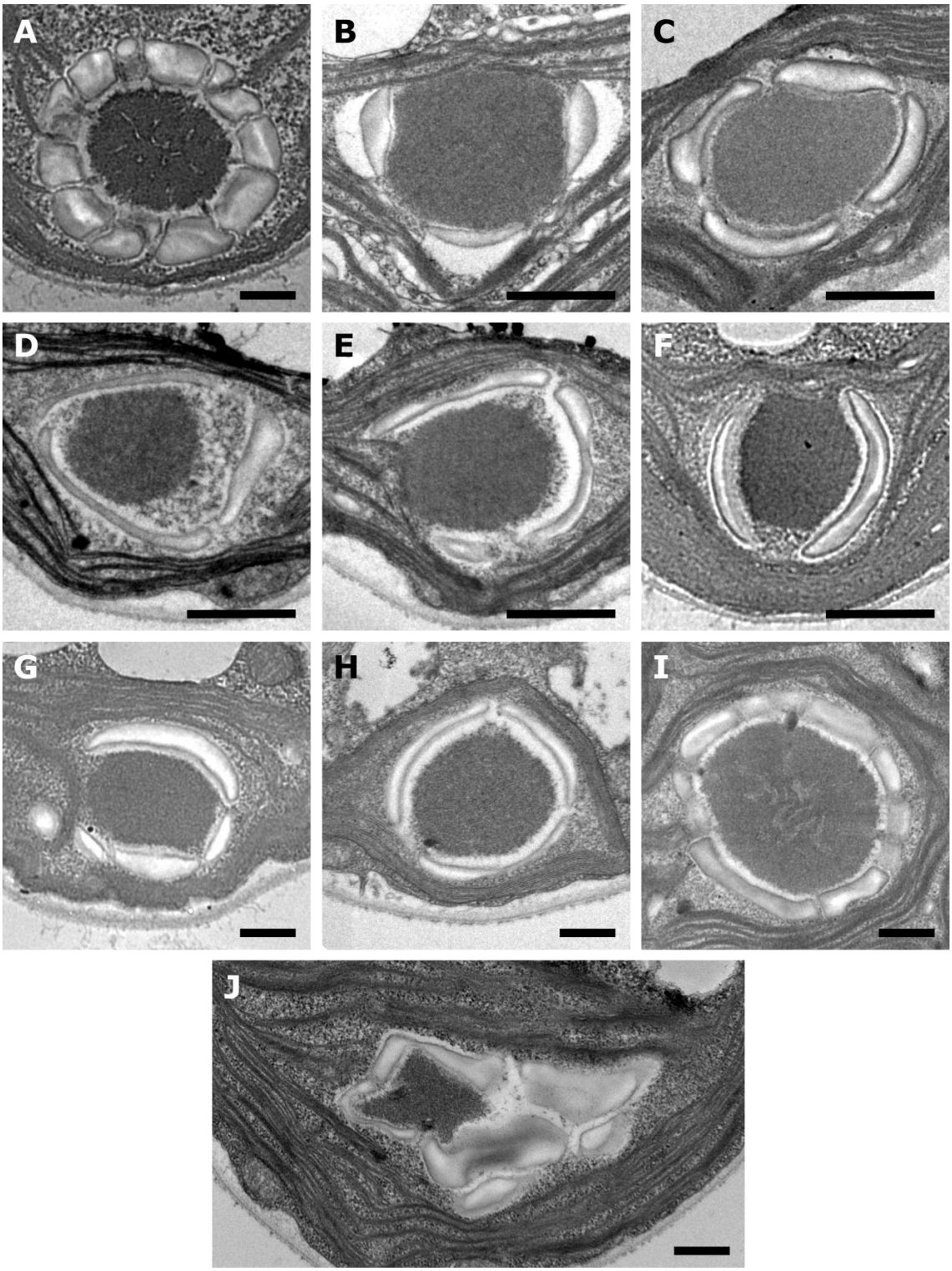



**Figure 3.6 Electron micrographs of the pyrenoid of the CCM defective mutants grown photoautotrophically at room temperature with illumination of about 50 µmol photons $m^{-2}s^{-1}$ under ambient level of $CO_2$ for 24 hours. (A) Wild type, (B) A5, (C) A7, (D) A8, (E) A11, (F) C4, (G) C8, (H) D4, (I) F8, (J) H10. Scale bars 500 nm.**

**Table 3.3 The number of cell profiles with and without pyrenoids in wild type *C. reinhardtii* and the CCM defective mutants.**

| Sample | No. of cells with pyrenoid | No. of cells without pyrenoid | Total Cells Profiles | Percentage with pyrenoid (%) |
|---|---|---|---|---|
| WT | 180 | 212 | 392 | 45.9 |
| A5 | 76 | 223 | 299 | 25.4 |
| A7 | 93 | 303 | 396 | 23.5 |
| A8 | 80 | 221 | 301 | 26.6 |
| A11 | 53 | 212 | 265 | 20.0 |
| C4 | 53 | 197 | 250 | 21.2 |
| C8 | 79 | 253 | 332 | 23.8 |
| D4 | 99 | 366 | 465 | 21.3 |
| F8 | 101 | 240 | 341 | 29.6 |
| H10 | 119 | 257 | 376 | 31.7 |

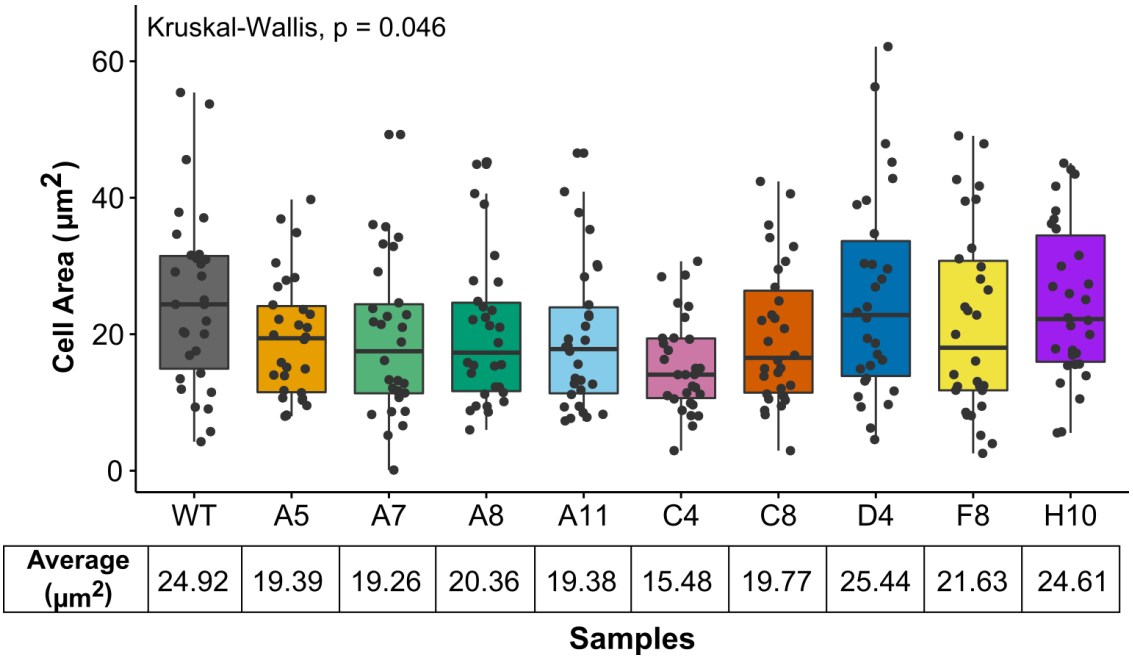



**Figure 3.7 Boxplot showing the measurement of the cell section area of the wild type (WT) and the nine CCM defective mutants (n=30).**

**Table 3.4 Pairwise comparisons between the cell sectional area of wild type and that of the mutants using Mann-Whitney-Wilcoxon test showed that all mutants except C4 had relatively similar cell area.**

| Pair | p-value | Significance |
|---|---|---|
| WT – A5 | 0.0748 | ns |
| WT – A7 | 0.0877 | ns |
| WT – A8 | 0.1393 | ns |
| WT – A11 | 0.0657 | ns |
| WT – C4 | 0.0018 | ** |
| WT – C8 | 0.1315 | ns |
| WT – D4 | 0.9 | ns |
| WT – F8 | 0.2643 | ns |
| WT – H10 | 0.9941 | ns |

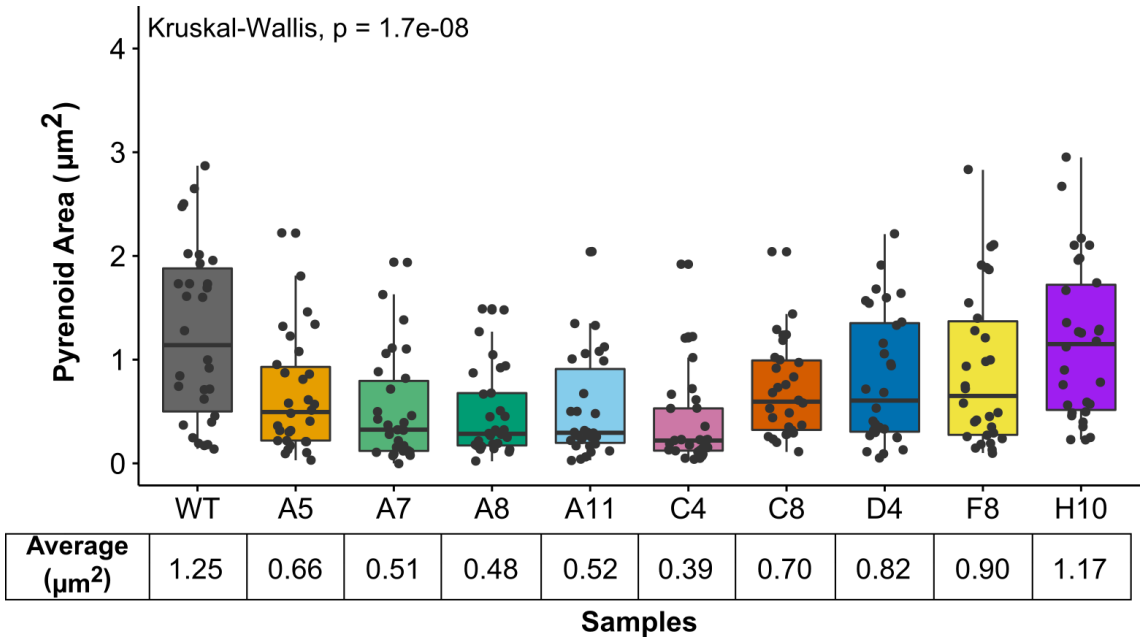



**Figure 3.8 Measurement of the pyrenoid sectional area of the CCM defective mutants displayed significant differences from the corresponding wild type (WT) with Kruskal-Wallis p-value of 1.7e-08 (n=30).**

**Table 3.5 Pairwise comparisons between the pyrenoid section area of wild type and that of the mutants using Mann-Whitney-Wilcoxon test showed that all mutants except F8 and H10 had significantly different pyrenoid size compared to wild type.**

| Pair | p-value | Significance |
|---|---|---|
| WT – A5 | 0.00569 | ** |
| WT – A7 | 0.0002 | *** |
| WT – A8 | 0.00012 | *** |
| WT – A11 | 0.00065 | *** |
| WT – C4 | $8.0 \times 10^{-6}$ | **** |
| WT – C8 | 0.01948 | * |
| WT – D4 | 0.02918 | * |
| WT – F8 | 0.11194 | ns |
| WT – H10 | 0.89413 | ns |

As the pyrenoid size appeared to be smaller, the amount of Rubisco was determined by quantitative fluorescence Western blot to test whether the amounts were reduced in mutants. There appeared to be no difference between strains (Figure 3.9) suggesting there might be a defect in Rubisco trafficking into the pyrenoid.

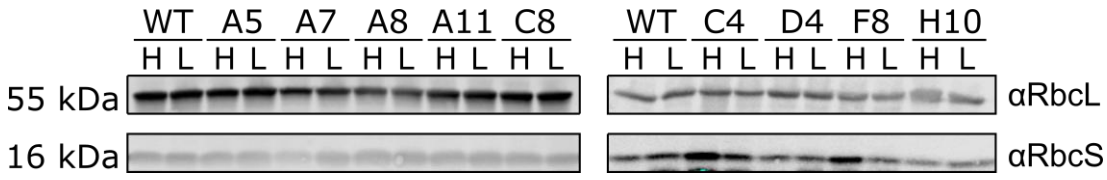



**Figure 3.9 Fluorescence Western blot using anti-Rubisco antibody of wild type and CCM defective mutants grown at 5% (v/v) and 0.04% (v/v) $CO_2$ conditions. H, 5% (v/v) $CO_2$; L, 0.04% (v/v) $CO_2$.Total soluble protein of 10 µg was loaded.**

In order to test the hypothesis that Rubisco is trafficked into the pyrenoid, immunogold labelling using an anti-Rubisco antibody was performed. From the observation made (Figure 3.10 and Table 3.6), it was clearly shown that Rubisco was preferentially aggregated into the pyrenoid of the cells (Lacoste-Royal and Gibbs, 1987; Süss et al., 1995; Morita et al., 1997; He et al., 2003; Mitchell et al., 2014). Upon elimination of the

background density, ~94% of gold particles localised into the pyrenoid of WT cells under LC conditions, while the mutants displayed values ranging from 74% (A5) to 100% (F8) (Table 3.6). Tukey's test indicated that there was no significant difference ($p<0.05$) between the percentage of Rubisco aggregating in the mutants and the wild type (Table 3.6) thus the hypothesis was rejected. The number of particles normalised to the area of the pyrenoid in the wild type was 140 ± 17.0 particles $nm^{-2}$ while the density varied from 28 ± 5.0 to 169 ± 25.2 particles $nm^{-2}$ in the mutants. The density of gold particles was lower in the mutants possessing smaller pyrenoids thus the difference in the size of the pyrenoid appeared to correlate to the volume of Rubisco in the pyrenoid. This suggested the possibility of missing component(s) in the pyrenoid of the mutants.

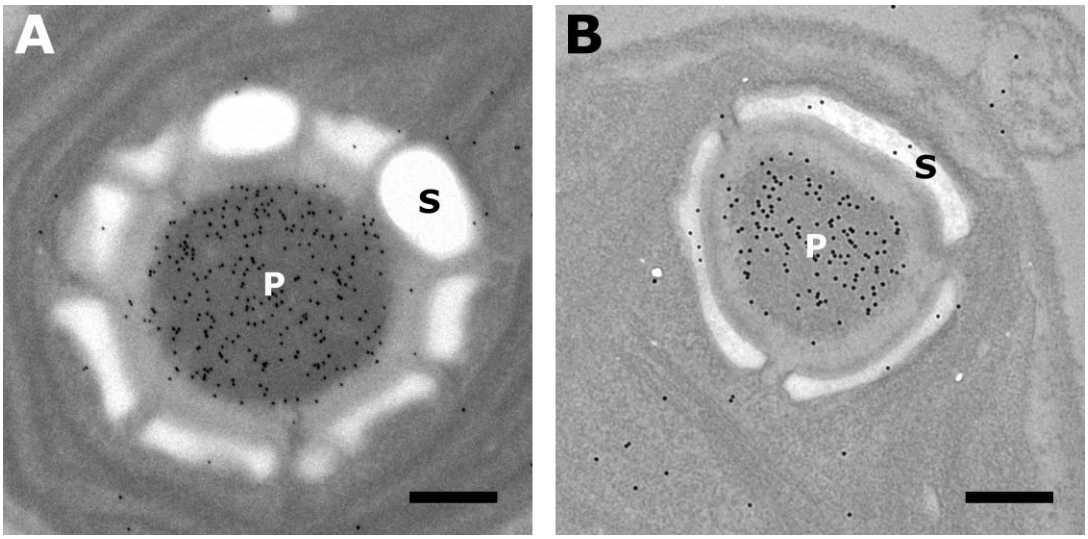



**Figure 3.10 Immunolocalisation of anti-RuBisCO antibody of *C. reinhardtii* wild type and mutant A8 cells grown on TP medium after 24 h CCM induction showed similar RuBisCO localisation pattern in the pyrenoid between wild type and the mutant. (A) Wild type cMJ030. (B) Example of a CCM defective mutant, A8. Sections were probed with an antibody raised against *C. reinhardtii* RuBisCO, as describe in "Materials and Methods". P, Pyrenoid; S, starch sheaths.**

The data in Table 3.6 showed that the lower immunogold particles per pyrenoid area of the mutants probed with anti-Rubisco antibody could be an effect of smaller pyrenoid size in the mutants as a cause of missing pyrenoidal components. In furtherance of the hypothesis that smaller pyrenoid is caused by missing pyrenoidal components, the initial step was to perform Western blots probing for the CCM-related proteins known to localise to the pyrenoid, namely CAH3, LCIB and LCIC (Figure 3.11). The corresponding wild type and all the CCM defective mutants displayed similar protein expression profiles of CAH3 under both high and low $CO_2$ conditions as shown by

Karlsson et al. (1995), Karlsson et al. (1998) and Blanco-Rivero et al. (2012). However, it was plausible that although protein amount was similar, there might be defect in post-translational modifications since CAH3 is activated by phosphorylation during CCM induction (Blanco-Rivero et al., 2012). The mutants possessed similar accumulation profiles of LCIB and LCIC as the corresponding wild type where the proteins were expressed at higher amounts under low $CO_2$ conditions, similar to the accumulation profiles reported by Yamano et al. (2010) and Yamano et al. (2014). This indicates that the gene knockouts possessed by the mutants did not affect the expression level of the CCM proteins known to accumulate in or around the pyrenoid.

**Table 3.6 Calculation of RuBisCO fraction in both the pyrenoid and stroma of wild type cells and the mutants acclimatised for 24 h at 0.04% $CO_2$ to induce CCM (n=20). All cells were grown in TP (tris phosphate) liquid medium at constant irradiance of about 50 µmol $m^{-2}s^{-1}$. The fraction of RuBisCO in the pyrenoid and chloroplast stroma was normalised to the density of the particles in the background. The data shown are the averages ± SE. The adjusted p-value (Tukey's test) of the percentage of aggregation of RuBisCO in pyrenoid for each mutant against wild type showed no significant difference ($p<0.05$).**

| Sample | Average size of pyrenoid cross section ($nm^2$) | Normalised Immunogold Particle Count ($particles.nm^{-2}$) | | RuBisCO in Pyrenoid (%) | Comparison to WT (Adjusted p-value) |
|---|---|---|---|---|---|
| | | Pyrenoid | Stroma | | |
| WT | $9.25 \times 10^6$ | 140 ± 17.0 | 6 ± 1.3 | 93.8 | - |
| A5 | $4.99 \times 10^6$ | 62 ± 8.8 | 16 ± 2.5 | 74.4 | $4.63 \times 10^{-3}$ |
| A7 | $6.98 \times 10^6$ | 28 ± 5.0 | 3 ± 0.8 | 83.6 | $3.21 \times 10^{-6}$ |
| A8 | $4.53 \times 10^6$ | 84 ± 13.7 | 14 ± 1.7 | 80.6 | $1.54 \times 10^{-1}$ |
| A11 | $5.62 \times 10^6$ | 100 ± 14.0 | 12 ± 3.4 | 88.8 | $6.23 \times 10^{-1}$ |
| C4 | $5.85 \times 10^6$ | 39 ± 5.1 | 6 ± 1.5 | 83.5 | $4.03 \times 10^{-5}$ |
| C8 | $7.91 \times 10^6$ | 50 ± 8.6 | 4 ± 0.8 | 90.3 | $5.27 \times 10^{-4}$ |
| D4 | $6.78 \times 10^6$ | 123 ± 18.6 | 8 ± 1.2 | 91.4 | $9.98 \times 10^{-1}$ |
| F8 | $8.44 \times 10^6$ | 100 ± 11.4 | 0 ± 0.7 | 100.0 | $6.17 \times 10^{-1}$ |
| H10 | $7.23 \times 10^6$ | 169 ± 25.2 | 16 ± 3.4 | 90.5 | $8.96 \times 10^{-1}$ |

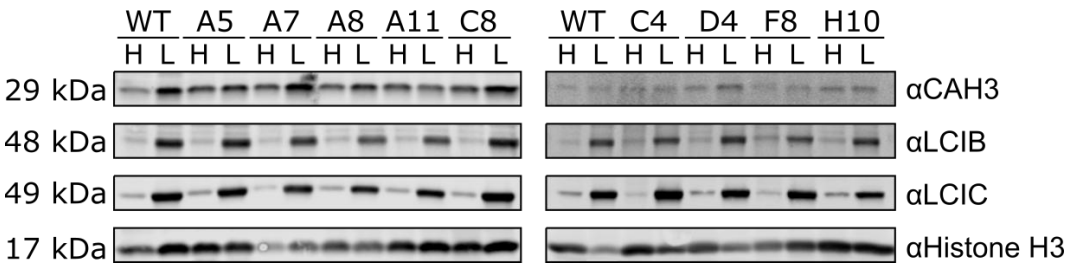



**Figure 3.11 Fluorescence Western blots of CCM defective mutants grown under 5% and 0.04% $CO_2$ conditions with antibodies of CCM-related proteins linked to the pyrenoid: anti-CAH3, anti-LCIB, anti-LCIC, with anti-Histone H3 as loading control. H, 5% (v/v) $CO_2$; L, 0.04% (v/v) $CO_2$. Total soluble protein of 5 µg was loaded.**

### 3.2.4 Phenotypic Aberrations of CCM Defective Mutants

As Rubisco aggregation into the pyrenoid appeared to be functional despite smaller pyrenoid size in the mutants with CCM defects, TEM images were assessed to investigate potential causes to the defective inorganic carbon transport of the mutants.

In the initial transmission electron microscopy analysis, the mutants were shown to possess some form of phenotypic aberrations from the wild type. A typical wild type electron micrograph at median plane (Figure 3.12) would possess a pyrenoid in the middle of a cup-shaped chloroplast, and is surrounded by starch plates with fenestrations where the thylakoid tubules penetrate into the pyrenoid and form a network of tubules.

Qualitative assessments of 50 electron micrographs of each of the mutants revealed that several possessed similar phenotypic aberrations including: 1) the lack of visible thylakoid tubule network in the pyrenoid, 2) the occurrence of multiple pyrenoids, 3) the starch sheaths which are non-wild-type-like, and 4) the presence of starch granules in the chloroplast (Table 3.7). There was only one out of 50 individual wild type cell images examined which possessed two pyrenoids (Figure 3.14A), which might be an artefact of asynchronous culture. However, the higher frequency of up to about 20% of the mutants possessing multiple pyrenoids (Figure 3.14B-F) might be a causal effect of the gene knockout rather than an artefact of asynchronous culture. The phenotype of multiple pyrenoids in *Chlamydomonas* has been documented in mutant strains, e.g. the the master regulator mutant, *cia5* (Fukuzawa et al., 2001), the putative methyltransferase mutant, *cia6* (Ma et al., 2011), the mutant of vacuole membrane

protein, *vmp-1* (Tenenboim et al., 2014), or when colchicine, a chemical which prevents microtubule assembly is added to the cells (Walne, 1966, 1967).

The observation of thylakoid tubule networks in the pyrenoids of all 50 of the individual wild type cell images used in this study was consistent with the three dimensional reconstruction of the thylakoid tubule network in the pyrenoid by Engel et al. (2015) and Meyer et al. (2017) which demonstrated the convergence of thylakoid tubules forming interconnected network towards the centre of the pyrenoid (Figure 3.13A). Most of the CCM defective mutants, except F8, showed low frequency ranging from 1 (A5 and A7) to a maximum of 24 individual cells (C8) with some form of visible pyrenoid tubules (Table 3.7). Mutant F8 showed a pyrenoidal tubule network highly similar to that of wild type indicating absence of physical defect to the pyrenoid tubule network in this mutant (Figure 3.13B). On the other hand, though thylakoid tubules could sometimes be observed in the mutant C8 (Figure 3.13C), the tubules did not form a network of tubules as described in literature (Ohad et al., 1967a; Ohad et al., 1967b; Engel et al., 2015) indicating the CCM defect might be associated with defects in thylakoid tubule network in the pyrenoid.

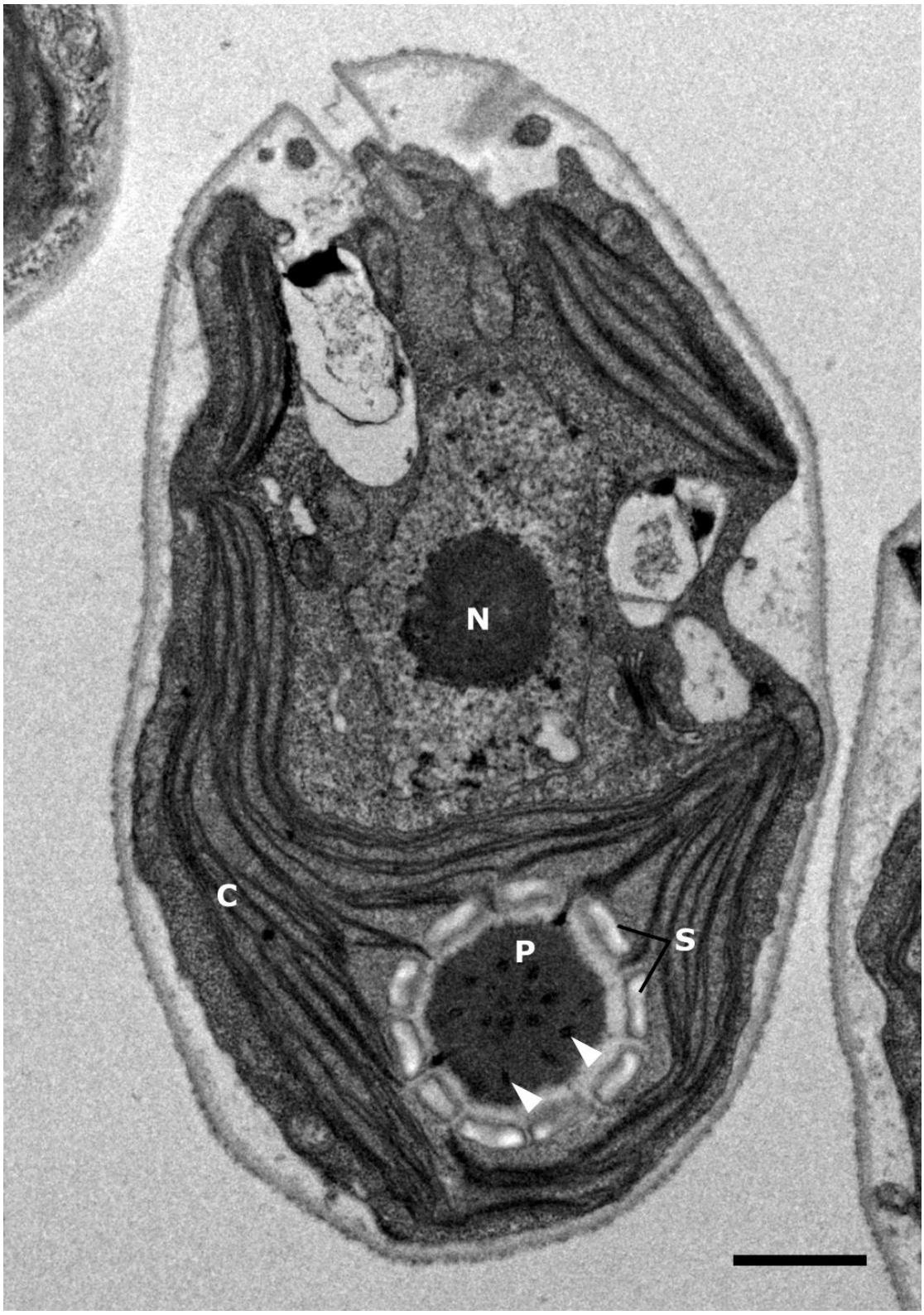



**Figure 3.12 Electron micrograph of a wild type cell, cMJ030, at the median plane. N, nucleus; P, pyrenoid; S, starch sheaths; C, chloroplast; arrows indicating pyrenoidal tubules. Bar 500 nm.**

**Table 3.7 Qualitative observations on 50 electron micrographs of the CCM defective mutants and the corresponding wild type based on four characteristics: 1) the presence of multiple pyrenoids, 2) the presence of thylakoid tubule network in the pyrenoid and 3) the presence of pyrenoidal starch plates with abnormal shapes and arrangement 4) the presence of chloroplastic starch granules.**

| Sample | Frequency (n=50) | | | |
|---|---|---|---|---|
| | Multiple Pyrenoids | Pyrenoidal Tubules | Abnormal Pyrenoidal Starch Sheath | Chloroplastic Starch Granules |
| WT | 1 | 50 | - | 11 |
| A5 | 11 | 1 | 50 | 23 |
| A7 | 11 | 1 | 50 | 18 |
| A8 | 10 | 3 | 47 | 8 |
| A11 | 11 | 10 | 48 | 16 |
| C4 | 13 | 4 | 50 | 46 |
| C8 | 11 | 24 | 48 | 31 |
| D4 | 7 | 3 | 49 | 10 |
| F8 | 1 | 49 | 6 | 49 |
| H10 | 8 | 9 | 27 | 16 |

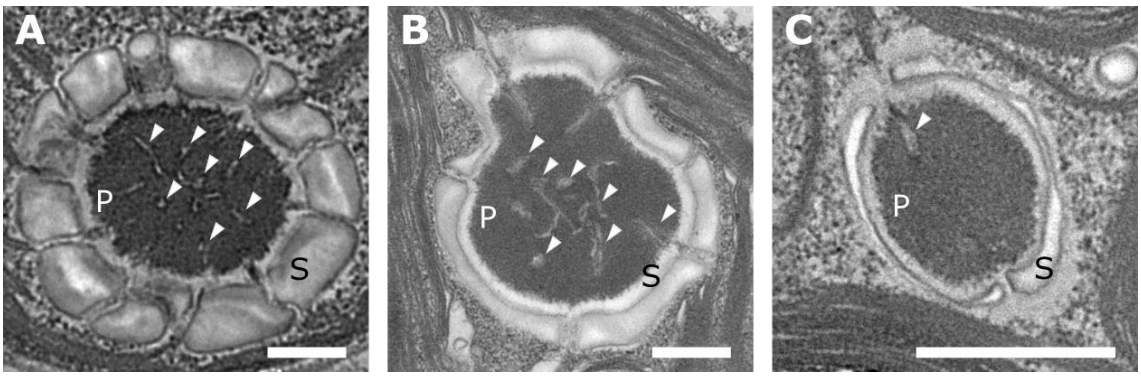



**Figure 3.13 Electron micrographs of (A) wild type, (B) F8 and (C) C8 that showed presence of thylakoid tubule(s) in the pyrenoid. White arrows indicated some of the thylakoid tubules in the pyrenoid. Scale bar 500 nm.**

Since it is hypothesised that the pyrenoid is formed as a result of external forces generated by starch sheath formation (Engel et al., 2015), the dimensions of the starch sheath in mutants was assessed both qualitatively (Figure 3.14B-F) and quantitatively (Figure 3.15 and Figure 3.16). The classical starch sheath in WT cells (Figure 3.6A, Figure 3.12, Figure 3.13A, Figure 3.14A) surrounded the pyrenoid and fenestrations enabled the thylakoid membranes to penetrate into the pyrenoid core (Engel et al., 2015). Aside from F8 which retained a starch sheath arrangement similar to wild type (Figure 3.6I), the rest of the mutants had a starch sheath which was abnormal in shape and arrangement. For example, most of the mutants displayed thin, crescent-shaped starch plates which did not surround the whole pyrenoid (Figure 3.6B-H, J).

From the total of 30 individual cells measured, the median of starch area per pyrenoid of wild type cells was about 1 $\mu m^2$ while the mutants ranged between 0.15 to 0.70 $\mu m^2$ (Figure 3.15). All mutants except H10 had a starch area which was significantly different from wild type. As for the number of starch plates per pyrenoid (Figure 3.16), wild type cells were generally surrounded by a median of 7 starch plates, followed by F8 with 6 starch plates, and H10 with 5 starch plates. All other mutants, A5, A7, A8, A11, C4, C8 and D4) had a significantly lower number of starch plates, between 2 to 3 per pyrenoid.

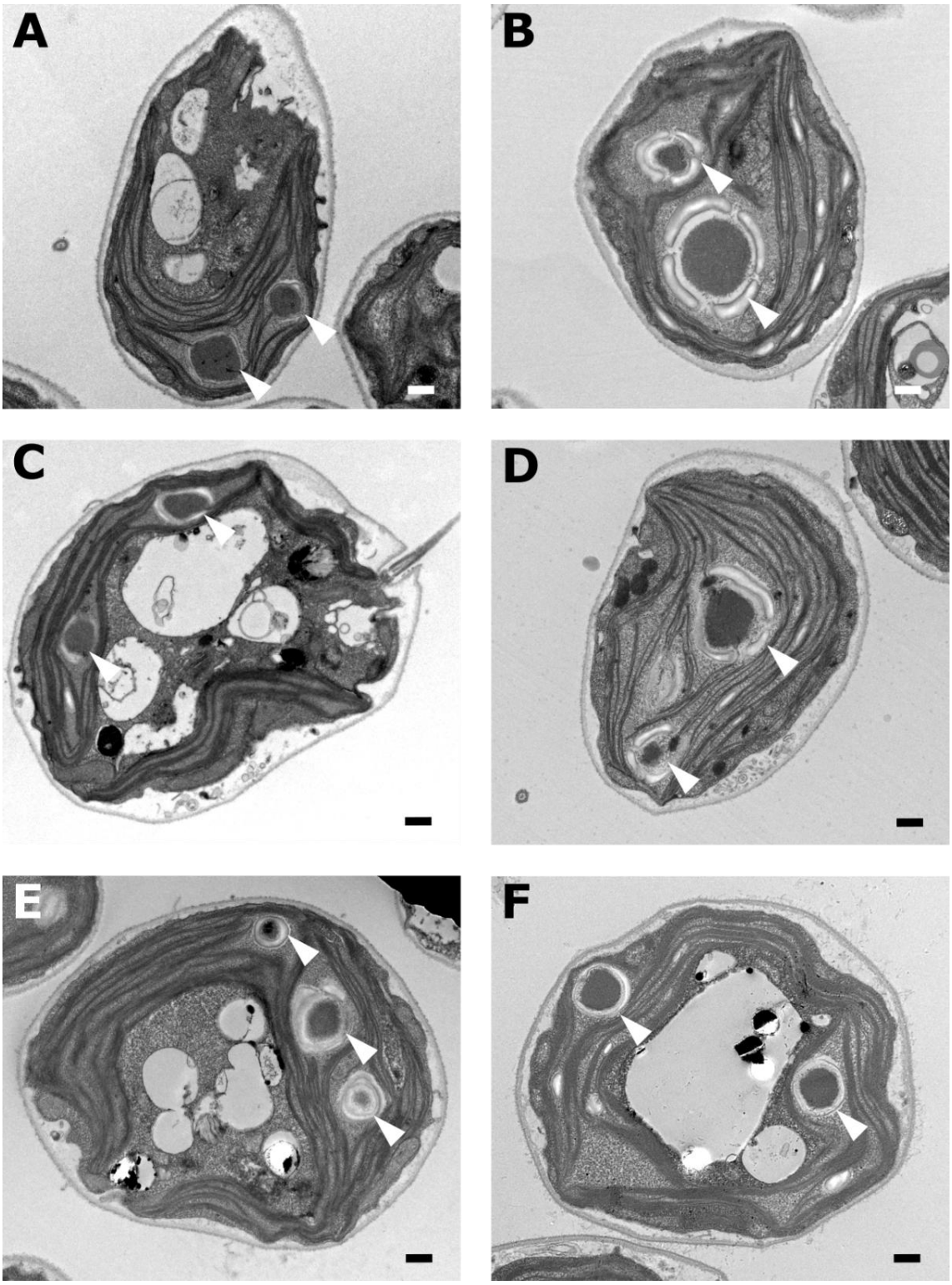



**Figure 3.14 Electron micrographs of (A) wild type and mutants (B) A5, (C) A7, (D) A8, (E) A11 and (F) C8 that showed multiple pyrenoids per cell. White arrows indicate the position of the pyrenoids. Scale bar 500 nm.**

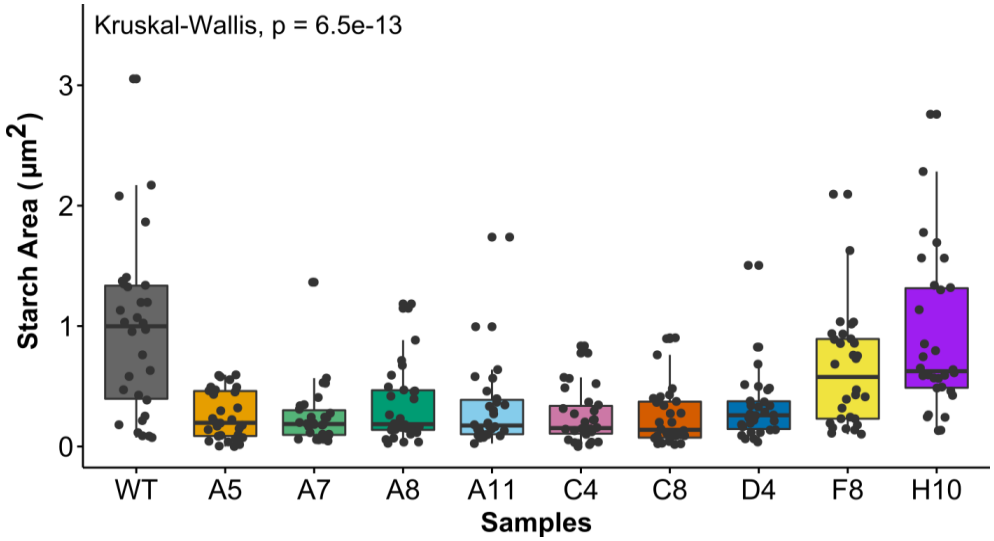



**Figure 3.15 Total pyrenoidal starch area of wild type and the mutants with p-value =6.5e-13, n=30.**

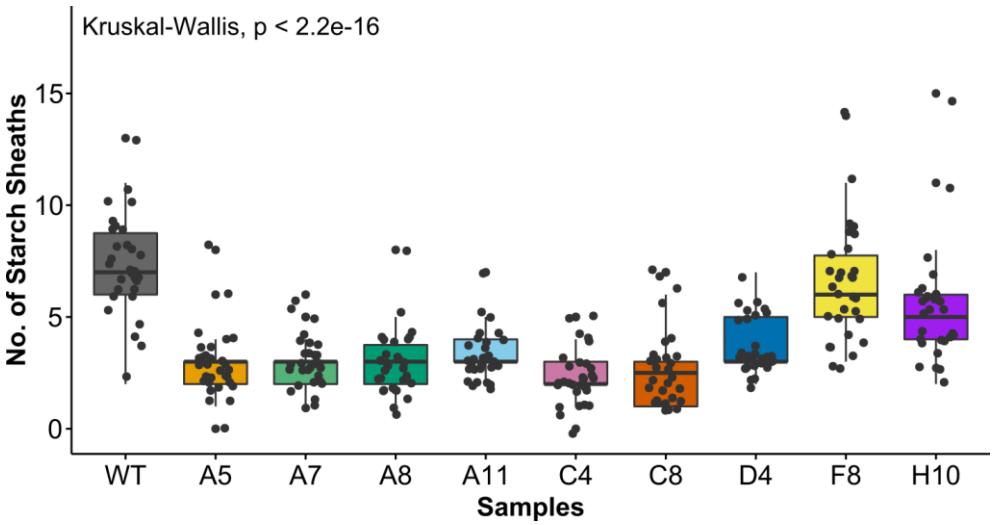



**Figure 3.16 Total number of starch plates per pyrenoid of mutants differed significantly from wild type with p-value<2.2e-16, n=30.**

WT and mutants were found to contain chloroplastic starch granules which were not pyrenoidal starch sheath (Figure 3.17). Out of the 50 images assessed for each sample, there were 49 images of mutant F8 displaying multiple chloroplastic starch granules within the cell (Figure 3.17B). Other mutants which were scored positive for chloroplastic starch granules would normally possess only one or two granules per cell. For example, 46 out of 50 individual cells of mutant C4 contained up to two starch granules per cell (Figure 3.17C). Another phenotypic aberrance observed while

assessing the electron micrographs was mislocalisation of pyrenoid(s) from the canonical position at the centre of the chloroplast. This occurred more often in samples with high frequency of multiple pyrenoids, such as mutants A5, A7, A8, A11, C8 (Figure 3.14B-F) and C4 (Figure 3.17C).

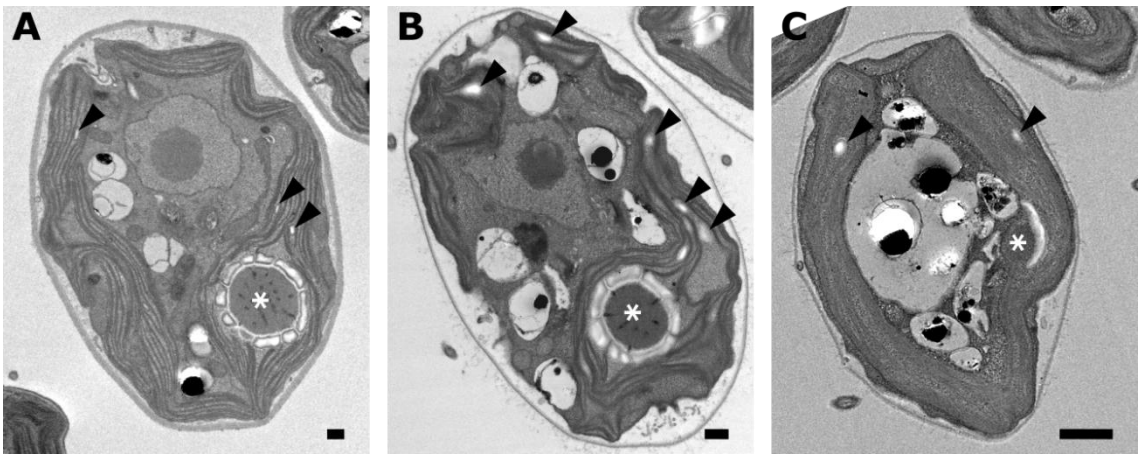



**Figure 3.17 Examples of (A) wild type and mutants (B) F8 and (C) C4 with the presence of chloroplastic starch granules. Starch granules were marked with arrows. Scale bar 500 nm.**

The qualitative assessment of the phenotypic aberrances of these nine CCM defective mutants acclimated at 0.04% $CO_2$ provided an insight on how the mutants differ from the corresponding wild type based on observation of electron micrographs. The radar charts (Figure 3.18) summarised these traits for each mutant: the presence of thylakoid tubule network in the pyrenoid, the occurrence of multiple pyrenoids, the presence of aberrant starch sheath and the presence of chloroplastic starch granules, dividing the mutants into several groups based on similar patterns of the phenotypes scored. Five mutants, A5, A7, A8, A11 and D4, were placed in the same group. It also suggested that mutant F8 exhibited the most similar phenotypic profile as WT, aside from the presence of chloroplastic starch granules at very high frequency. As most mutants lacked pyrenoidal tubules with aberrant pyrenoidal starch plates, this inferred the possibility that a gene or a group of genes in the same pathway might have been knocked-out, causing the similar phenotypic aberrances.

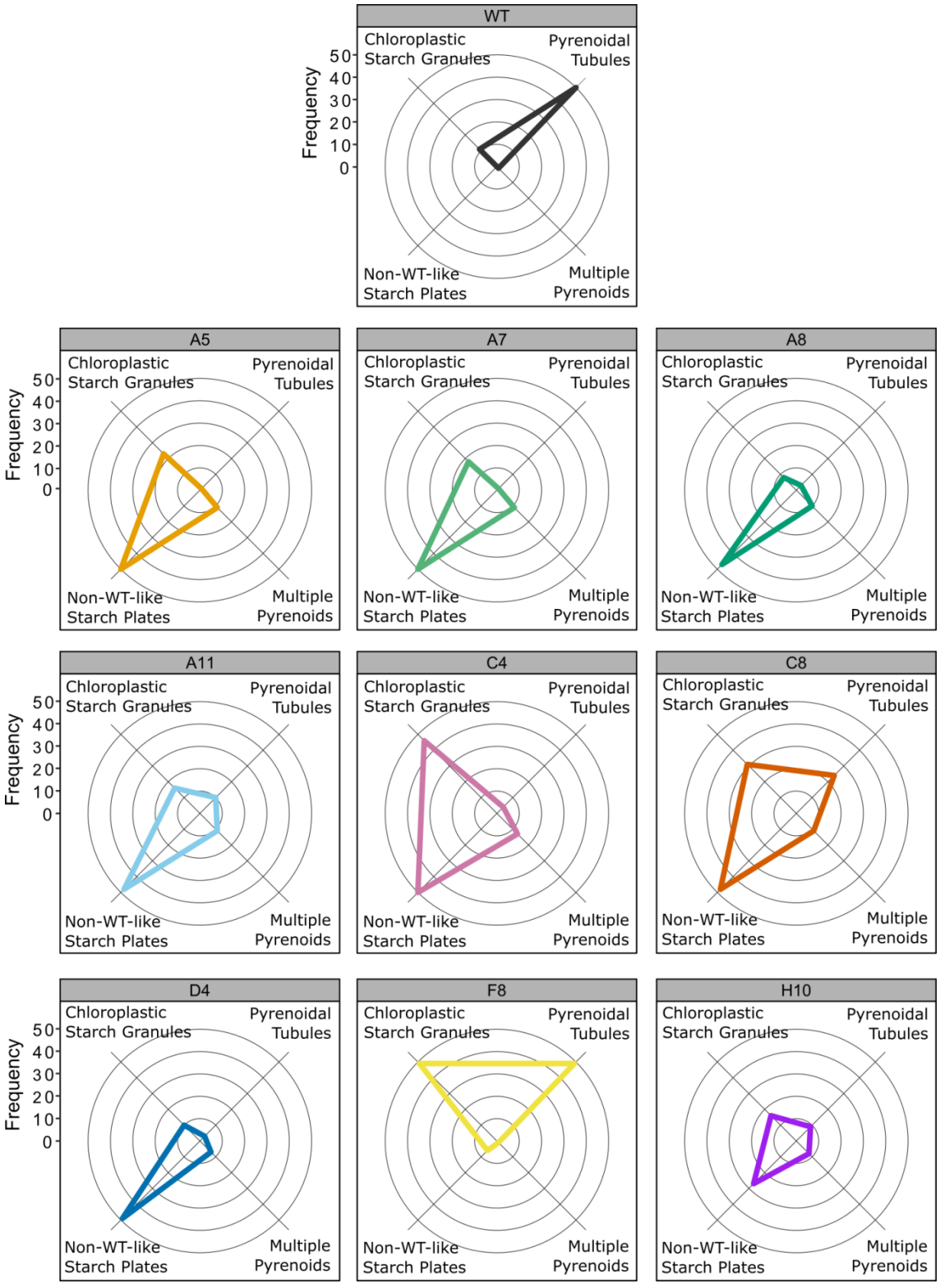



**Figure 3.18 Summary of the qualitative assessments from the data of Table 3.7 using the four phenotypes observed on the electron micrographs of the CCM defective mutants.**

### 3.2.5 Identification of Candidate Genes by Whole Genome Sequencing

Whole genome sequencing was carried out in collaboration with the Jonikas lab (previously Stanford University, now Princeton University), and Sequence alignment on the raw data was performed by Ms. Weronika Patena (Stanford University) using the scripts of Ms. Mia Jaffe (Stanford University). The data generated by the Jonikas lab was used to identify gene knockouts responsible for the CCM defects displayed by the mutants (Table 3.8). Five mutants, A5, A7, A8, A11 and C8, were predicted to be candidate knockouts on gene Cre11.g467712 encoding an unknown protein with a putative starch binding domain and structural maintenance of chromosomes (SMC) N terminal domain. The candidate knockout for mutant F8 was on Cre03.g199000, which encodes the phototropin protein and G6 had a potential knockout on Cre14.g616600. No annotation was found for mutant C4 (Cre02.g143387) and H10 (Cre06.259100).

**Table 3.8 Summary of the top candidate gene for each CCM defective mutant used in this study showed that five mutants contained knockout of gene Cre11.g467712. Whole genome sequencing was performed by the Stanford Functional Genomics Facilities and sequence alignment for all mutants was performed by Ms. Weronika Patena of Jonikas Lab using Ms. Mia Jaffe's scripts.**

| Mutant | Phytozome ID | Description | Gene Location | Highest no. of hits |
|---|---|---|---|---|
| A5 | Cre11.g467712 | PTHR18937 – Structural maintenance of chromosomes SMC family member | chromosome_11:1315112..1333013 forward | 9/14 |
| A7 | Cre11.g467712 | PTHR18937 – Structural maintenance of chromosomes SMC family member | chromosome_11:1315112..1333013 forward | 15/48 |
| A8* | Cre11.g467712 | PTHR18937 – Structural maintenance of chromosomes SMC family member | chromosome_11:1315112..1333013 forward | 7/20 |
| A11 | Cre11.g467712 | PTHR18937 – Structural maintenance of chromosomes SMC family member | chromosome_11:1315112..1333013 forward | 8/18 |
| C4 | Cre02.g143387 | Unannotated | chromosome_2:9114756..9117427 | 9/22 |
| C8 | Cre11.g467712 | PTHR18937 – Structural maintenance of chromosomes SMC family member | chromosome_11:1315112..1333013 forward | 17/22 |
| D4 | NA | Intergenic | chromosome_6:4414218..4414467 | 10/16 |
| F8 | Cre03.g199000 | Phototropin | chromosome_3:8321420..8329663 reverse | 10/20 |
| G6 | Cre14.g616600 | PTHR11649:SF36 - FZL | chromosome_14:1274369..1283459 forward | 9/12 |
| H10 | Cre06.g259100 | Unannotated | chromosome_6:1418964..1430024 reverse | 3/10 |

## 3.3 Discussion

The pyrenoid, where Rubisco is localised (Lacoste-Royal and Gibbs, 1987; Süss et al., 1995; Morita et al., 1997; Borkhsenious et al., 1998; He et al., 2003), acts as the central site of carbon fixation when the CCM is induced. A functional pyrenoid enables the cell to survive under low-$CO_2$ conditions while the lack of a pyrenoid or functional Rubisco complexes contribute to the inability to thrive in the same environment (Rawat et al., 1996; Genkov et al., 2010). The CCM mutants with pyrenoid defects provided by the Jonikas Lab were characterised based on the main hypothesis that the pyrenoid is formed by multiple components and the results of this study pointed towards this direction with the identification of knockouts of at least three different genes.

### 3.3.1 Localisation of Rubisco in the Chloroplast of *Chlamydomonas reinhardtii*

The initial experiment of growth spot tests under three different conditions: mixotrophic in the dark and photoautotrophic in constant illumination under high- and low-$CO_2$, confirmed that most mutants had a high-$CO_2$-requiring phenotype indicating a dysfunctional CCM (Van et al., 2001), consistent with the outputs of the original Stanford screen (Zhang et al., 2014). Further validation using oxygen evolution showed that the mutants had reduced photosynthetic affinity for inorganic carbon similar to the profiles exhibited by known CCM mutants, such as the CCM master regulator, *ccm1* mutant (Fukuzawa et al., 2001), a CAH1 regulator, *lcr1* mutant (Yoshioka et al., 2004), the LCIB protein, *abl* mutants (Yamano et al., 2014), the Rubisco linker protein, *epyc1* mutant (Mackinder et al., 2016) and a solute carrier protein, *cia8* mutant (Machingura et al., 2017).

Based on the immunofluorescence screen using anti-Rubisco antibody by Caspari (2015), the CCM defective mutants used in this study possessed pyrenoid defects. Caspari's experiment used Rubisco as a key indicator for the position of the pyrenoid in wild type cells, since Rawat et al. (1996) demonstrated that 90% of Rubisco is transported into the pyrenoid following CCM induction. Thus Rubisco is also a suitable gauge of potential defects in pyrenoid formation in mutants. The work of Rawat et al. (1996), Genkov et al. (2010) and Caspari et al. (2017) demonstrated that mutation to Rubisco of *C. reinhardtii* leads to the lack of an observable pyrenoid matrix in the

mutants The initial question which originated from the pyrenoid screens was "Is the process of Rubisco localisation affected by the insertional mutagenesis?"

The initial scoring of visible pyrenoids in the cells using low magnification electron micrographs verified the presence of pyrenoids in the mutant, though the frequency was lower than that of the wild type. The probability of visible pyrenoids in wild type cell profiles in this study was higher than the observation made by Rawat et al. (1996). The frequency of observing visible pyrenoid in cells is dependent on various factors, including the ratio of pyrenoid-to-cell size, the angle of the section, fixation parameters and the physiological conditions of the cells (Rawat et al., 1996), so the higher frequency of wild type cells with visible pyrenoid in this study compared to what was observed by Rawat et al. (1996) might be caused by the induction of the CCM, since the inducible fraction of Rubisco increases the size of the pyrenoid matrix (Meyer et al., 2017). In the work of Rawat et al. (1996), cells were grown in tris-acetate phosphate (TAP) medium, where the CCM is repressed due to the presence of an alternative carbon source. The size of the pyrenoid is expected to be larger under CCM-induced conditions, as supported by Genkov et al. (2010). The authors reported that 66% of wild type cell profiles showed visible pyrenoids under CCM-induced conditions. As the frequency of observing cell profiles with visible pyrenoids were lower for all mutants compared to the wild type, it is suggested that the mutants in this study might have a smaller pyrenoid or an increased cell size.

Measurements of cell and pyrenoid areas performed using 30 high magnification electron micrographs of cells grown under low $CO_2$ conditions showed that wild type and mutants had similar cell areas but the pyrenoids of the mutants were significantly smaller. A smaller pyrenoid could be due to various reasons. Moroney et al., (1989) and Fukuzawa et al. (2001) both demonstrated that defects in signal transduction pathway for acclimation to limiting $CO_2$ could lead to dysfunctional CCM and pyrenoid defects as present in *cia5* mutants. Mislocalised pyrenoid-linked protein, such as LCIB (Yamano et al., 2014) and lack of Rubisco linker protein, as in *epyc1* mutant (Mackinder et al., 2016), could also lead to significantly smaller pyrenoids.

Despite containing smaller pyrenoids, Western blots with an anti-Rubisco antibody showed similar amounts of Rubisco between wild type and the mutants. Immunogold labelling with the same antibody revealed that percentage of gold particles localised within the pyrenoid of wild type and mutants remained relatively similar. However, the particle density in the pyrenoid for some mutants was significantly different from the

corresponding wild type. It was noted in the work of Spreitzer et al. (1985) and Rawat et al. (1996) that mutation to Rubisco could affect the formation of a pyrenoid. This was first demonstrated using Rubisco large subunit mutants, 18-7G and 68-4PP, whereby 18-7G mutant does not possess pyrenoid matrix under any conditions, while 68-4PP mutant had no pyrenoid at higher temperature (Rawat et al., 1996). Similarly, Rubisco small subunit substitutions with higher plant RbcS2 caused the absence of pyrenoid matrix in the cells (Genkov et al., 2010). The initial hypothesis that there is defect in Rubisco trafficking of the mutants was rejected since the data pointed to the fact that the trafficking mechanism was intact.

### 3.3.2 The Biophysical Components of the Pyrenoid of *Chlamydomonas reinhardtii*

Considering that more than 80% of Rubisco was trafficked into the pyrenoid of the mutants but the size of the pyrenoid of the CCM defective mutants was significantly smaller, the second hypothesis was the lack of other pyrenoid-related CCM components causes the mutants to possess smaller pyrenoids. The consistent protein profiles of Western blots using anti-LCIB, anti-LCIC and anti-CAH3 antibodies between wild type and the mutants in this study led to the inference that the mutants did not possess knockouts in the tested genes. This experiment was designed based on the knowledge that aside from Rubisco, LCIB/LCIC complex is shown to localise at the peripyrenoidal region (Yamano et al., 2010, 2014; Jin et al., 2016) while CAH3 is present in the thylakoidal network of the pyrenoid (Karlsson et al., 1998; Benlloch et al., 2015; Engel et al., 2015; Mackinder, 2017). Mackinder et al. (2017) recently demonstrated through their microscopy and proteomic screens that there are four distinct layers within the pyrenoid structure, with a total of 89 pyrenoid-related proteins (Figure 5.1). This study showed that the lack of any layer of the pyrenoid structures could lead to smaller pyrenoids, as displayed by the mutants. Although there are a large amount of proteins which could have been tested, the design of the experiment was limited by the availability of the antibodies in the lab at the time of the study.

### 3.3.3 Importance of the Biophysical Components in the Pyrenoid

Phenotypic observations revealed a high occurrence of multiple pyrenoids in some mutants, which could also explain the reduced size of pyrenoids. Usually every wild type cell possesses a single pyrenoid (Figure 3.19A), but multiple pyrenoids (Figure 3.19B) can sometimes be observed in asynchronous cultures due to the fact that

pyrenoid divisions are completed prior to cell division (Buffaloe, 1958; Warr, 1968; Goodenough, 1970). The cultures in this study were asynchronous so it was possible that some cells had completed pyrenoid division but yet to start cell division leading to observation of multiple pyrenoids. However, Tenenboim et al. (2014) suggested that delayed cytokinesis might cause multiple pyrenoids to occur within a cell, as demonstrated using knockdown mutants of VMP1 (Vacuole Membrane Protein 1) homologue. Some CCM defective mutants are shown to possess multiple pyrenoids, such as the *cia5* mutant of master regulator, CCM1/CIA5 (Fukuzawa et al., 2001) and the *epyc1* mutant of putative Rubisco linker protein (Mackinder et al., 2016). Interestingly, cells treated with colchicine (Walne, 1966, 1967), a reagent preventing microtubule assembly, and the multi-eyespot mutant, *mlt2* (Boyd et al., 2011) both exhibited multiple pyrenoid phenotypes suggesting the possibility that the presence of this phenotype is linked to microtubule or cytoskeleton development of the cell.

The thylakoid tubule network within the pyrenoid is proposed to be associated with the delivery of $CO_2$ into the pyrenoid (Figure 3.19A) (Raven, 1997). Most of the mutants in this study appeared to lack an observable pyrenoidal tubule network, with wide spaces between the pyrenoid and the thylakoid membranes, instead of possessing an arrangement of thylakoid membranes around the pyrenoid (Figure 3.19B). The lack of tubules may result in the smaller pyrenoid and also a dysfunctional CCM, as this would prevent effective delivery of $CO_2$, energy and substrates for carbon fixation (Engel et al., 2015). This is a novel CCM defect since the published mutants include those lacking pyrenoids and aberration in starch synthesis (Goodenough, 1970; Süss et al., 1995; Villarejo et al., 1996; Morita et al., 1997; Borkhsenious et al., 1998; He et al., 2003; Inwood et al., 2008; Genkov et al., 2010; Ma et al., 2011) but there has not been any published mutant which lacks pyrenoidal tubules, which occur even in the SSU mutants of Meyer et al 2012, (see Caspari et al 2017; Meyer et al 2017). Both Engel et al. (2015) and Meyer & Griffiths (2015) stressed on the importance of pyrenoidal tubule network in a functional CCM.

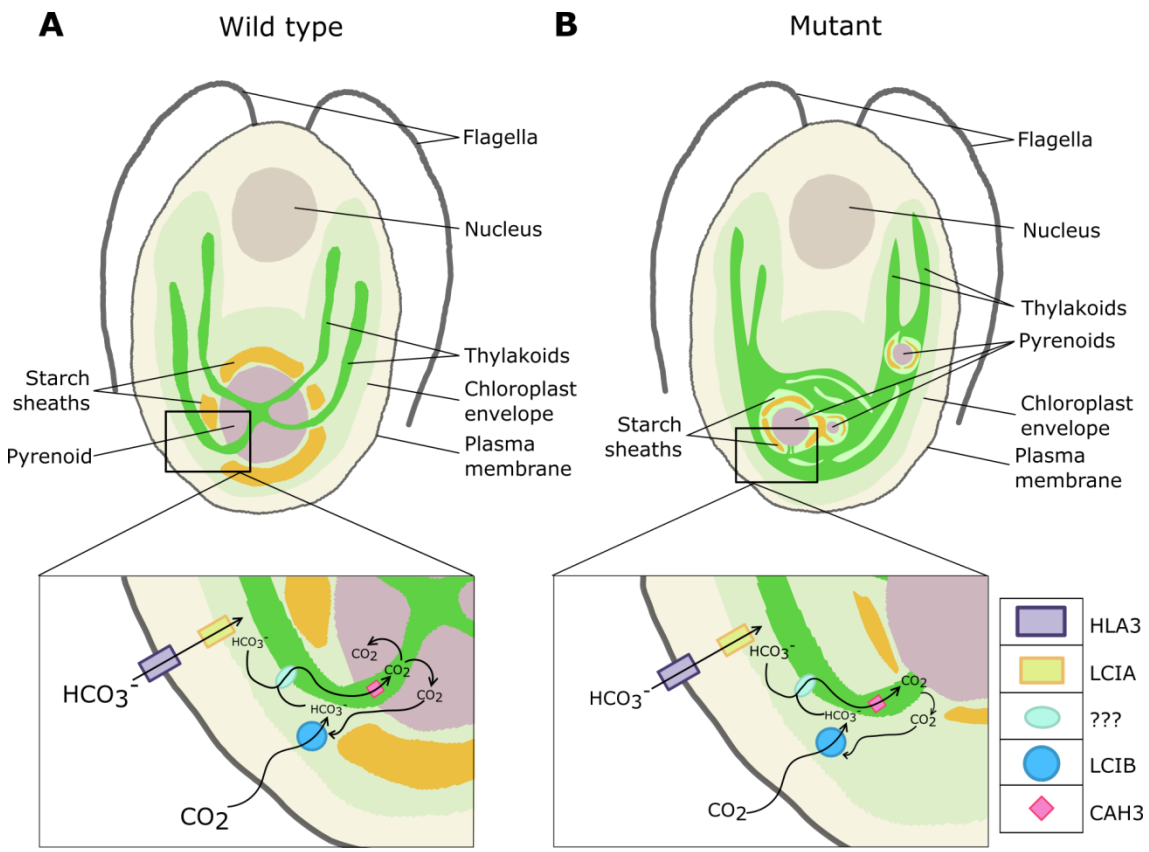



**Figure 3.19 (A) Current model of carbon concentrating mechanism of *C. reinhardtii* based on Mackinder (2017). (B) Hypothetical model of the pyrenoid of mutants that lack pyrenoidal tubules causing smaller pyrenoid size and defects of $CO_2$ and other components of CCM.**

Most of the mutants in this study were found with aberrances to the pyrenoidal starch plates (Figure 3.6, Figure 3.14 and Figure 3.19B). These significant differences in the pyrenoidal starch plates of the mutants compared to wild type could cause a lower packing force of Rubisco in the pyrenoid. Engel et al. (2015) suggested that the starch plates at the periphery of the pyrenoid help to increase the packing forces when Rubisco complexes are trafficked into the fixed volume which then drives the formation of the pyrenoid. It is also worth noting that work of Kuchitsu et al. (1988) and Geraghty & Spalding (1996) showed changes in starch localisation within the chloroplast when CCM activity is induced. Ramazanov et al. (1994) showed that pyrenoidal starch is formed during CCM induction leading to the hypothesis that the pyrenoidal starch sheath prevents diffusion of $CO_2$ from the pyrenoid. However, the work of Villarejo et al. (1996) using a starchless mutant, *BAJ-6*, disputed the direct role of pyrenoidal starch plates in the CCM. Although Thyssen et al. (2001) validated the suggestion that

pyrenoidal starch sheath might not be involved directly with CCM but is linked to the reduction of starch degrading enzymes, Mackinder et al. (2017) indicates that the starch sheath plays a role in structural formation of the pyrenoid where STA2 (starch synthase) and SBE3 (starch branching enzyme) are localised to the starch sheath.

In summary, the lack of a structural component of the pyrenoid, for example, the thylakoid tubule network, could be a strong cause of the dysfunctional CCM observed in most of the mutants in this study (Figure 3.19B). The morphological aberrances observed included smaller pyrenoids with a lack of a visible pyrenoidal tubule network and abnormal starch sheath, leading to to the next hypothesis that these defects are caused by a knockout or knockdown of known CCM genes controlling the formation of the pyrenoid.

### 3.3.4 Multiple Genes Are Involved in the Biogenesis of the Pyrenoid

The morphological and physiological characterisation of the mutants in this study suggested that a component is missing from the pyrenoid which led to defects in CCM. Based on these data, whole genome sequencing was performed to confirm which gene was knocked-out in the mutants. The hypothesis that the aberrant phenotypes of the mutants were caused by knockout of a known CCM gene was rejected. The results of the whole genome sequencing (Table 3.8) suggested knockouts in Cre11.g467712 (mutants A5, A7, A8, A11 and C8), which is a gene encoding an unknown protein with a starch-binding domain (SBD) and structural maintenance of chromosomes (SMC) N terminal domain (Chapter 5), two unannotated proteins, Cre02.g143387 (mutant C4) (Chapter 4) and Cre06.g259100 (mutant H10), Cre03.g199000 (mutant F8) coding for the phototropin gene and Cre14.g616600 (mutant G6) which is a FZO-like protein homologue. All mutants were paromomycin-resistant, thus confirming that they contained either a full, partial or multiple insertional cassettes (Zhang et al., 2014).

The phenotypic summary of mutant H10 suggested this CCM defective mutant did not possess a thylakoid tubule network within the pyrenoid with a clear area between the pyrenoid and starch plates (Figure 3.6H and Appendix 2), but retained other similarities to wild type in term of cell size, pyrenoid area and starch area. Interestingly, mutant H10 is a knockout of the gene Cre06.g259100, which is grouped as part of the pyrenoid matrix by Mackinder et al. (2017). This protein is also detected in the co-immunoprecipitation (CoIP) products of isolated chloroplasts and chloroplast (cp) nucleoids, and has been demonstrated to localise within the pyrenoid (Kobayashi et al.,

2015). ChlamyNET clustered this gene in the purple cluster, which is suggested to be involved in translation, RNA processing and lipid metabolism (Romero-Campero et al., 2016), while Hamel et al. (2017) and  suggested that this protein is a candidate for centriolar protein with a high percentage of coiled-coil proteins. This suggested that the protein is potentially involved in translation of a certain chloroplast protein required for the structural formation of the pyrenoid. The hypothesis can be tested by complementing the mutant with the paromomycin-resistant, YFP-tagged plasmid constructed by Kobayashi et al. (2015) with some modification to the antibiotic selection as the mutant is paromomycin-resistant.

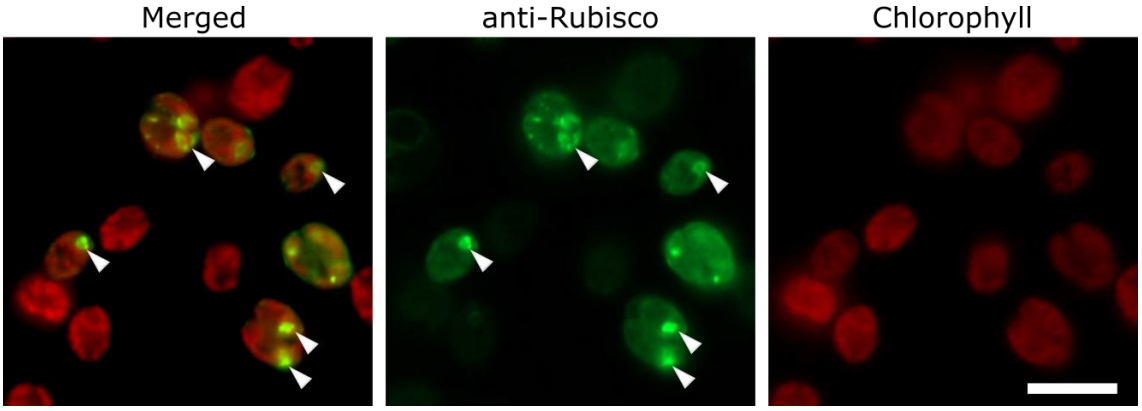



**Figure 3.20 Indirect immunofluorescence with anti-Rubisco antibody on mutant F8 (Caspari, 2015) showed dispersed anti-Rubisco antibody signal as well as wild-type-like signal.**

On the other hand, mutant F8 is a knockout of a known gene, Cre03.g199000 (*Phot*) which encodes for phototropin. There is only a copy the blue-light receptor *Phot* gene in *C. reinhardtii* which consists of two highly-conserved LOV (light, oxygen and voltage) domains in the N-terminal half and a Serine/Threonine kinase domain in the C-terminal half (Huang et al., 2002; Kasahara et al., 2002; Trippens et al., 2012). In higher plants, phototropins are involved in phototropism and also in the opening of stomatal pores activated by blue light (Christie, 2007). In *C. reinhardtii*, this gene is reported to play a role to complete the hatching process of the sexual life cycle (Huang and Beck, 2003) and in transcription of gene targets (Huang et al., 2002). The insertional cassette of mutant F8 was suggested to be in the kinase domain of the gene, which was demonstrated to regulate light-independent eye spot size (Trippens et al., 2012). Although no CCM defect was suggested based on the subsequent physiological analyses of this mutant, the indirect immunofluorescence screen using anti-Rubisco antibody (Caspari, 2015) had showed variable results (Figure 3.20) whereby some cells showed

dispersed signal indicating pyrenoidal defects while some displayed wild-type-like signal. In relation to CCM, blue light is reported to increase the activity of carbonic anhydrases (Dionisio, 1989; Dionisio-Sese et al., 1990; Borodin, 2008). It might be a possibility that F8 was a false positive of the initial CCM screen. Slower growth could occur as an effect of lower carbonic anhydrase activity caused by this mutation. In order to confirm whether the mutation in this region of *Phot* gene causes potential CCM defects, this mutant could be grown alongside with the corresponding wild type under different light and $CO_2$ conditions, first to compare the growth rate and then to detect the expression of carbonic anhydrases, for example, the blue-light dependent *Cah1* (Fujiwara et al., 1996).

With regards to mutant D4, the site of insertional cassette was matched to an intergenic region of Chromosome 6 between two genes with no functional annotation, Cre06.g278273 and Cre06.g278274. According to the *C. reinhardtii* genome in Phytozome 12, these two genes are homologues to various RNA-directed DNA polymerase/revertase with similarity of more than 80%, respectively (Goodstein et al., 2012). In general, most of the intergenic region of the genome is assumed to be non-functional (Nelson et al., 2004) but recent studies have shown that this region is also home to up to half of the DNase I hypersensitive sites (DHS) in humans (Thurman et al., 2012) and in crop plants (Burgess et al., 2017). DHSs are regions in the genome which are enriched by regulatory elements, such as transcriptional binding sites, promoters, enhancers, silencers, insulators and locus control regions (Thurman et al., 2012; Burgess et al., 2017). In *C. reinhardtii*, Winck et al. (2013) reported that the DHSs are normally found in the intergenic region near to transcription start sites (TSS) or transcription termination sites (TTS). This suggested the potential that mutant D4 might be a knockout of a regulatory element involved in the formation of thylakoid tubule network which had yet to be characterised. On the other hand, the phenotypic aberrances observed in this mutant could be the effect of other mutations caused by DNA rearrangement during the process of transformation (Jinkerson and Jonikas, 2015). Confirmation if this knockout causes regulatory defect on pyrenoidal tubule network formation could be performed using DNase-Seq, whereby comparison of DHSs between wild type and mutant could infer whether the mutant possesses a DHS.

In the duration of this study, only physiological analysis and whole genome sequencing were performed on mutant G6, while further analyses are currently being conducted by Dr. Moritz Meyer in Princeton University. Both spot test and oxygen evolution

suggested that the knockout effect on the CCM of the mutant was not significant but there were morphological aberrances compared to the wild type (personal communication, Dr. Moritz Meyer). This mutant was suggested to contain a knockout on the gene Cre14.g616600 encoding for FZL protein which has a dynamin family domain and a thiamine monophosphate synthase (TENI) domain. In general, dynamin is a family of GTPases involved in eukaryotic endocytosis (Urrutia et al., 1997). The dynamin proteins discovered in plants are usually involved in chloroplast biogenesis and thylakoid organisation (Park et al., 1998; Gao et al., 2006). This plant FZO-like GTPase found in *Arabidopsis thaliana* is also reported to be conserved across green plant and algal lineage, including *Chlamydomonas* (Gao et al., 2006). A cytoskeletal dynamin GTPase family, the FtsZ proteins, are reported to be a key component in chloroplast division, where the FtsZ protein complexes will form a Z-ring in the centre of the plastid and eventually separate the daughter chloroplasts (TerBush et al., 2013; Terbush et al., 2017). The cell and pyrenoid sizes of mutant G6 were variable; some were similar to wild type cells while some were relatively smaller (personal communication, Dr. Moritz Meyer) thus leading to the hypothesis that the knockout of the gene causes asymmetrical organelle division in *C. reinhardtii*. Further work which could be done to validate this hypothesis includes complementing the knockout mutant with a tagged-FZL protein to observe the localisation of the gene and restoration of the phenotype.

Mutants A5, A7, A8, A11 and C8 were knockout candidates of the gene, Cre11.g467712, which is suggested to be part of the structural maintenance chromosome (SMC) family, with a putative starch binding domain. With the presence of 5 mutants possessing knockouts of the same gene, the gene was subjected to further characterisation (Chapter 5) using mutant C8. Further discussion about mutant C4 which contained knockout of the gene Cre02.g143387, with no functional annotation will be presented in Chapter 4, where this mutant was used for the proof of concept experiment using Golden Gate Cloning system.

## 3.4 Conclusions and Future Work

The CCM defective mutants in this study which also possess pyrenoid defects could be an important resource to the community to understand the biogenesis of the pyrenoid in *C. reinhardtii*. There are at least three candidate genes which are suggested by whole genome sequencing to be linked to the structural formation of the pyrenoid, including a

structural maintenance chromosome (SMC) family member (Cre11.g467712), FZL (FZO-like) protein (Cre14.g616600), an intergenic region and two unannotated proteins (Cre02.g143387 and Cre06.g259100). Coupled with the physiological and morphological analyses on the mutants, this has led to the hypothesis that multiple genes are controlling the structural formation of the pyrenoid. In order to follow up with this hypothesis, complementation of the genes upon validation of knockout mutants could be performed to confirm the genes are indeed responsible for the biogenesis of pyrenoidal tubules. Other potential experiments to further understand the physiological characteristics of the mutants include carbon isotope discrimination and TIC pool size measurement to measure the carbon utilisation of the mutants.

# 4 MODULAR CLONING IN *CHLAMYDOMONAS REINHARDTII*

## 4.1 Introduction

### 4.1.1 The Use of *Chlamydomonas reinhardtii* in Biotechnology

Microalgae have a great potential for biotechnological applications. Besides having a wide variation in terms of cellular architecture and biosynthetic capacity, most algal species are unicellular, photosynthetic and have rapid cell cycle, enabling microalgae to be more productive per unit land area than any plant-based system (Scaife et al., 2015). In terms of industrial biotechnology, photosynthesis and the carbon assimilation pathways in microalgae provide a more sustainable avenue for the production of high value compounds. Commercially, microalgae have been used to produce pigments such as astaxanthin (Han et al., 2013; Shah et al., 2016) and beta-carotene (Borowitzka, 2013; Wichuk et al., 2014). Some microalgae are also used for animal feedstocks (Becker, 2013; Belay, 2013; Liu and Hu, 2013) and biofuel production (Wen and Johnson, 2009; Zittelli et al., 2013; Scranton et al., 2015).

With the development of molecular biology tools for in algal research, delivery of foreign DNA into algal cells has often been done using particle bombardment,

commonly known as biolistic transformation (Boynton et al., 1988; Sanford, 1990; Remacle et al., 2006) or electroporation (Shimogawara et al., 1998; Johnson et al., 2010; Rasala et al., 2012; Kobayashi et al., 2015). Biolistic transformation involves the introduction of DNA-coated small gold or tungsten particles into the cells using accelerated pressurised gas (Boynton et al., 1988) while electroporation uses electric current to perforate the cell membrane temporarily. Shimogawara et al. (1998) reported the transformation efficiency of *C.reinhardtii* was about 2 x $10^5$ transformants per microgram of DNA using electroporation. Another popular transformation method for *C. reinhardtii* is cell agitation in the presence of glass beads or silicon carbide whiskers, though these approaches require minimal equipment, they also result in lower transformation rates (Kindle, 1990; Hall et al., 1993; Tam and Lefebvre, 1993; Shimogawara et al., 1998). In order to improve glass-bead transformation in walled cells of *C. reinhardtii*, autolysin was used to remove cell wall prior to transformation (Neupert et al., 2012). In addition Barrera and Mayfield (2013) remarked on the existence of the *Agrobacterium tumefaciens*-mediated transformation of algal cells (Kathiresan et al., 2009; Cheng et al., 2012; Rathod et al., 2013). Though this method has been used extensively in higher plants, this method is not applied extensively in algal biotechnology research (Barrera and Mayfield, 2013). The first successful *Agrobacterium*-mediated transformation was demonstrated on *Haematococcus pluvialis* Flot., into which Kathiresan et al. (2009) introduced *hpt* (hygromycin phosphotransferase) gene with GFP (green fluorescence protein) and GUS (β-glucuronidase) as reporter genes in a single plasmid.

### 4.1.2 Synthetic Biology in *Chlamydomonas reinhardtii*

The progress and potential benefits from genetic engineering in terms of genome editing and synthetic biology have led to a common hypothesis among scientists working in the area of food security that crop plants can be engineered to improve photosynthetic efficiency and yield potential to meet the global food demand (Long et al., 2015). In their review on the applications of synthetic biology, Khalil and Collins (2010) explained that synthetic biology aims to engineer or "wire" biological circuitry to manifest logical forms of cellular control. The progress of recombinant DNA technologies simplifies the construction and assembly of biomolecular parts, while the ideas of modularisation, rationalisation and modelling have led to the production of a suite of genetic devices and biological modules (Khalil and Collins, 2010). A well-known example of the output of incorporating synthetic biology in plant research is the

development of the modular cloning tool box for plants (Engler et al., 2014). This originated from the previous research on a one-step, one-pot, precision cloning method by Engler et al. (2008), now known as Golden Gate Cloning. The power of Golden Gate cloning comes from the standardisation of genetic parts designed for the one-step, one-pot approach to assemble the recombinant plasmids and assembly of multi-gene constructs, with the 96 standard parts from the kit available on Addgene (Engler et al., 2008, 2014; Weber et al., 2011).

The standard genetic parts developed for modular cloning in plants need to be tailored for specific species in order to increase the efficiency of the expression of transgenes. For example, in *Chlamydomonas*, due to the fact that *C. reinhardtii* has different GC content for nuclear genome (62%), mitochondria (48%) and chloroplast (34%), it is necessary to perform codon optimisation because the ability for the transgenes to be expressed and successfully generate product are both related to the codon bias being shared with the host microalgae (Rasala et al., 2013). The codon usage database, CUTG (Nakamura et al., 2000), is therefore a useful tool, as it contains tables on codon usage generated from GenBank (http://www.kazusa.or.jp/codon/). Therefore, the synthetic biology concept of Design-Build-Test-Learn (Khalil and Collins, 2010; Scaife and Smith, 2016) is crucial in developing and domesticating the different parts for modular cloning in *C. reinhardtii* (Figure 4.1).

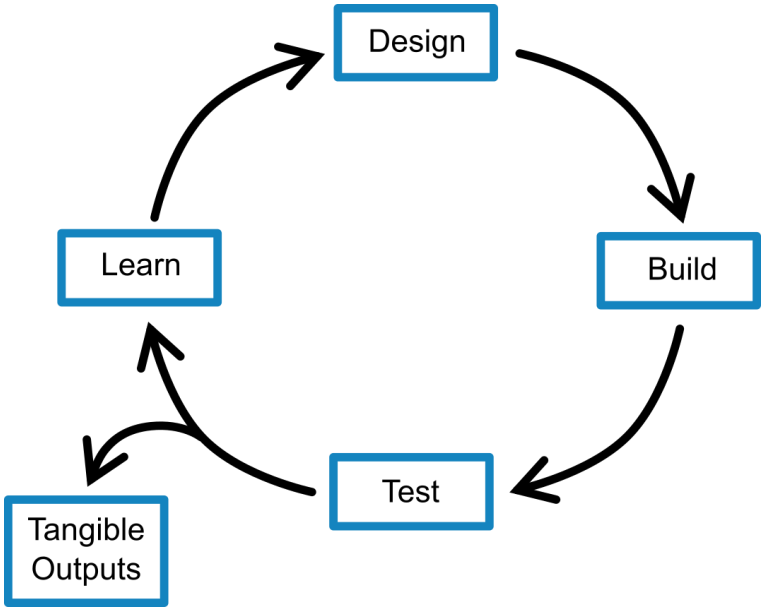



**Figure 4.1 The basis of synthetic biology is formed by iterative cycles of a Design-Build-Test-Learn process carried out using standardised workflows which could lead to tangible outputs.**

### 4.1.3 Development of Parts in Synthetic Biology

The basic parts in the 96-part MoClo Kit for Plants consist of: promoters (Pro), 5' untranslated sequences (5U), coding sequences (CDS1), 3' untranslated sequences (3U) and transcription terminator (Ter). When DNA is introduced into the cells, reporter gene assays can be used to optimise transformation, differentiate between transformed and untransformed cells, and analyse the strength and specificity of regulatory elements, e.g. promoters (Matthews et al., 1995). The two widely used and complementary reporter gene assays are the green fluorescent protein (GFP) (Allard, 2008), which is developed from aequorin, a bioluminescence protein isolated from jelly fish (Bronstein et al., 1994), and the β-glucuronidase gene fusion system (GUS), which is an enzymatic assay (Jefferson, 1989). In *Chlamydomonas* research, the widely-used GFP has been codon-optimised for nuclear transformation (Fuhrmann et al., 1999) and chloroplast transformation (Franklin et al., 2002). The combination of the PsaD promoter and PsaD terminator for transformation in *C. reinhardtii* was reported to improve gene expression significantly (Kumar et al., 2013). The years of research into the various promoters, UTRs, terminators and antibiotic resistance in *C. reinhardtii* provided the knowledge to produce a *Chlamydomonas*-specific MoClo Tool Kit (publication in preparation) (Dr. Payam Mehrshahi and Dr. Francisco Navarro, University of Cambridge, personal communication).

### 4.1.4 Selection of Mutant with CCM Defect for Golden Gate Cloning

The set of mutants characterised in Chapter 3 could be classified into two main groups: CCM defective, and non-CCM defective. Physiological and morphological analysis of mutant F8 suggested the knockout did not affect the CCM and the mutant retained a pyrenoid structure which resembled the pyrenoid in wild type cells. The second group of mutants with CCM defects also possessed pyrenoidal defects. WGS classified the mutants into knockouts of Cre11.g467712 (A5, A7, A8, A11 and C8), Cre06.259100 (H10), a gene with unknown function, Cre02.g143387 (C4) and an intergenic region with no known gene identified (D4). Localisation of Cre11.g467712 will be discussed in the next chapter while pyrenoid-localisation of Cre06.259100 is demonstrated by Kobayashi et al. (2015). In the absence of knowledge about the intergenic region where the cassette in mutant D4 was located, mutant C4 was selected as the best option for further characterisation.

Aside from possessing a defective CCM, mutant C4 also had morphological aberrances, e.g. occurrence of multiple pyrenoids, absence of visible thylakoid tubule network in the pyrenoid and abnormal thylakoid stacking. With regards to the molecular background of the gene (Cre02.g143387), no known function could be ascribed from the *Chlamydomonas reinhardtii* v5.5 genome in Phytozome12 (Goodstein et al., 2012). CircadiaNET suggested no orthologs with similar circadian expression patterns, but there are potential orthologs with different circadian expression patterns in *Arabidopsis thaliana* (AT1G68360, AT4G02715, AT3G06890, AT1G55220, AT3G55000) and *Ostreococcus* (ostta15g00770) (de los Reyes et al., 2017). This gene met the basic requirement of level zero parts as the relatively short sequence is absent of recognition sequence for *Bsa*I, *Bpi*I and *Bsm*BI within the internal sequence.

In this chapter, modules from the Chlamy MoClo project were provided by Dr. Francisco Navarro (University of Cambridge) to be used with a newly-constructed level zero CDS module to test the hypothesis that Cre02.g143387 is a gene involved in the formation of a functional pyrenoid in order to maintain the carbon concentrating mechanism in *C. reinhardtii* cells. The physiological analysis of the transformants indicated the lack of CCM phenotype restoration in the mutant, hence the chapter acted as a methodological exercise of genetic complementation assay in *C. reinhardtii* while the discussion would focus on the analysis of potential confounding factors leading to the absence of phenotype restoration in the mutant.

## 4.2 Results

### 4.2.1 Morphological Observation of the Mutant

As mentioned in Chapter 3, physiological assessment using oxygen-evolution assay supported the notion that this mutant was CCM defective. Under both high and low $CO_2$ conditions, the photosynthetic $K_{0.5}$ for Ci of the mutant was ~270 μM Ci compared to the $K_{0.5}$ of ~50 μM Ci for wild type (Table 3.2). This showed that the mutant possessed a lower affinity for inorganic carbon which led to a higher $K_{0.5}$ value. The occurrence of multiple pyrenoids and absence of thylakoid tubule network in the pyrenoid were observed. The percentage of Rubisco localised in the pyrenoid of mutant C4 was significantly lower at 83.5% compared to the wild type at 93.8% (p-value = $4.03 \times 10^{-5}$) with confirmation that the pyrenoid area was also smaller than the corresponding wild type (Figure 3.8).

In order to test the hypothesis that the lack of thylakoid tubule network in the mutant causes multiple pyrenoids in the chloroplast of the mutant, a time-course experiment was designed to observe wild type and mutant cells grown under high $CO_2$, then 1h and 24h after acclimation to low $CO_2$ (ambient air) conditions. Interestingly, wild type cells grown under different conditions retained a single electron-dense pyrenoid structure with the presence of a thylakoid tubule network within the pyrenoid under all conditions (Figure 4.2A-F). The starch granules around the pyrenoid at high $CO_2$ formed a starch sheath surrounding the pyrenoid when cells were acclimated at low $CO_2$ (Figure 4.2D-F). However, the mutant C4 cells under all conditions possessed multiple pyrenoids lacking a visible thylakoid tubule network with a thin starch sheath that did not fully surround the pyrenoids (Figure 4.2G-O).

Whole genome sequencing of mutant C4 identified multiple sites where the paromomycin cassette might have inserted, with the highest number of hits mapping to a location within Cre02.g143387 (**Table 4.1**). Further analysis on the raw WGS data confirmed the multiple potential insertion sites of paromomycin cassette (Dr. Steven Burgess, personal communication). Based on Phytozome v5.5, this gene possesses three introns and four exons. The size of the genomic sequence of Cre02.g143387 is about 2,672 bp including the 5' UTR and 3'UTR (about 1.7 kb without the UTRs) while the protein is a 151 amino-acid peptide sequence. The Compute pI/Mw tool on ExPASy server estimated the size of the protein to be ~14 kDa (Gasteiger et al., 2005). Based on the WGS data, the insertional mutagenesis cassette was located in intron 1 of the gene (Figure 4.3). Secondary structure, as predicted by Phyre2 (Kelley et al., 2015), suggested with a relatively high confidence level that this protein was comprised of 50% disordered regions, 35% alpha helices and 8% beta strands (Figure 4.4). Prediction using NetPhos 3.1a server identified potential serine and threonine phosphorylation sites in the disordered region of the protein, and a tyrosine phosphorylation site in the ordered region (Figure 4.5), similar to the observation made by Kurotani and Sakurai (2015). Phyre2 suggested a protein model for the gene, but it was relatively inaccurate as it was based on *ab initio* modelling in the absence of similar structures being available (Figure 4.6). The protein was predicted to be targeted to the nucleus using DeepLoc-1.0.

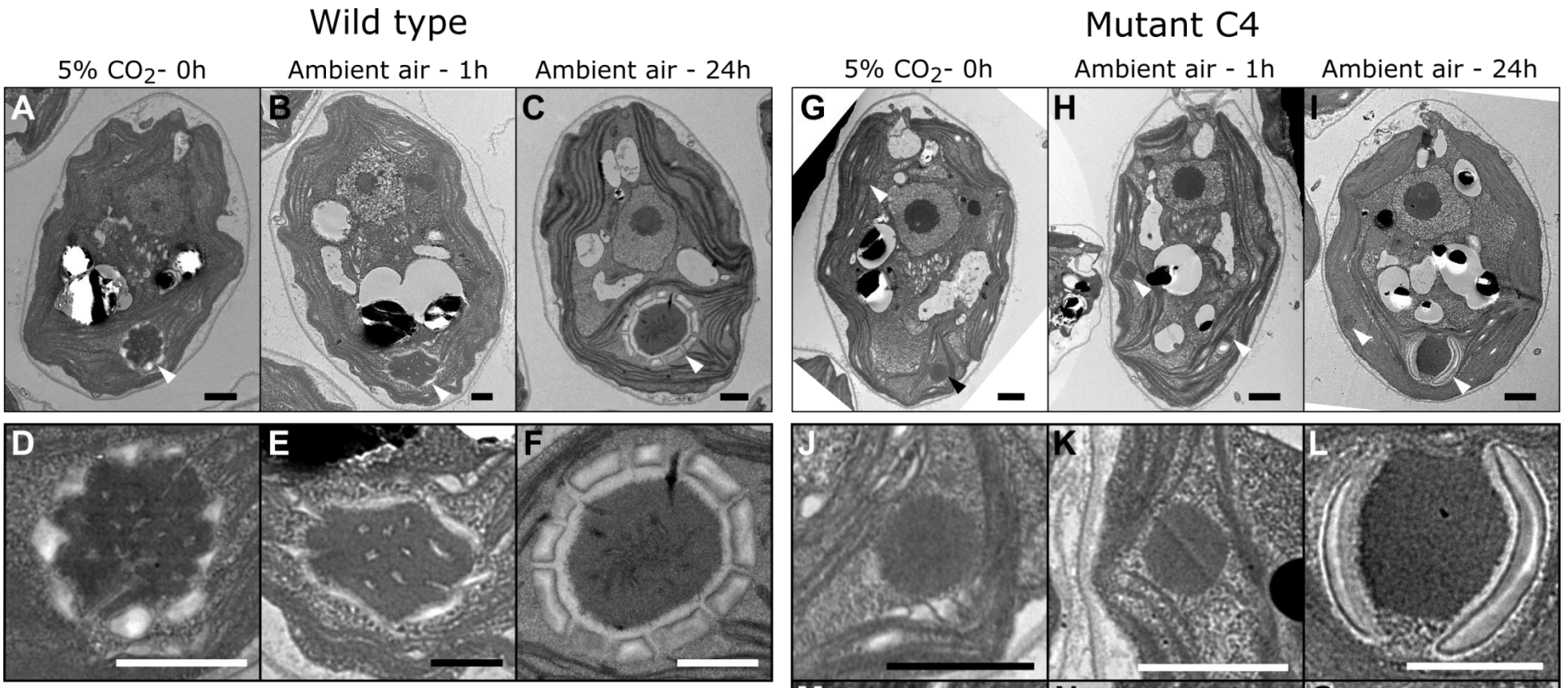



**Figure 4.2 Electron micrographs of wild type and mutant C4 whole cells and zoomed out images of the pyrenoids grown under different conditions to observe pyrenoid phenotypes. (A-F) WT whole cells and pyrenoids grown under high $CO_2$ (A, D), 1 h (B, E) and 24 h (C, F) after acclimation to low $CO_2$. (G-O) Mutant whole cells and the pyrenoids grown under high $CO_2$ (G, J, M), 1 h (H, K, N) and 24 h (I, L, O) after acclimation to low $CO_2$. Scale bars 500 nm.**

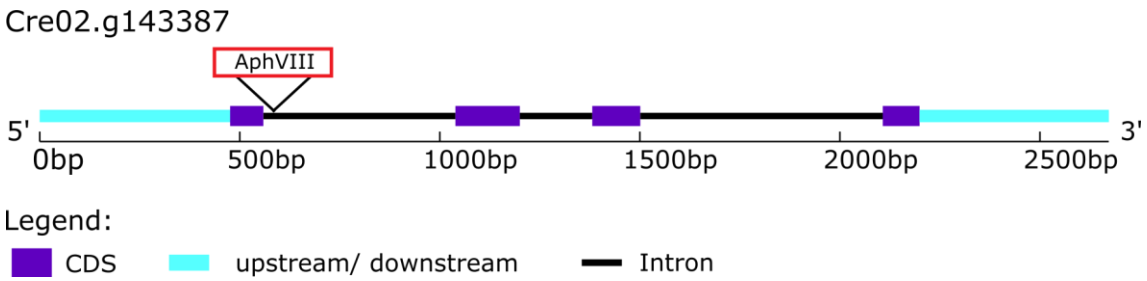



**Figure 4.3 Genomic structure of Cre02.g143387 showing the position of *AphVIII* insertion (red box).**

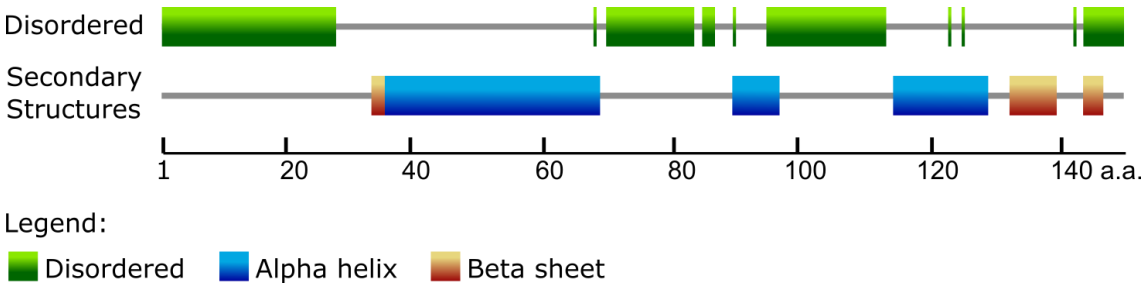



**Figure 4.4 Secondary structure prediction of Cre02.g143387 using Phyre2 suggested that this protein was comprised of 50% disordered protein, 35% alpha helices and 8% of beta strands.**

**Table 4.1 Whole genome sequencing showed potential locations of the insertional mutagenesis cassette in the genome of mutant C4.**

| Chromosome | Location | Gene | Gene Location | Description |
|---|---|---|---|---|
| chromosome_1 | 6613293 | Cre01.g047350 | chromosome_1:6611336..6617929 forward | Unannotated |
| chromosome_11 | 1219982 | Cre11.g467697 | chromosome_11:1219972..1221943 forward | Unannotated |
| chromosome_11 | 1220028 | | | |
| chromosome_11 | 1220028 | | | |
| chromosome_16 | 5993959 | Cre16.g676600 | chromosome_16:5991432..5994829 forward | Unannotated |
| chromosome_16 | 5994078 | | | |
| chromosome_16 | 5994085 | | | |
| chromosome_16 | 5994085 | | | |
| chromosome_16 | 5994085 | | | |
| chromosome_2 | 9115221 | Cre02.g143387 | chromosome_2:9114756..9117427 | Unannotated |
| chromosome_2 | 9115221 | | | |
| chromosome_2 | 9115263 | | | |
| chromosome_2 | 9115427 | | | |
| chromosome_2 | 9115427 | | | |
| chromosome_2 | 9115427 | | | |
| chromosome_2 | 9115438 | | | |
| chromosome_2 | 9115438 | | | |
| chromosome_2 | 9115438 | | | |
| chromosome_5 | 2978173 | Cre05.g238332 | chromosome_5:2978242..2979235 reverse | PSAD (belong to the cassette) |
| chromosome_5 | 2987804 | NA | | |
| chromosome_6 | 8668788 | Cre06.g309550 | chromosome_6:8662125..8670223 reverse | Unannotated |
| mitochondrial | 9665 | | | |

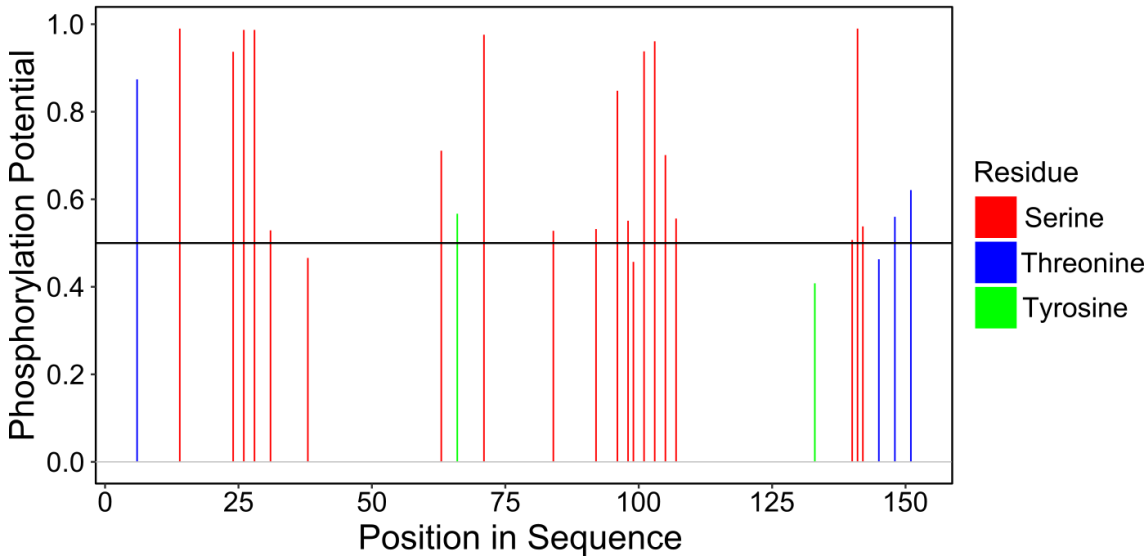



**Figure 4.5 NetPhos 3.1a predicted multiple phosphorylation sites in the protein encoded by the gene Cre02.g143387.**

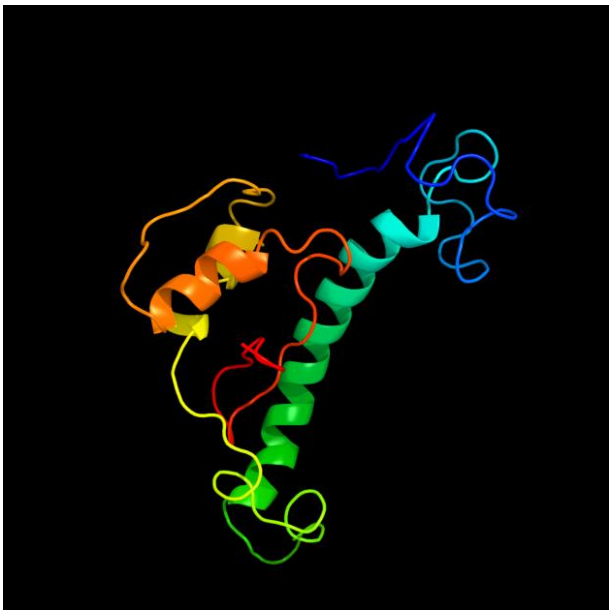


**Figure 4.6 The protein model for Cre02.g143387 was predicted by Phyre2 using *ab initio* modelling.**

### 4.2.2 Construction of Transformation Vector using Golden Gate Cloning

In order to test the hypothesis that Cre02.g143387 was responsible for pyrenoid defects in C4, the Golden Gate cloning system, which is a DNA assembly system that utilises Type IIS restriction enzymes, was used to create a recombinant (Engler et al., 2008).

With regard to the steps taken to assemble the final pFPC2-1C4-HygR plasmid for transformation, first, the coding sequence for Cre02.g143387 (CDSns) was amplified

from genomic DNA of wild type cells and then inserted into a level 0 acceptor module (Figure 4.7).

The level 0 CDSns module for this study was constructed based on the requirements suggested in the MoClo Tool Kit for Plants (Engler et al., 2014). The *Bpi*I sites required to flank the coding sequence were created by including the sequence to the primers designed to amplify the gene of interest from genomic DNA. Since no recognition sequences for *Bsa*I, *Bpi*I and *Bsm*B1 were found in the coding sequence of Cre02.g143387, the amplified and purified fragment of about 1.7 kb (Figure 4.8) was cloned directly into pAGM1827 forming the pCCN3 level 0 plasmid (Figure 4.9).

Next, all the domesticated level 0 modules containing sequences for Pro+5U (PsaD promoter and 5' untranslated region), C-YFP (C-terminus YFP), 3U+Ter (PsaD 3' untranslated region and terminator) and the CDSns (coding sequence for Cre02.g143387) were combined with a level 1 acceptor module, pICH47811, from the Plant MoClo Kit which was carbenicillin resistant to form a level 1 transcriptional unit. The validated transcriptional unit, TU_1C4, contained the gene of interest fused with C-terminal YFP. Finally, pCCN3, TU_4 (transcriptional unit containing hygromycin resistance gene), a domesticated level 2 acceptor plasmid, pFPC2, and an end-linker, pICH41744, were ligated to form a level 2 plasmid, pFPC2-1C4-HygR, which would be used for transformation.

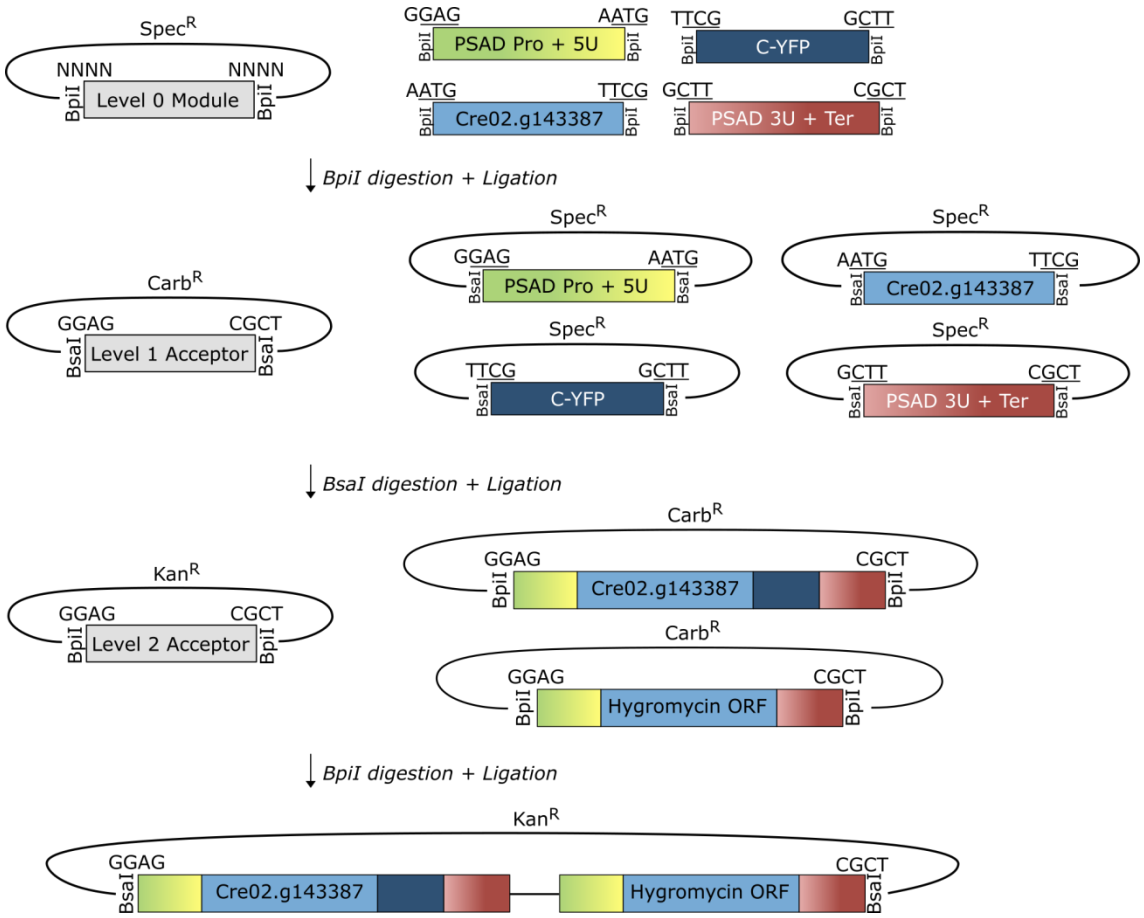



**Figure 4.7 Illustration showing the steps to transform the gene Cre02.g143387 into mutant C4 using Golden Gate Cloning system**

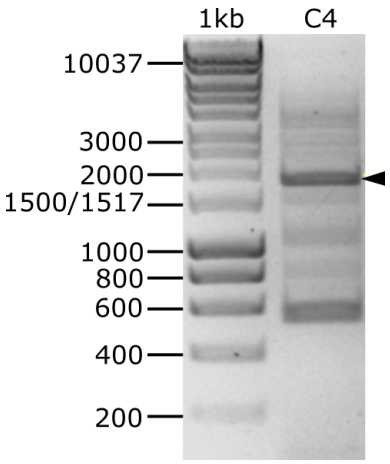



**Figure 4.8 Cre02.g143387 genomic DNA was amplified from wild type *C. reinhardtii* cells using primer pair CC021 and CC022. The size of the amplicon was estimated to be about 1.7kb as indicated in the gel with a black arrow.**

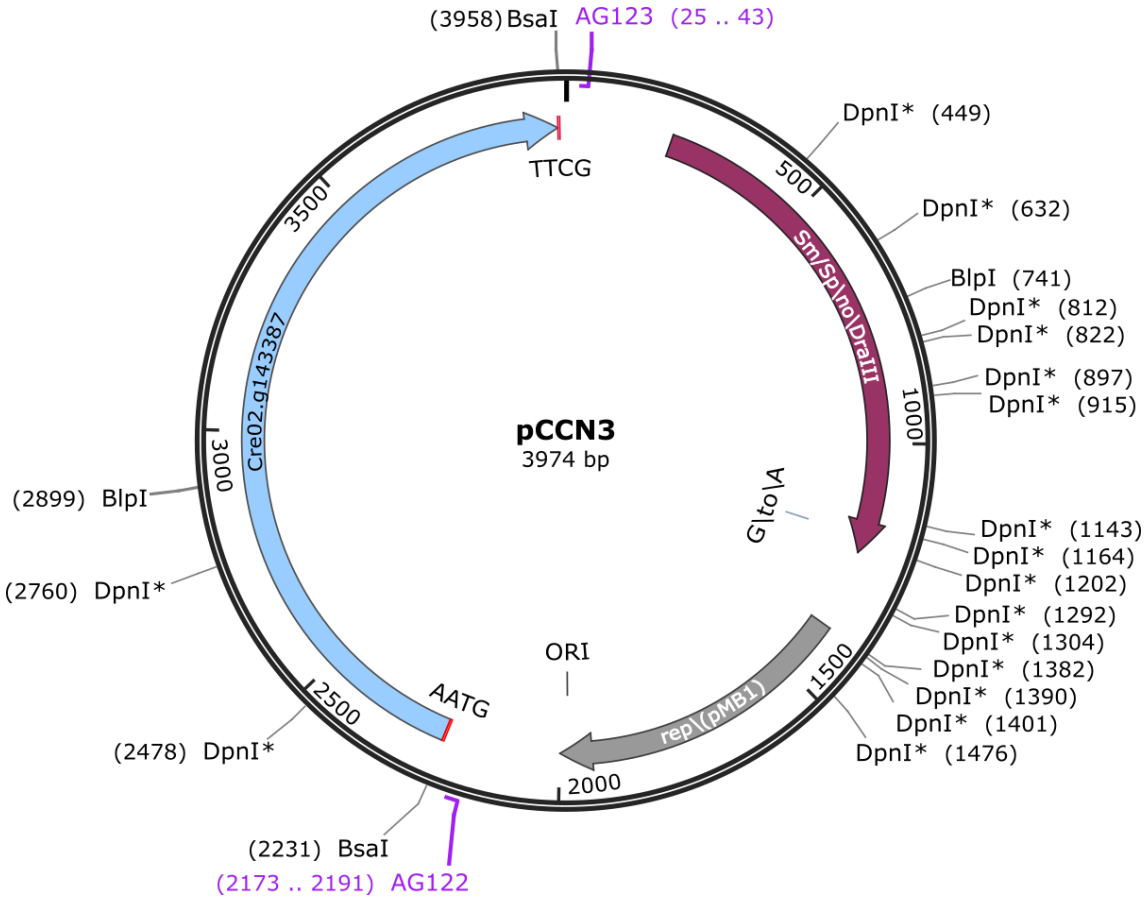



**Figure 4.9 Level 0 module containing the coding sequence of gene Cre02.g143387 indicating the location of the cutting sites of *Bsa*I, *Blp*I and *Dpn*I in the module.**

With the knowledge of the predicted locations of the cutting sites for commonly-used restriction enzymes, *in silico* restriction enzyme digestion of *Bsa*I, *Blp*I and *Dpn*I suggested distinct patterns for undigested pAGM1827, digested pAGM1827 with no insert and pAGM1827 with insert (Figure 4.10A). The experimental results showed that two out of the three colonies which grew on selective LB agar containing spectinomycin possessed the correct insert (Figure 4.10B). *Bsa*I was the cutting site flanking both ends of the CDSns fragment, and the level 0 vector, pCCN3, when digested with the enzyme resulted in two fragments of about 2.2kb and 1.7kb. *Blp*I enzyme digestion of pCCN3 yielded two bands of about 2.1kb and 1.8kb. Although *in silico* digestion using *Dpn*I was predicted to result up to 5 different bands, the two bands of higher molecular weight would clearly differentiate the module with and without the gene insert, with the pCCN3 digestion yielding a band at 1.6 kb and 1kb, respectively.

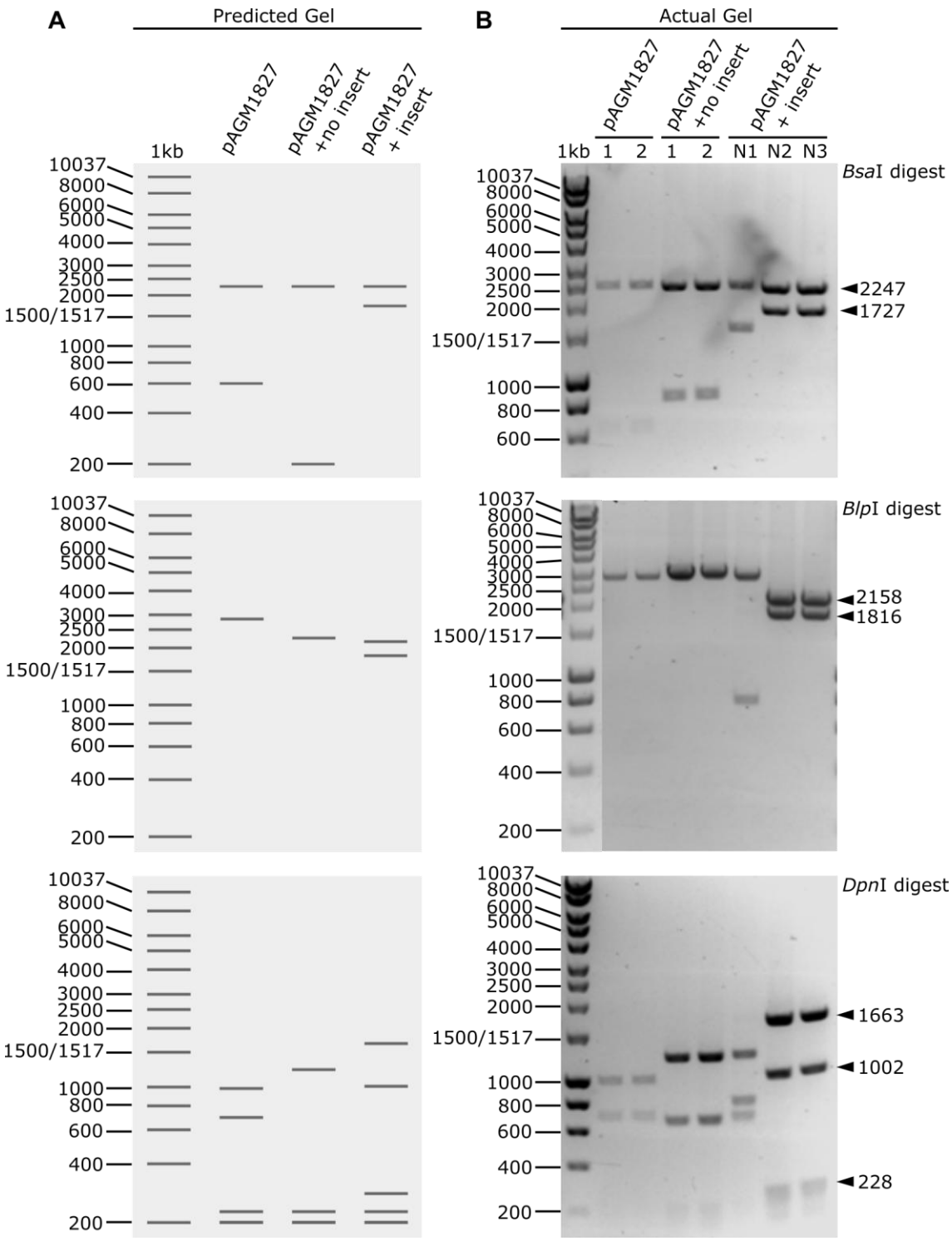



**Figure 4.10 (A) *In silico* restriction enzyme digestion with *Bsa*I, *Blp*I and *Dpn*I on Benchling and (B) the actual gel of the restriction enzyme digestions validated the presence of the CDS insert in the level 0 module.**

The transcriptional unit, TU_1C4, carrying the gene of interest was assembled from a set of compatible sequenced level 0 modules, which included pFJN43 (PsaD Pro+5U), pCCN3, pFJN70 (C-YFP tag) and pFJN45 (PsaD 3U+Ter) using the enzyme *Bsa*I to release the parts, followed by ligation to a level 1 module which conferred ampicillin resistance (Figure 4.11). The module was transformed into DH5α *E. coli* competent cells to be screened on selective LB agar containing 100 µg/ml (w/v) carbenicillin and four positive colonies were picked for downstream experiments. Orientation-specific primers were designed to validate that the module was assembled in the correct order: Pro+5U_CDSns_CT_3U+Ter. The successful amplification of the colonies using these two primer pairs indicated that the orientation of parts assembled in TU_1C4 was in the correct order and therefore the module was ready for the next step (Figure 4.12).

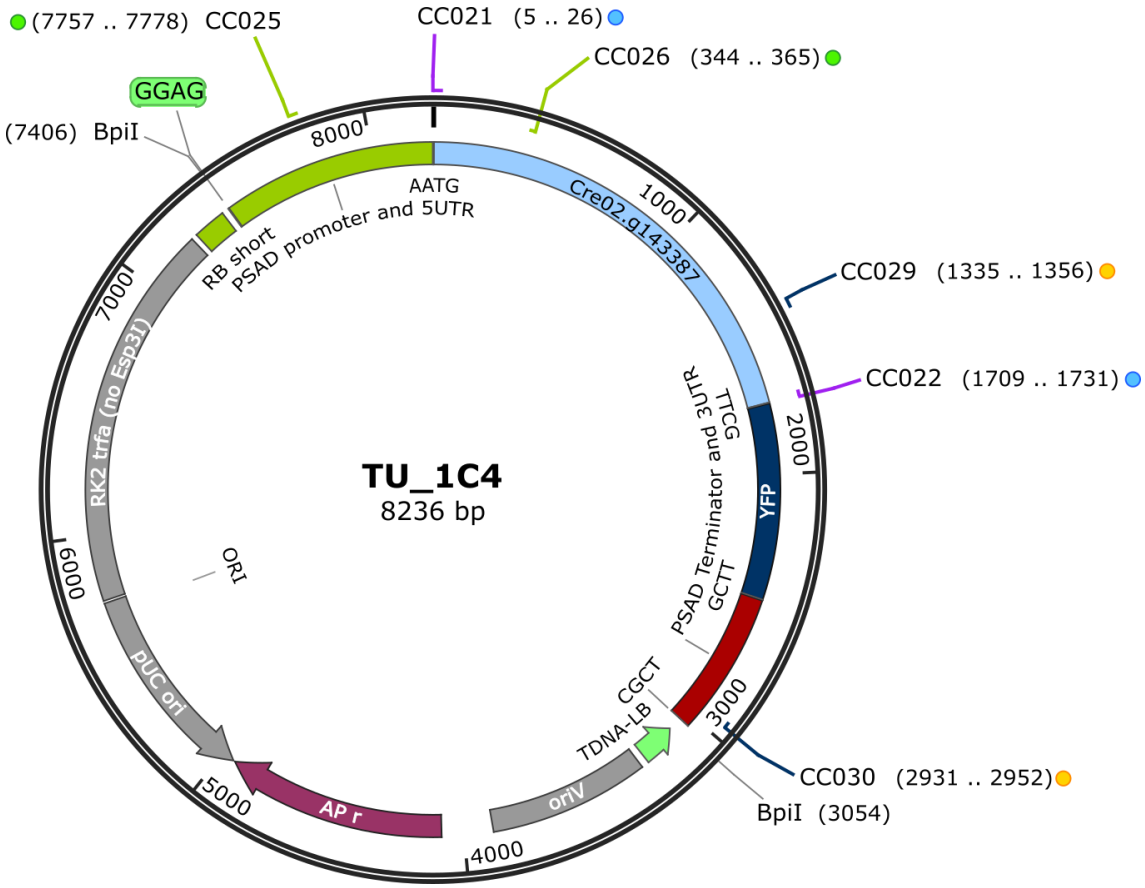



**Figure 4.11 Level 1 transcriptional unit containing Cre02.g143387 gene and a C-terminus YFP tag.**

The level 2 module used in complementation was a multigene construct which contained transcription units for the gene Cre02.g143387 and the gene encoding for hygromycin resistance. The level 2 destination vector, pFPC2, conferred resistance to kanamycin and encoded a red colour selectable marker. By utilising *Bpi*I to release the transcription units and the end-linkers to be ligated to pFPC2, the level 2 module was assembled (Figure 4.13). The resulting colonies upon transformation appeared white instead of orange/red, and were validated using the orientation specific primer pairs mentioned previously. Out of the 3 white colonies selected, only 1 colony, N3212, was positive in both PCR reactions (Figure 4.14).

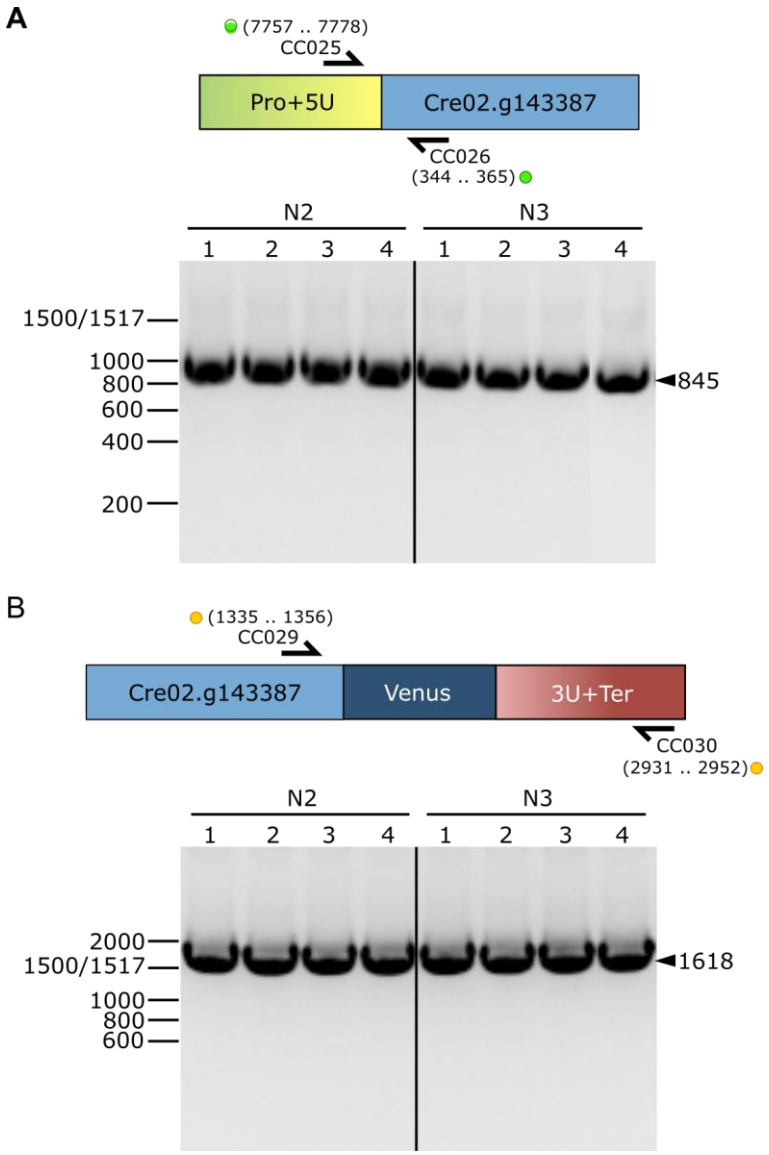



**Figure 4.12 PCR validation of the level 1 transcriptional unit using orientation-specific primer pairs: (A) CC025+CC026 amplified the region between the 5'UTR and the gene of interest resulting a fragment of about 845bp and (B) CC029+CC030 amplified the fragment about 1.6 kb between the CDS and 3'UTR/Ter to confirm that the orientation and arrangement of the CDS, C-terminal YFP and 3'UTR/Ter parts were correct.**

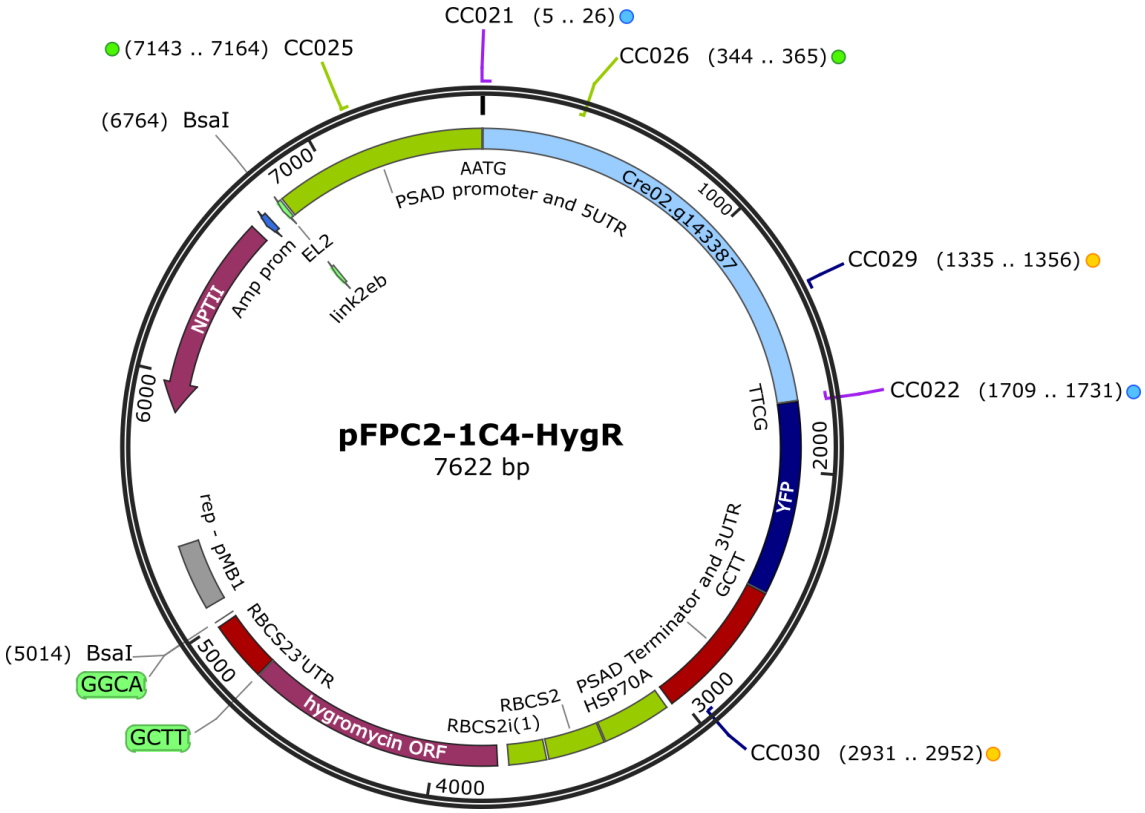



**Figure 4.13 The assembled level 2 plasmid contained the gene of interest with a YFP tag and hygromycin resistant gene (NPTII).**

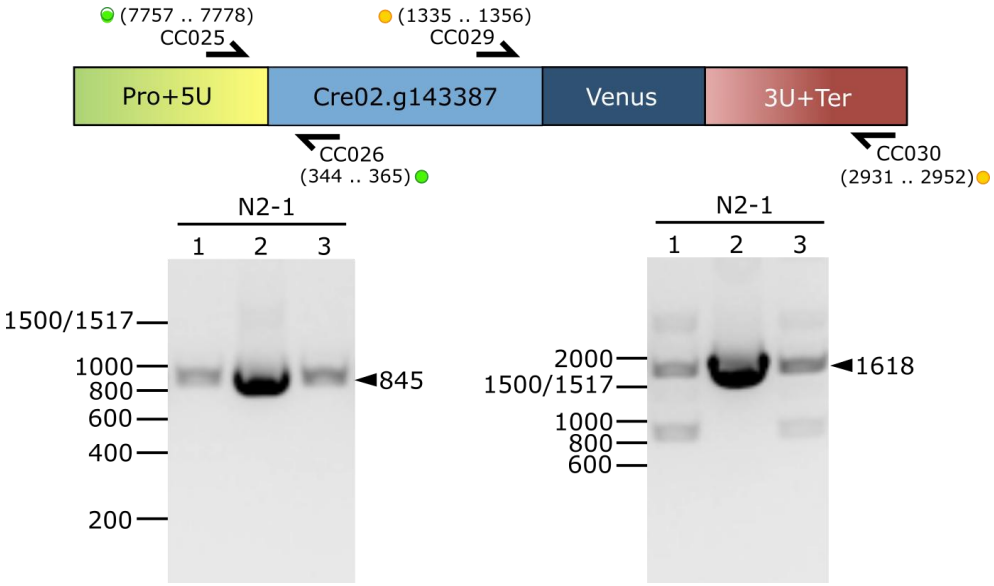



**Figure 4.14 PCR validation was performed using primer pairs CC025+CC026 and CC029+CC030 to confirm that level 2 plasmids assembled using level 1 transcriptional unit contained the gene of interest in the correct orientation.**

**Figure 4.15 Sequence alignment using Clustal Omega of the Sanger sequencing results to confirm the full sequence of the gene of interest in the assembled level 2 plasmid. The exons of the gene were underlined on the reference sequence, Cre02.g143387. Two mutations are found on the introns of the insert (highlighted in yellow): a nucleotide substitution (C>T) at position 420 and an insertion of a cytosine (C) at position 1190. The stop codon TAA was detected at the end of the insert (red letters, highlighted in yellow)**

```
Cre02.g143387      ATGCGCCCGAGGCTGACAGGCGCGGCCGCTGGCGCAGACAGCGACGGTGAAGGGGCAGCG 60
seq_results        ATGCGCCCGAGGCTGACAGGCGCGGCCGCTGGCGCAGACAGCGACGGTGAAGGGGCAGCG 60
                   ************************************************************

Cre02.g143387      GGCGGCAGGTCGCGGAGCAGGAGGTGAGGCGGCGGGCAGGGCCAAGCTGCGGGGAGGGAG 120
seq_results        GGCGGCAGGTCGCGGAGCAGGAGGTGAGGCGGCGGGCAGGGCCAAGCTGCGGGGAGGGAG 120
                   ************************************************************

Cre02.g143387      GAGGGGTCGGAGGGACGGGTTGCGTGTGGGGTGTGGCGTGTGGCGTGGCCTGCCTGTGTC 180
seq_results        GAGGGGTCGGAGGGACGGGTTGCGTGTGGGGTGTGGCGTGTGGCGTGGCCTGCCTGTGTC 180
                   ************************************************************

Cre02.g143387      GTTGCGCCTTCCCCGCTCTGCTTCCGCGTTGTGAACGCGGTCGCGTCCTGACCACCAGAG 240
seq_results        GTTGCGCCTTCCCCGCTCTGCTTCCGCGTTGTGAACGCGGTCGCGTCCTGACCACCAGAG 240
                   ************************************************************
```

```
Cre02.g143387      AGATCTCTGCGCCTGAAGTGCTTTACCGTATGCACCGGTGGGTCCGAGGCCTGTGGGTCG 300
seq_results        AGATCTCTGCGCCTGAAGTGCTTTACCGTATGCACCGGTGGGTCCGAGGCCTGTGGGTCG 300
                   ************************************************************

Cre02.g143387      CCGATGGCGCGCCGACCGATGATGACATCGCGCTTCAGCAACACACATTTGGCTAGCTCG 360
seq_results        CCGATGGCGCGCCGACCGATGATGACATCGCGCTTCAGCAACACACATTTGGCTAGCTCG 360
                   ************************************************************

Cre02.g143387      CGTATTGCTACCCTGGCTGCAACAACATTGTCCGAGGAAAGGGACTCGCGAGCGCGGGGC 420
seq_results        CGTATTGCTACCCTGGCTGCAACAACATTGTCCGAGGAAAGGGACTCGCGAGCGCGGGGT 420
                   ************************************************************

Cre02.g143387      CAGGCGCAGCCGCAGCATAGGCCTTCCCTCTGCTGCTGACGCCGCTGCGGAGGCTGGTGA 480
seq_results        CAGGCGCAGCCGCAGCATAGGCCTTCCCTCTGCTGCTGACGCCGCTGCGGAGGCTGGTGA 480
                   ************************************************************

Cre02.g143387      ACTTGGCAGTTGCAGTTGCACATGCGCCCGCGGCATTGTGTGCGATCTATTGCGTGCCGC 540
seq_results        ACTTGGCAGTTGCAGTTGCACATGCGCCCGCGGCATTGTGTGCGATCTATTGCGTGCCGC 540
                   ************************************************************
```

```
Cre02.g143387      CGCCCCATTTGCCACCCCCGCAGCTGCCCGAGCCACGGCTGCGTGTGCTGCAGCTGCTGG 600
seq_results        CGCCCCATTTGCCACCCCCGCAGCTGCCCGAGCCACGGCTGCGTGTGCTGCAGCTGCTGG 600
                   ************************************************************

Cre02.g143387      TTGCAACTTCGCGTTGGTGTACATGAGCAAGTGCATCAGCATCAGCAACTGGCAGCCAGG 660
seq_results        TTGCAACTTCGCGTTGGTGTACATGAGCAAGTGCATCAGCATCAGCAACTGGCAGCCAGG 660
                   ************************************************************

Cre02.g143387      GCGCTGAGCAATGTGTACGCCCGCGGCGACAGTGGTGGCGGTGGCGCTGCCGGCGTGGCA 720
seq_results        GCGCTGAGCAATGTGTACGCCCGCGGCGACAGTGGTGGCGGTGGCGCTGCCGGCGTGGCA 720
                   ************************************************************

Cre02.g143387      GCAGGTGTGTATGGGCTGTGGGGCTGCGGGCGCAGATGTAGGCATTGGTGTGGGCGTGTG 780
seq_results        GCAGGTGTGTATGGGCTGTGGGGCTGCGGGCGCAGATGTAGGCATTGGTGTGGGCGTGTG 780
                   ************************************************************

Cre02.g143387      GGTGATGGCGGTGGCGCCGGCGGCCTTGGGCATTCGCGCCTCCCTTCCCCCGGCCACCCT 840
seq_results        GGTGATGGCGGTGGCGCCGGCGGCCTTGGGCATTCGCGCCTCCCTTCCCCCGGCCACCCT 840
                   ************************************************************
```

```
Cre02.g143387      GCAACCACTGACGTTTGCTGCTCTGTCATCACATGCCATGCTCCCCTCCCTCTCCCACCC 900
seq_results        GCAACCACTGACGTTTGCTGCTCTGTCATCACATGCCATGCTCCCCTCCCTCTCCCACCC 900
                   ************************************************************

Cre02.g143387      GTCAGGACCATCCCGGCCTCCACAGCCGCTGCCAAGCGAGGTGATGTCGCTGTCAAGTGA 960
seq_results        GTCAGGACCATCCCGGCCTCCACAGCCGCTGCCAAGCGAGGTGATGTCGCTGTCAAGTGA 960
                   ************************************************************

Cre02.g143387      CAGCTTGAGCGCCAGCGGCAGCGGTGGCAATGGTGGTGGCGGCGGCGCTGCCTGCGTGGT 1020
seq_results        CAGCTTGAGCGCCAGCGGCAGCGGTGGCAATGGTGGTGGCGGCGGCGCTGCCTGCGTGGT 1020
                   ************************************************************

Cre02.g143387      GGCAGGTGTGTACAGGCTGTGGGGCTGCACGTGCAGATGTAGGCTCTTGAATGGGTGTGT 1080
seq_results        GGCAGGTGTGTACAGGCTGTGGGGCTGCACGTGCAGATGTAGGCTCTTGAATGGGTGTGT 1080
                   ************************************************************

Cre02.g143387      GGGTGGTAGCAGTGCTCGTATCTTGTAGATGCCCTCAGGCACCCTGCATCTATCTACACA 1140
seq_results        GGGTGGTAGCAGTGCTCGTATCTTGTAGATGCCCTCAGGCACCCTGCATCTATCTACACA 1140
                   ************************************************************
```

```
Cre02.g143387      TGTGTAGATGACCCTGACCGCGGCATCTACACTTAAGCCCCCAAGCCCCC-CCCCCCCGA 1199
seq_results        TGTGTAGATGACCCTGACCGCGGCATCTACACTTAAGCCCCCAAGCCCCCCCCCCCCGA 1200
                   ********************************************** *********

Cre02.g143387      ACGTCTACACTTCCCCCAGTTTTGCGCCAAAACCGGCCAAAAACGTCTACACCTGCGTAG 1259
seq_results        ACGTCTACACTTCCCCCAGTTTTGCGCCAAAACCGGCCAAAAACGTCTACACCTGCGTAG 1260
                   ************************************************************

Cre02.g143387      ATGCGCGGCACCCAAACCCAAACATTTTTCTCCGCCTTCGGCGGATATTGGGCAGTTATC 1319
seq_results        ATGCGCGGCACCCAAACCCAAACATTTTTCTCCGCCTTCGGCGGATATTGGGCAGTTATC 1320
                   ************************************************************

Cre02.g143387      TACAGTGTAGATGGAAGATACGAGCACTGGGTGGTAGCTGCAAGGCCCTTGGGCATTCAC 1379
seq_results        TACAGTGTAGATGGAAGATACGAGCACTGGGTGGTAGCTGCAAGGCCCTTGGGCATTCAC 1380
                   ************************************************************

Cre02.g143387      GCCTCCCTCCCCTAGCCATCCTAGTGATGTATATGTTGAGAGGACTGTACTGTAGCAGTG 1439
seq_results        GCCTCCCTCCCCTAGCCATCCTAGTGATGTATATGTTGAGAGGACTGTACTGTAGCAGTG 1440
                   ************************************************************
```

```
Cre02.g143387     ATTGGACACGGGCTTGGGCCTGGGCACTGACAGCGTGGGCACCTGAGTAGGTGTGTGCGC 1499
seq_results       ATTGGACACGGGCTTGGGCCTGGGCACTGACAGCGTGGGCACCTGAGTAGGTGTGTGCGC 1500
                  ************************************************************

Cre02.g143387     TTCACGAGTGCTGTCCTGTTGCGAGTGCGTGCTGTCCTCTGCAAACACTGTCTGTGCCTC 1559
seq_results       TTCACGAGTGCTGTCCTGTTGCGAGTGCGTGCTGTCCTCTGCAAACACTGTCTGTGCCTC 1560
                  ************************************************************

Cre02.g143387     ACCGCCACCCACCCACCCTCACTACCAAAACCTGCCCTCTCTCTTCCATTCCTCTCCATT 1619
seq_results       ACCGCCACCCACCCACCCTCACTACCAAAACCTGCCCTCTCTCTTCCATTCCTCTCCATT 1620
                  ************************************************************

Cre02.g143387     CCCATTCAATAGGTGCGGCCGTGCAGCGGCTCCAACCGCGCCCCTATGCGATTGAGCTGC 1679
seq_results       CCCATTCAATAGGTGCGGCCGTGCAGCGGCTCCAACCGCGCCCCTATGCGATTGAGCTGC 1680
                  ************************************************************

Cre02.g143387     TGAACAGCAGCTCTGGGGACACCGTGCACACCCAAGGCACCTAA---------------- 1723
seq_results       TGAACAGCAGCTCTGGGGACACCGTGCACACCCAAGGCACCTAATTCGNGAGACCACGAN 1740
                  ****************************************
```

Sanger sequencing was used to validate the full sequence of the gene of interest. A nucleotide substitution of C>T occurred at position 420 on the first intron while an additional cytosine (C) was inserted at position 1190 on the 3rd intron (Figure 4.15). No nucleotide change was observed on the exons but the stop codon (TAA) was detected at the end of the sequence. Since no mutation was detected in the exons, the inferred amino acid sequence was the same as the protein sequence of Cre02.g143387 found on the Phytozome database (Figure 4.16).

```
Cre02.g143387      MRPRLTGAAAGADSDGEGAAGGRSRSRSCPSHGCVCCSCWLQLRVGVHEQVHQHQQLAAR    60
seq_results        MRPRLTGAAAGADSDGEGAAGGRSRSRSCPSHGCVCCSCWLQLRVGVHEQVHQHQQLAAR    60
                   ************************************************************

Cre02.g143387      ALSNVYARGDSGGGGAAGVAAGPSRPPQPLPSEVMSLSSDSLSASGSGGNGGGGGAACVV    120
seq_results        ALSNVYARGDSGGGGAAGVAAGPSRPPQPLPSEVMSLSSDSLSASGSGGNGGGGGAACVV    120
                   ************************************************************

Cre02.g143387      AGAAVQRLQPRPYAIELLNSSSGDTVHTQGT*  151
seq_results        AGAAVQRLQPRPYAIELLNSSSGDTVHTQGT*  151
                   ********************************
```

**Figure 4.16 Sequence alignment showed no difference between the gene product of Cre02.g143387 with the inferred amino acid sequence of the gene of interest in the level 2 plasmid.**

### 4.2.3 Transformation and Validation of Transformants

The validated pFPC2-1C4-HygR level 2 construct was transformed into mutant C4 via electroporation after cells were treated with GeneArt MAX Efficiency Transformation Reagent for Algae (Invitrogen). The control plate, cells electroporated without plasmid, yielded no colonies on TAP-Hygromycin, while the transformant plate had >1000 colonies upon transformation with the level 2 construct, equating to transformation efficiency of about 1.2 x $10^4$ transformants/µg of DNA (Figure 4.17). A total of 288 colonies were picked and cultured in liquid TAP media with 10 µg/ml of hygromycin (Figure 4.18).

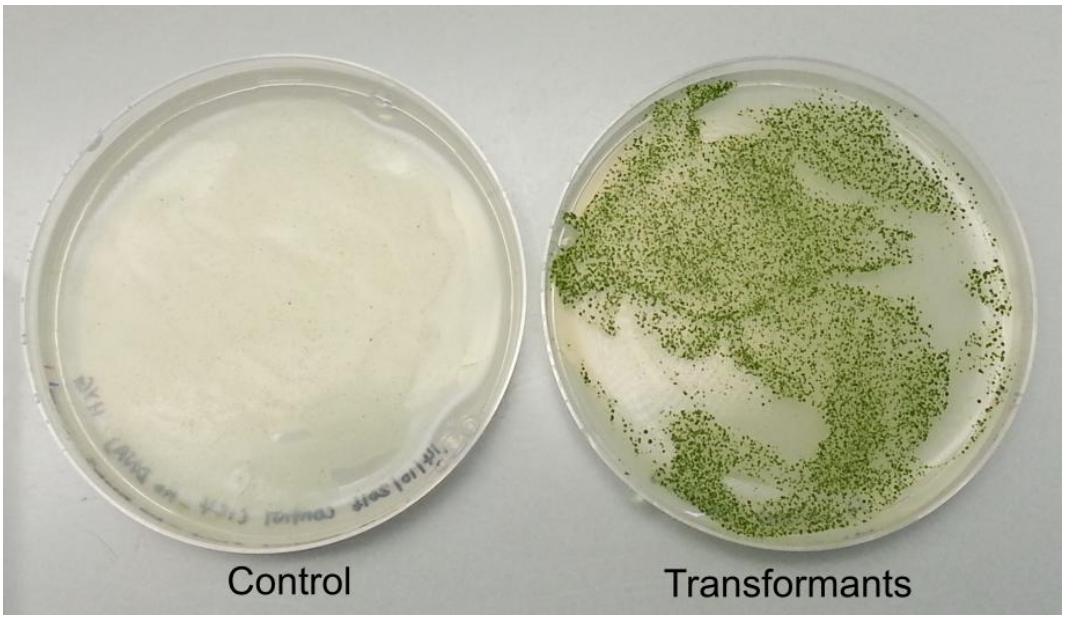



**Figure 4.17 Seven-day old transformants plated on TAP agar plates with 10 µg/ml hygromycin grown under low light conditions at 25°C. Left: Control plate of mutant 1C4 cells electroporated with no DNA yielded no colony. Right: Transformant plate of mutant 1C4 cells transformed with pFPC2-1C4-HygR.**

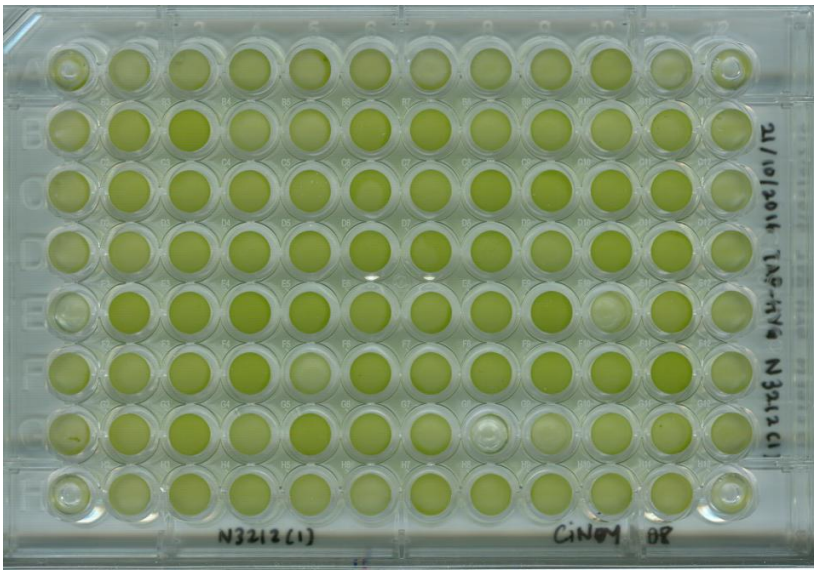


**Figure 4.18 The selected colonies were grown in selective TAP medium containing 10 µg/ml hygromycin for three consecutive generations.**

After three subcultures to TAP-Hygromycin medium, the cells were subcultured into two different plates: one with TAP medium as control and another with selective TP medium for a week to mimick CCM-induced condition. Those in TP medium that were of $OD_{750}$>0.3 from each plate were selected for the next experiments (Table 4.2). The

transformants were renamed according to the order of which they were in the plate used for screening of potential restoration of CCM function (Figure 4.19).

Colony PCR of the region between the CDS and 5'UTR showed no amplification in the controls (wild type and mutant colonies) while 16 out of 49 selected transformants showed positive amplification (Figure 4.20). The 16 transformants were subcultured on TAP-Hygromycin medium (Figure 4.21A) for PCR check and growth spot test. The wild type and mutant yielded no amplicon at the expected size while 10 out of the 16 transformants retained the construct (Figure 4.21B). On the contrary, growth spot test showed that the transformants seemed to grow slightly better than the mutant but did not grow as well as the corresponding wild type (Figure 4.21C) indicating that CCM function was not restored. Confocal microscopy, on the other hand, suggested that when compared to the positive (Figure 4.22A) and negative controls (Figure 4.22B & C), very weak YFP signal which co-localised with the chlorophyll autofluorescence was detected in the transformants (Figure 4.22 Validation of the transformants by observation of the subcellular localisation of the gene product of pFPC2-1C4-HygR in transformed mutant C4 cells compared to positive and negative controls.D). Since neither spot tests nor confocal microscopy suggested restoration of phenotype, it remained uncertain whether the gene Cre02.g143387 was responsible for the morphological and physiological defects observed in mutant C4.

**Table 4.2 $OD_{750}$ measurement for each colony grown on TP medium with 10 µg/ml hygromycin under CCM induced conditions. Colonies selected for downstream validation were highlighted in yellow and possessed $OD_{750}$ > 0.3.**

| Plate 1 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.1007 | 0.3107 | 0.2258 | 0.3279 | 0.3387 | 0.3886 | 0.3887 | 0.3765 | 0.3345 | 0.3313 | 0.3249 | 0.1228 |
| B | 0.3759 | 0.1335 | 0.1306 | 0.1325 | 0.1561 | 0.17 | 0.1909 | 0.1524 | 0.2011 | 0.1207 | 0.1551 | 0.232 |
| C | 0.3078 | 0.1453 | 0.1283 | 0.16 | 0.1751 | 0.1418 | 0.1574 | 0.1697 | 0.1431 | 0.1759 | 0.1601 | 0.2151 |
| D | 0.2052 | 0.1337 | 0.1186 | 0.1263 | 0.138 | 0.1303 | 0.1558 | 0.161 | 0.1556 | 0.1436 | 0.1582 | 0.1956 |
| E | 0.2835 | 0.1248 | 0.1228 | 0.1536 | 0.1237 | 0.128 | 0.1439 | 0.1371 | 0.1581 | 0.1427 | 0.1771 | 0.1994 |
| F | 0.2541 | 0.1141 | 0.1402 | 0.1373 | 0.1311 | 0.155 | 0.1212 | 0.1438 | 0.1541 | 0.1397 | 0.1752 | 0.1747 |
| G | 0.2971 | 0.1192 | 0.1392 | 0.1403 | 0.1563 | 0.1095 | 0.1279 | 0.1148 | 0.1617 | 0.1392 | 0.1296 | 0.2383 |
| H | 0.1662 | 0.4692 | 0.6455 | 0.4712 | 0.3358 | 0.3125 | 0.489 | 0.3975 | 0.3624 | 0.2182 | 0.5942 | 0.1408 |

| Plate 2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.2236 | 0.3293 | 0.1831 | 0.2421 | 0.3448 | 0.2506 | 0.2262 | 0.2269 | 0.3438 | 0.2207 | 0.4331 | 0.1278 |
| B | 0.1796 | 0.1436 | 0.1936 | 0.2123 | 0.2104 | 0.2143 | 0.2125 | 0.1961 | 0.1819 | 0.222 | 0.2254 | 0.2118 |
| C | 0.1642 | 0.2117 | 0.1789 | 0.2383 | 0.2253 | 0.2081 | 0.1902 | 0.2124 | 0.2261 | 0.2027 | 0.1801 | 0.2343 |
| D | 0.1631 | 0.1683 | 0.3273 | 0.3352 | 0.354 | 0.2953 | 0.2842 | 0.2505 | 0.2132 | 0.2084 | 0.1898 | 0.2273 |
| E | 0.1618 | 0.1916 | 0.3239 | 0.383 | 0.3541 | 0.3221 | 0.312 | 0.276 | 0.2181 | 0.1764 | 0.1073 | 0.1988 |
| F | 0.1367 | 0.1389 | 0.2199 | 0.3246 | 0.309 | 0.266 | 0.1834 | 0.1327 | 0.1393 | 0.133 | 0.1781 | 0.2275 |
| G | 0.1527 | 0.1374 | 0.1404 | 0.1265 | 0.1093 | 0.1377 | 0.1108 | 0.1224 | 0.1539 | 0.1454 | 0.1487 | 0.2577 |
| H | 0.1151 | 0.3197 | 0.2307 | 0.3669 | 0.3451 | 0.4187 | 0.2604 | 0.2615 | 0.3255 | 0.3562 | 0.3188 | 0.1247 |

| Plate 3 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.1339 | 0.2938 | 0.1887 | 0.3062 | 0.3975 | 0.4269 | 0.4779 | 0.3625 | 0.2915 | 0.4407 | 0.5078 | 0.1015 |
| B | 0.1844 | 0.1121 | 0.1101 | 0.114 | 0.1251 | 0.1247 | 0.1496 | 0.1421 | 0.1379 | 0.131 | 0.1606 | 0.2762 |
| C | 0.1719 | 0.1369 | 0.1312 | 0.1124 | 0.13 | 0.1146 | 0.1299 | 0.1692 | 0.1439 | 0.1171 | 0.1514 | 0.5171 |
| D | 0.2333 | 0.1603 | 0.119 | 0.1156 | 0.1506 | 0.1602 | 0.1423 | 0.1917 | 0.14 | 0.1726 | 0.1302 | 0.2403 |
| E | 0.1734 | 0.1366 | 0.1491 | 0.1243 | 0.1477 | 0.1372 | 0.1239 | 0.1514 | 0.1293 | 0.1402 | 0.1414 | 0.2836 |
| F | 0.1591 | 0.1258 | 0.1403 | 0.1323 | 0.1396 | 0.1436 | 0.0964 | 0.1529 | 0.1365 | 0.1538 | 0.1282 | 0.2603 |
| G | 0.2215 | 0.1314 | 0.1566 | 0.1533 | 0.1359 | 0.1534 | 0.149 | 0.162 | 0.1339 | 0.1436 | 0.1576 | 0.3289 |
| H | 0.1295 | 0.6852 | 0.683 | 0.4048 | 0.2779 | 0.3797 | 0.4105 | 0.494 | 0.3318 | 0.2897 | 0.2222 | 0.1489 |

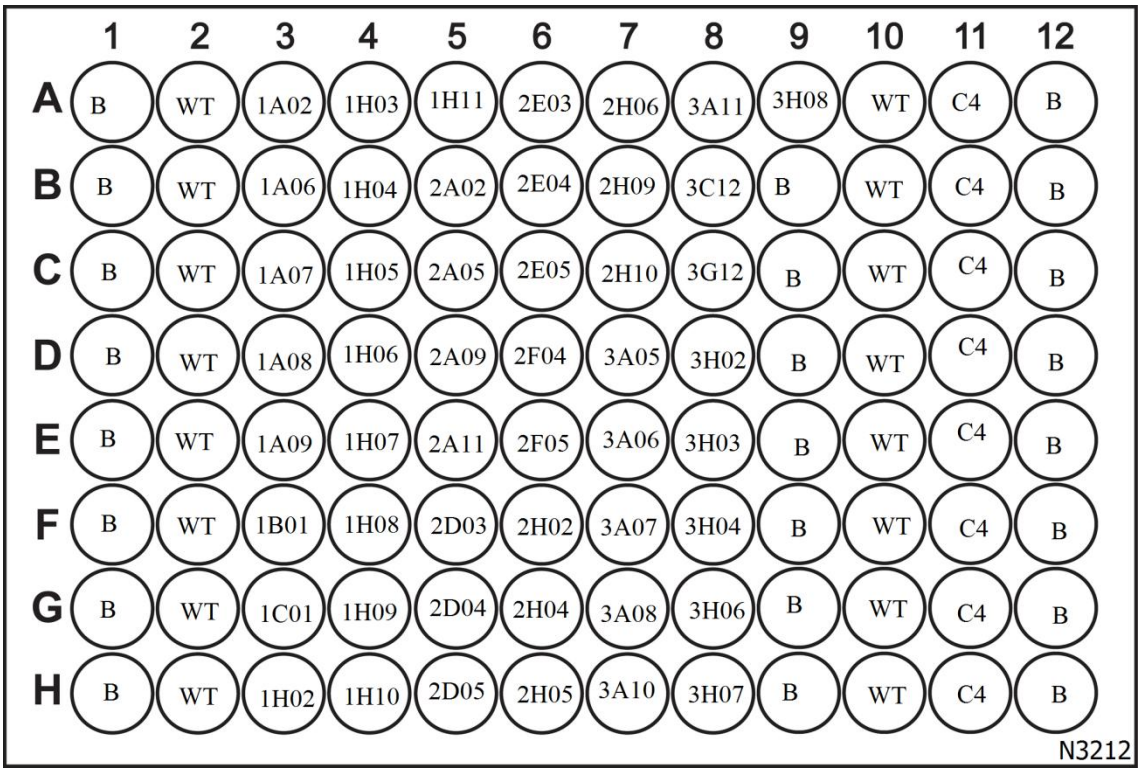



**Figure 4.19 The selected colonies from previous generation were grown in this arrangement in TAP medium with 10 µg/ml hygromycin. B, blank; WT, corresponding wild type; C4, CCM mutant with pyrenoidal defects.**

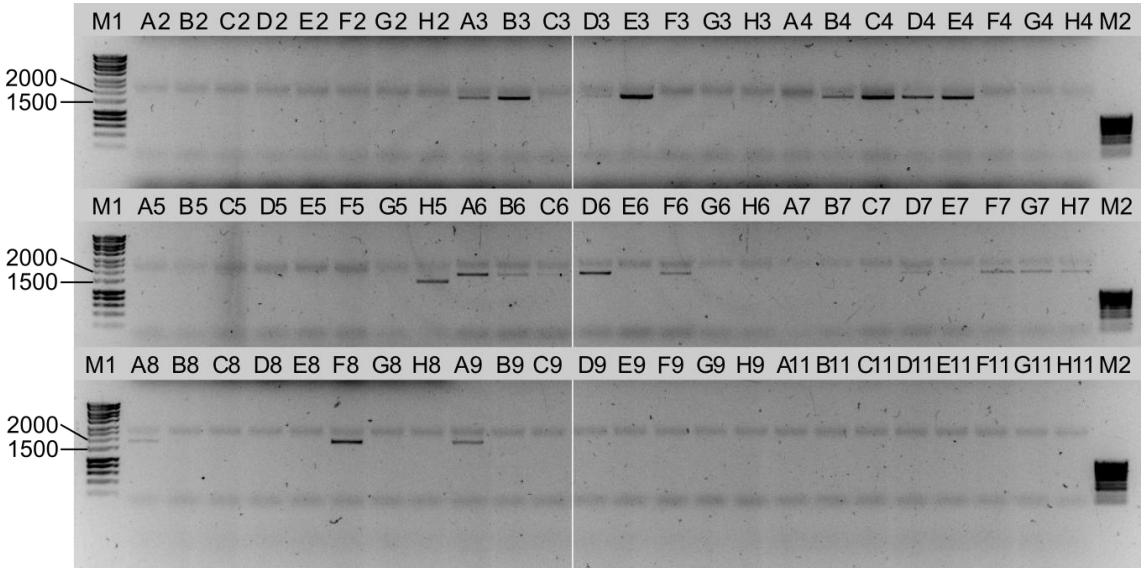



**Figure 4.20 Colony PCR was performed using primer pair CC027+CC030 which amplified the region between CDS and 3'UTR in pFPC2-1C4-HygR plasmid with estimated amplicon size of about 1.8 kb. The label of each colony corresponded to the plate arrangement in Figure 5.19.**

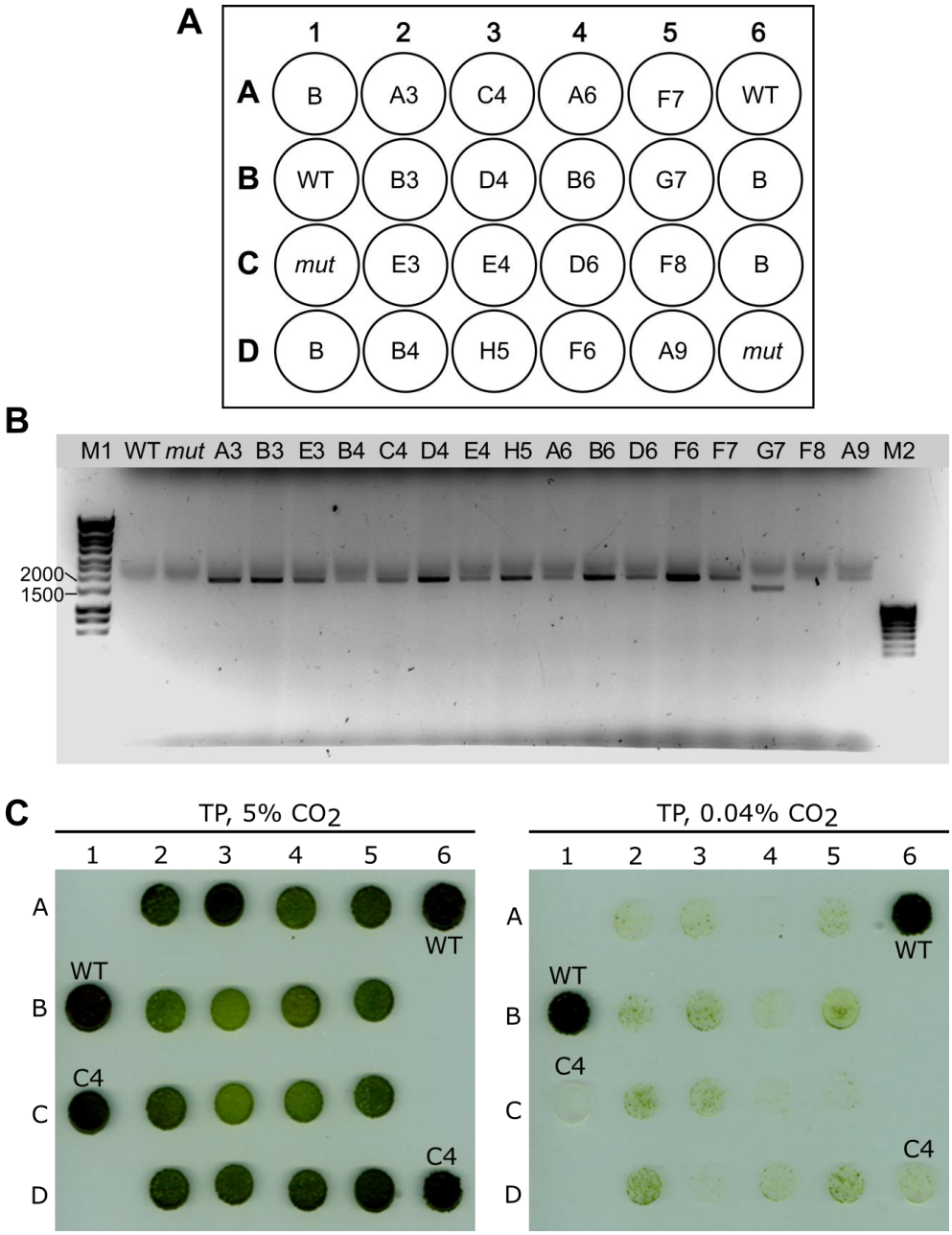



**Figure 4.21 (A) The transformants with positive amplification were cultured in TAP medium with 10 µg/ml hygromycin in a 24-well plate. (B) Confirmation that the transformants retained pFPC2-1C4-HygR using PCR amplification with primer pair CC027+CC030. (C) Growth spot test performed on the transformants to confirm the restoration of CCM phenotype. B, blank; WT, wild type; *mut*, mutant C4; M1, Bioline Hyperladder 1kb; M2, Bioline Hyperladder 100bp; TP, Tris-phosphate medium.**

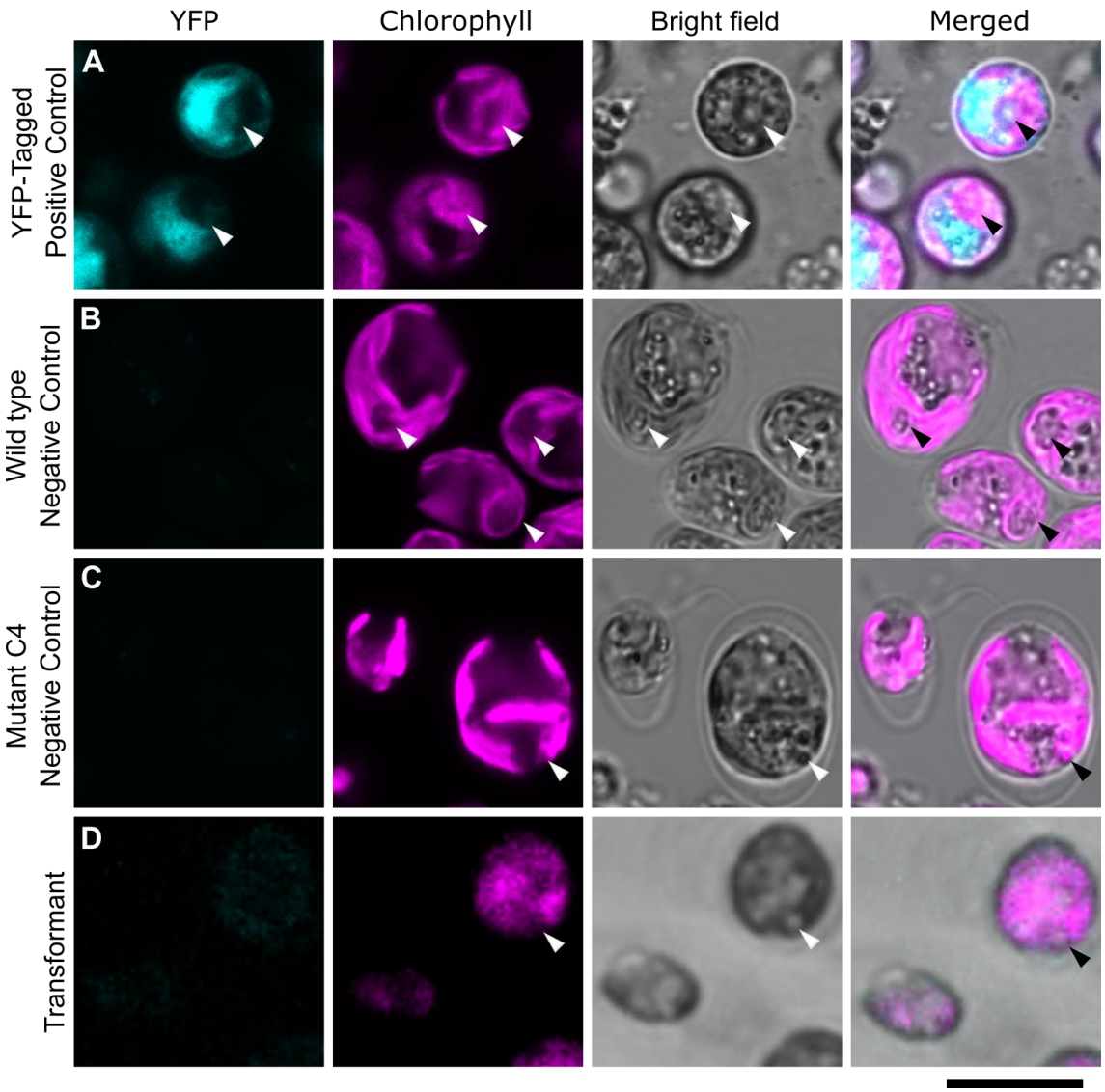



**Figure 4.22 Validation of the transformants by observation of the subcellular localisation of the gene product of pFPC2-1C4-HygR in transformed mutant C4 cells compared to positive and negative controls. The YFP signal, chlorophyll autofluorescence, bright field and composite images of (A) a positive control with YFP localised in the nucleus, (B) untransformed wild type cells, (C) untransformed mutant C4 cells, (D) an example of transformants of mutant C4 transformed with the pFPC2-1C4-HygR plasmid showing weak YFP signals which seemed to co-localised with chlorophyll autofluorescence. Black and white arrows indicate the position of the pyrenoids. Scale bar 10 µm.**

## 4.3 Discussion

In this study, the gene disruption in mutant C4 was suggested to be in Cre02.g143387 based on whole genome sequencing data. A search on Phytozome suggested that this *Chlamydomonas*-specific gene is one of the many genes with unknown functions found in the genome (Merchant et al., 2007) despite the development of the tools for molecular genetics and genome sequencing (Jinkerson and Jonikas, 2015; Zones et al., 2015). Specifically in the area of CCM, many $CO_2$-responsive genes have yet to be characterised and these might be candidate genes playing roles in the CCM regulation and the transportation of inorganic carbon from the stroma into the pyrenoid (Mackinder, 2017; Mackinder et al., 2017). When the cells are shifted from high $CO_2$ to low $CO_2$, up to about 5880 genes presented major shifts in gene expression, with about 1000-2000 genes showing significant upregulation when CCM is induced (Brueggeman et al., 2012; Fang et al., 2012; Wang et al., 2015). The authors also showed that there is a high proportion of differentially-expressed genes with unknown functions based on their high throughput sequencing (Brueggeman et al., 2012; Fang et al., 2012) The main objective of this chapter was to explore whether Cre02.g143387 could lead to functional and morphological restoration of the CCM in mutant C4.

When a mutation is complemented using genomic complementation assay, restoration of phenotype would be observed in the mutant cells (Rochaix, 1995). The usual methods to determine the restoration of CCM phenotype in transformants include the growth spot test (Spreitzer and Mets, 1981), fluorescence assays, e.g. confocal microscopy, if the vector expresses a protein fused to a fluorescence tag or Western blotting if the protein is fused to a smaller epitope tag, such as the complementation of *epyc1* mutant with EPYC1-mVenus-3xFLAG (Mackinder et al., 2016), oxygen evolution to determine the affinity of the transformants towards inorganic carbon, such as the restoration of CCM in *cia5* (Fukuzawa et al., 2001), *cia6* (Ma et al., 2011), Spinach RbcS hybrid (Meyer et al., 2012) and *cia8* (Machingura et al., 2017).

In this study, the genetic complementation of mutant C4 did not restore the CCM phenotype when the transformants were subjected to spot test and confocal microscopy. The results of Sanger sequencing of the construct suggested that the gene insert possessed two mutations in the introns, but none was detected in the exons of the gene. This indicated that a functional protein would have been translated from the final mRNA. However, the presence of a stop codon between the gene and the YFP sequence

in the construct would cause the translation to stop before the YFP sequence, thus YFP would not be present in the final product. Assuming that the final gene product was translated, mutant C4 would have been complemented. Since no complementation was observed on the transformants, the failure to rescue the mutant might be due to the possibility that some other gene was knocked-out instead of Cre02.g143387.

This complementation assay which was set up based on Golden Gate cloning system (Engler et al., 2014) would be an excellent platform to test the rest of the CCM mutants using in this research (Chapter 3). The Golden Gate (GG) cloning system was used as the cloning strategy as this method possesses the ability to assemble a multi-gene construct in a single-step reaction (Engler et al., 2008) and to avoid inefficient, time-consuming and expensive steps such as designing and ordering customised primers, PCR amplification and gel purification (Engler et al., 2014). For the purpose of restoring the CCM phenotype of mutant C4, the gene Cre02.g143387 was the choice guided by WGS data suggesting the gene was knocked out in the mutant strain. Aside from amplifying the gene from the genomic DNA of wild type *C. reinhardtii* cells, the primers were also designed to include the flanking region carrying *Bsa*I recognition sites at opposing orientations and the same pair of fusion sites for any level zero CDS non-stop module (AATG – TTCG) (Engler et al., 2014). The pair of fusion sites is different for each part, with one being compatible with the part directly next to it, creating a defined linear order for single-step restriction-ligation step (Engler and Marillonnet, 2014; Engler et al., 2014).

Since the GG system enables the stacking of multiple genes in a single construct (Weber et al., 2011; Binder et al., 2014; Engler et al., 2014), the strategy taken for this study was to use two different transcriptional units, one carrying the gene fused to a fluorescent protein reporter gene and another with the hygromycin resistance gene. The inclusion of an antibiotic selection to the cells would ascertain that the construct was delivered into the cells. As mutant C4 was a mutant of insertional mutagenesis using paromomycin cassette (Zhang et al., 2014), hygromycin resistance was chosen as the antibiotic selection to ascertain the transformation success in the preliminary screen of transformants.

These two transcription units were driven by different promoters and terminators. In TU_1C4, the psaD promoter and terminator used to drive expression of Cre02.143387 were domesticated specifically for *C. reinhardtii*. The combination was chosen because reporter gene assay using psaD promoter and psaD terminator was demonstrated to

possess the best performance (Kumar et al., 2013). The transformation of YFP gene driven by psaD promoter and terminator into the nucleus of *C. reinhardtii* (UVM11 strain) cells also showed higher YFP signal under confocal microscopy (Dr. Francisco Navarro, personal communication). The transcription unit carrying hygromycin resistance gene was driven by a *HSP70A*/*RBCS2* fusion promoter of which the *HSP70* promoter activates the transcription of *RBCS2* promoter when placed upstream of it (Schroda et al., 2000).

In this study, pFJN70 which carried the YFP gene (Dr. Francisco Navarro) originated from the YFP domesticated by Neupert et al. (2009) showed a weak signal when observed at excitation wavelength of 514 nm. BLASTN suggested that the sequence for this YFP gene was 100% similar to the mVenus used in (Mackinder et al., 2016). Both Neupert et al. (2009) and Rasala et al. (2013) suggested that YFP/Venus protein is the second most used fluorescent protein (FP) in algal transformation. Venus is excellent for individual protein tagging or reporter gene assay for promoter studies due to the high signal-to-noise ratio but the fluorescence signal overlaps with GFP and bleeds into the orange channels (excitation 532 nm, emission 585 nm) thus this FP might not work well for multi-tagging experiments (Rasala et al., 2013). The best alternate FP to be considered for future work of this experiment would be the *Chlamydomonas*-expressed mCherry. This FP is a true monomer easily detectable above auto-fluorescence where Rasala et al. (2013) also demonstrated that it works well as a fusion protein and has very efficient transformation.

Normally, the attachment of a fluorescent protein tag to the protein of interest would have no interfering effect on the function, structure and localisation of the protein. However, in the case of this study, it was possible that fusing the YFP tag to the gene of interest had affected the protein structure, protein folding or post-translational modifications leading to the loss of protein function (Snapp, 2005, 2009; Costantini and Snapp, 2013) and therefore no restoration of CCM in the mutant. In order to overcome this issue, a smaller epitope tag, such as FLAG tag, influenza virus hemagglutinin (HA) tag, polyhistidine (His) tag, metal affinity (MAT) tag or Myc tag (Fritze and Anderson, 2000; Jinkerson and Jonikas, 2015), could be an alternative to fluorescent protein tags. In the case where the protein of interest is not fluorescently-tagged, the protein could be visualised with indirect immunofluorescence assay using high quality antibody to avoid nonspecific signal from secondary antibody (Fritze and Anderson, 2000). For downstream application purposes, the addition of fluorescent protein tag or epitope tag

is an option for identification of potential protein interactors by performing co-immunoprecipitation and proteomics analysis (Rothbauer et al., 2008; Desmyter et al., 2015; Xing et al., 2016).

Both the insertional mutagenesis technique used to for the CCM mutant screen of which mutant C4 originated from, and the transformation of the newly assembled multi-gene construct, pFPC2-1C4-HygR were performed using electroporation technique. This technique was preferred over glass bead method (Hall et al., 1993) or biolistic transformation (Sanford, 1990) as electroporation is known to be more efficient to transform DNA into the nucleus of *C. reinhardtii* (Brown et al., 1991; Shimogawara et al., 1998; Jinkerson and Jonikas, 2015) while multiple integrations of transforming DNA using nuclear biolistic transformation is not desirable (Rochaix, 1995). The walled cells were treated with MAX Efficiency Transformation reagent (Invitrogen) instead of using autolysin to induce cell wall removal (Kindle, 1990) or treating with aminoclay nanoparticle (Kim et al., 2014) prior to transformation. This commercial reagent had been demonstrated to improve transformation of walled *C. reinhardtii* cells to up to >1000 more transformants per µg of DNA (Life Technologies Corporation, 2013).

## 4.4 Conclusion and Future Work

Mutant C4 was part of the insertional mutagenesis screen which inserted a cassette containing paromomycin resistance into the genome (Zhang et al., 2014). In this study, Golden Gate cloning was demonstrated to be a useful cloning tool for nuclear transformation of *C. reinhardtii* given that it is a simple one-step, one-pot cloning technique with standard parts for plants available via Addgene repository (Engler et al., 2008, 2014). Due to the flexibility to multiplex transcriptional units for transformation, and the standard syntax of this system, it was shown in this study that level zero parts domesticated by one lab could be shared among different labs working with the same organism to answer different biological questions.

Since the work in this study could not discount the possibility that errors in preparation of the construct might have caused the lack of phenotype restoration, it would be crucial that the level 0 CDS construct is fully sequenced and examined before proceeding to assemble the next level of construct. It would also be crucial to a few other preliminary tests to validate the transformants besides confocal microscopy, such as Western blots to determine the presence of the gene product, quantitative PCR to determine mRNA

level of the gene and oxygen evolution to determine the affinity of the cells towards inorganic carbon.

It is equally possible that the CCM defect was not caused by the gene transformed into the mutant, since WGS suggested multiple potential insertional sites. Therefore, it would be crucial to run a PCR check on all the suggested locations to identify the insertional sites. Backcrossing (Sager, 1954) is also an option to isolate the relevant mutation which caused the CCM defects in this mutant which might have been masked by pleiotropic effects. Another alternative to resolve the necessity of backcrossing a mutant with multiple insertions is to isolate multiple knockout mutants of the same gene (Dent et al., 2005), which is discussed further in the next chapter (Chapter 5). Aside from the possibility that the cassette was inserted in another location in the nuclear genome, the loss of CCM function and the pyrenoid defects in mutant C4 might be due to an insertion of "junk" DNA (Li et al., 2016), deletion of genomic DNA (Tam and Lefebvre, 1993) or genomic rearrangement (Jinkerson and Jonikas, 2015) during the process of transformation. Sometimes, the insertional mutagenesis cassettes are inserted in multiple locations in the genome (Zhang et al. 2014; Dr. Martin Jonikas, personal communication).

# 5 STRUCTURAL FORMATION OF THE PYRENOID

## 5.1 Introduction

### 5.1.1 The Structure of the Pyrenoid of *Chlamydomonas reinhardtii*

The presence of a pyrenoid in *Chlamydomonas* was first described by Dill (1895), and a comprehensive morphological and taxonomical characteristation of the pyrenoid in algae was published more than 70 years later (Griffiths, 1970), where the structure of the thylakoid network within the pyrenoid was described in detail based on the work of Ohad et al (1967). The pyrenoid was suggested to play a role in the carbon metabolism of *Chlamydomonas* (Griffiths, 1980) upon discovery of Rubisco in the pyrenoid (Holdsworth, 1971). Following this, immunogold labelling of anti-Rubisco antibody demonstrated that Rubisco is localised in the pyrenoid of *C. reinhardtii* during induction of the $CO_2$-concentrating mechanism (Lacoste-Royal and Gibbs, 1987; K. Kuchitsu et al., 1988; Süss et al., 1995; Morita et al., 1997; Borkhsenious et al., 1998; He et al., 2003). Using *in situ* cryo-electron tomography (cryo-ET), the three dimensional structure of the thylakoid tubule network and the arrangement of Rubisco in the pyrenoid of *C. reinhardtii* was validated by Engel et al in 2015.

Recent studies have identified new components associated with the pyrenoid, such as LCIB (Miura et al., 2004; Yamano et al., 2010), EPYC1 (Mackinder et al., 2016) and

Cre06.g259100 (Kobayashi et al., 2015), while a carbonic anhydrase, CAH3 (Duanmu et al., 2008; Sinetova et al., 2012) and a $Ca^{2+}$-binding protein, CAS (Wang et al., 2016) are localised in the thylakoid tubule network of the pyrenoid. The spatial interactome and microscopy data of Mackinder et al (2017) defined at least four different layers of the pyrenoid: 1) the pyrenoid matrix, 2) starch plates, 3) proteins arranged near the stroma puncta and 4) proteins arranged in a mesh-like pattern and 5) proteins in the thylakoid lumen puncta (Figure 5.1). The biophysical structure of the pyrenoid also includes the thylakoid tubule network which plays the role of delivering high concentration of $CO_2$ to the pyrenoid matrix (Wang et al., 2015).

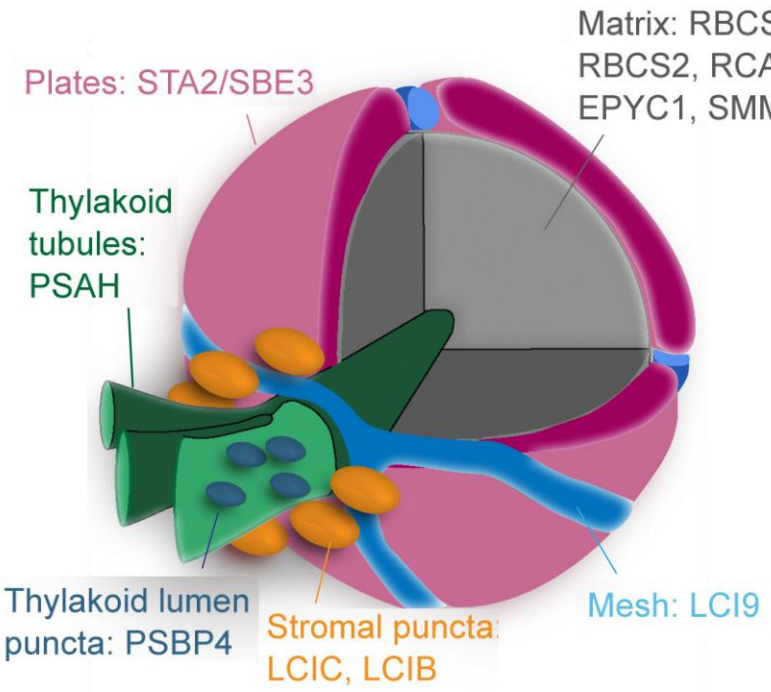



**Figure 5.1 A proposed model by Mackinder et al (2017) which shows the different protein-containing layers of the pyrenoid based on spatial interactome, microscopy and past literature.**

### 5.1.2 Biogenesis of the Pyrenoid

In *Chlamydomonas*, the pyrenoid matrix was first suggested to consist of crystalline protein (Griffiths, 1970) but recent work by Freeman Rosenzweig et al (2017) suggests that the pyrenoid is a more fluid-like structure, with Rubisco, Rubisco activase and EPYC1/LCI5 being the main components in the matrix. Mackinder et al (2016) reported the presence of the linker protein in the pyrenoid, EPYC1, which interacts with Rubisco small subunit. The increase in size of the pyrenoid during CCM induction was suggested to be facilitated by a phase change caused by the interaction between EPYC1

and Rubisco as both are incorporated into the pyrenoid (Freeman Rosenzweig et al., 2017). Based on modelling prediction, the authors suggested that this phase change leading to the formation of the pyrenoid matrix occurs when the eight binding sites of Rubisco heterodimer interact with the four binding sites of two EPYC1 proteins forming a stable dimer pyrenoid (Freeman Rosenzweig et al., 2017).

Aside from the increased size of the pyrenoid when the cells are shifted from high to low $CO_2$ conditions, the size of the starch plate enveloping the pyrenoid also increases (Ramazanov et al., 1994; Rawat et al., 1996). Engel et al (2015) hypothesised that the pyrenoidal starch sheath provides a rigid structure that helps to increase the density of Rubisco packing. The work of Kuchitsu et al (1988) and Ramazanov et al (1994) indicated that the formation of pyrenoidal starch sheath is independent of the degradation of chloroplastic starch granules during CCM induction. Kuchitsu et al. (1988a) demonstrated this by showing that the starch degradation enzymes involved in starch metabolism remain similar before and after CCM induction indicating the semi-independence of the carbon metabolism in the pyrenoid. This suggestion was backed up by the evidence that the formation of pyrenoidal starch sheath is due to CCM induction, where external source of carbon is provided to the cell (Ramazanov et al., 1994).

In contrast, the thylakoid tubule network, which is part of the pyrenoid structure of *Chlamydomonas*, does not seem to undergo visible structural changes under different $CO_2$ conditions based on the electron micrographs published by Ramazanov et al (1994) and Rawat et al (1996). The coalescing chloroplast thylakoids that form the thylakoid network in the pyrenoid were first observed in electron micrographs (Sager and Palade, 1957; Ohad et al., 1967a) and affirmed 60 years later when a 3D reconstruction was performed using cryo-ET (Engel et al., 2015). This thylakoid tubule network was observed to form at a faster rate compared to the chloroplast thylakoids (Golberg and Ohad, 1970; Griffiths, 1980) and the work of Goodenough and Levine (1970) suggests that the development of a typical pyrenoid core is closely related to the transport of protein material through the tubule network. The reduced pigment concentration in pyrenoidal thylakoids (Gibbs, 1962) was found to be caused by a lack of segregation of photosystems in the pyrenoidal tubule network (Bertos and Gibbs, 1998; Gunning and Schwartz, 2000), suggesting that this network plays a role different from the chloroplast thylakoids. The differences found in the pyrenoidal tubules and chloroplast thylakoids led to the hypothesis that the thylakoid tubule network plays a central role in algal

metabolism different from that of the chloroplast thylakoids which they are connected to, in terms of remobilisation of protein to and from the pyrenoid (Griffiths, 1980).

The relationship between thylakoid biogenesis and the pyrenoid was further studied by Uniacke and Zerges (2007) where they proposed the existence of a T (translation) zone at the periphery of pyrenoid, where photosynthetic complexes and CCM-related proteins are synthesized, such as the large subunit (LSU) of Rubisco. Rast et al (2015) pointed out in their review that this thylakoid network in the pyrenoid may be a structural nucleation point, a starting point for thylakoid membrane biogenesis.

### 5.1.3 Mutants with Defective Pyrenoidal Tubule Network

In Chapter 3, characterisation of CCM mutants with pyrenoidal defects confirmed the work of Mackinder et al (2017) that a mature pyrenoid is made up of various protein components forming the pyrenoid matrix, as well as a developed core of the pyrenoid which is made up of thylakoid nucleation site. Five of the characterised CCM mutants lacking thylakoid nucleation site (Chapter 3) were found to be knockouts of the same gene, Cre11.g467712, which is named *SAGA* (**S**tarch-pyrenoid **Ag**gregation **A**nchor) in this chapter for the convenience of the readers. The mutant line of Cre11.g467712 is referred as *saga* while the gene product is labeled as SAGA in this chapter. Morphological observation of all *saga* mutants (A5, A7, A8, A11 and C8) showed obvious defects to the pyrenoid (Figure 5.2): the lack of visible thylakoid tubule network within the pyrenoids, anomaly in the pyrenoidal starch plates, multiple pyrenoids and mislocalised pyrenoids. The work performed in this chapter is focussed on mutant C8.

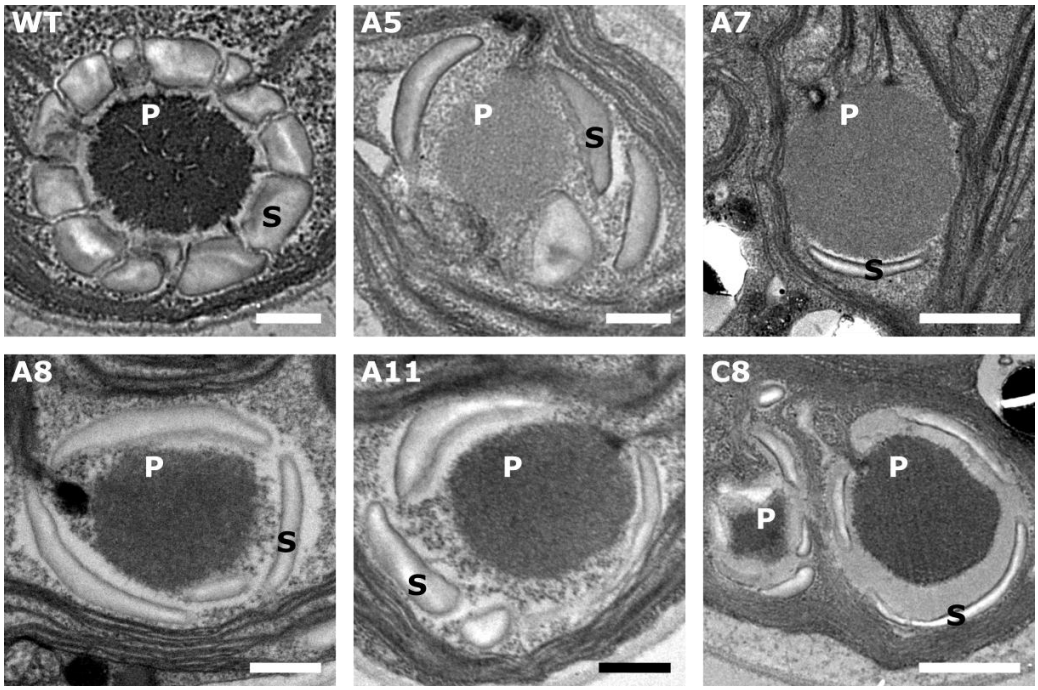



**Figure 5.2 Electron micrographs of the pyrenoid of the wild type and *saga* mutants grown photoautotrophically at room temperature with illumination of about 50 µmol photons $m^{-2}s^{-1}$ under ambient level of $CO_2$ for 24 hours. Scale bar 500 nm.**

The summary of Griffiths (1980) with regards to the development of the thylakoid core of the pyrenoid gave rise to the main hypotheses of this chapter:

1. The formation of the core of the pyrenoid has direct impact on the carbon concentrating mechanism of *C. reinhardtii*.
2. The development of a mature pyrenoid during the induction of CCM is facilitated by a protein or a group of proteins leading to the interaction between the thylakoid nucleation site and the pyrenoid matrix.

## 5.2 Results

### 5.2.1 Prediction of Structure and Domains of *SAGA* gene

Comparison to the genome sequence in Phytozome showed that *SAGA* was located on Chromosome 11 and possessed 35 exons (Figure 5.3). When translated, SAGA protein is predicted to contain 1626 amino acids and is about 168 kDa in size. PHYRE2 and PREDATOR were used to predict the secondary structure of the protein, of which the consensus suggested 67.2% of alpha helices, 30.9% of random coil and about 2% of extended strand (Figure 5.4). Using whole genome sequencing, the location of the paromomycin cassette in mutant C8 was identified to be between the 28th exon and 29th exon of the gene (Figure 5.4). The disruption of the gene was confirmed by Western blot using anti-SAGA antibody, where SAGA was expressed in the wild type but not in the mutant..

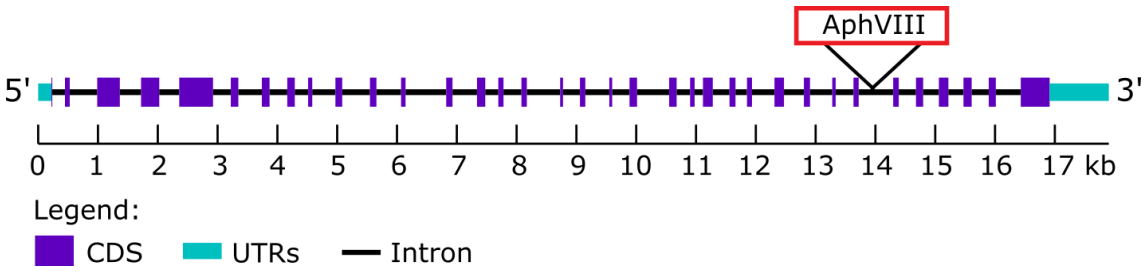



**Figure 5.3 Genomic structure of *the SAGA* gene locus (ID: Cre11.g467712) showing the position of the *AphVIII* insertion (red box).**

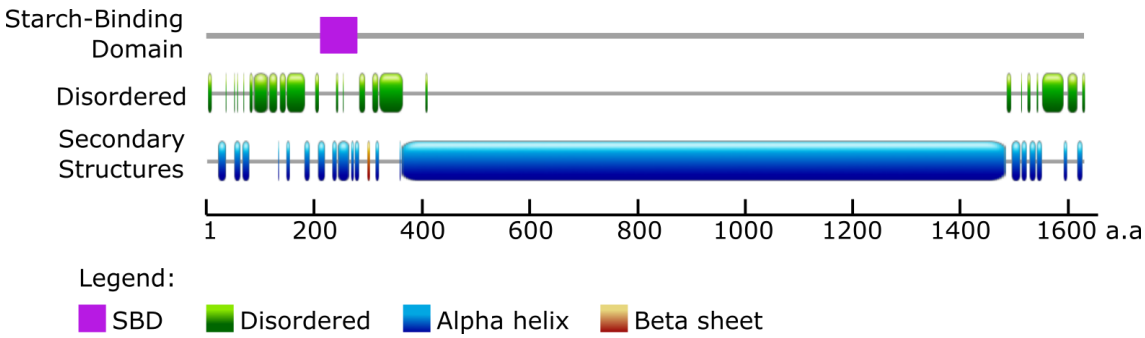



**Figure 5.4 Secondary structure prediction of *SAGA* using PHYRE2 suggested high degree of alpha helices and a starch-binding domain.**

BLASTP predicted the presence of two domains in SAGA: a starch binding domain (SBD) from the CBM20 (carbohydrate-binding module 20) family and a SMC (structural maintenance of chromosomes) N terminal domain (Figure 5.5A). According to the Conserved Domain Database (Marchler-Bauer et al., 2017), the putative CBM20

family domain in SAGA was 100% aligned with the highly conserved residues in the starch recognition site (feature 1) in the known CBM20 domain structure (Figure 5.5B). Although the SBD in SAGA was also aligned to the feature of CBM20 which is known to undergo significant structural change upon ligand binding (feature 2) (Figure 5.5C), the protein was more likely to be of feature 1. Interestingly, 14 proteins were suggested to be homologous to SAGA, mainly around the starch-binding domain region, with 6 from the same species and the rest were of higher plants origin (Figure 5.6). Most of the alpha helices were in the SMC_N domain region while the random coil was present around the CBM20 domain region. PCOILS predicted that almost 100% of the alpha helical region of SAGA was in a coiled-coil structure (Figure 5.7). The top two predicted protein models of SAGA aligned to different regions of the protein, the SMC_N family domain (Figure 5.8A) with a confidence score of 98.9% and the CBM20 domain (Figure 5.8B) with a confidence score of 98.7%. The confidence score of more than 90% indicated that the suggested models should adopt the overall fold as shown in Figure 5.8 even though the sequence identities were relatively low at 11%. In general, a total of 106 models were suggested by PHYRE2 with confidence level of more than 90%.

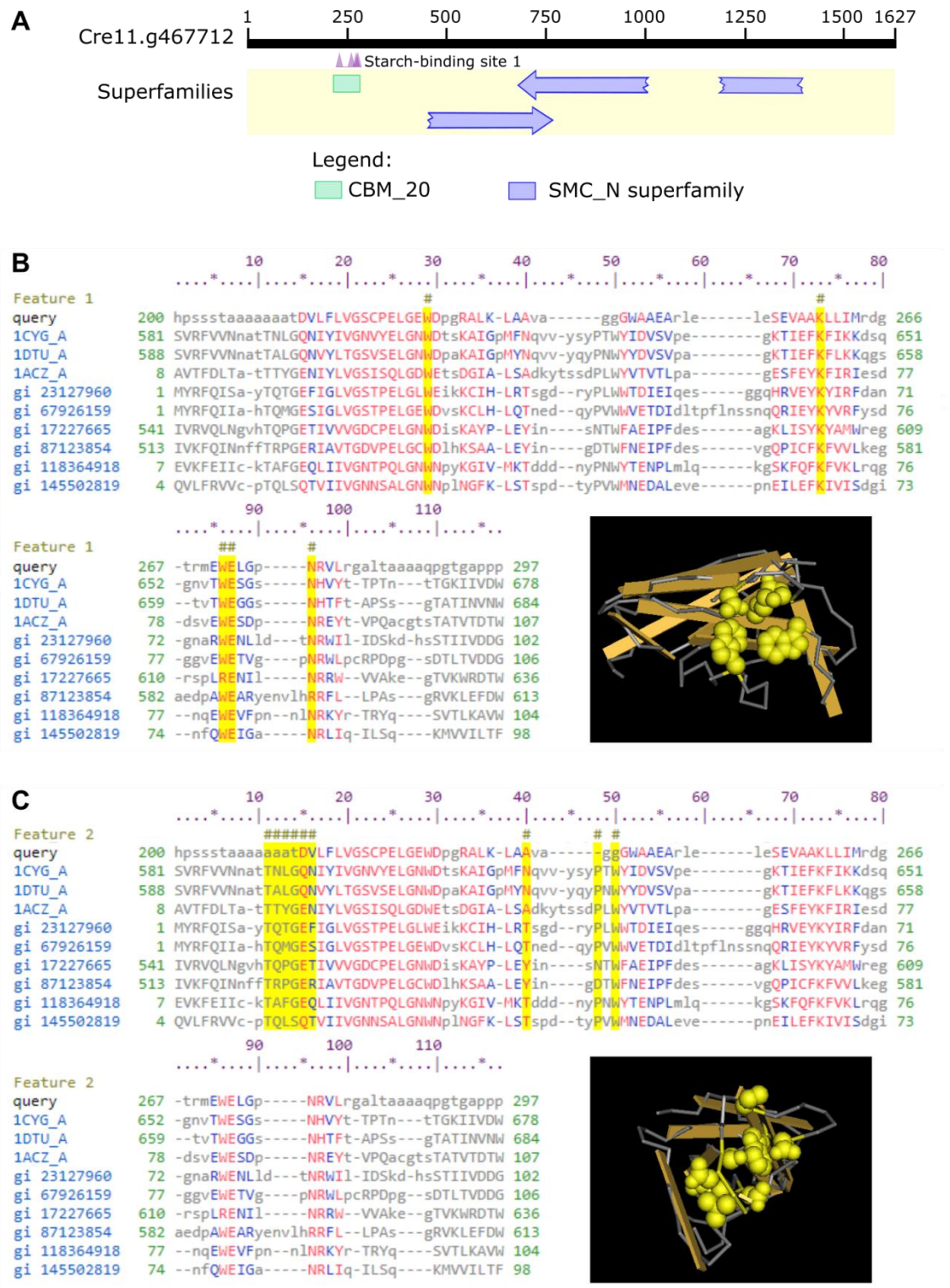



**Figure 5.5 (A) BLASTP of SAGA suggested a starch binding domain from CBM20 family and a SMC (Structural Maintenance of Chromosome) N terminal domain in the protein. NCBI Conserved Domain Database search showed that (B) SAGA (query) possessed all the highly conserved sites (#) of the CBM20 starch-binding feature 1 which acts as starch recognition site. (C) SAGA (query) aligned with the SBD feature 2 but did not possess the conserved active site of this feature.**

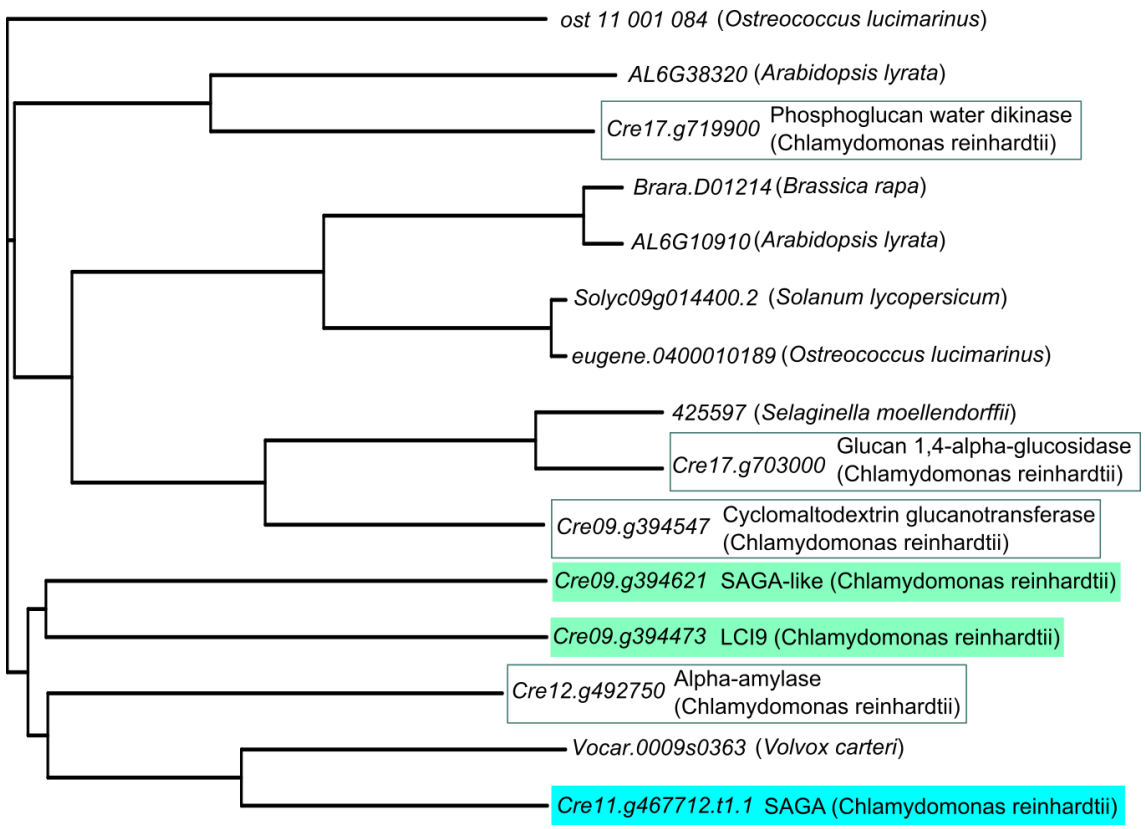



**Figure 5.6 A dendrogram of pairwise relationships between SAGA and the suggested homologues in *Chlamydomonas reinhardtii*, *Volvox carteri*, *Ostreococcus lucimarinus*, *Arabidopsis lyrata*, *Brassica rapa*, *Solanum lycopersicum* and *Selaginella moellendorffii*. The alignment was performed using Clustal Omega on EMBL-EBI on the protein sequences downloaded from Phytozome database.**

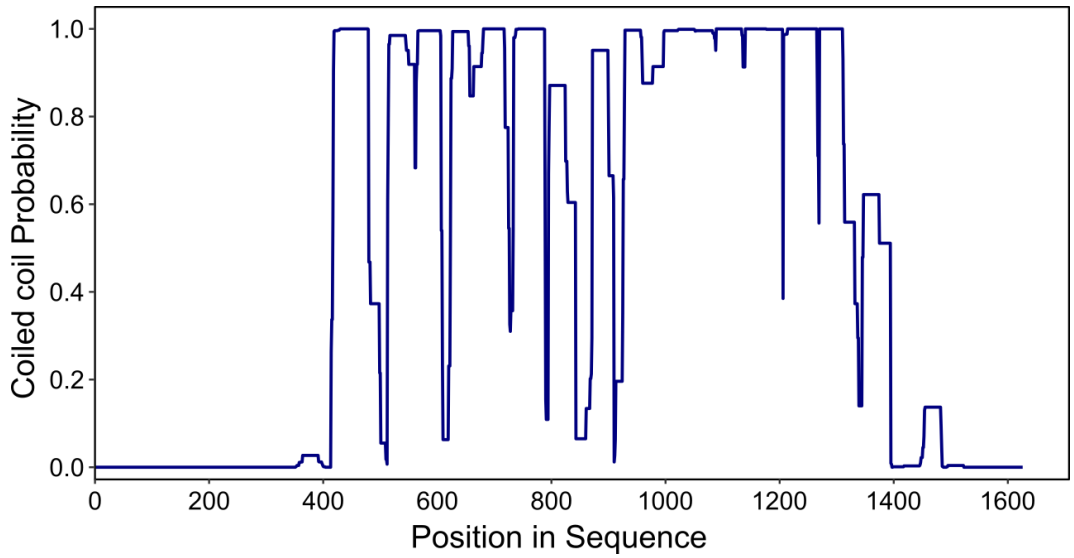



**Figure 5.7 PCOILS predicted high probability of coiled-coil secondary structure within the alpha helix region of SAGA.**

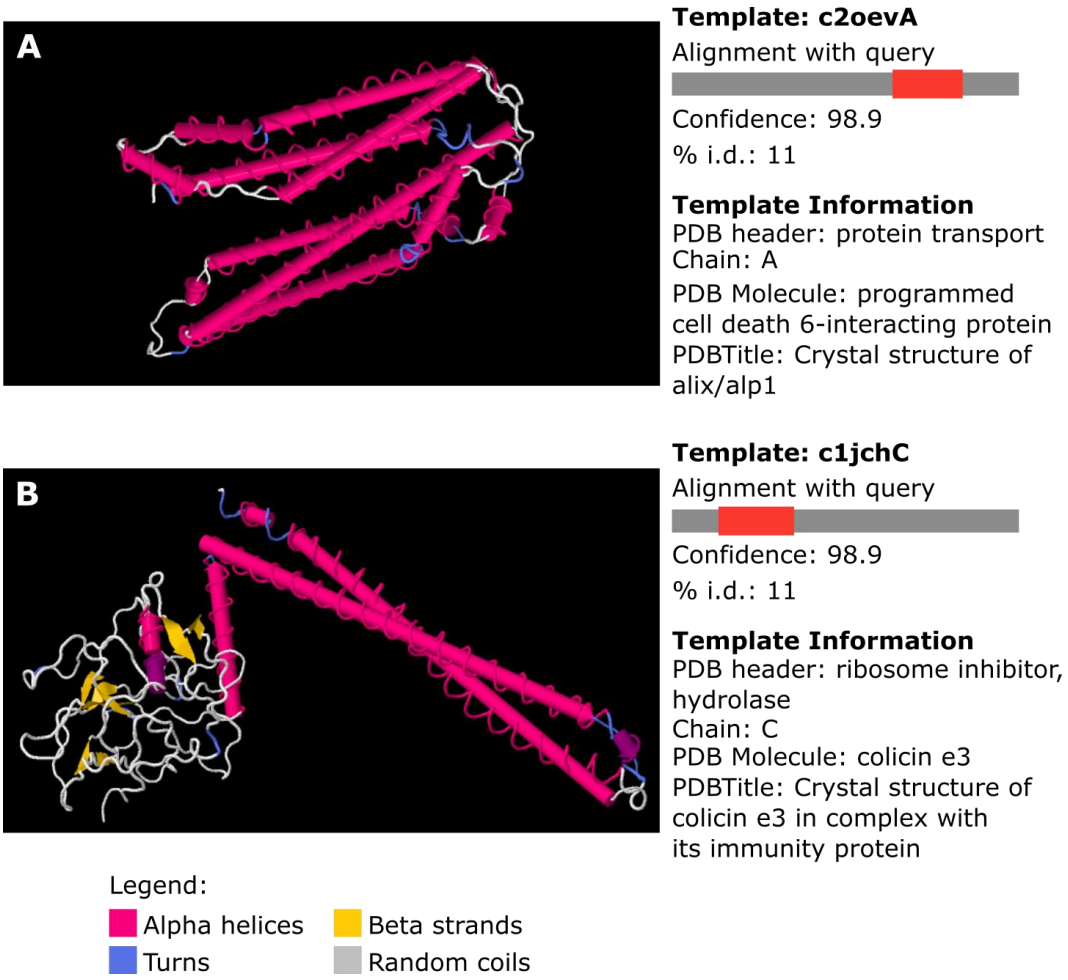



**Figure 5.8 Top two predicted protein structures of SAGA based on alignment at different locations of the protein to (A) crystal structure of ALIX/AIP1 and (B) crystal structure of colicin e3.**

### 5.2.2 Localisation of SAGA in the Pyrenoid

Mutant C8 (renamed as *saga*) was complemented using pRAM118_SAGA that was assembled using cDNA (Figure 5.9) which expressed the SAGA-Venus-3xFLAG protein. Two stable complemented lines kindly provided by Alan Itakura (University of Stanford), *saga::SAGA-Venus-3xFLAG A7* (SVA7) and *saga::SAGA-Venus-3xFLAG B2* (SVB2), were used in this study to further understand the role of *SAGA* in the structural formation of the pyrenoid and the carbon concentrating mechanism of *C. reinhardtii*. SAGA was predicted by ChloroP to contain a chloroplast transit peptide of about 73 residues while both PredAlgo and TargetP suggested that the transit peptide would direct the protein to the mitochondria (Table 5.1).

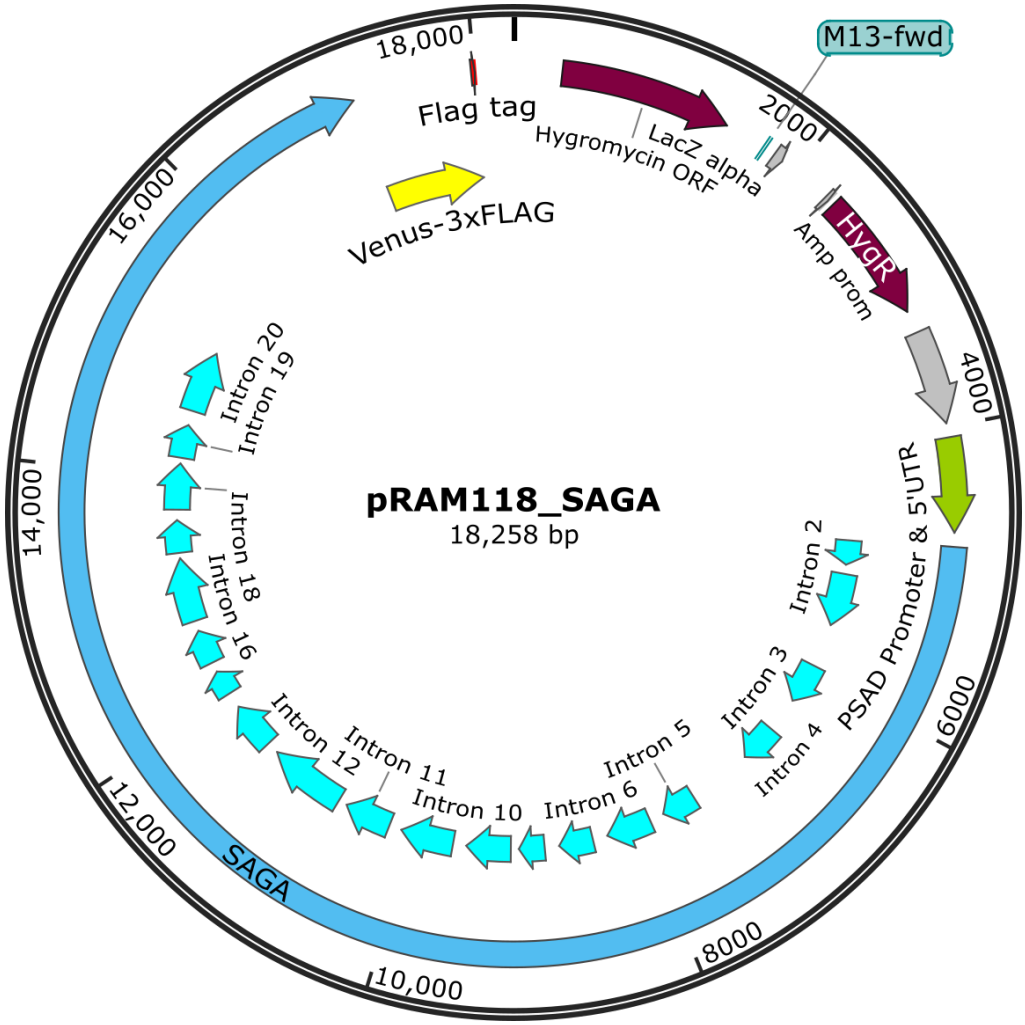



**Figure 5.9 Schematic representation of pRAM118_SAGA plasmid constructed by Alan Itakura, Stanford University, using Gibson assembly. The expression vector had a *psaD* promoter to drive constitutive expression of *SAGA-Venus-3xFLAG* protein and hygromycin resistance.**

In order to validate the expression of the tagged SAGA protein in the strains, Western blots using anti-GFP (Figure 5.10A) and anti-FLAG (Figure 5.10B) antibodies were performed using total soluble protein of cells grown under 5% (v/v) $CO_2$ and upon acclimation to 0.04% (v/v) $CO_2$ with Rubisco as loading control (Figure 5.10C). The SAGA-Venus-3xFLAG protein was expressed under both high- and low-$CO_2$ conditions in the complemented lines but was not detected in the wild type protein samples (Figure 5.10A & B). No changes to the amount of Rubisco were detected (Figure 5.10C).

**Table 5.1 Prediction of target and location of SAGA transit peptide using ChloroP, PredAlgo and TargetP suggested different locations and lengths.**

| | ChloroP | PredAlgo | TargetP |
|---|---|---|---|
| Target | Chloroplast | Mitochondria | Mitochondria |
| M score | NA | 1.7484547 | 0.566 |
| mTP length | NA | 165.12017 | NA |
| C score | 8.575 | 0.9437804 | 0.166 |
| cTP length | 73 | 2967.8654 | NA |
| SP score | NA | 0.0193037 | 0.000 |
| spTP length | NA | 80.525488 | NA |

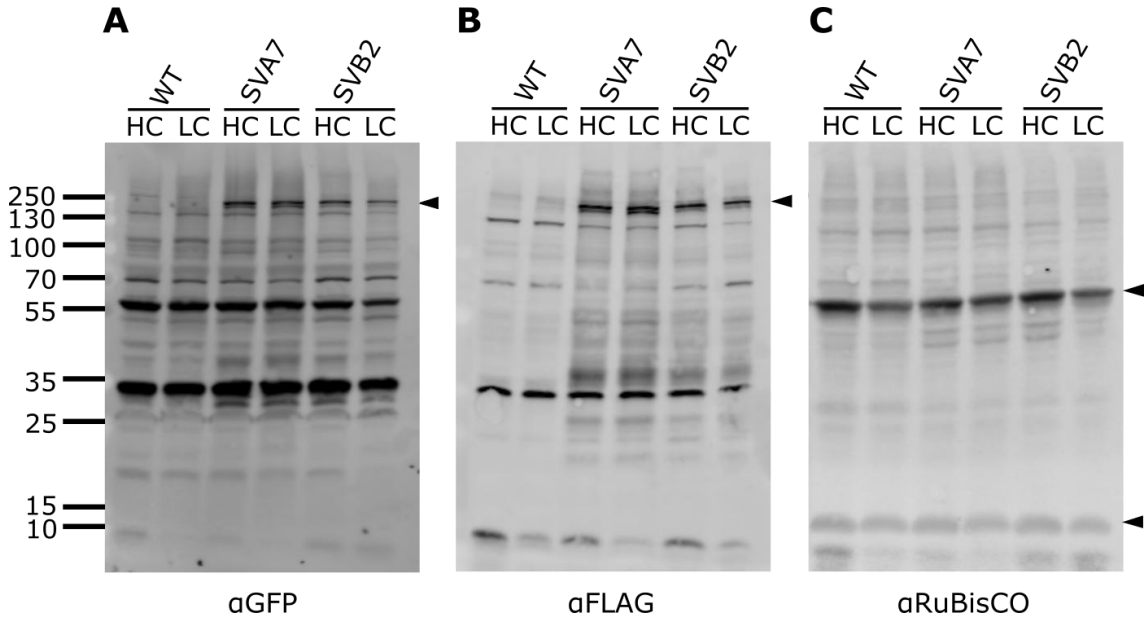



**Figure 5.10 Fluorescence Western blots of the wild type and the two complemented lines, SVA7 and SVB2, grown under 5% (v/v) $CO_2$ and 0.04% (v/v) $CO_2$ conditions with (A) anti-GFP and (B) anti-FLAG antibodies to validate that the *SAGA-Venus-3xFLAG* protein was expressed in the mutant (*saga*) transformed with pRAM118_SAGA. (C) Anti-Rubisco antibody was used as loading control. Arrows indicate correct band in each blot. SVA7, *saga::SAGA-Venus-3xFLAG A7*; SVB2, *saga::SAGA-Venus-3xFLAG B2*; HC, 5% (v/v) $CO_2$; LC, 0.04% (v/v) $CO_2$. Total soluble protein of 5 µg was loaded.**

The complemented strains, SVA7 and SVB2, were able to grow on TAP agar with 25 µg/ml (v/v) hygromycin and 20 µg/ml (v/v) paromomycin confirming that the complemented *saga* strains contained at least a copy of hygromycin resistance gene and retained the paromomycin resistance gene from the original mutagesis cassette (data not shown). Spot tests were used as initial growth test to confirm the restoration of CCM in the complemented lines (Figure 5.11). The controls in this experiment include the corresponding wild type, cMJ030 (Zhang et al., 2014) as positive control, and *saga* as negative control. A control was set up under heterotrophic conditions with cells grown on TAP agar in the dark to test whether the strains in this study possessed additional growth defects unrelated to $CO_2$ requirement (Figure 5.11A).

The cells were also grown on TP agar under high $CO_2$ with illumination of about 50 µmol photons $m^{-2}s^{-1}$ to assess light-adapted growth phenotypes. As the CCM would be repressed, growth of the mutant and the complemented lines should be indistinguishable from the corresponding wild type. However, the results showed that the complemented lines grew better than *saga* but not as well as the wild type (Figure 5.11B). In order to assess the restoration of functional CCM in SVA7 and SVB2, cells were grown on TP agar at 0.04% (v/v) $CO_2$ under similar light intensity. The growth of the corresponding wild type indicated the cells possessed functional CCM while *saga* displayed slower growth as demonstrated in Chapter 3. SVA7 and SVB2 grew better than *saga* but not as well as wild type (Figure 5.11C) suggesting the presence of a CCM. It was possible that the initial concentration of the wild type cells was slightly higher causing the growth difference between the wild type, *saga*, SVA7 and SVB2 under the different $CO_2$ conditions. When measured directly as photosynthetic efficiency, there was no difference in performance between the complemented lines and the wild type (Table 5.2).

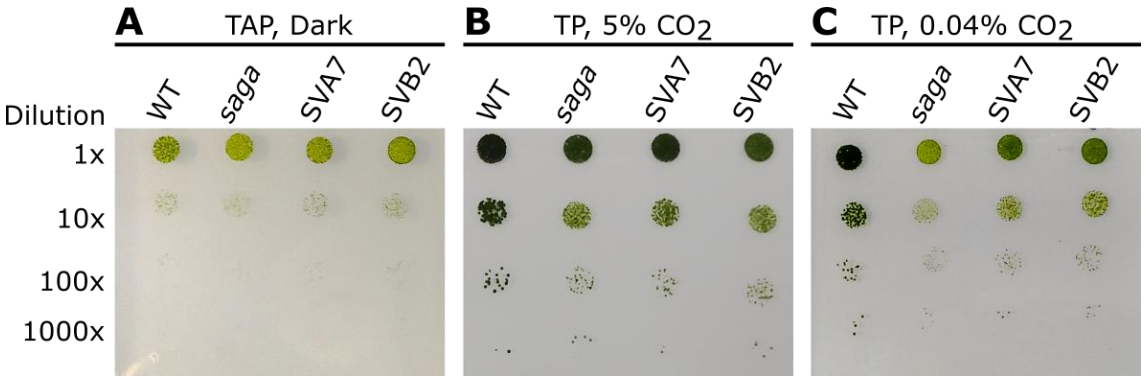



**Figure 5.11 Spot tests investigating the growth response of the wild type, cMJ030, *saga* and the complemented lines, SVA7 and SVB2. (A) Control TAP plate was kept in the dark at room temperature. TP plates were kept were in constant light of about 50 µmol photon $m^{-2}s^{-1}$ at room temperature either under (B) 5% (v/v) $CO_2$ or (C) 0.04% (v/v) $CO_2$ for 7 days. Complemented lines grew better than mutant under CCM induced condition indicating restoration of functional CCM.**

**Table 5.2 Photosynthetic $K_{0.5}$ (Ci) of wild type, *saga* and two complemented lines, *saga::SAGA-Venus-3xFLAG A7* (SVA7) and *saga::SAGA-Venus-3xFLAG B2* (SVB2) indicating that the complemented lines had similar affinity for inorganic carbon (Ci) to the corresponding wild type. Data are shown as average ± standard error, which were obtained from at least three independent experiments. Paired t-test was used to determine the significance compared to wild type. Ci, inorganic carbon; HC, high $CO_2$; LC, low $CO_2$; ****, p<0.0001; ***, 0.0001<p<0.01; *, 0.01<p<0.1; ns, p>0.1.**

| Sample | Conditions | $K_{0.5}$ (Ci) (µM) | p-value | Significance |
|---|---|---|---|---|
| WT | HC | 115.8 ± 24.2 | - | - |
| | LC | 38.7 ± 11.8 | - | - |
| *saga* | HC | 775.5 ± 88.7 | $1.0 \times 10^{-5}$ | **** |
| | LC | 160.0 ± 19.1 | 0.00015 | *** |
| SVA7 | HC | 267.4 ± 84.3 | 0.20929 | ns |
| | LC | 41.9 ± 26.3 | 0.91083 | ns |
| SVB2 | HC | 211.8 ± 71.3 | 0.41823 | ns |
| | LC | 40.7 ± 12.4 | 0.94406 | ns |

The rate of photosynthetic oxygen evolution under different concentrations of inorganic carbon was used to infer the existence of a functional $CO_2$-concentrating mechanism in the complemented lines (Figure 5.12 & Table 5.2). When the CCM was repressed, the photosynthetic $K_{0.5}$ (Ci) of wild type was about 116 µM Ci while *saga* had a $K_{0.5}$ of 776 µM Ci. The $K_{0.5}$ (Ci) values of SVA7 and SVB2 were at the range of 212 µM Ci to 267 µM Ci, which were higher compared to wild type but paired t-test suggested that the difference was insignificant. On the other hand, under CCM-induced conditions, the photosynthetic $K_{0.5}$ (Ci) value of wild type was about 40 µM Ci, which is within the range of 10-100 µM Ci (Wang et al., 2014) while *saga* showed higher photosynthetic $K_{0.5}$ (Ci) value of about 160 µM Ci. The $K_{0.5}$ (Ci) values of both SVA7 and SVB2 were 42 and 41 µM Ci, respectively, suggesting that the ability to concentrate inorganic carbon by the complemented lines was similar to wild type.

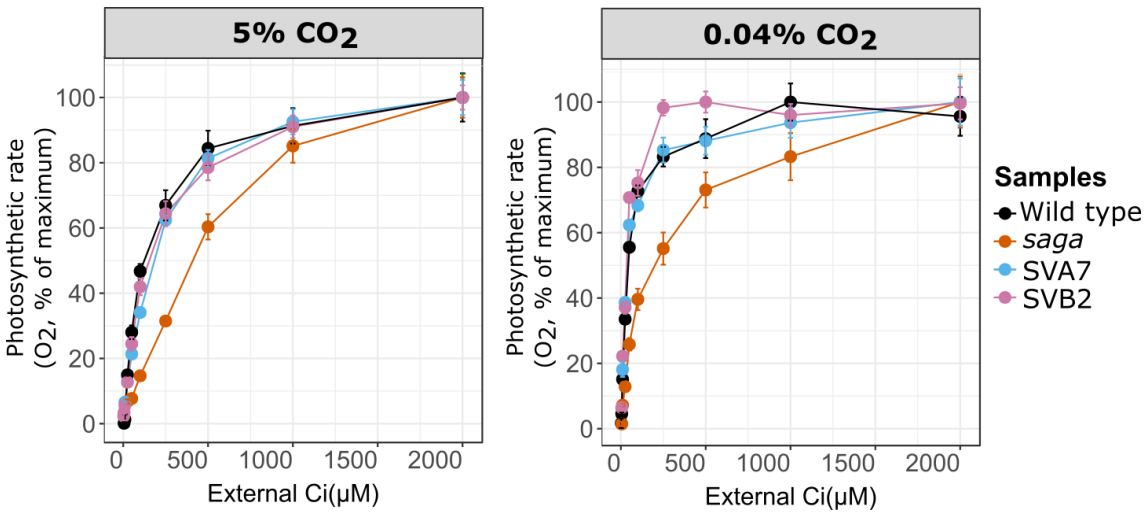



**Figure 5.12 Oxygen-evolving activity of WT, *saga* and the complemented lines in response to external dissolved inorganic carbon concentrations. All strains were grown under 5% $CO_2$ and 0.04% $CO_2$ conditions. Photosynthetic oxygen evolution was measured using a Clark-type oxygen electrode in the presence of different concentrations of $NaHCO_3$ and light intensity of about 300 µmol photons $m^{-2}s^{-1}$.**

The pyrenoid of wild type *C. reinhardtii* (Figure 5.13A, Appendix 3) consists of an electron-dense central matrix, where Rubisco is localised (K. Kuchitsu et al., 1988; Rawat et al., 1996; Mackinder et al., 2017) and the interconnected thylakoid tubule network (usually seen in cross section) which extends from the thylakoids into the pyrenoid through fenestrations in the starch sheath (Engel et al., 2015). The pyrenoid of *saga* displayed an electron-dense matrix and a starch sheath which did not fully

surround the pyrenoid matrix and with fewer fenestrations. In addition, there was an absence of a visible thylakoid tubule network within the pyrenoid (Figure 5.13B, Appendix 4). The complemented lines, SVA7 (Figure 5.13C, Appendix 6) and SVB2 (Figure 5.13D, Appendix 7), regained the typical features of a wild type pyrenoid, with visible presence of pyrenoidal tubule network and starch sheath which surrounded the pyrenoid. Electron micrographs of cells fixed under high $CO_2$, 1h after switching to ambient air (low $CO_2$) and 24h air acclimation showed obvious biophysical changes to the pyrenoid (Figure 5.14). During CCM induction, wild type, SVA7 and SVB2 possessed a thylakoid tubule network within the pyrenoid with the starch plates surrounding the pyrenoid increased in size, while no visible thylakoid nucleation site was observed in *saga* under the same conditions.

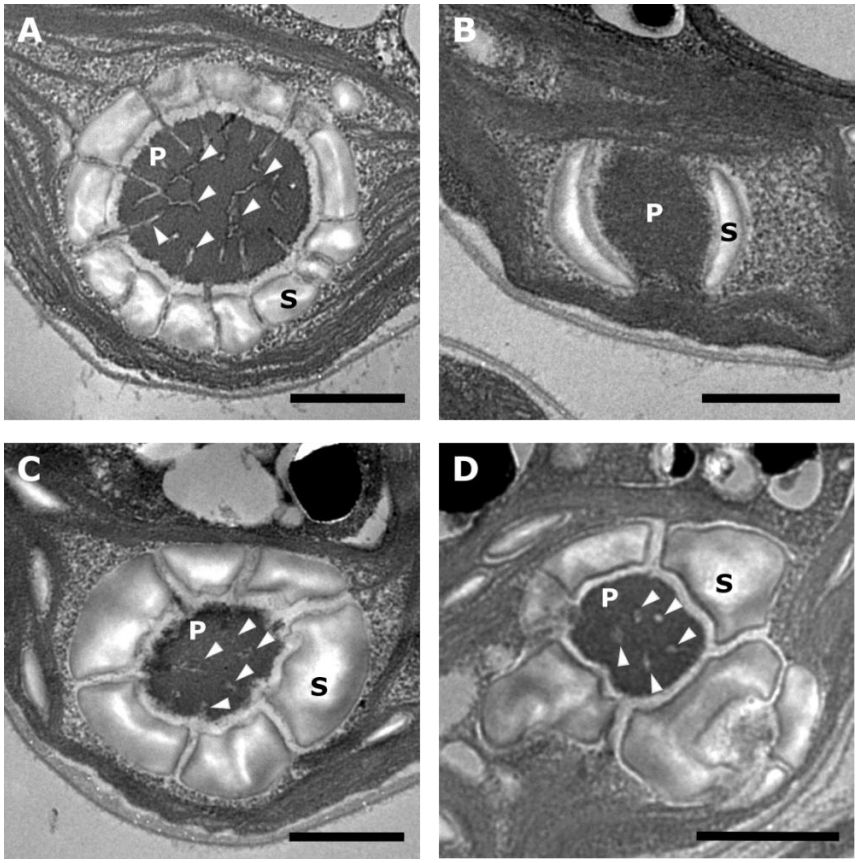



**Figure 5.13 The electron micrographs of the pyrenoid of (A) wild type, (B) *saga* and the two complemented lines, (C) SVA7 (*saga::SAGA-Venus-3xFLAG A7)* and (D) SVB2 (*saga::SAGA-Venus-3xFLAG B2*) grown photoautotrophically at room temperature with illumination of about 50 µmol photons $m^{-2}s^{-1}$ under ambient level of $CO_2$ for 24 hours. Both complemented lines showed restoration of thylakoid tubules within the pyrenoids. White arrows indicate thylakoid tubules in the pyrenoid. P, pyrenoid; S, starch plate. Scale bar 500 nm.**

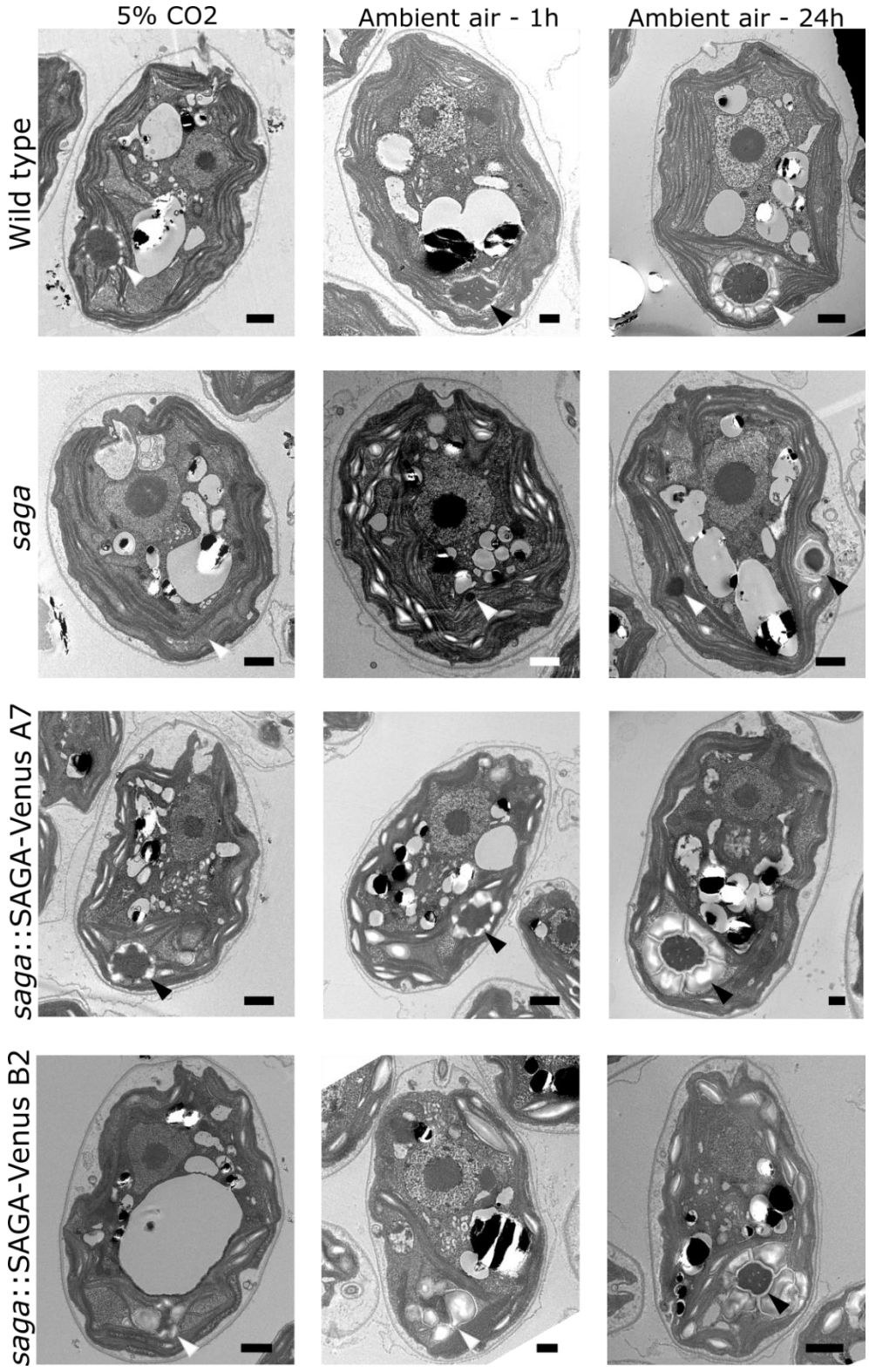



**Figure 5.14 Electron micrographs of the wild type, *saga* and the two complemented lines grown photoautotrophically under 5% $CO_2$ (high $CO_2$), 1h and 24h after acclimation to ambient air (low $CO_2$). All cells were grown at room temperature with illumination of about 50 µmol photons $m^{-2}s^{-1}$. Scale bar 500 nm.**

Using a sample size of 20 individual cells, the pyrenoid area of wild type, *saga*, SVA7 and SVB2 were measured to determine whether the restoration of the thylakoid tubule network within the pyrenoid would affect the overall size of the pyrenoid (Figure 5.15).When grown under CCM repressed conditions, the pyrenoid area of all samples including the corresponding wild type was on average below 1 $\mu m^2$ (Figure 5.15A). The pyrenoid area of *saga* was significantly different from that of the wild type with the average pyrenoid area about 0.04 $\mu m^2$ (p-value<0.0001) while the average pyrenoid area for SVA7 and SVB2 was 0.35 and 0.36 $\mu m^2$, respectively (Figure 5.15A). When CCM was induced, all cells showed increase in the size of the pyrenoid to the average range between 0.38 to 0.72 $\mu m^2$, though the pyrenoid area of *saga* remained significantly smaller than the wild type (0.72 $\mu m^2$), SVA7 (0.59 $\mu m^2$) and SVB2 (0.47 $\mu m^2$) (Figure 5.15B).

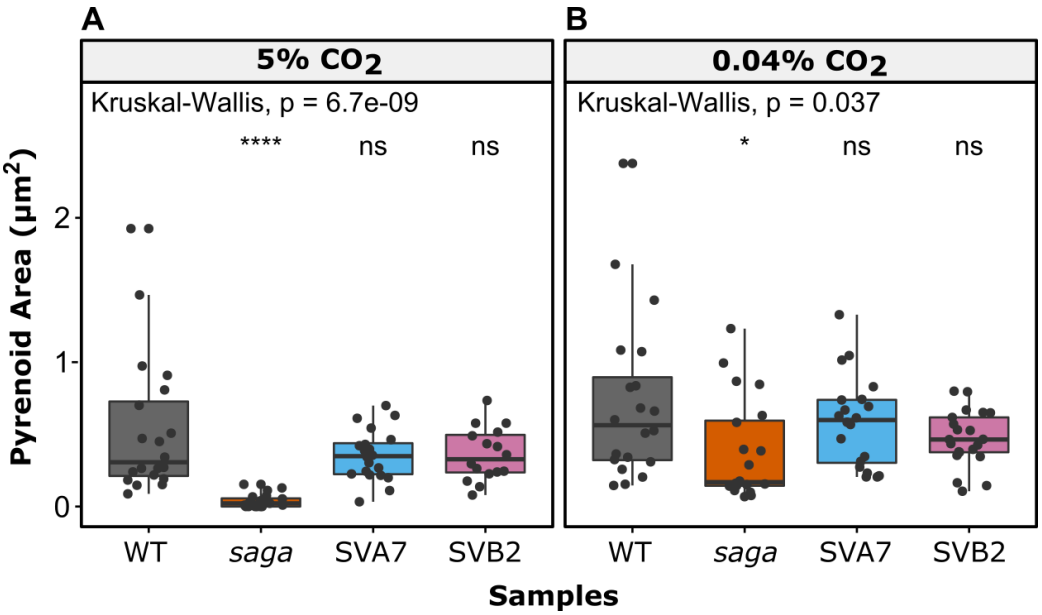



**Figure 5.15 Measurement of the pyrenoid area of wild type, *saga* and the two complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) under (A) 5% (v/v) $CO_2$ (high $CO_2$) and (B) upon 24h acclimation to 0.04% (v/v) $CO_2$ (low $CO_2$). Pairwise comparison using Mann-Whitney-Wilcoxon test showed that the complemented lines were not significantly different from wild type. ****, p<0.0001; ***, 0.0001<p<0.01; *, 0.01<p<0.1; ns, p>0.1.**

The general characterisation of *saga* under low $CO_2$ conditions in Chapter 3 showed that *saga* possessed fewer starch plates which had a reduced area. These phenotypic aberrances remained similar in cells grown under high $CO_2$ conditions (Figure 5.16A

and Figure 5.17A). Since some of the *saga* cells showed a lack of pyrenoid at high $CO_2$, the pyrenoid area was considered as zero in the measurement. The pyrenoidal starch area per pyrenoid of the complemented lines was significantly larger than the wild type under both high- and low $CO_2$ conditions (Figure 5.16) indicating the expression of SAGA-Venus-3xFLAG protein had affected the pyrenoidal starch area. The average starch area per pyrenoid of wild type under high $CO_2$ was about 0.18 $\mu m^2$ while SVA7 and SVB2 were about 0.52 and 0.69 $\mu m^2$ respectively. Mutant cells possessed pyrenoidal starch granules measured to be on average about 0.095 $\mu m^2$ (Figure 5.16A). Under low $CO_2$, wild type had a 5-fold increase of average starch area per pyrenoid to about 0.96 $\mu m^2$, while SVA7 had increase to 1.62 $\mu m^2$ and SVB2 to 1.42 $\mu m^2$. The mutant cells had a large increment of 21-fold to an average of 0.2 $\mu m^2$ starch area per pyrenoid when CCM was induced (Figure 5.16B)

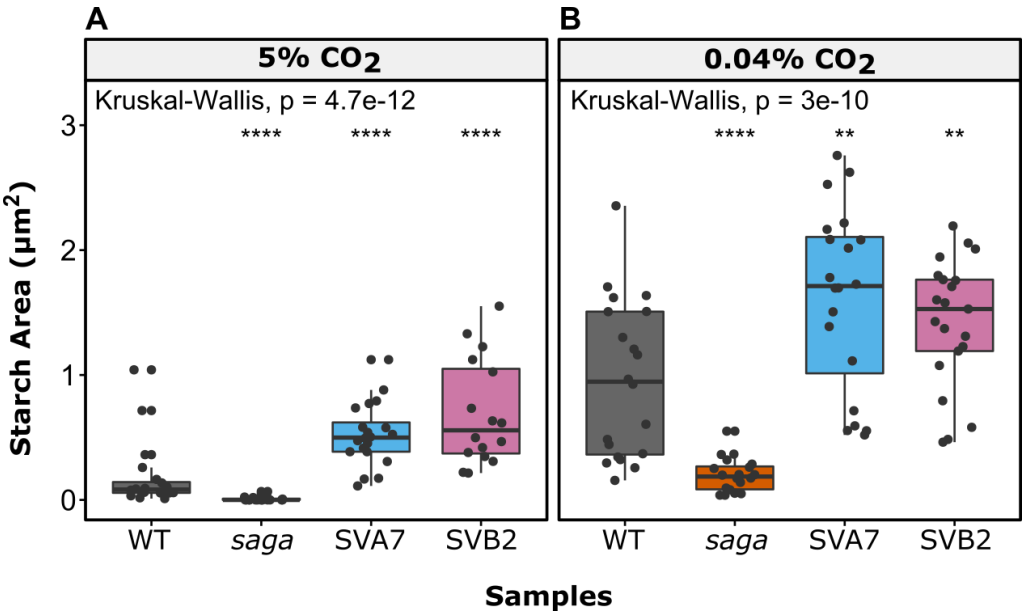



**Figure 5.16 Measurement of the starch area per pyrenoid of wild type, *saga* and the two complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) under (A) 5% (v/v) $CO_2$ (high $CO_2$) and (B) upon 24h acclimation to 0.04% (v/v) $CO_2$ (low $CO_2$). Pairwise comparison using Mann-Whitney-Wilcoxon test showed that the starch area per pyrenoid of *saga* and the complemented lines were significantly different from wild type. ****, p<0.0001; ***, 0.0001<p<0.01; *, 0.01<p<0.1; ns, p>0.1.**

In terms of the average number of starch plates per pyrenoid, wild type and the two complemented lines grown under high $CO_2$ were relatively similar to one another, with most cells possessing about 5 to 6 starch plates while there was an absence of starch

plates in most *saga* cells (Figure 5.17). Although the number of pyrenoidal starch plates of *saga* increased upon CCM induction, the number remained significantly lower than wild type (Figure 5.17B). The data also showed that the complemented lines had lower number of starch plates per pyrenoid compared to wild type under low $CO_2$ (Figure 5.17B). Interestingly, the number of starch plates per pyrenoid of the complemented lines remained relatively constant of around 5 to 6 starch plates per pyrenoid before and after CCM induction (Figure 5.17B).

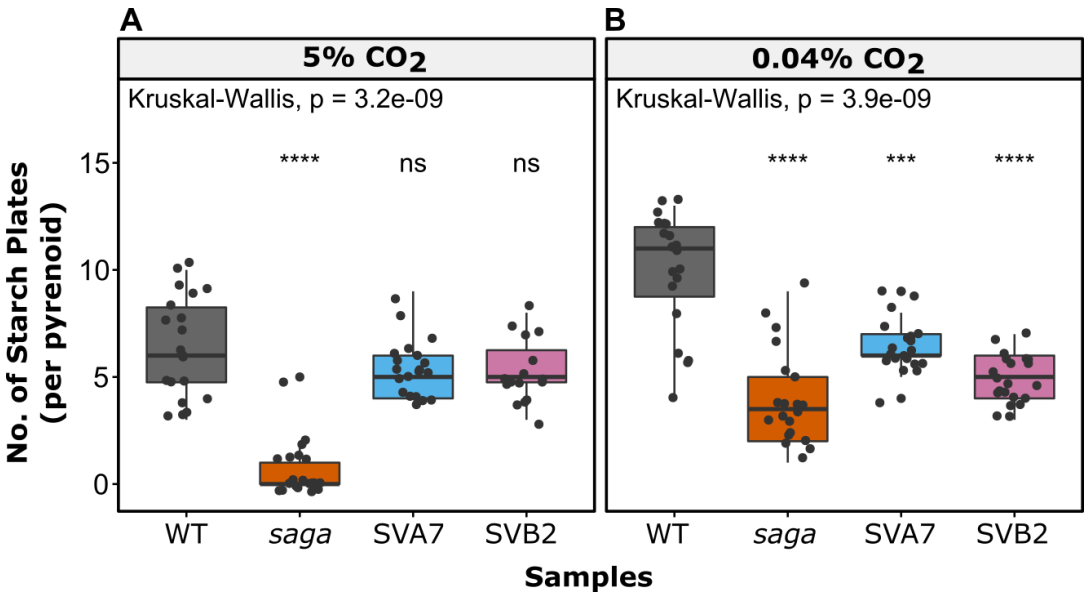



**Figure 5.17 Starch plates count per pyrenoid of wild type, *saga* and the two complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) under (A) 5% (v/v) $CO_2$ (high $CO_2$) and (B) upon 24h acclimation to 0.04% (v/v) $CO_2$ (low $CO_2$). Pairwise comparison using Mann-Whitney-Wilcoxon test showed that the complemented lines were not significantly different from wild type under 5% $CO_2$ but differed significantly under 0.04% $CO_2$. ****, p<0.0001; ***, 0.0001<p<0.01; *, 0.01<p<0.1; ns, p>0.1.**

The observation from the electron micrographs of wild type, *saga* and the two complemented lines inferred that SAGA played multiple roles in the structure of the pyrenoid: 1) to form a visible thylakoid tubule network, 2) to induce the growth of the pyrenoid during CCM induction and 3) to aid the formation of a starch sheath enveloping the pyrenoid matrix. Under the assumption that SAGA plays the role in the structural biogenesis of the pyrenoid, SAGA was hypothesised to localise in the pyrenoid.

Using the *saga::SAGA-Venus-3xFLAG* lines, SVA7 and SVB2, the subcellular localisation of SAGA was assessed using confocal microscopy, with wild type and *saga* as negative controls (Figure 5.18). No Venus signal was observed in the controls, while fluorescent signals of SAGA-Venus protein were detected in the pyrenoid of both SVA7 and SVB2 cells acclimated for 24h under CCM induced conditions, appearing as clusters of signal dispersed along the thylakoid tubules (Figure 5.18). This supported the hypothesis that SAGA plays a role in the structural formation of the pyrenoid.

In order to assess the localisation of SAGA in wild type cells and to ensure that SAGA localisation signal detected in the complemented lines was not an artefact resulting from overexpression, indirect immunofluorescence assay was performed. The location of the pyrenoid was determined with an anti-Rubisco antibody as a control assay (Figure 5.19). Under high $CO_2$ conditions (Figure 5.19A), Rubisco was localised in the pyrenoid of wild type, SVA7 and SVB2, though a weak fluorescence signal could also be observed within the whole cell indicating that not all Rubisco was aggregated into the pyrenoid under high $CO_2$ conditions, similar to the IF results of Caspari (2015). Rubisco in *saga* was dispersed around the whole cell without an obvious aggregation site (Figure 5.19A). Under low $CO_2$ conditions, most Rubisco was localised in the pyrenoid for all cell lines, except in *saga*, where the signal was dispersed in a few locations in each cell (Figure 5.19B). This was consistent with the suggestion of Caspari (2015) that some of the CCM mutants with pyrenoidal defects displayed delocalised pyrenoids and/or additional Rubisco aggregates.

The subcellular localisation of SAGA *in vivo* was performed using an indirect immunofluorescence assay with an anti-SAGA antibody (Figure 5.20). Under both high and low $CO_2$, the fluorescent signal in wild type were mainly clustered within the pyrenoid (Figure 5.20) similar to the Venus fluorescent signals of the confocal images of the complemented lines in Figure 5.18. The fluorescent signals of SAGA in the two complemented lines were aggregated in a manner similar to wild type under both growth conditions indicating restoration of SAGA localisation. In contrast, the *saga* mutant showed a diffuse distribution of fluorescence signal throughout the chloroplast stroma under both high and low $CO_2$ conditions. One of the cells clearly showed a starch sheath microcompartment where a pyrenoid would be expected to form (Figure 5.20). There was, however, no clear fluorescence signal associated with that formation.

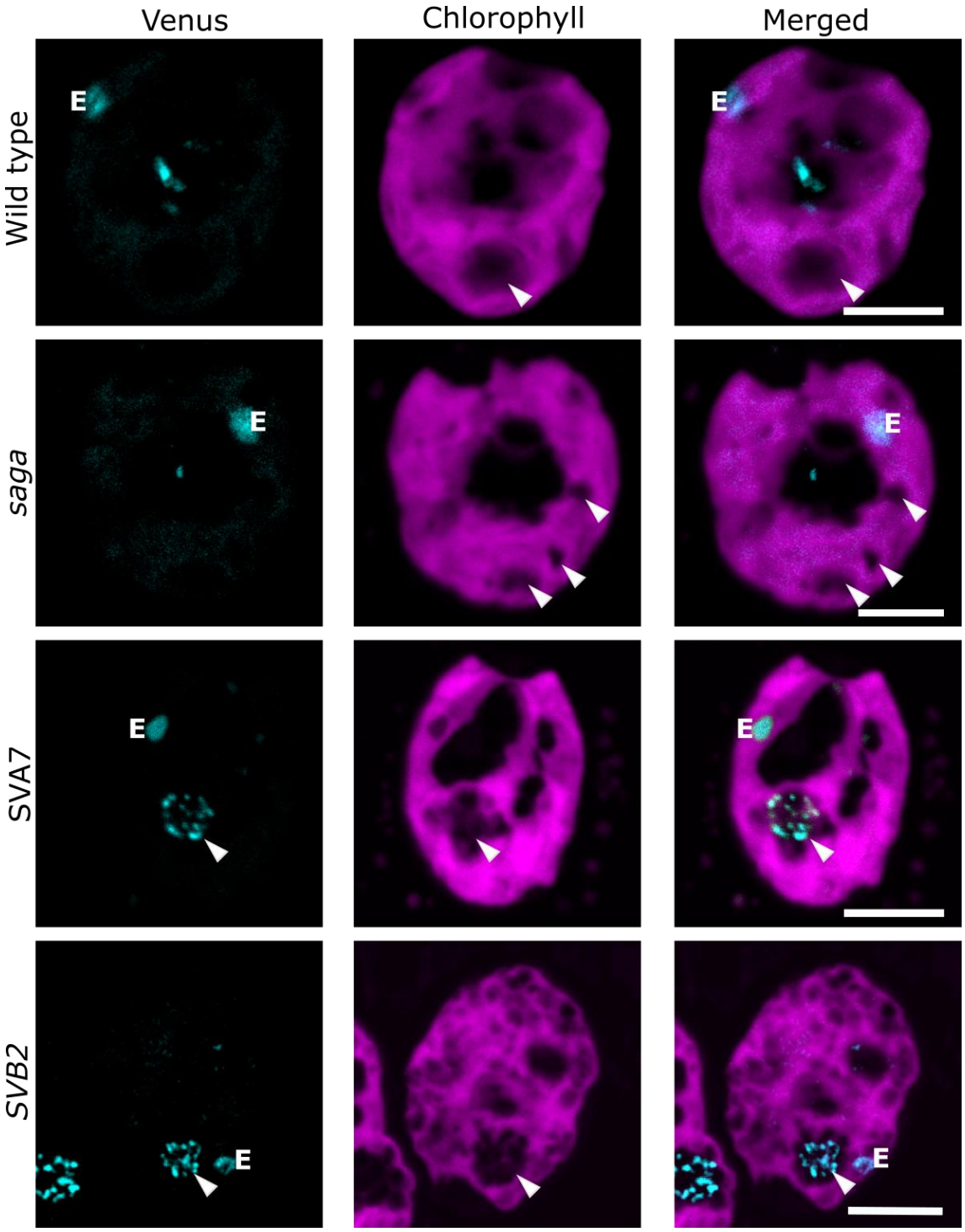



**Figure 5.18 SAGA-Venus-3xFLAG protein was localised within the pyrenoid of the complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) after 24h acclimation to ambient air (low $CO_2$) to induce CCM with wild type and *saga* as controls. All cells were grown at room temperature with illumination of about 50 µmol photons $m^{-2}s^{-1}$. E: eyespot. White arrows indicated the location(s) of the pyrenoid(s). Scale bar 5 µm.**

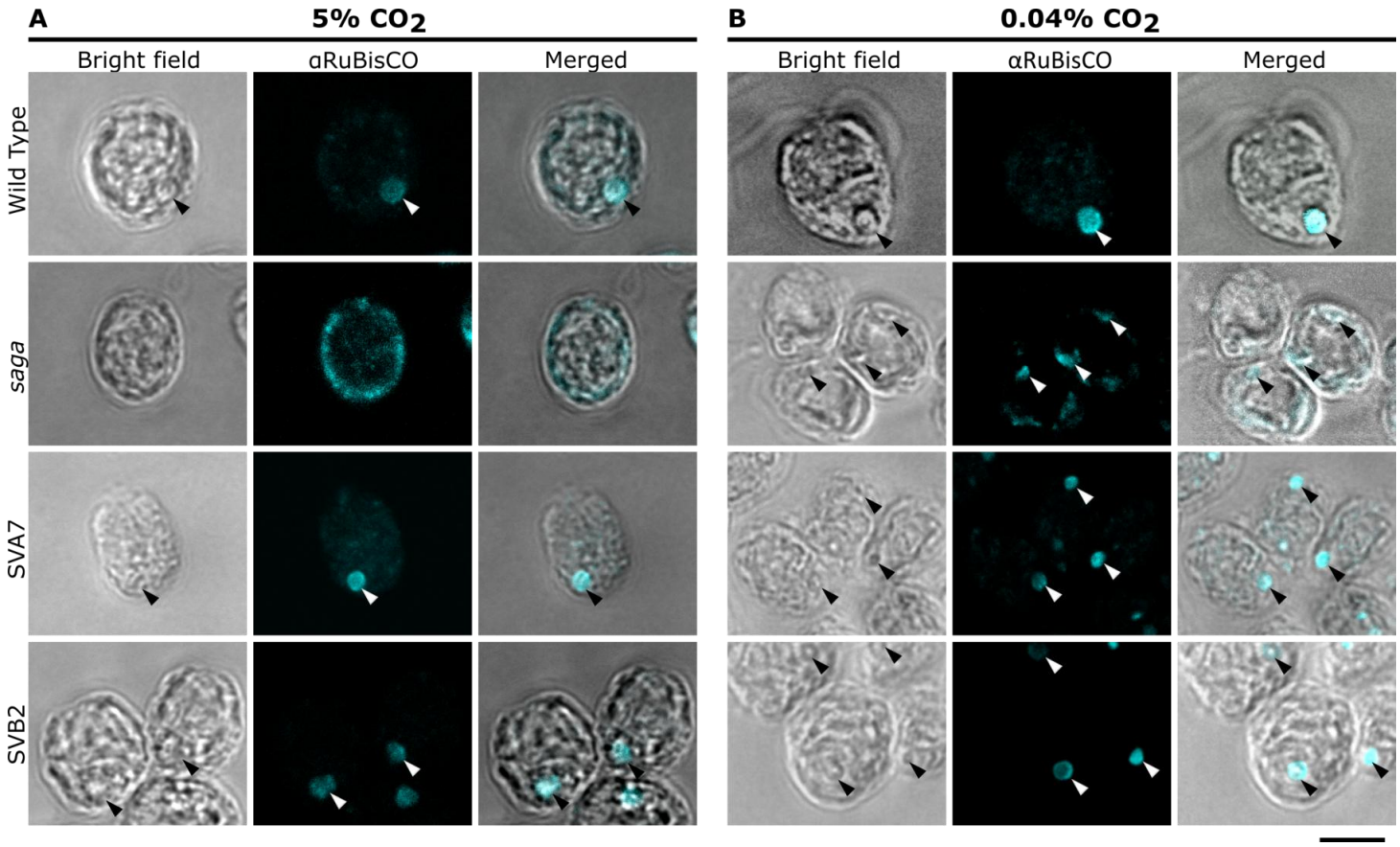



**Figure 5.19 Subcellular localisation of Rubisco in wild type, *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) by indirect immunofluorescence assay with an anti-Rubisco antibody. All cells were grown in photoautotrophically under (A) 5% $CO_2$ (high $CO_2$) and (B) acclimated for 24 h under 0.04% $CO_2$ (low $CO_2$). Arrows show suggested position of the pyrenoid. Scale bar 5 µm.**

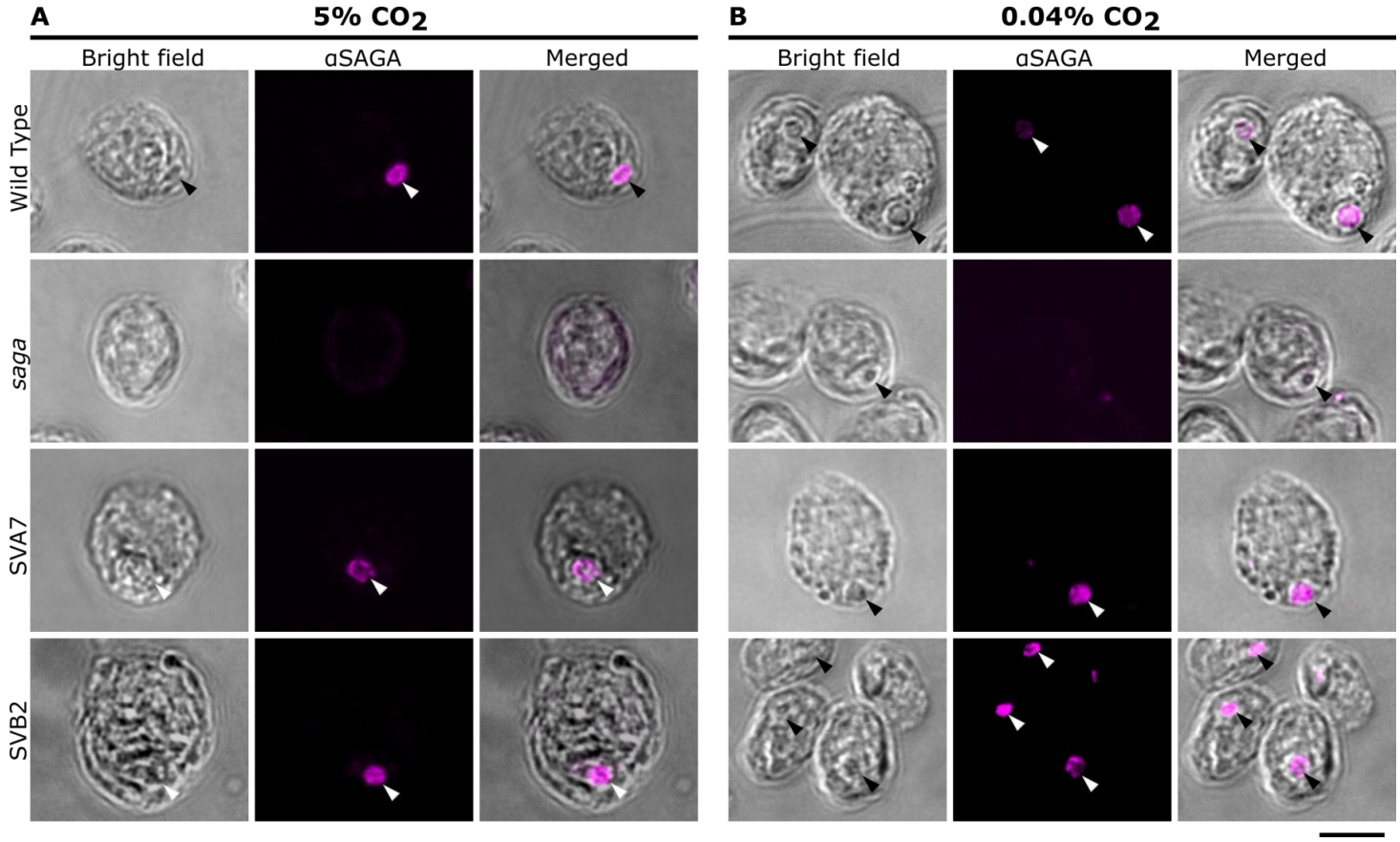



**Figure 5.20 Subcellular localisation of SAGA in wild type, *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) by indirect immunofluorescence assay with anti-SAGA antibody. All cells were grown photoautotrophically under (A) 5% $CO_2$ (high $CO_2$) and (B) acclimated for 24 h under 0.04% $CO_2$ (low $CO_2$). Arrows show suggested position of the pyrenoid. Scale bar 5 µm**

### 5.2.3 .Candidate SAGA Protein Interactors

To identify potential protein interaction partners, total soluble proteins of wild type, *saga*, SVA7 and SVB2 were isolated from cells grown under high $CO_2$ and after 24 h acclimation to ambient air and subjected to native PAGE (Figure 5.21A) and subsequent Western blots with anti-GFP (Figure 5.21B), anti-EPYC1 (Figure 5.21C) and anti-Rubisco (Figure 5.21D) antibodies. Western blot with anti-GFP antibody showed that SAGA-Venus-3xFLAG protein formed two high molecular weight complexes (Figure 5.21B) between 480 and 1048 kDa. On the other hand, Western blots probed with anti-EPYC1 and anti-Rubisco antibodies showed that Rubisco and EPYC1 were co-migrating to the range between 480 and 720 kDa (Figure 5.21C & D) which supported the results of Mackinder et al. (2016) that Rubisco and EPYC1 interacts with each other.

In order to gain further understanding of the biological function of SAGA, coimmunoprecipitation was performed to identify proteins that interact with SAGA. For this, native total soluble protein of wild type, *saga* and the two complemented lines, SVA7 and SVB2 were incubated with ChromoTek GFP-Trap® agarose beads. Western blots of the eluates using anti-GFP (Figure 5.22A) and anti-SAGA (Figure 5.22B) antibodies showed the specificity of GFP-Trap agarose beads binding to the Venus tag on the SAGA-Venus-3xFLAG protein, whereas no trace of SAGA was observed in the eluate from the *saga* mutant. Total soluble protein was used as comparison to confirm the enrichment SAGA protein in elute of GFP-Trap.

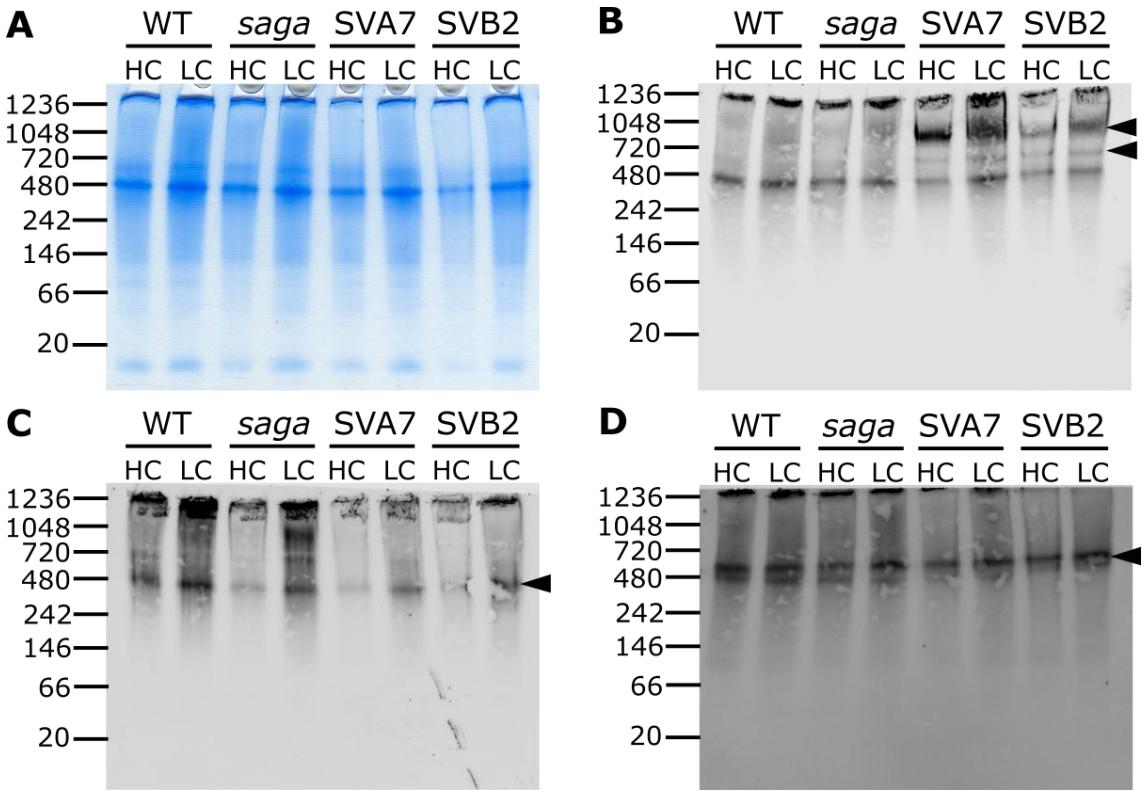



**Figure 5.21 (A) Native PAGE of total soluble protein of wild type, *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*). (B) Western blot using anti-GFP antibody showed that SAGA-Venus-3xFLAG protein formed high molecular weight bands (black arrows) with wild type and *saga* as negative controls. Western blots using (C) anti-EPYC1 and (D) anti-Rubisco antibodies showed that these proteins were complexes of about the same molecular weight but did not form complexes with SAGA-Venus-3xFLAG.**

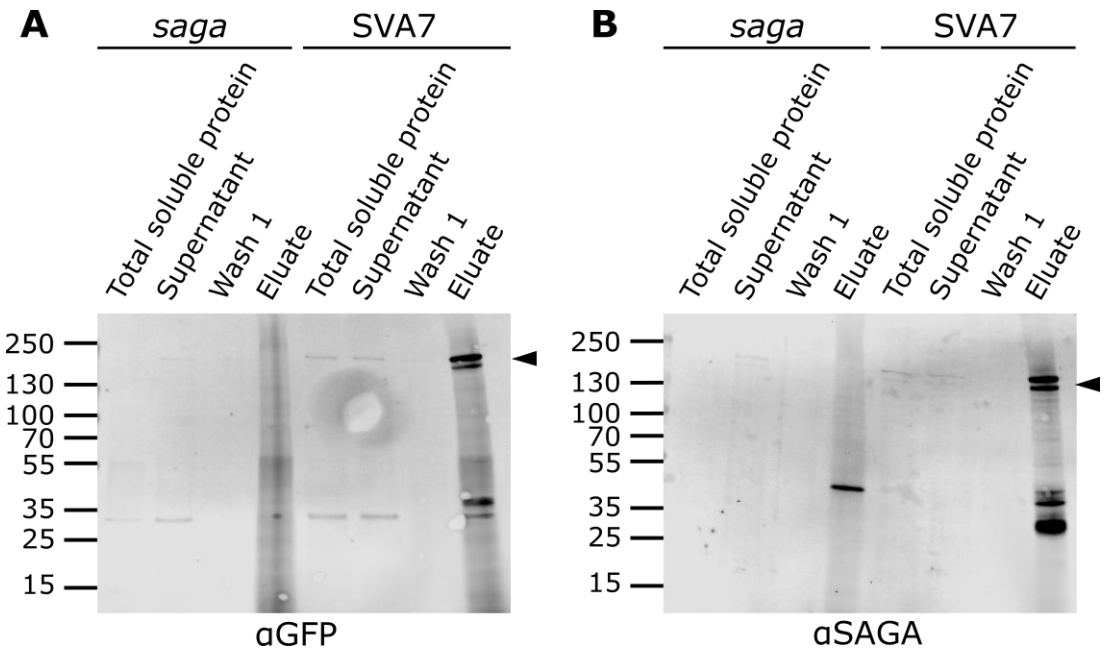



**Figure 5.22 Example of Western blots using (A) anti-GFP and (B) anti-SAGA antibodies to confirm that SAGA-Venus-3xFLAG protein was pulled down only in the complemented lines during the co-immunoprecipitation assay performed with GFP-Trap_A (ChromoTek, Germany). Arrows show the size of the enriched SAGA-Venus-3xFLAG protein.**

A co-immunoprecipitation (CoIP) assay coupled with LC-MS/MS analysis was used to identify candidate protein interactors of SAGA during CCM induction. Total soluble protein of wild type, *saga* and the two complemented lines, SVA7 and SVB2, were used for CoIP assay usng GFP-Trap agarose beads (ChromoTek, Germany) and the eluates were submitted for LC-MS/MS analysis in the Cambridge Centre for Proteomics. A total of 16 proteins were identified in at least two out of three biological replicates of the complemented lines (Figure 5.23), including SAGA (Cre11.g467712) and green fluorescent protein (GFP) but undetected in the wild type and *saga* indicating a high possibility of interaction with SAGA (Table 5.3). Out of these proteins, only five proteins were present in all three of the biological replicates of the complemented lines: protein disulfide isomerase 2 (Cre01.g033550), a serine/threonine kinase domain (Cre12.g549750), Cre12.g544450, Cre13.g566700 and Cre18.g750047 (Table 5.3). Nine other proteins were detected from at least two out of three biological replicates of each complemented lines. From the total of 14 potential candidate interactors, three were predicted to be transmembrane proteins using Phobius and PHYRE2 prediction servers (Figure 5.24). Despite the detection of CCM-related proteins known to localise in the pyrenoid region, such as Rubisco small subunits Cre02.g120100/Cre02.g120150),

Rubisco activase (Cre04.g229300), LCIB (Cre10.g452800), LCIC (Cre06.g307500), EPYC1/LCI5 (Cre10.g436550), these proteins were present in all the samples including the controls thus eliminating the potential that SAGA was interacting with them. The carbonic anhydrase known to localise in the pyrenoidal thylakoid network, CAH3 (Sinetova et al., 2012), was detected albeit in only one replicate in each of the complemented lines but not in wild type and *saga* thus was not considered a potential protein interactor of SAGA.

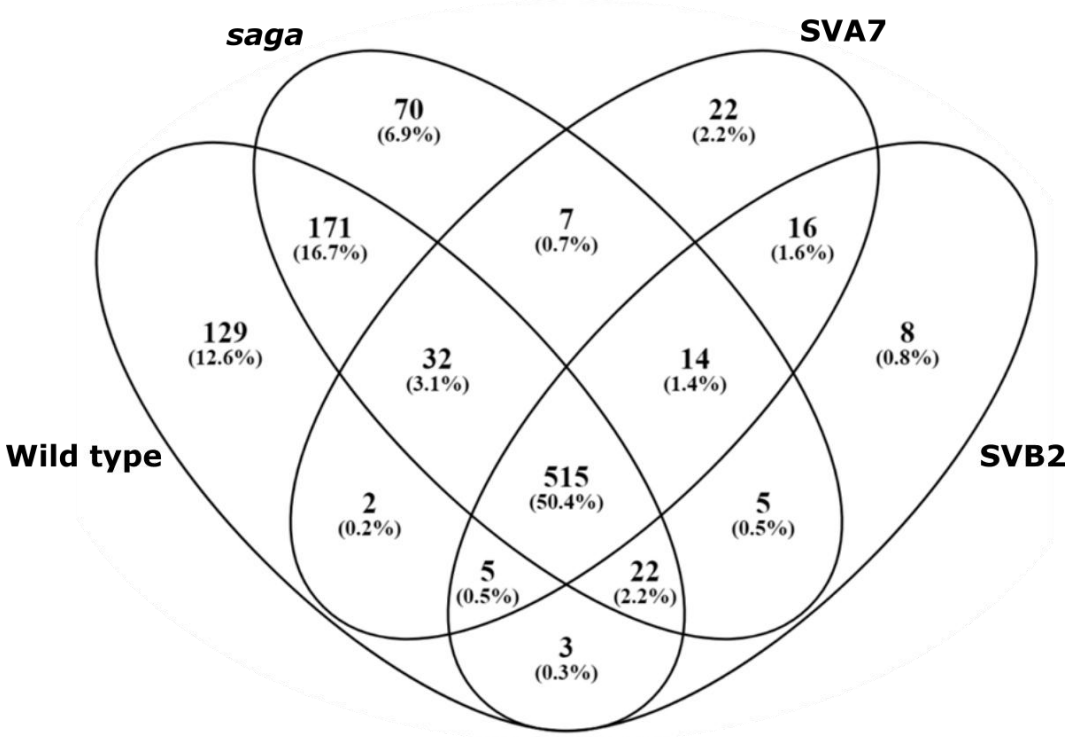



**Figure 5.23 Venn diagram showed the number of proteins identified from the co-immunoprecipitation assay using GFP-Trap_A. A total of 16 proteins (equivalent to 1.6% of the total proteins detected) were only found in all three replicates of the two complemented lines, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) with wild type and *saga* as negative controls.**

**Table 5.3 Summary list of potential protein interactors of SAGA identified by LC-MS/MS of the GFP-Trap_A coimmunoprecipitation assay. Three biological replicates of wild type, *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) were grown photoautotrophically under 5% $CO_2$ (high $CO_2$) for three days and then acclimated for 24 h under 0.04% $CO_2$ (low $CO_2$) for CCM induction. SAGA and green fluorescent protein were identified in all the replicates of the complemented lines but none was detected in the negative controls. 0, absent in all biological replicates; 1, present in one biological replicate; 2, present in two biological replicates; 3 present in all biological replicates.**

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | saga | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| Cre01.g033550.t1.2 | Cre01.g033550.t1.1 | GO:0005783, GO:0045454 | Protein Disulfide Isomerase 2 (CrPDI2), PDI-D subfamily, PMID: 16143836, circadian expressed, possible targeting to the ER. | *PDI2* | 27 | 0 | 0 | 3 | 3 |
| Cre11.g467712.t1.1 | Cre18.g743450.t1.1 | GO:2001070 | | *SAGA* | 168 | 0 | 0 | 3 | 3 |
| Cre12.g544450.t1.2 | Cre12.g544450.t1.1 | | No functional annotation | | 81 | 0 | 0 | 3 | 3 |

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | saga | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| Cre12.g549750.t1.2 | Cre12.g549750.t1.1 | GO:0006468, GO:0005524, GO:0004672 | Ser-/Thr-kinase domain; homology to casein kinase 1 of higher plants and animals where it is involved in phosphorylation of components of the circadian clock; found in the flagellar proteome. Previously annotated as CK1 | *CK1 / CSK1* | 38 | 0 | 0 | 3 | 3 |
| Cre13.g566700.t1.1 | | GO:0003676, GO:0017069, GO:0000398 | U2 small nuclear ribonucleoprotein B | *SNRPB2* | 25 | 0 | 0 | 3 | 3 |
| Cre18.g750047.t1.1 | Cre31.g780700.t1.1 | | Flagella membrane glycoprotein, minor form. . FMG1-A; belongs on chr9. This gene model is longer than Genbank entry AY2067 | *FMG1-A* | 447 | 0 | 0 | 3 | 3 |

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | saga | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| Cre01.g050850.t1.2 | Cre01.g050850.t1.1 | GO:0016787 | Serine/threonine protein phosphatase 1 Beta (PP1b); has a Kelch repeat at the N-terminus; found in the flagellar proteome [PMID: 15998802]. This is a calcineurin-like phosphoesterase. Previously annotated as PKL1 | | 96 | 0 | 0 | 3 | 2 |
| Cre07.g322176.t1.1 | | GO:0006351, GO:0003899, GO:0003677, GO:0005634, GO:0032549 | DNA-directed RNA polymerase I subunit RPA2 | *RPA2*, *POLR1B* | 114 | 0 | 0 | 3 | 2 |

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | saga | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| Cre12.g511050.t1.2 | Cre12.g511050.t1.1 | | A structural HSP40 (DnaJ) in the radial spoke stalk [PMID: 16507594, PMID: 15563613; GI:56404248] | *RSP16* | 39 | 0 | 0 | 3 | 2 |
| Cre02.g104650.t1.1 | | GO:0003676 | nucleolar protein 4 | *NOP4 / RBM28* | 126 | 0 | 0 | 2 | 3 |
| Cre03.g190850.t1.2 | Cre03.g190850.t1.1 | | EnoylCoA hydratase/isomerase; most similar to animal isoforms, probably mitochondrial. Identified by comparative genomics as being present only in organisms having motile (MOT) cilia | *ECH1* | 30 | 0 | 0 | 2 | 2 |
| Cre04.g218150.t1.2 | Cre04.g218150.t1.1 | GO:0003824, GO:0006470, GO:0004722 | Serine/threonine protein phosphatase | *PPP18* | 45 | 0 | 0 | 2 | 2 |

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | saga | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| Cre05.g238687.t1.1 | | | Pherophorin | | 83 | 0 | 0 | 2 | 2 |
| Cre07.g325762.t1.1 | Cre07.g325750.t1.2 | GO:0005515, GO:0005524 | ATPase family associated with various cellular activities (AAA) | *DRC11* | 95 | 0 | 0 | 2 | 2 |
| Cre12.g522550.t1.2 | Cre12.g522550.t1.1 | GO:0008017, GO:0007018, GO:0005524, GO:0003777, GO:0005871 | Kinesin motor protein, kinesin-2 (KRP85/95) family. Anterograde motor for intraflagellar transport. Previously annotated as FLA8. Upregulated upon deflagellation | *FLA8* | 82 | 0 | 0 | 2 | 2 |
| sp\|cRAP032\|P42212 \|GFP_AEQVI | | | Green fluorescent protein | | | 0 | 0 | 3 | 3 |

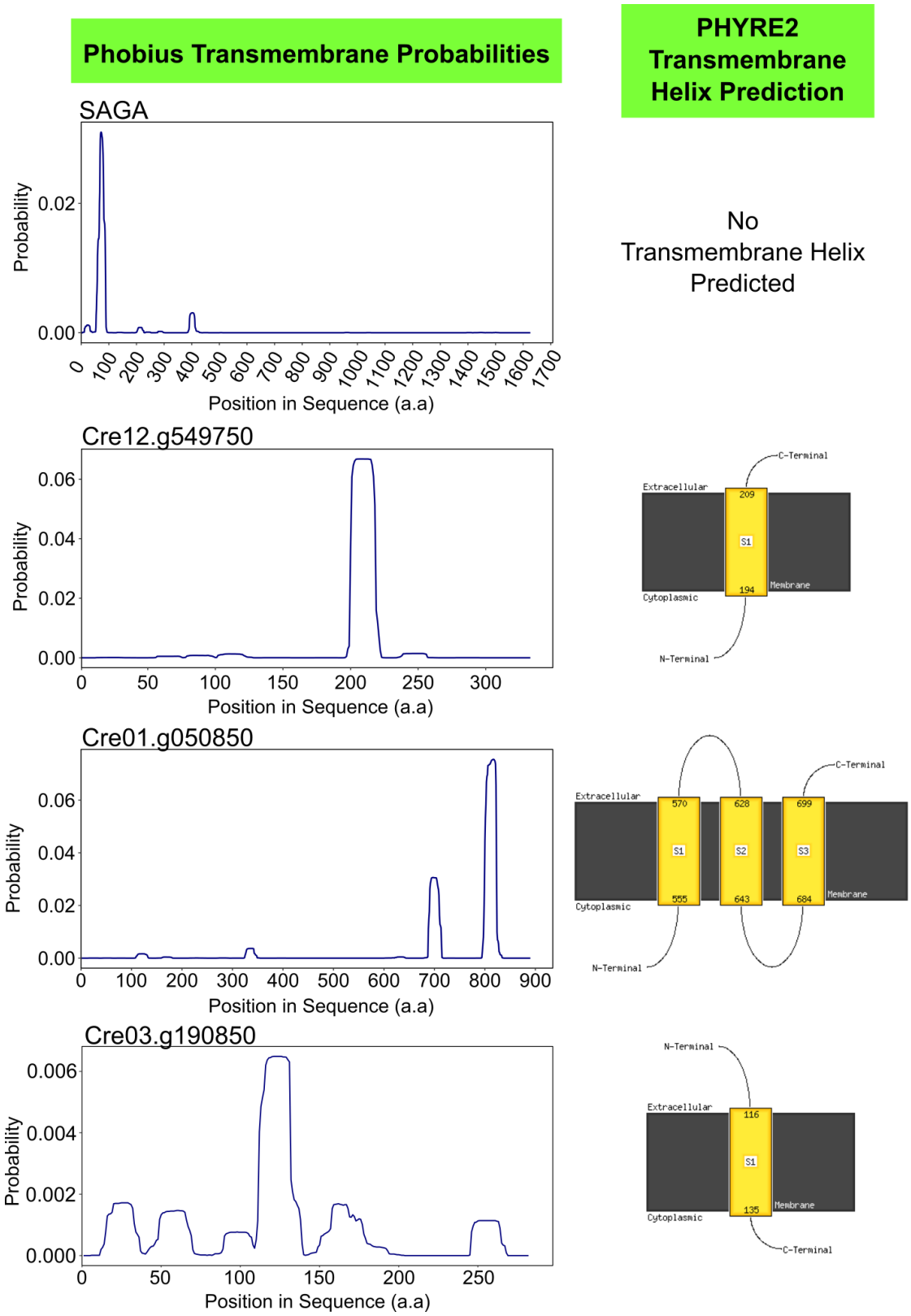



**Figure 5.24 Phobius and PHYRE2 predicted that while SAGA was not a transmembrane protein, three out of the 14 candidate protein interactors, Cre12.g549750, Cre01.g050850 and Cre03.g190850, possessed transmembrane helices.**

Since SAGA was localised in the pyrenoid under both high and low $CO_2$ conditions as demonstrated by the aggregation of SAGA-Venus-3xFLAG protein in the complemented lines (Figure 5.18) and indirect immunofluorescence with anti-SAGA antibody in wild type cells of *C. reinhardtii* (Figure 5.20), it was hypothesised that post-translational modifications could be involved in the mechanism of SAGA remobilisation during CCM induction. Post-translational modifications (PTMs) are known to be involved in many aspects of the CCM pathway, such as the phosphorylation of CAH3 during the shift from high to low $CO_2$ (Blanco-Rivero et al., 2012), the proposed PTM involved in dissociation/association of LCIB/LCIC complex (Y. Wang and Spalding, 2014; Yamano et al., 2014; Yingjun Wang and Spalding, 2014; Wang et al., 2015) and the suggested PTM in the C-terminus of CIA5/CCM1 (Fukuzawa et al., 2001; Xiang et al., 2001). Bioinformatics analysis using NetPhos 3.1a server (Blom et al., 1999, 2004) predicted multiple phosphorylation sites in SAGA protein (Figure 5.25). SAGA protein sequence was rich in serine, threonine and tyrosine residues known to be phosphorylated in eukaryotic proteins (Cieśla et al., 2011).

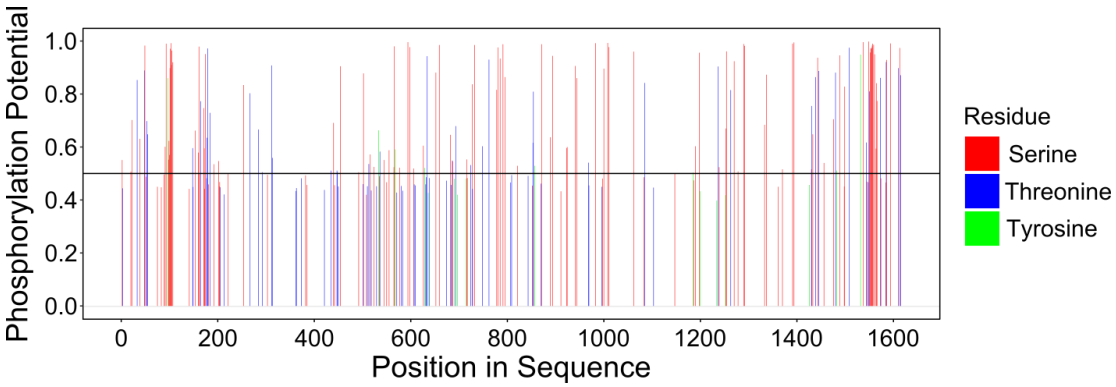



**Figure 5.25 NetPhos 3.1a predicted multiple phosphorylation sites in SAGA protein.**

A preliminary experiment to test the hypothesis that SAGA undergoes phosphorylation during CCM induction was performed using Pro-Q® Diamond phosphoprotein gel stain (Invitrogen, USA) on a SDS-PAGE gel of total soluble protein samples of wild type, *saga*, SVA7 and SVB2 (Figure 5.26, all lanes labeled "ProQ"). The total protein of each sample was visualised upon subsequent staining with SYPRO® Ruby protein gel stain (Invitrogen, USA) (Figure 5.26, all lanes labeled "Sypro"). A band of about 168 kDa was observed in both phosphoprotein- and total protein-stained lane of wild type protein sample, which was according to the ExPASy prediction (Gasteiger et al., 2005) of untagged SAGA protein size (Figure 5.26A). A band of about 180 kDa was observed in

both stains for the complemented lines (Figure 5.26), consistent with the additional of mVenus-3xFLAG tag, whilst no equivalent band was observed for *saga* (Figure 5.26B). The presence of a band on the ProQ lanes for wild type and the complemented lines (Figure 5.26) could be correlated with the proteomics data that there was a potential phosphorylation event in the C-terminus of SAGA protein (Table 5.4). However, this set of results needs to be validated in future work which is focused on post-translational modifications as the phosphorylation data (Table 5.4) was extracted from the raw data of the GFP-Trap elutes. The protein samples used for proteomics analysis were prepared only with the addition of phosphatase inhibitors targeting tyrosine/alkaline phosphatases and serine/threonine phosphatases. Although Phot protein (Cre03.g199000) was detected in all samples, a phosphorylation event was only detected in the two out of three biological replicates of each complemented line and none in the controls (Table 5.4).

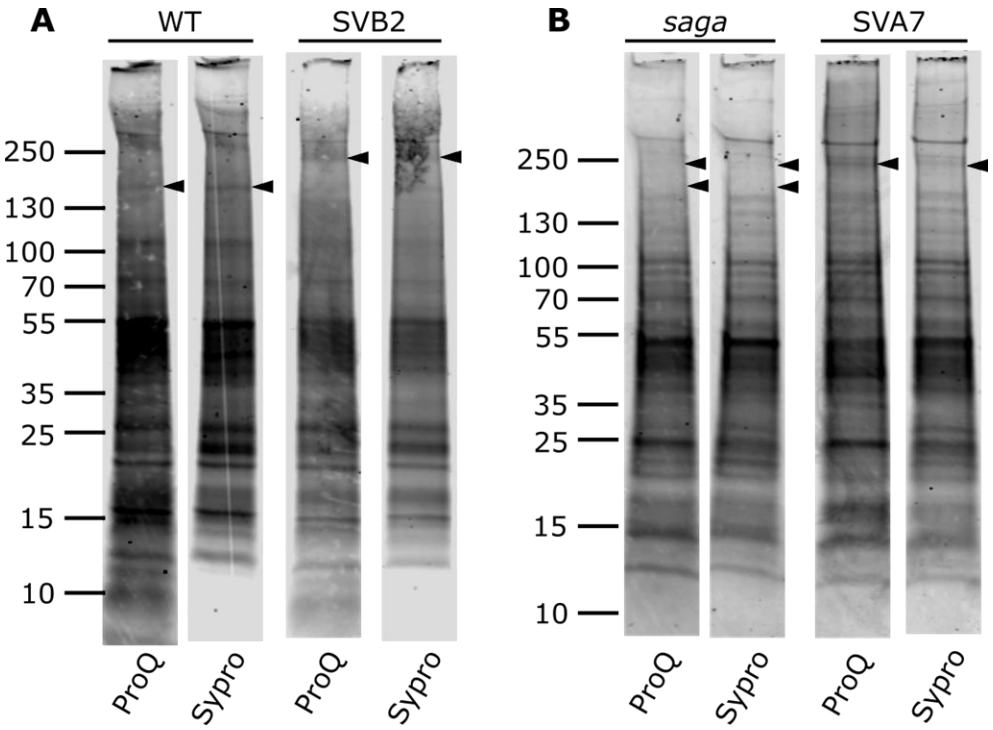



**Figure 5.26 PAGE gel of total soluble proteins of (A) wild type, SVB2 (*saga::SAGA-Venus-3xFLAG B2*), (C) *saga* and SVA7 (*saga::SAGA-Venus-3xFLAG A7*) were treated first with Pro-Q® Diamond phosphoprotein gel stain to observe phosphorylated proteins and then with SYPRO® Ruby total protein stain for comparison to total soluble protein. Arrows indicate SAGA or SAGA-Venus-3xFLAG proteins on the sample lanes of wild type and the complemented lines, respectively. The two arrows by the sample lanes of *saga* total soluble protein indicate the possible locations where the band should be.**

**Table 5.4 List of phosphorylated proteins identified in the samples by LC-MS/MS of the GFP-Trap_A coimmunoprecipitation assay. Three biological replicates of wild type, *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) were grown photoautotrophically under 5% $CO_2$ (high $CO_2$) for three days and then acclimated for 24 h under 0.04% $CO_2$ (low $CO_2$) for CCM induction. 0, absent in all biological replicates; 1, present in one biological replicate; 2, present in two biological replicates; 3, present in all biological replicates.**

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| Cre03.g199000.t1.2 | Cre03.g199000.t1.1 | GO:0006468, GO:0004672 | Phototropin-like, blue light receptor [PMID: 12121468; independent sequence: AJ416556] Found in the flagellar proteome [PMID: 15998802]; LOV1 and LOV2 domains. Characterized by Lanzl et al. 2008 [PMID: 18330851] and by Kutta et al. [PMID: 18604833] | *PHT1* | 81 | 0 | 0 | 2 | 2 |

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| Cre10.g436550.t1.2 | Cre10.g436550.t1.1 | | Regulated by CCM1 [PMID: 15235119]. Acclimation to changing CO2 concentrations and light intensities was studied by Yamano et al. 2008 [PMID: 18322145]. | *LCI5* | 32 | 1 | 0 | 3 | 1 |
| Cre03.g143887.t1.1 | Cre01.g056050.t1.1 | GO:0006420, GO:0005737, GO:0005524, GO:0004814, GO:0000166 | arginine-tRNA ligase activity ATP binding cytoplasm arginyl-tRNA aminoacylation nucleotide binding translation | | 65 | 0 | 0 | 0 | 1 |
| Cre11.g467712.t1.1 | Cre18.g743450.t1.1 | GO:2001070 | | *SAGA* | 168 | 0 | 0 | 0 | 1 |
| Cre01.g054800.t1.2 | Cre01.g054800.t1.1 | | PALADIN | *SF54* | 135 | 1 | 0 | 1 | 0 |
| Cre07.g325736.t1.1 | Cre26.g773100.t1.1 | | PAP_fibrillin | | 56 | 1 | 0 | 1 | 0 |
| Cre12.g516200.t1.2 | Cre12.g516200.t1.1 | GO:0005525 | Putative elongation factor 2 | *EFG2* | 94 | 1 | 0 | 0 | 0 |

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| | | | (EF-2), EF-TU; found in the flagellar proteome [PMID: 15998802] | | | | | | |
| Cre06.g250200.t1.2 | Cre06.g250200.t1.1 | GO:0006556, GO:0004478, GO:0005524 | S-adenosylmethionine (AdoMet) synthetase (SAS1); Methionine Adenosyl-transferase; there is a single EST, with an intron, on the other strand: could be a regulatory mechanism. Identified in the eyespot phosphoproteome by Wagner et al. 2008 [PMID: 18065559] | | 43 | 0 | 1 | 1 | 1 |
| Cre09.g401515.t1.1 | Cre02.g139200.t1.1 | | | | 41 | 1 | 1 | 0 | 0 |
| Cre17.g698000.t1.2 | Cre17.g698000.t1.1 | GO:0046034, | ATP synthase F1F0 beta | *ATP2* | 62 | 0 | 1 | 0 | 0 |

| Phytozome ID | Alternate | GO | Gene Description | Gene Name | MW (kDa) | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|---|---|---|---|
| | | GO:0015992, GO:0033178, GO:0016820, GO:0015991, GO:0005524, GO:0046933, GO:0015986 | chain, mitochondrial precursor [PMID: 1386535; GI:584796]. Previously annotated as ATPBM | | | | | | |
| Cre08.g360900.t1.2 | Cre08.g360900.t1.1 | GO:0006412, GO:0005840, GO:0003735 | Cytosolic 80S ribosomal protein S15; Cytosolic 40S small ribosomal subunit protein S15 | *RPS15* | 17 | 3 | 2 | 2 | 2 |

## 5.3 Discussion

In *Chlamydomonas reinhardtii*, the demonstration that Rubisco is localised in the pyrenoid under low $CO_2$ conditions (Lacoste-Royal and Gibbs, 1987; K. Kuchitsu et al., 1988; Süss et al., 1995; He et al., 2003) followed by subsequent efforts to understand regulation of the CCM (CCM1 and LCR1), and the role of the pyrenoidal components (e.g. Rubisco small subunit, CAH3, LCIB/LCIC, EPYC1 and CAS) during CCM induction (Fukuzawa et al., 2001; Miura et al., 2004; Yoshioka et al., 2004; Kohinata et al., 2008; Duanmu et al., 2009; Yamano et al., 2010; Genkov et al., 2010; Ma et al., 2011; Blanco-Rivero et al., 2012; Meyer et al., 2012; Sinetova et al., 2012; Yingjun Wang and Spalding, 2014; Chen, 2016; Wang et al., 2016; Mackinder et al., 2016, 2017; Freeman Rosenzweig et al., 2017) form a foundation which led to this study. In Chapter 3, mutants with normal Rubisco aggregation but defective pyrenoid morphology showed defects in CCM activity suggesting that a fully functional CCM would require the assembly of a multi-component pyrenoid structure. The main hypotheses explored in this study focused on the development of the thylakoid tubule core of the pyrenoid and the maturation process of the pyrenoid during the induction of CCM.

### 5.3.1 Bioinformatics of SAGA

From whole genome sequencing performed on the mutants characterised in Chapter 3, five mutants (A5, A7, A8, A11 and C8) all were shown to be knockouts of the gene Cre11.g467712. The significant decreased affinity to inorganic carbon of these mutants indicated defective CCM, as demonstrated in known CCM mutants such as *cia5*, a knockout of the master regulator for $CO_2$ sensing (Fukuzawa et al., 2001; Xiang et al., 2001; Miura et al., 2004), *cah3*, a mutant lacking carbonic anhydrase proposed to localised in the thylakoid tubule network (Karlsson et al., 1998), *lcib*, a knockout of the LCIB/LCIC complex which localises at the periphery of the pyrenoid (Yamano et al., 2010), a Rubisco linker protein knockout, *epyc1* (Mackinder et al., 2016) and a membrane-bound solute carrier protein knockout, *cia8* (Machingura et al., 2017).

From the phylogenetic analysis of the list of homologues downloaded from Phytozome, SAGA possesses homology to 6 proteins in *C. reinhardtii* which contain at least one starch-binding domain. The putative starch-binding domain (SBD) predicted in SAGA belonged to the CBM (carbohydrate-binding module) 20 family which is found to conserved in some amylolytic enzymes which bind granular starch (Goto et al., 1994;

Chen et al., 1995). Similar to that of SAGA, the phosphoglucan, water dikinase (PWD) in *Arabidopsis thaliana* contains a CBM20 domain at the N-terminus which plays a role in the initial stage of starch degradation (Mahlow et al., 2016), while LCI9 in *C. reinhardtii* with two CBM20 domains, shared a close phylogenetic relationship to a SAGA-like protein (Cre09.g394621), alpha-amylase (Cre12.g492750) and SAGA. LCI9 was suggested by Mackinder et al. (2017) to function in starch degradation as a glucan-1,4-α-glucosidase and is localised in a mesh structure in the gaps between the fenestrations of pyrenoidal starch sheath.

The putative starch-binding domain in SAGA might be involved in the initial starch recognition process which is thought to cause minimal structural change during the interaction with starch (Strokopytov et al., 1996; Sorimachi et al., 1997) The other feature of which SAGA only possessed two out of the nine residues in the conserved region which would have structural change when the domain binds to the starch strands (Strokopytov et al., 1996; Sorimachi et al., 1997). From this prediction, it would be more likely that the CBM20 domain in SAGA potentially plays the role of the initial starch recognition in the pyrenoidal starch metabolism.

On the hand, the SMC_N domain in SAGA resided in the coiled coil protein region. Proteins from the SMC superfamily normally contain ATP-binding domains at the N- and C- termini, with extended coiled-coil domains. Coiled-coil protein structure had been found to play a role in the formation of microtubule apparatus in the non-contractile system I fibers of flagellate green alga *Spermatozopsis similis* (Weber et al., 1993) while Kuhn et al. (2014) suggested that coiled-coil proteins are involved in expansion of the eukaryotic centriosome. Although BLASTP predicted that there was a SMC_N domain within the SAGA protein sequence, detailed examination of the sequence revealed the lack of the signature Walker A box with a consensus nucleotide-binding motif, GxxGxGKS/T (Jessberger, 2002). SMC proteins usually play the role of molecular motors, which control chromosome structures (Jessberger, 2002). For example, SMC and the molecular chaperone DnaJ are reported to have co-immunoprecipitated suggesting that SMC protein is involved in rhodopsin trafficking in verterbrates and in *C. reinhardtii* (Awasthi et al., 2016). However, the current results were insufficient to link SAGA to the suggested functions of SMC_N domain in flagella and chromosome organisation. The SAGA protein in *C. reinhardtii* is more likely to be involved in thylakoid tubule organisation in the pyrenoid and as a part of starch sheath formation under CCM induced conditions based on the data collected in this study.

### 5.3.2 Localisation of SAGA in the Pyrenoid Aids the Carbon Concentrating Mechanism in *Chlamydomonas reinhardtii*

The main hypothesis of this chapter is that SAGA is one of the components of the pyrenoid core. This was first explored by introducing the pRAM_118 vector expressing SAGA-Venus-3xFLAG to complement *saga* (work undertaken by Alan Itakura in collaboration with the Jonikas Lab, then at Stanford University). We hypothesised that the SAGA-Venus-3xFLAG protein would be expressed and localised in the pyrenoid. A preliminary fluorescence screen carried out by Alan Itakura (Stanford University) on the complemented lines grown on TAP suggested that despite the transit peptide having been predicted to direct the protein into the mitochondria, the SAGA-Venus-3xFLAG protein localised in the pyrenoid. This indicated the transit peptide prediction was inaccurate, and Mackinder et al. (2017) point out that up to 10% of PredAlgo and nearly 70% of TargetP predictions are false negatives.

In this study, the decreased affinity for inorganic carbon under both high and low $CO_2$ conditions shown by *saga* was restored when an intact *SAGA* gene fused with Venus-3xFLAG tag was introduced. The CCM defect of *saga* was similar to that of known CCM mutants such as the master regulator, *cia5/ccm1* (Fukuzawa et al. 2001), *lcib* (Yamano et al. 2014), *epyc1* (Mackinder et al. 2016) and *cia8* (Machingura et al. 2017). As in this study, complementation of mutants has been used by various authors to study the function of candidate genes involved in the CCM of *C. reinhardtii* . For example, the authors were able to detect expression of $CO_2$-responsive genes controlled by the master regulator when *cia5* was complemented with a truncated version of the Cia5 gene (Fukuzawa et al., 2001; Xiang et al., 2001) where Fukuzawa et al. (2001) also demonstrated that the complemented line possessed improved affinity to inorganic carbon and photosynthetic rate similar to the corresponding wild type. A knockout of the $Ca^{2+}$-binding protein (CAS), which is localised in the pyrenoidal tubule network during CCM induction, exhibited reduced affinity to inorganic carbon which was then restored when complemented with the gene (Wang et al., 2016).

In terms of pyrenoid morphology, the phenotype of multiple smaller pyrenoids was shared between *saga*, *cia6* (Ma et al., 2011), and the multiple-eyespot mutant, *mlt2* (Boyd et al., 2011), while a single pyrenoid with visible thylakoid tubule network was restored upon complementation for both SAGA and Cia6 (Ma et al., 2011). Although the localisation of Cia6, which is a putative methyltransferase is currently unknown (Mackinder et al., 2017), it is interesting to keep in mind that the lack of this protein

affects the pyrenoid formation (Ma et al., 2011). Aside from *cia6*, *epyc1*, a mutant of the Rubisco-linker protein known to localise in the pyrenoid (Mackinder et al., 2016) and *vmp*-1, a vacuole membrane protein homologue knockout also possesses higher frequency of multiple pyrenoids (Tenenboim et al., 2014). Such studies have inferred that the phenotype of multiple pyrenoids leading to defects in CCM function is a multi-gene phenotype. The introduction of SAGA-Venus-3xFLAG protein in *saga* enabled the cells to form a single pyrenoid with visible tubules in the pyrenoid, as shown in the electron micrographs, indicating the possibility that SAGA is a protein which has multiple functions, which include forming a single functional pyrenoid and the genesis of a pyrenoid nucleation site.

### 5.3.3 SAGA Plays a Role in the Formation of Thylakoid Tubule Network in the Pyrenoid

The lack of a fully functional CCM in *saga* suggested that the missing pyrenoidal component, SAGA, is one of the protein-containing layers of the pyrenoid as suggested in Mackinder et al. (2017). The hypothesis that SAGA is involved as part of the components to construct the thylakoid tubule network was supported through the data shown in this study. The pyrenoid core which is made up of a network of coalescing thylakoid tubules (Sager and Palade 1957; Ohad et al. 1967a, 1967b; Golberg and Ohad 1970; Griffiths 1970) was not visible in transmission electron micrographs of *saga* but was restored in the complemented lines, supporting the hypothesis that SAGA is one of the structural pyrenoidal components which formed the tubule network in the pyrenoid. Although the pyrenoid in *saga* cells mostly did not possess visible tubule network (Chapter 3), some cells possessed multiple pyrenoids with one possessing an incomplete form of a thylakoid tubule traversing in or close to the pyrenoid matrix (Figure 5.27, Appendix 5). As the lack of SAGA in *saga* cells caused disruption in forming the thylakoid tubule network, and the complementation restored the phenotype, it affirmed the hypothesis that SAGA is involved as part of the components needed to construct the thylakoid tubule network.

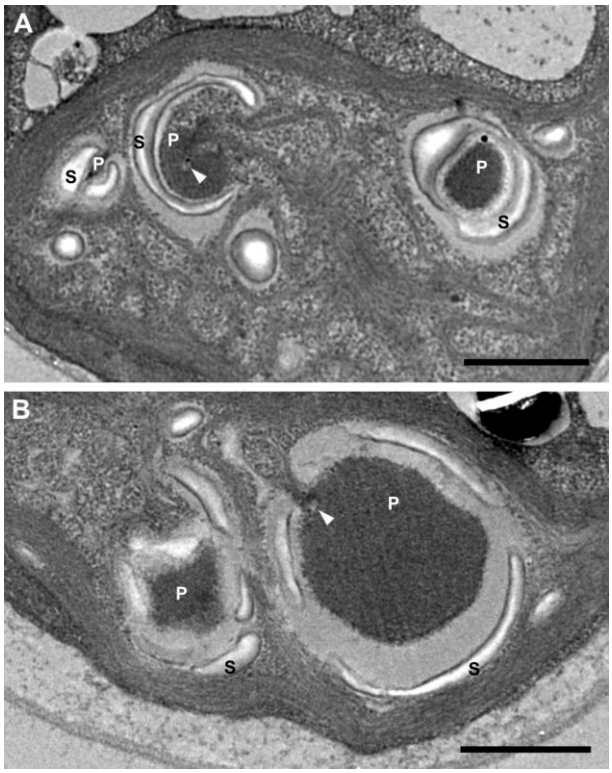



**Figure 5.27 Two examples of incomplete formation of thylakoid tubule (white arrow) penetrated into the pyrenoid matrix of a pyrenoid in the *saga* mutant although there were at least two pyrenoids per cell. Scale bar 500 nm.**

The results which showed a significant reduction in areas of pyrenoid matrix and pyrenoidal starch sheath in *saga* led to the hypothesis that SAGA also plays a role in the maturation of the pyrenoid. In this study, the wild type *C. reinhardtii* cells grown under high $CO_2$ conditions exhibited morphological similarity to the electron micrographs published by Ramazanov et al. (1994) and subsequently by Rawat et al. (1996) that the cells contain smaller pyrenoids and less starch surrounding the pyrenoid. When grown under high concentrations of $CO_2$, the affinity for $CO_2$ in *C. reinhardtii* cells, measured as the $K_{0.5}$, are consistent with the absence of a CCM and more typical of that found in $C_3$ plants (Badger et al., 1980; Rawat et al., 1996). Observations made on cells transferred to ambient air conditions (low $CO_2$) after 1h and 24h suggested that the increase in both the size of the pyrenoid and the area of starch plates were linked to the induction of the carbon concentrating mechanism, as recorded by both Ramazanov et al. (1994) and Rawat et al. (1996). It is also worth noting that the wild type possessed a constant thylakoid tubule network under different $CO_2$ conditions while *saga* did not seem to form a visible core, which was restored in the *saga::SAGA-Venus-3xFLAG* lines. The constant presence of a thylakoid tubule network under different $CO_2$ conditions had been demonstrated previously by various authors including Ramazanov

et al. (1994), Rawat et al. (1996), Blanco-Rivero et al. (2012) and Mackinder et al. (2016), suggesting the importance of this pyrenoid core as an anchor for the pyrenoid matrix.

The SAGA-Venus-3xFLAG protein introduced into *saga* cells was used to determine the localisation of SAGA in the cell. The hypothesis that SAGA is localised in the pyrenoid was guided by the inference from various results that the lack of SAGA caused morphological defects to the pyrenoid. Using fluorescence microscopy, SAGA was shown to localise in the pyrenoid in a pattern which was a cross between a mesh-like structure and a thylakoid puncta as demonstrated by Mackinder et al. (2017) for LCI9 and PSBP4, respectively (Figure 5.28).

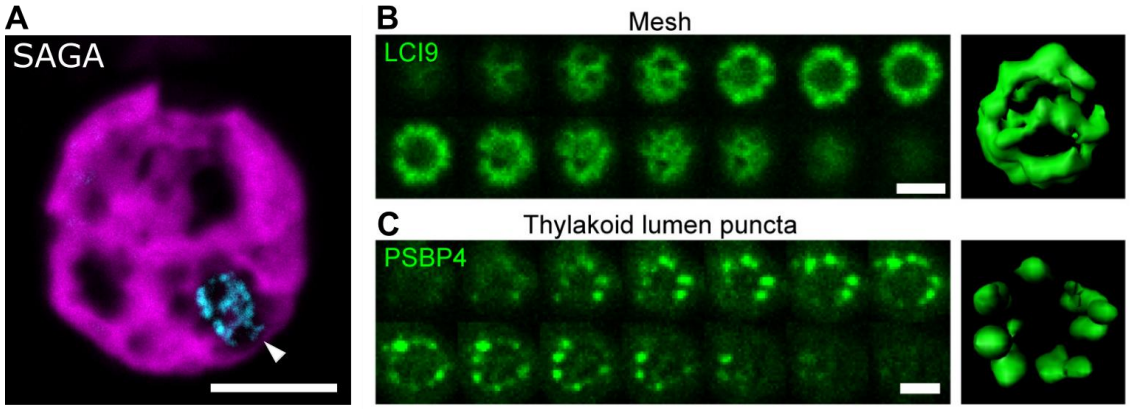



**Figure 5.28 (A) The pattern of SAGA-mVenus fluorescence signal localisation. Scale bar 5 µm. Confocal sections of the localisation pattern of (B) LCI9 and (C) PSBP4 within the pyrenoid, extracted from Mackinder et al. (2017). Scale bars 2 µm.**

In order to confirm the localisation of SAGA in wild type *C. reinhardtii* cells, the indirect immunofluorescence assay was carried out using an anti-SAGA antibody raised against the 19 amino acid on the C-terminus of the protein (courtesy of Alan Itakura & Dr. Martin Jonikas). The results affirmed the hypothesis that SAGA was localised in the pyrenoid under both high and low $CO_2$ conditions. An indirect immunofluorescence assay has also been used to determine the localisation of proteins in the cells in the following studies: the position of Rubisco LSU in the pyrenoid when CCM is induced (Morita et al., 1997); location of the Myb-binding transcription factor, LCR1, which regulates Ci uptake and intracellular Ci accummulation under low $CO_2$ conditions (Ohnishi et al., 2010); the localisation of various CCM proteins including LCIB which is found at the peripheral region of the pyrenoid (Yamano et al., 2014); LCIA and

HLA3 which are regulated by LCR1 and are localised in the chloroplast membrane (Yamano et al., 2015); the $Ca^{2+}$-binding protein CAS protein aggregating in the thylakoid membrane in the pyrenoid (Wang et al., 2016). The current proposal for the location of SAGA in the pyrenoid was along the pyrenoidal tubule network within the pyrenoid matrix (Figure 5.29).

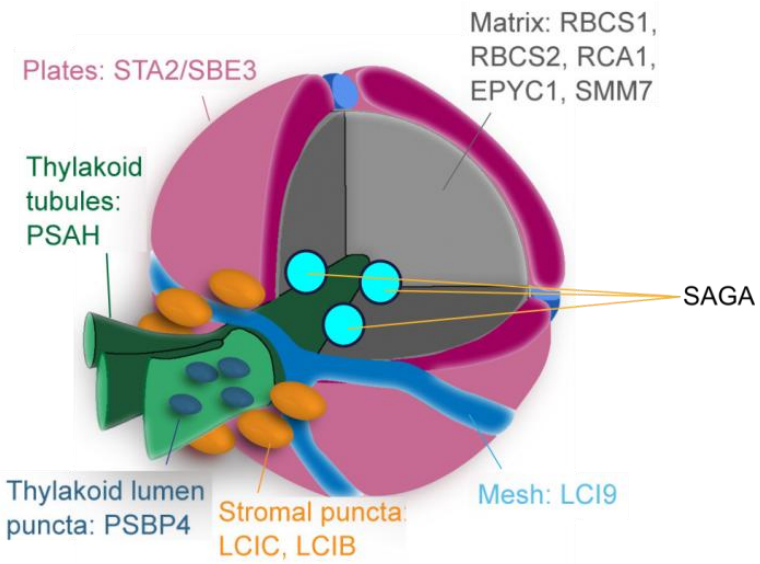



**Figure 5.29 Proposed position of SAGA in the pyrenoid layer based on the original image taken from Mackinder et al. (2017).**

Based on Phobius (Käll et al. 2007) and Phyre2 (Kelley et al. 2015) analyses, SAGA did not possess any transmembrane helices, while DeepLoc (Almagro Armenteros et al. 2017) predicted that SAGA was a soluble protein localised in the cytoplasm. Thus, total soluble protein was incubated with the GFP-Trap_A agarose beads (ChromoTek, Germany) in the co-immunoprecipitation assay protein interactors of SAGA. GFP-Trap_A is a ChromoTek product developed from camelid nanobody which is known to possess enhanced binding specificity with reduced incubation time (Xing et al., 2016). In *Arabidopsis thaliana*, GFP-Trap has been used to identify the interaction of HD2C histone deacetylase and SWI/SNF chromatin remodelling complex in mediating heat stress responses (Buszewicz et al., 2016) and to understand the regulation of trichome cell shape by the interaction between a microtubule-binding protein, TCS1, and a microtubule motor, KCBP/ZWICHEL (Chen et al., 2016). Co-immunoprecipitation assays have been routinely used to identify CCM protein interactions, for example the identification of EPYC1 which interacts with Rubisco small subunit using anti-FLAG CoIP assay (Mackinder et al., 2016).

A preliminary phosphoprotein search on the proteomic dataset was performed with the assistance of Dr. Marco Chiapello (Cambridge Center for Proteomics) and suggested that SAGA might undergo phosphorylation during CCM induction since the protein was shown to be present when cells were grown under high $CO_2$ and after acclimation to low $CO_2$, in line with the prediction of phosphorylation sites in SAGA. The importance of post-translational modifications in *C. reinhardtii* has been highlighted in regulation of chloroplast proteins (Bienvenut et al., 2011; Lehtimäki et al., 2015), in function specialisation of microtubules (Wloga and Gaertig, 2011), in starch metabolism (Mahlow et al., 2016) and in the regulation of CCM (Fukuzawa et al., 2001; Houtz et al., 2008; Blanco-Rivero et al., 2012; Yingjun Wang and Spalding, 2014; Recuenco-Muñoz et al., 2015). As shown by the localisation of Venus-tagged SAGA and indirect immunofluorescence with anti-SAGA antibody on the wild type under different $CO_2$ conditions, SAGA was constantly expressed. Thus, it is likely that the involvement of SAGA in CCM regulation and pyrenoid formation is through post-translational modifications. However, it is crucial that specific experiments to study the phosphorylation of SAGA are carried out in future to follow up on this work. The chances of detecting phosphorylation in the samples in this study were very low as these samples were prepared for normal proteomic analysis with addition of only protease inhibitors, and two different phosphatase inhibitors (sodium fluoride and sodium orthovanadate) (Dr. Marco Chiapello, personal communication). Improved detection of post-translational modifications in protein samples could be done by including additional phosphatase inhibitors, such as the ones used in the sample preparation of eyespot fraction for phosphoproteomic analysis which include microcystin LR, cantharidin, (-)-*p*-bromo-tetramisole, vanadate, molybdate, tartrate and imidazole to reduce potential dephosphorylation (Wagner et al., 2007).

Based on the current results, SAGA is hypothesised to function as an interface protein that anchors the starch sheath, the pyrenoid matrix and the thylakoid tubule network, giving rise to a functional pyrenoid (Figure 5.30). Due to potential presence of post-translational modification sites on SAGA, it is possible that starch-binding domain in SAGA would actively interact with starch only when starch strands are phosphorylated (Mahlow et al., 2016). When no post-translational modifications occur, then SBD would be inactive and other parts of SAGA protein would interact with thylakoid membrane proteins. Aside from post-translational modification as part of the regulation model of SAGA, yeast-two-hybrid assay suggest that SAGA has a strong interaction with the

Rubisco large subunit (Dr. Nicky Atkinson & Dr. Alistair McCormick, University of Edinburgh, personal communication) which suggests the possibility that there are other binding sites within SAGA. The formation of the thylakoid tubules in the pyrenoid might be driven by multiple components besides the regulation of the thylakoid membrane proteins. There is a possibility that SAGA undergoes post-translational modification during CCM induction to anchor the three different layers of the pyrenoid as proteomic analysis detected the presence of a membrane glycoprotein, a serine/threonine kinase and a protein disulfide isomerase in the CoIP of SAGA-Venus-3xFLAG. However, this hypothesis can only be confirmed when more phosphoproteomic experiments are conducted on the same set of samples, not only under CCM induction but also when the CCM is repressed.

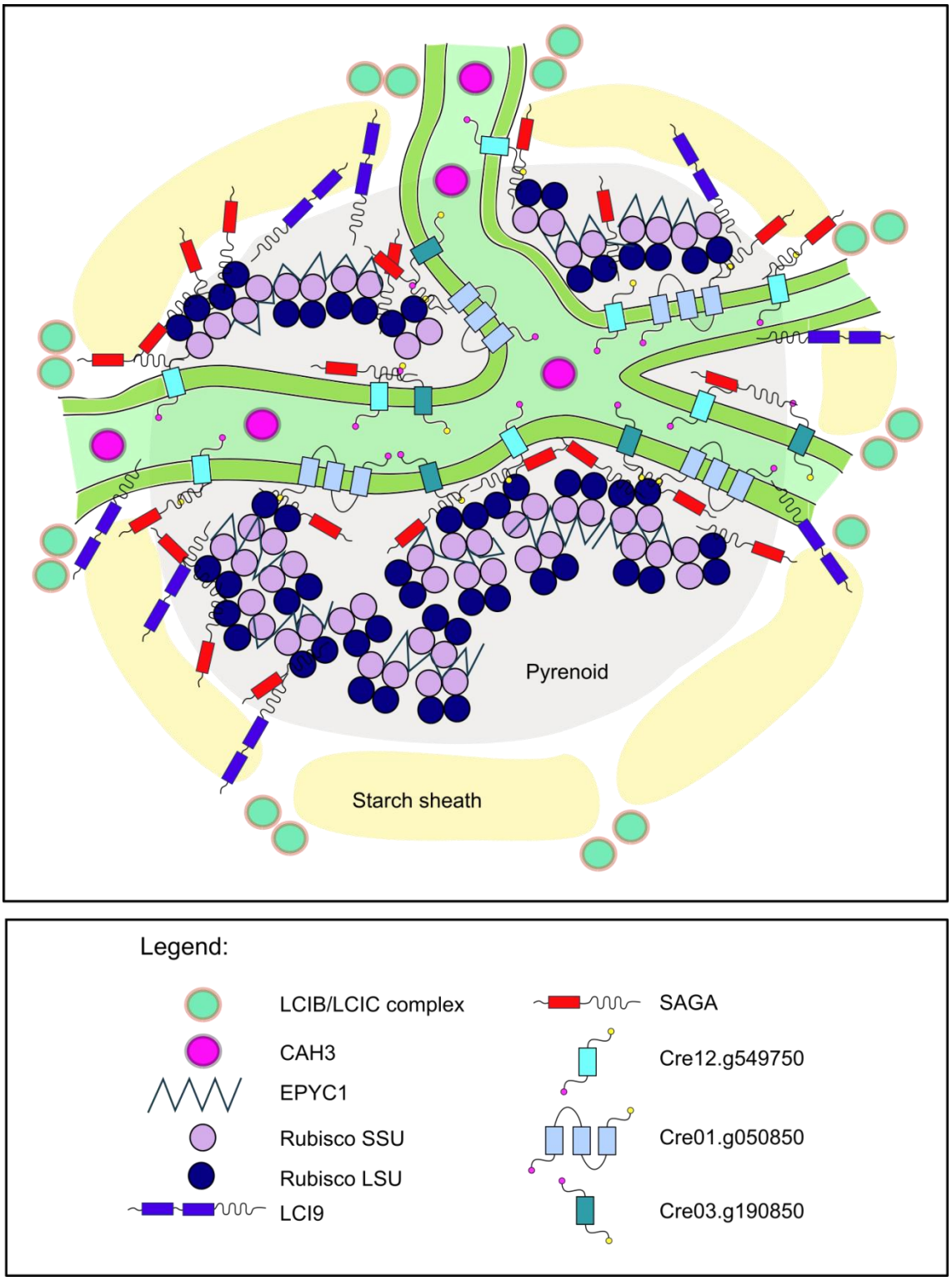



**Figure 5.30 A model of SAGA interaction in the pyrenoid. SAGA is proposed to act as an interface which anchors the pyrenoid matrix, the thylakoid tubule network and the starch sheath to form a functional pyrenoid.**

## 5.4 Conclusion and Future Work

The importance of SAGA in the structural formation of the pyrenoid was clearly observed when the morphological and physiological defects of *saga* mutant were restored with the complementation of SAGA-Venus-3xFLAG protein. The presence of SAGA protein in the mutant enabled the cells to operate a functional CCM as the pyrenoid matrix and starch sheath were anchored to a well-formed thylakoid tubule network. Based on personal communications with Dr. Martin Jonikas (University of Princeton), Dr. Alistair McCormick (University of Edinburgh) and Dr. Luke Mackinder (University of York), there are at least two other proteins with similar properties to SAGA and it would be crucial to identify if these proteins perform similar function as SAGA as an interface protein for functional pyrenoid formation by anchoring the starch sheath, the pyrenoid matrix and the thylakoid tubule network. The McCormick group has started with yeast-two-hybrid work to understand the interaction of SAGA with the pyrenoid matrix with preliminary results indicating the interaction between Rubisco LSU and SAGA.

With regards to the proteomics results obtained in this study, it would be crucial to compare the candidate protein interactors with the list of CCM mutants. As this master list from the Jonikas Lab is still in preparation for publication, further work can be carried out when the dataset is published. If any of the candidate proteins were to be a CCM mutant upon confirmation, and share similar phenotype with *saga*, strategies such as complementation, yeast-two-hybrid, proteomic and phosphoproteomic analyses should be carried out to validate the interaction. It would be interesting to also try to obtain the native SAGA protein crystal structure and to capture the interaction structure between SAGA and other proteins. This will further improve the proposed model of SAGA interaction and the role that SAGA plays in the structural formation of the pyrenoid.

In summary, SAGA plays an important role as an anchor protein to enable pyrenoid matrix and starch sheath to assemble at the thylakoid nucleation site in the canonical position in the chloroplast. Further understanding of this protein and the SAGA-like proteins would likely advance the vision of engineering an algal pyrenoid into higher plants with the hope of improving photosynthesis in $C_3$ plants in future.

# 6 CONCLUSIONS

## 6.1 The Pyrenoid is a Multi-Component Subcompartment of the Chloroplast

The $CO_2$-concentrating mechanism (CCM) in *Chlamydomonas reinhardtii* consists of an inducible system of inorganic carbon transporters facilitating the uptake of carbon from the environment into the chloroplast, where it is then dehydrated into $CO_2$ by carbonic anhydrase and accumulated in the pyrenoid where Rubisco is aggregated (Moroney and Ynalvez, 2007; Jungnick et al., 2014; Wang et al., 2015; Mackinder, 2017). The importance of the Rubisco small subunit (SSU) of *C. reinhardtii* in the formation of the pyrenoid matrix was demonstrated with the hybrid mutants with *RbcS* from higher plants (Arabidopsis, sunflower and spinach) and *rbcL* of *C. reinhardtii* that lacked a clear pyrenoid matrix in the cells, and the absence of any CCM activity (Genkov et al., 2010; Meyer et al., 2012). Aside from the fact that the alpha helices of the Rubisco SSU are crucial in pyrenoid matrix aggregation (Meyer et al., 2012), perhaps linked via the EPYC1 protein, the SSU is also suggested to be a "$CO_2$ reservoir" (van Lun et al., 2014) suggesting that Rubisco might be playing an additional role of retaining $CO_2$ which has been converted by carbonic anhydrase within the pyrenoid of *C. reinhardtii*.

The mutants characterised in this study (Chapter 3) were selected from the insertional mutagenesis screen (Zhang et al., 2014) and shown to possess a defective CCM via growth spot tests and oxygen evolution during inorganic carbon response analyses.

Although morphological assessments highlighted that the mutants had a smaller pyrenoid area similar to that of the Rubisco linker protein mutant, *epyc1* (Mackinder et al., 2016), immunogold labelling with anti-Rubisco antibody determined that the underlying mechanism of Rubisco aggregation was functional and thus not a contributing cause of the defective CCM found in the mutants. The results (Chapter 3) were similar to previous immunolocalisation work showing that up to 99% of Rubisco was localised into the pyrenoid under CCM-induced conditions (Lacoste-Royal and Gibbs, 1987; K. Kuchitsu et al., 1988; Süss et al., 1995; Morita et al., 1997; Borkhsenious et al., 1998; He et al., 2003). This led to the hypothesis that there are multiple components needed to build a structurally-functional pyrenoid.

Following on to the validation of a functional Rubisco aggregating mechanism in the mutant, a new hypothesis set out to explore whether the structural-functional pyrenoid is a multi-component subcompartment of the chloroplast. Whole-genome sequencing of the mutants (Chapter 3) suggested the presence of at least 3 proteins with no actual functional annotation in the structural formation of the pyrenoid: Cre11.g467712 (SAGA, Chapter 5), Cre02.g143387 (Chapter 4) and Cre06.g259100 (Kobayashi et al., 2015). A spatial interactome study using 38 proteins associated with the pyrenoid and CCM as baits to pull down potential interactors, provided insight into the dynamic interaction among proteins within and between different layers of the pyrenoid (Mackinder et al., 2017). Interestingly, Cre06.g259100 (Chapter 3), though demonstrated to localise within the pyrenoid (Kobayashi et al., 2015), was not a candidate interactor of the 38 protein baits used in (Mackinder et al., 2017). The fact that this pyrenoid-localised protein is pulled down in the chloroplast and chloroplast nucleiod co-immunoprecipitation (CoIP) experiments (Kobayashi et al., 2015) indicated that some pyrenoidal proteins interact with proteins unrelated to CCM induction to form the structure of the pyrenoid.

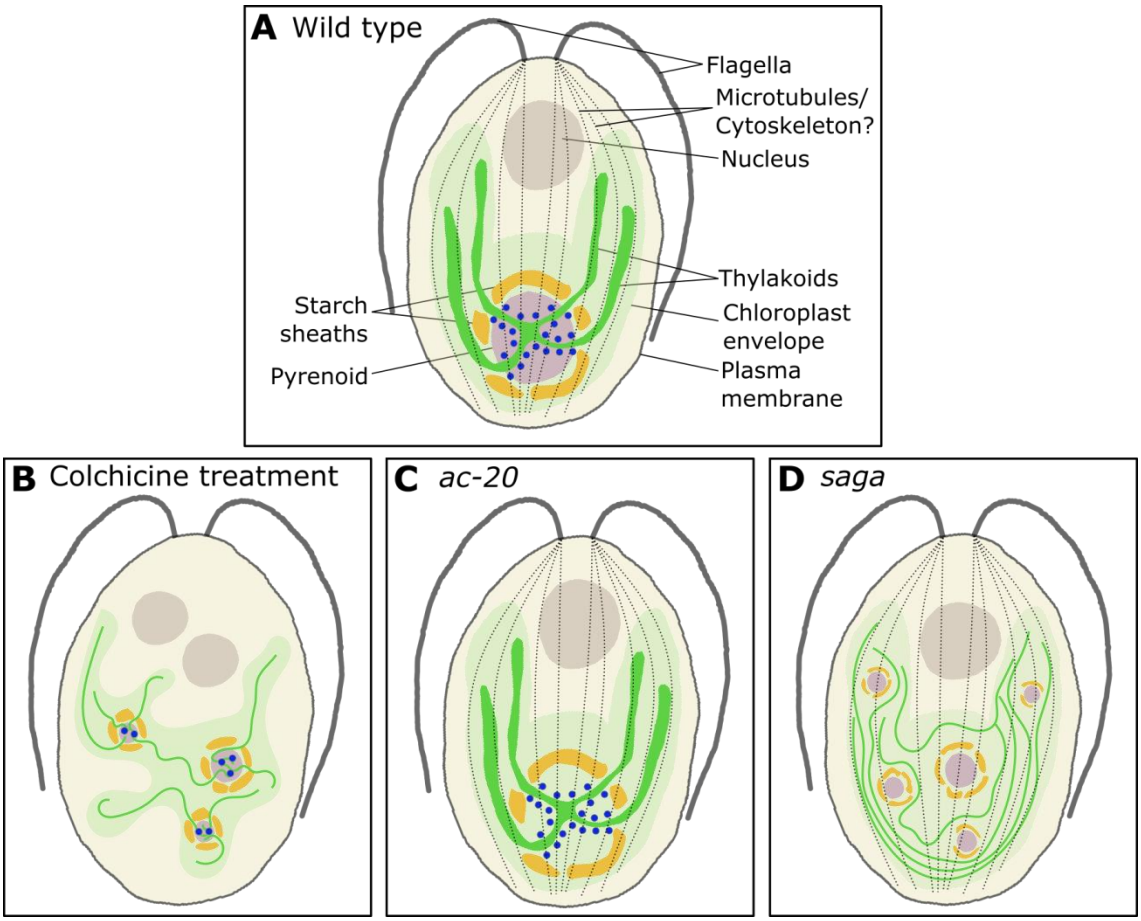



**Figure 6.1 (A) Proposed mode of action of SAGA with the cytoskeleton and/or microtubules in wild type *C. reinhardtii*, where SAGA (blue spheres) located near the thylakoid tubule network would interact with cytoskeleton proteins and fixed the position of the pyrenoid core to the centre of the chloroplast. (B) Colchicine treatment disrupts the formation of microtubules disabling cell division** (Walne, 1966, 1967)**. The potential presence of SAGA in the cells led to pyrenoids with some form of thylakoid tubule network in the pyrenoid. (C) The absence of the pyrenoid matrix in *ac-20* did not affect the position and formation of the thylakoid tubule network** (Goodenough and Levine, 1970) **suggesting that SAGA remains present in the pyrenoid core acting as an anchor protein. (D) The lack of SAGA in *saga* caused an almost complete absence of the thylakoid tubule network (pyrenoid core) causing multiple pyrenoid matrices to form in the chloroplast, indicating that SAGA plays a role in anchoring the pyrenoid matrix to the thylakoid tubule network.**

## 6.2 The Pyrenoid Matrix is anchored to a Pyrenoid Core

The work by Feldman et al. (2007) suggested that the centrioles in *C. reinhardtii* define the internal architecture of the cell. This implies that the positioning of the organelles within the cell is determined by the positioning of the centrioles in the cells. For example, the nucleus is attached by the centrosome, which comprises of spatially distinct structures of a mother centriole, a daughter centriole and percentriolar material connected by fibers (Feldman et al., 2007). This led to a general hypothesis that the pyrenoid matrix which is formed on the basis of hydrophobic interactions between Rubisco complexes (Engel et al., 2015) is anchored to the thylakoid tubule network by a thylakoid anchor protein or a group of proteins which also interacts with the cytoskeleton (Figure 6.1A). The use of colchicine is known to prevent microtubules formation (Flavin and Slaughter, 1974), so cells treated with colchicine exhibited disrupted cell division leading to multiple pyrenoids in a single cell (Buffaloe, 1959; Walne, 1966, 1967) (Figure 6.1B). In the *ac-20* mutant, the cell was devoid of a pyrenoid matrix due to the absence of Rubisco, but starch plates could be observed to surround a thylakoid tubules network (Goodenough and Levine, 1970) indicating the presence of an anchor protein or a group of anchor proteins which determine the localisation of the thylakoid network and starch plates. Assuming that no disruption of cytoskeleton and microtubule formation is present in *ac-20*, the thylakoid anchor protein would enable the pyrenoidal tubule network and the surrounding starch sheath to form at its canonical position (Figure 6.1C).

Based on the observations in Chapters 3 and 5, SAGA fits to the description of an anchor protein which fixes the pyrenoid matrix onto the thylakoid tubule network and the cytoskeleton (Figure 6.1A), since the absence of SAGA not only caused the almost complete lack of a visible thylakoid tubule network in the pyrenoid but also the random positioning of multiple pyrenoid matrices in the chloroplast (Figure 6.1D). On the other hand, the presence of SAGA enabled the cells to form a functional pyrenoid where SAGA was observed within the pyrenoid trailing along the thylakoid tubules and was expressed both under high and low $CO_2$ conditions (Chapter 5), suggesting its importance to the structural formation of the pyrenoid. The general outcome of this study suggested that aside from SAGA, there are at least two other genes, Cre06.g259100 (Chapter 3 and Kobayashi et al. 2015) and Cre02.g143387 (Chapter 4), which are involved in the structural formation of the pyrenoid. This is shown in the results that the mutants of these genes also possessed similar morphological defects as

*saga* indicating that the genes involved might be associated with the same mechanistic pathway.

The co-immunoprecipitation experiment performed using GFP-Trap enabled a glimpse into potential interactors with SAGA using the complemented lines (Chapter 5). A casein kinase 1 that is found both in eyespot and flagella proteomes (Pazour et al., 2005; Eitzinger et al., 2015) was also present in all three biological replicates of the two complemented SAGA lines (Chapter 5), suggesting that CK1 is one of the potential linker proteins between other cell structures and the pyrenoid, with SAGA as one of the linkers to the thylakoid tubule network. Another potential SAGA interactor was the flagella membrane glycoprotein, minor form, FMG1-A (Cre18.g750047) (Chapter 5) which is also found in other interactome datasets of CCM-related proteins: HLA3, LCI1, CAH6 and CAH2 (Mackinder et al., 2017). HLA3 and LCI1 are localised in the plasma membrane (Ohnishi et al., 2010; Wang et al., 2014; Mackinder et al., 2017), forming a complex that acts as a bicarbonate transporter activated by both light and $CO_2$, while CAH2, an isozyme of CAH1 (Fujiwara et al., 1990; Fukuzawa et al., 1990; Rawat and Moroney, 1991; Tachiki et al., 1992), is localised in the periplasmic space of the cell (Fujiwara et al., 1990). CAH6, found at the tip of the flagella (Mackinder et al., 2017), is shown to be an abundant intraflagellar transport (IFT) cargo protein (Engel et al., 2012). IFT is an active transport process needed to assemble and maintain the dynamic structures of the flagella in *C. reinhardtii* (Kozminski et al., 1993; Engel et al., 2012; Kozminski, 2012). A pull-down experiment using a Rubisco column against protein samples of wild type cells, acclimated to low $CO_2$, showed an association between α- and β-tubulin with Rubisco suggesting potential involvement of microtubule/cytoskeleton in pyrenoid formation (Caspari, 2015).

Recently, Wang et al. (2016) demonstrated that a $Ca^{2+}$-binding protein, CAS, which is transported to the pyrenoidal tubule network during CCM induction, plays a role as a chloroplast-mediated regulator of the CCM. In the spatial interactome study conducted by Mackinder et al. (2017), CAS protein is suggested to interact with some photosystem proteins (FDX1, PSAH, PSAK, PSBP3, PSBP4 and HST1) as well as CCM/low-$CO_2$-induced proteins, e.g. LCIA, LCI24, LCI1, HLA3, and CAH3, as well as CCP1 which is known to localise into the mitochondria (Atkinson et al., 2016). It has also been pointed out that $Ca^{2+}$ signalling is important for the motility of *C. reinhardtii* (Wheeler and Brownlee, 2008) while some IFT proteins, such as IFT139, are shown to localise homogenously in the cell (Engel et al., 2012) potentially improving the rapidity of

signal transduction. With many of the proteins associated to SAGA are cytoskeleton-related and flagella associated proteins (Chapter 5), there is a possibility that SAGA might also play a dual function linker protein role which is responsive to the signalling pathway of inorganic carbon sensing (CAH6), or the uptake inorganic carbon (HLA3/LCIA), which connects the pyrenoid to the other organelles via the cytoskeleton.

## 6.3 SAGA is a Multi-Functional Anchor Protein

According to the morphological observation in this study (Chapters 3 and 5), SAGA played a role in maintaining the starch sheath surrounding the pyrenoid, in coordination with other proteins, such as LCI9 and SAGA-like (Dr. Lianyong Wang and Dr. Martin Jonikas, personal communication). The constitutive expression of SAGA-mVenus protein in the cells led to the presence of larger starch sheath area surrounding the pyrenoid (Chapter 5) under different conditions. At the same time, the observations of *saga* transformed with RBCS2-mVenus (Alan Itakura, personal communication) and immunogold labelling with anti-Rubisco antibody (Chapter 3) both showed that the pyrenoid matrix of *saga* remain intact and a complete starch sheath was formed around these pyrenoids. SAGA possesses a CBM20 starch binding domain, similar to that of another protein, LCI9, which localises as a distinctive stromal mesh outer layer of the pyrenoid (Mackinder et al., 2017). The localisation of LCI9 in *saga* in the same location as Rubisco (Alan Itakura, personal communication) suggested that it is one of the proteins that assist in the formation of starch sheath surrounding the pyrenoid matrix. Interestingly, the cDNA microarray performed by Miura et al. (2004) showed the upregulation of a low-$CO_2$-inducible gene with a CBM20 domain, indicating a possibility that this group of proteins is involved in starch accumulation related to CCM function. With a CBM20 starch-binding domain found on SAGA, SAGA-like and LCI9 proteins, it is plausible that there is interaction between starch and these proteins, either by direct interaction, linker proteins or post-translational modifications.

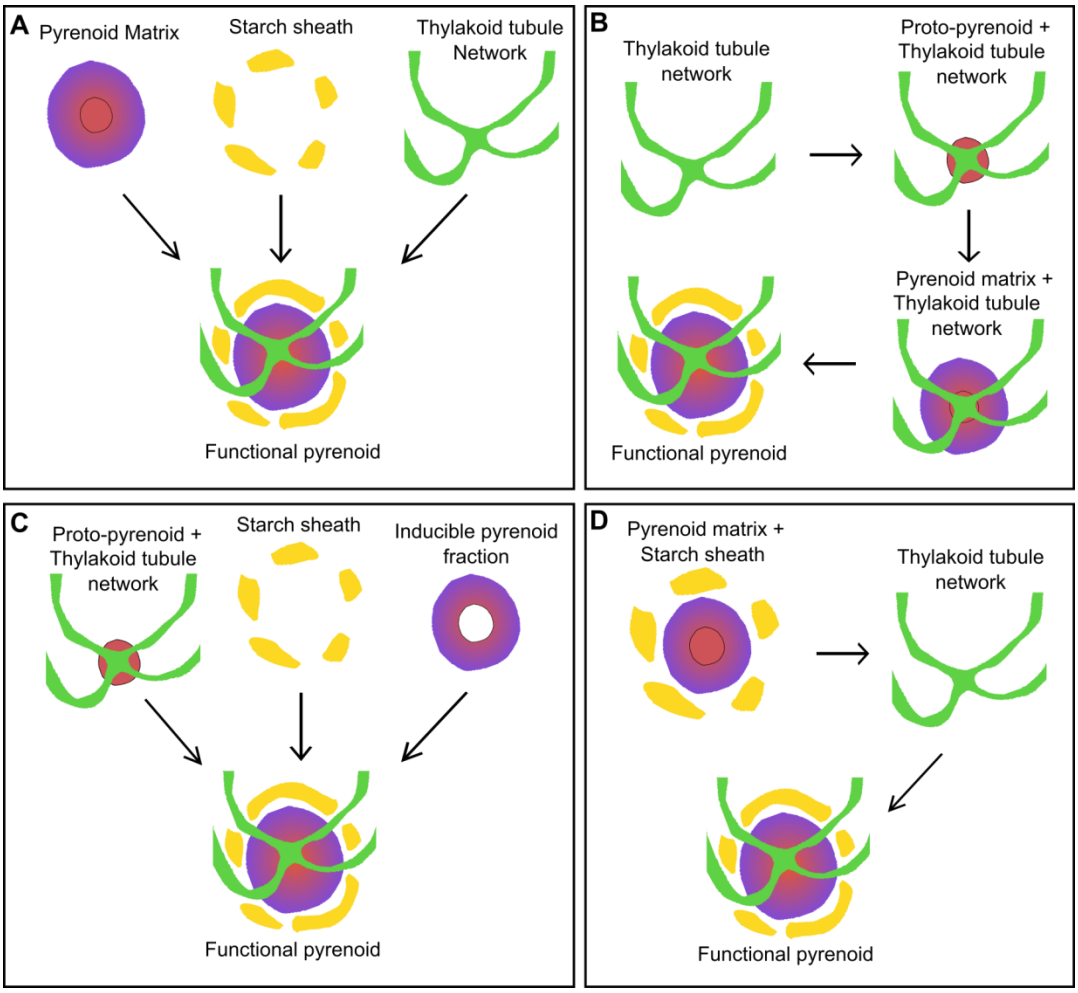



**Figure 6.2 (A to C) Models of the formation of an algal pyrenoid proposed by Meyer et al. (2017). (A) An interacting inter-dependencies model assumes that the major components of algal pyrenoid (pyrenoid matrix, starch sheath and thylakoid tubule network) are regulated by independent molecular pathways. (B) A hierachical model suggests a unidirectional and sequential formation of the pyrenoid, initiated by the formation of a thylakoid tubule network and gradually surrounded by the peripheral elements of the pyrenoid. (C) Hybrid model integrates the previous two models based on the observations of Borkhsenious et al. (1998) and Mitchell et al. (2014) that the $CO_2$-independent fraction of Rubisco is constantly present at the thylakoid tubule network while the $CO_2$-inducible fraction of the pyrenoid and starch sheath gradually build around the pyrenoid. (D) A revised model of pyrenoid biogenesis in *C. reinhardtii* based on the work of this doctoral research, where the formation of the pyrenoid matrix surrounded by the starch sheath assumes an interacting inter-dependencies model while the assembly of the pyrenoid matrix and starch sheath to the pyrenoid core is a more sequential formation.**

In pursuit of the further understanding of the biogenesis of the pyrenoid in *C. reinhardtii*, Meyer et al. (2017) suggested the possible mode of high-level pyrenoid biogenesis and regulation for this green alga species is an "Interacting inter-dependencies" model (Figure 6.2A). This model suggests that the three major components of the pyrenoid: the matrix, the starch sheath and the thylakoid tubules, have respective molecular determinants but require a tightly-regulated interplay between the structures to form a functional pyrenoid. The authors also proposed other modes of algal pyrenoid biogenesis, which includes a hiearchical or sequential mode (Figure 6.2B) where the formation of the pyrenoid is unidirectional, starting from the biogenesis of a thylakoid tubule network where the proto-pyrenoid is anchored and gradually the structure of a functional pyrenoid is built, while the hybrid mode (Figure 6.2C) is a merge of the first and second models (Meyer et al., 2017). Previously, the pyrenoids in algae were suggested to form by *de novo* assembly or by fission (Goodenough, 1970; Meyer et al., 2017).

However, Freeman Rosenzweig et al. (2017) redefined the formation of the pyrenoid matrix of *C. reinhardtii*, demonstrating the ability of this organism to form the pyrenoid by both *de novo* assembly and by fission due to the behaviour of the matrix being liquid-like rather than a rigid proteinaceous subcompartment (Holdsworth, 1968). Whilst the interacting inter-dependencies model was the most likely model of the biogenesis of the pyrenoid based on what was then the published literature, the results of this doctoral research and the observation of Freeman Rosenzweig et al. (2017) suggest that a hybrid model is more plausible (Figure 6.2D): the difference being that the formation of the pyrenoid matrix, surrounded by a starch sheath, is separate from the biogenesis of the thylakoid tubule network, which acts as a pyrenoid core (Griffiths, 1970). The outcome of this work suggested that the biogenesis of the pyrenoid is semi-hierarchical since the absence of the pyrenoid core causes multiple pyrenoid matrices surrounded by incomplete starch sheaths, rather than a single functional pyrenoid (Chapter 5).

In summary, the results of this doctoral research suggested that there are multiple genes involved in the structural formation of the pyrenoid, with SAGA validated as a potential anchor protein that assists the formation of a pyrenoid core. This multi-functional protein also potentially plays a role in maintaining the starch sheath surrounding the pyrenoid and interacting with the cytoskeleton to anchor the pyrenoid. This protein may not be involved directly in the $CO_2$-concentrating mechanism as an inducible protein,

but the mode of action of SAGA and the related proteins (SAGA-like and LCI9) needs to be determined in future work, such as co-localisation, biomolecular fluorescence study and yeast-two-hybrid. Although the assembly of a pyrenoid structure in higher plants remains a vision for the future, the discovery of this anchor protein at the heart of the pyrenoid matrix leads us a step closer to fulfilling that vision.

# 7 REFERENCES


Addgene, 2017. Plasmids 101: A Desktop Resource, 3rd Edition. ed. https://doi.org/10.1128/genomeA.00232-16

Allard, S.T.M., 2008. Bioluminescent reporter genes. Promega Notes. URL http://www.promega.co.uk/resources/pubhub/enotes/bioluminescent-reporter-genes/

Almagro Armenteros, J.J., Sønderby, C.K., Sønderby, S.K., Nielsen, H., Winther, O., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33, 3387–3395. https://doi.org/10.1093/bioinformatics/btx431

Alva, V., Nam, S.-Z., Söding, J., Lupas, A.N., 2016. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res. 44, W410–W415. https://doi.org/10.1093/nar/gkw348

Andersson, I., 2008. Catalysis and regulation in Rubisco. J. Exp. Bot. 59, 1555–68. https://doi.org/10.1093/jxb/ern091

Atkinson, N., Feike, D., Mackinder, L.C.M., Meyer, M.T., Griffiths, H., Jonikas, M.C., Smith, A.M., Mccormick, A.J., 2016. Introducing an algal carbon-concentrating mechanism into higher plants: Location and incorporation of key components. Plant Biotechnol. J. 14, 1302–1315. https://doi.org/10.1111/pbi.12497

Atkinson, N., Leitão, N., Orr, D.J., Meyer, M.T., Carmo-Silva, E., Griffiths, H., Smith, A.M., McCormick, A.J., 2017. Rubisco small subunits from the unicellular green

alga *Chlamydomonas* complement Rubisco-deficient mutants of *Arabidopsis*. New Phytol. 214, 655–667. https://doi.org/10.1111/nph.14414

Awasthi, M., Ranjan, P., Sharma, K., Veetil, S.K., Kateriya, S., Armitage, J.P., Holland, E.M., Harz, H., Uhl, R., Hegemann, P., Hardie, R.C., Franze, K., Deretic, D., LaVail, M.M., Basinger, S., Bok, D., Hall, M., Hegemann, P., Sudo, Y., Yuasa, Y., Shibata, J., Suzuki, D., Homma, M., Keady, B.T., Le, Y.Z., Pazour, G.J., Bhowmick, R., Kozminski, K.G., Johnson, K.A., Forscher, P., Rosenbaum, J.L., Binder, L.I., Dentler, W.L., Rosenbaum, J.L., Rosenbaum, J.L., Child, F.M., Kozminski, K.G., Beech, P.L., Rosenbaum, J.L., Pazour, G.J., Wilkerson, C.G., Witman, G.B., Cole, D.G., Walther, Z., Vashishtha, M., Hall, J.L., Gibbons, B.H., Asai, D.J., Tang, W.J., Hays, T.S., Gibbons, I.R., Porter, M.E., Bower, R., Knott, J.A., Byrd, P., Dentler, W., Perrone, C.A., Hou, Y., Pazour, G.J., Witman, G.B., Wren, K.N., Craft, J.M., Harris, J.A., Hyman, S., Kner, P., Lechtreck, K.F., Cole, D.G., Taschner, M., Kotsis, F., Braeuer, P., Kuehn, E.W., Lorentzen, E., Follit, J.A., Xu, F., Keady, B.T., Pazour, G.J., Taschner, M., Bhogaraju, S., Lorentzen, E., Bhogaraju, S., Engel, B.D., Lorentzen, E., Hou, Y., Ahmed, N.T., Gao, C., Lucker, B.F., Cole, D.G., Mitchell, D.R., Bhogaraju, S., Ishikawa, H., Evans, R.J., Waters, A.M., Beales, P.L., Fliegauf, M., Benzing, T., Omran, H., Kateriya, S., Nagel, G., Nagel, G., Holland, E.M., Braun, F.J., Nonnengässer, C., Harz, H., Hegemann, P., Harz, H., Hegemann, P., Kröger, P., Hegemann, P., Nonnengässer, C., Holland, E.M., Harz, H., Hegemann, P., Trippens, J., Suzuki, T., Berthold, P., Mittelmeier, T.M., Boyd, J.S., Lamb, M.R., Dieckmann, C.L., Engel, B.D., Pazour, G.J., Adams, G.M., Huang, B., Luck, D.J., Piperno, G., Mead, K., Henderson, S., Pazour, G.J., Dickert, B.L., Witman, G.B., Lim, Y.S., Chua, C.E.L., Tang, B.L., Krock, B.L., Perkins, B.D., Tsujikawa, M., Malicki, J., Iomini, C., Li, L., Esparza, J.M., Dutcher, S.K., Pazour, G.J., Papermasrer, D.S., Schneider, B.G., Besharse, J.C., Deretic, D., Wang, J., Chuang, J.-Z., Hsu, Y.-C., Sung, C.-H., Khanna, H., Gühmann, M., MOREL-LAURENS, N.M.L., FEINLEIB, M.E., Ueki, N., Jékely, G., Finetti, F., Paccani, S.R., Rosenbaum, J., Baldari, C.T., Lechtreck, K.-F., 2016. The trafficking of bacterial type rhodopsins into the *Chlamydomonas* eyespot and flagella is IFT mediated. Sci. Rep. 6, 34646. https://doi.org/10.1038/srep34646

Badger, M.R., Kaplan, A., Berry, J.A., 1980. Internal Inorganic Carbon Pool of *Chlamydomonas reinhardtii*: Evidence for a Carbon Dioxide-Concentrating Mechanism. Plant Physiol. 66, 407–419.

Barrera, D., Mayfield, S., 2013. High-value Recombinant Protein Production in Microalgae, in: Handbook of Microalgal Culture: Applied Phycology and Biotechnology. pp. 532–544. https://doi.org/10.1002/9781118567166.ch27

Bauwe, H., Hagemann, M., Fernie, A.R., 2010. Photorespiration: players, partners and origin. Trends Plant Sci. 15, 330–336. https://doi.org/10.1016/j.tplants.2010.03.006

Becker, W.E., 2013. Microalgae in Human and Animal Nutrition. Handb. Microalgal Cult. Appl. Phycol. Biotechnol. 461–503. https://doi.org/10.1002/9780470995280.ch18

Belay, A., 2013. Biology and Industrial Production of Arthrospira (Spirulina). Handb. Microalgal Cult. Appl. Phycol. Biotechnol. 339–358. https://doi.org/10.1002/9781118567166.ch17

Benlloch, R., Shevela, D., Hainzl, T., Grundström, C., Shutova, T., Messinger, J., Samuelsson, G., Sauer-Eriksson, A.E., 2015. Crystal Structure and Functional Characterization of Photosystem II-Associated Carbonic Anhydrase CAH3 in *Chlamydomonas reinhardtii*. Plant Physiol. 167, 950–62. https://doi.org/10.1104/pp.114.253591

Bennoun, P., Levine, R.P., 1967. Detecting mutants that have impaired photosynthesis by their increased level of fluorescence. Plant Physiol. 42, 1284–7. https://doi.org/10.1104/pp.42.9.1284

Benson, A.A., Calvin, M., 1950. Carbon Dioxide Fixation By Green Plants. Annu. Rev. Plant Physiol. 1, 25–42. https://doi.org/10.1146/annurev.pp.01.060150.000325

Berry, J., Boynton, J., Kaplan, A., Badger, M., 1976. Growth and photosynthesis of *Chlamydomonas reinhardtii* as a function of $CO_2$ concentration. Carnegie Inst. Washington, Year B. 75, 423–432.

Bertos, N.R., Gibbs, S.P., 1998. Evidence for a Lack of Photosystem Segregation in *Chlamydomonas reinhardtii* (Chlorophyceae). J. Phycol. 34, 1009–1016. https://doi.org/10.1046/j.1529-8817.1998.341009.x

Betti, M., Bauwe, H., Busch, F.A., Fernie, A.R., Keech, O., Levey, M., Ort, D.R., Parry, M.A.J., Sage, R., Timm, S., Walker, B., Weber, A.P.M., 2016. Manipulating photorespiration to increase plant productivity: Recent advances and perspectives for crop improvement. J. Exp. Bot. 67, 2977–2988. https://doi.org/10.1093/jxb/erw076

Bienvenut, W. V., Espagne, C., Martinez, A., Majeran, W., Valot, B., Zivy, M., Vallon, O., Adam, Z., Meinnel, T., Giglione, C., 2011. Dynamics of post-translational modifications and protein stability in the stroma of *Chlamydomonas reinhardtii* chloroplasts. Proteomics 11, 1734–1750. https://doi.org/10.1002/pmic.201000634

Binder, A., Lambert, J., Morbitzer, R., Popp, C., Ott, T., Lahaye, T., Parniske, M., 2014. A modular plasmid assembly kit for multigene expression, gene silencing and silencing rescue in plants. PLoS One 9. https://doi.org/10.1371/journal.pone.0088218

Blaby, I.K., Blaby-Haas, C.E., Tourasse, N., Hom, E.F.Y., Lopez, D., Aksoy, M., Grossman, A., Umen, J., Dutcher, S., Porter, M., King, S., Witman, G.B., Stanke, M., Harris, E.H., Goodstein, D., Grimwood, J., Schmutz, J., Vallon, O., Merchant, S.S., Prochnik, S., 2014. The *Chlamydomonas* genome project: A decade on. Trends Plant Sci. 19, 672–680. https://doi.org/10.1016/j.tplants.2014.05.008

Blanco-Rivero, A., Shutova, T., Román, M.J., Villarejo, A., Martinez, F., 2012. Phosphorylation controls the localization and activation of the lumenal carbonic anhydrase in *Chlamydomonas reinhardtii*. PLoS One 7, e49063. https://doi.org/10.1371/journal.pone.0049063

Blom, N., Gammeltoft, S., Brunak, S., 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. J. Mol. Biol. 294, 1351–1362. https://doi.org/10.1006/jmbi.1999.3310

Blom, N., Sicheritz-Pontén, T., Gupta, R., Gammeltoft, S., Brunak, S., 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4, 1633–1649. https://doi.org/10.1002/pmic.200300771

Bonfil, D.J., Ronen-Tarazi, M., Sültemeyer, D., Lieman-Hurwitz, J., Schatz, D., Kaplan, A., 1998. A putative HCO3- transporter in the cyanobacterium Synechococcus sp. strain PCC 7942. FEBS Lett. 430, 236–240.

Borkhsenious, O.N., Mason, C.B., Moroney, J. V., 1998. The intracellular localization of ribulose-1,5-bisphosphate Carboxylase/Oxygenase in *Chlamydomonas reinhardtii*. Plant Physiol. 116, 1585–91. https://doi.org/10.1104/pp.116.4.1585

Borland, A.M., Hartwell, J., Weston, D.J., Schlauch, K.A., Tschaplinski, T.J., Tuskan, G.A., Yang, X., Cushman, J.C., 2014. Engineering crassulacean acid metabolism to improve water-use efficiency. Trends Plant Sci. 19, 327–38.

https://doi.org/10.1016/j.tplants.2014.01.006

Borland, A.M., Yang, X., 2013. Informing the improvement and biodesign of crassulacean acid metabolism via system dynamics modelling. New Phytol. 200, 946–949. https://doi.org/10.1111/nph.12529

Borodin, V.B., 2008. Effect of red and blue light on acclimation of *Chlamydomonas reinhardtii* to $CO_2$-limiting conditions. Russ. J. Plant Physiol. 55, 441–448. https://doi.org/10.1134/S102144370804002X

Borowitzka, M.A., 2013. *Dunaliella*: Biology, Production, and Markets. Handb. Microalgal Cult. Appl. Phycol. Biotechnol. 359–368. https://doi.org/10.1002/9781118567166.ch18

Botany. A. GENERAL, including Anatomy and Physiology of the Phanerogamia, 1896. . J. R. Microsc. Soc. 16, 199–231.

Boyd, J.S., Lamb, M.R., Dieckmann, C.L., 2011. Miniature- and multiple-eyespot loci in *Chlamydomonas reinhardtii* define new modulators of eyespot photoreception and assembly. G3 1, 489–498. https://doi.org/10.1534/g3.111.000679

Boynton, J., Gillham, N., Harris, E., Hosler, J., Johnson, A., Jones, A., Randolph-Anderson, B., Robertson, D., Klein, T., Shark, K., et, al., 1988. Chloroplast transformation in *Chlamydomonas* with high velocity microprojectiles. Science 80. 240, 1534–1538. https://doi.org/10.1126/science.2897716

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254. https://doi.org/10.1016/0003-2697(76)90527-3

Bronstein, I., Fortin, J., Stanley, P.E., Stewart, G.S.A.., Kricka, L.J., 1994. Chemiluminescent and bioluminescent reporter gene assays. Anal. Biochem. 219, 169–181.

Brown, L.E., Sprecher, S.L., Keller, L.R., 1991. Introduction of exogenous DNA into *Chlamydomonas reinhardtii* by electroporation. Mol. Cell. Biol. 11, 2328–2332. https://doi.org/10.1128/MCB.11.4.2328

Brueggeman, A.J., Gangadharaiah, D.S., Cserhati, M.F., Casero, D., Weeks, D.P., Ladunga, I., 2012. Activation of the carbon concentrating mechanism by $CO_2$ deprivation coincides with massive transcriptional restructuring in *Chlamydomonas reinhardtii*. Plant Cell 24, 1860–1875.

https://doi.org/10.1105/tpc.111.093435

Buffaloe, N.D., 1959. Some effects of colchicine on cells of *Chlamydomonas eugametos* Moewus. Exp. Cell Res. 16, 221–231.

Buffaloe, N.D., 1958. A comparative cytological study of four species of *Chlamydomonas*. Bull. Torrey Bot. Club 85, 157–178.

Burgess, S.J., Reyna-Llorens, I., Jaeger, K., Hibberd, J.M., 2017. A transcription factor binding atlas for photosynthesis in cereals identifies a key role for coding sequence in the regulation of gene expression. bioRxiv. https://doi.org/http://dx.doi.org/10.1101/165787

Busch, F.A., Sage, R.F., Farquhar, G.D., 2017. Plants increase $CO_2$ uptake by assimilating nitrogen via the photorespiratory pathway. Nat. Plants 1. https://doi.org/10.1038/s41477-017-0065-x

Buszewicz, D., Archacki, R., Palusiński, A., Kotliński, M., Fogtman, A., Iwanicka-Nowicka, R., Sosnowska, K., Kuciński, J., Pupel, P., Olędzki, J., Dadlez, M., Misicka, A., Jerzmanowski, A., Koblowska, M.K., 2016. HD2C histone deacetylase and a SWI/SNF chromatin remodelling complex interact and both are involved in mediating the heat stress response in *Arabidopsis*. Plant Cell Environ. 39, 2108–2122. https://doi.org/10.1111/pce.12756

Canfield, D.E., 2005. The early history of atmospheric oxygen: Homage to Robert M. Garrels. Annu. Rev. Earth Planet. Sci. 33, 1–36. https://doi.org/10.1146/annurev.earth.33.092203.122711

Carvalho, J. de F., Madgwick, P.J., Powers, S.J., Keys, A.J., Lea, P.J., Parry, M.A., 2011. An engineered pathway for glyoxylate metabolism in tobacco plants aimed to avoid the release of ammonia in photorespiration. BMC Biotechnol. 11, 111. https://doi.org/10.1186/1472-6750-11-111

Caspari, O.D., 2015. Molecular physiology of the *Chlamydomonas* pyrenoid. University of Cambridge.

Caspari, O.D., Meyer, M.T., Tolleter, D., Wittkopp, T.M., Cunniffe, N.J., Lawson, T., Grossman, A.R., Griffiths, H., 2017. Pyrenoid loss in *Chlamydomonas reinhardtii* causes limitations in $CO_2$ supply, but not thylakoid operating efficiency. J. Exp. Bot. 68, 1–11. https://doi.org/10.1093/jxb/erx197

Chen, B., 2016. The function and regulation of CIA5/CCM1 in *Chlamydomonas*

*reinhardtii*. Iowa State University.

Chen, B., Lee, K., Plucinak, T., Duanmu, D., Wang, Y., Horken, K.M., Weeks, D.P., Spalding, M.H., 2017. A novel activation domain is essential for CIA5-mediated gene regulation in response to $CO_2$ changes in *Chlamydomonas reinhardtii*. Algal Res. 24, 207–217. https://doi.org/10.1016/j.algal.2017.03.006

Chen, L., Coutinho, P.M., Nikolov, Z., Ford, C., 1995. Deletion analysis of the starch-binding domain of *Aspergillus* glucoamylase. Protein Eng. Des. Sel. 8, 1049–1055. https://doi.org/10.1093/protein/8.10.1049

Chen, L., Peng, Y., Tian, J., Wang, X., Kong, Z., Mao, T., Yuan, M., Li, Y., 2016. TCS1, a microtubule-binding protein, interacts with KCBP/ZWICHEL to regulate trichome cell shape in *Arabidopsis thaliana*. PLoS Genet. 12, 1–21. https://doi.org/10.1371/journal.pgen.1006266

Cheng, R., Ma, R., Li, K., Rong, H., Lin, X., Wang, Z., Yang, S., Ma, Y., 2012. *Agrobacterium tumefaciens* mediated transformation of marine microalgae *Schizochytrium*. Microbiol. Res. 167, 179–186. https://doi.org/10.1016/j.micres.2011.05.003

Choi, H. Il, Young, J., Kim, H., Kwak, H.S., Sung, Y.J., Sim, S.J., Choi, H. Il, Hwan, Y., Kwak, S., 2016. Quantitative analysis of the chemotaxis of a green alga, *Chlamydomonas reinhardtii*, to bicarbonate using diffusion-based microfluidic device. Biomicrofluidics 10, 14121. https://doi.org/10.1063/1.4942756

Christie, J.M., 2007. Phototropin blue-light receptors. Annu. Rev. Plant Biol. 58, 21–45. https://doi.org/10.1146/annurev.arplant.58.032806.103951

Cieśla, J., Fraczyk, T., Rode, W., 2011. Phosphorylation of basic amino acid residues in proteins: Important but easily missed. Acta Biochim. Pol. 58, 137–148. https://doi.org/201114789 [pii]

Colman, B., Huertas, I.E., Bhatti, S., Dason, J.S., 2002. The diversity of inorganic carbon acquisition mechanisms in eukaryotic microalgae. Aust. J. Plant Physiol. 29, 261–270.

Costantini, L.M., Snapp, E.L., 2013. Fluorescent proteins in cellular organelles: Serious pitfalls and some solutions. DNA Cell Biol. 32, 622–627. https://doi.org/10.1089/dna.2013.2172

Covshoff, S., Hibberd, J.M., 2012. Integrating $C_4$ photosynthesis into $C_3$ crops to

increase yield potential. Curr. Opin. Biotechnol. 23, 209–214. https://doi.org/10.1016/j.copbio.2011.12.011

Dalal, J., Lopez, H., Vasani, N.B., Hu, Z., Swift, J.E., Yalamanchili, R., Dvora, M., Lin, X., Xie, D., Qu, R., Sederoff, H.W., 2015. A photorespiratory bypass increases plant growth and seed yield in biofuel crop *Camelina sativa*. Biotechnol. Biofuels 8, 175. https://doi.org/10.1186/s13068-015-0357-1

de los Reyes, P., Romero-Campero, F.J., Ruiz, M.T., Romero, J.M., Valverde, F., 2017. Evolution of daily gene co-expression patterns from algae to plants. Front. Plant Sci. 8, 1–22. https://doi.org/10.3389/fpls.2017.01217

Dent, R.M., Haglund, C.M., Chin, B.L., Kobayashi, M.C., Niyogi, K.K., 2005. Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of *Chlamydomonas reinhardtii*. Plant Physiol. 137, 545–556. https://doi.org/10.1104/pp.104.055244

DePaoli, H.C., Borland, A.M., Tuskan, G.A., Cushman, J.C., Yang, X., 2014. Synthetic biology as it relates to CAM photosynthesis: Challenges and opportunities. J. Exp. Bot. 65, 3381–3393. https://doi.org/10.1093/jxb/eru038

Desmyter, A., Spinelli, S., Roussel, A., Cambillau, C., 2015. Camelid nanobodies: Killing two birds with one stone. Curr. Opin. Struct. Biol. 32, 1–8. https://doi.org/10.1016/j.sbi.2015.01.001

Dill, E.O., 1895. Die Gattung *Chlamydomonas* und ihre nächsten Verwandten. Jahrb Wiss Bot 28, 323–358.

Dionisio-Sese, M.L., Fukuzawa, H., Miyachi, S., 1990. Light-induced carbonic anhydrase expression in *Chlamydomonas reinhardtii*. Plant Physiol. 94, 1103–10. https://doi.org/10.1104/pp.94.3.1103

Dionisio, M.L.T.M. and M.S., 1989. Blue light induction of carbonic anhydrase activity in *Chlamydomonas reinhardtii*. Plant Cell Physiol. 30, 215–219.

Duanmu, D., Miller, A.R., Horken, K.M., Weeks, D.P., Spalding, M.H., 2009. Knockdown of limiting- $CO_2$-induced gene HLA3 decreases $HCO_3^-$ transport and photosynthetic $C_i$ affinity in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. U. S. A. 106, 5990–5. https://doi.org/10.1073/pnas.0812885106

Duanmu, D., Wang, Y., Spalding, M.H., 2008. Thylakoid lumen carbonic anhydrase (CAH3) mutation suppresses air-dier phenotype of LCIB mutant in

*Chlamydomonas reinhardtii*. Plant Physiol. 149, 929–937. https://doi.org/10.1104/pp.108.132456

Eitzinger, N., Wagner, V., Weisheit, W., Geimer, S., Boness, D., Kreimer, G., Mittag, M., 2015. Proteomic analysis of a fraction with intact eyespots of *Chlamydomonas reinhardtii* and assignment of protein methylation. Front. Plant Sci. 6, 1–16. https://doi.org/10.3389/fpls.2015.01085

Ellis, R.J., 2010. Tackling unintelligent design. Nature 463, 164–165.

Emanuelsson, O., Nielsen, H., Brunak, S., von Heijne, G., 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J. Mol. Biol. 300, 1005–1016. https://doi.org/10.1006/jmbi.2000.3903

Emanuelsson, O., Nielsen, H., Heijne, G. Von, 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. 8, 978–984. https://doi.org/10.1110/ps.8.5.978

Engel, B.D., Ishikawa, H., Wemmer, K.A., Geimer, S., Wakabayashi, K.I., Hirono, M., Craige, B., Pazour, G.J., Witman, G.B., Kamiya, R., Marshall, W.F., 2012. The role of retrograde intraflagellar transport in flagellar assembly, maintenance, and function. J. Cell Biol. 199, 151–167. https://doi.org/10.1083/jcb.201206068

Engel, B.D., Schaffer, M., Kuhn Cuellar, L., Villa, E., Plitzko, J.M., Baumeister, W., 2015. Native architecture of the *Chlamydomonas* chloroplast revealed by *in situ* cryo-electron tomography. eLife 4, 1–29. https://doi.org/10.7554/eLife.04889

Engler, C., Kandzia, R., Marillonnet, S., 2008. A one pot, one step, precision cloning method with high throughput capability. PLoS One 3. https://doi.org/10.1371/journal.pone.0003647

Engler, C., Marillonnet, S., 2014. DNA Cloning and Assembly Methods, in: Valla, S., Lale, R. (Eds.), DNA Cloning and Assembly Methods. Methods in Molecular Biology (Methods and Protocols). Humana Press, Totowa, New Jersey, pp. 119–131. https://doi.org/10.1007/978-1-62703-764-8

Engler, C., Youles, M., Gruetzner, R., Ehnert, T.M., Werner, S., Jones, J.D.G., Patron, N.J., Marillonnet, S., 2014. A Golden Gate modular cloning toolbox for plants. ACS Synth. Biol. 3, 839–843. https://doi.org/10.1021/sb4001504

Fang, W., Si, Y., Douglass, S., Casero, D., Merchant, S.S., Pellegrini, M., Ladunga, I., Liu, P., Spalding, M.H., 2012. Transcriptome-wide changes in *Chlamydomonas*

*reinhardtii* gene expression regulated by carbon dioxide and the $CO_2$-Concentrating Mechanism Regulator CIA5/CCM1. Plant Cell 24, 1876–1893. https://doi.org/10.1105/tpc.112.097949

Feldman, J.L., Geimer, S., Marshall, W.F., 2007. The mother centriole plays an instructive role in defining cell geometry. PLoS Biol. 5, 1284–1297. https://doi.org/10.1371/journal.pbio.0050149

Flavin, M., Slaughter, C., 1974. Microtubule assembly and function in *Chlamydomonas*: inhibition of growth and flagellar regeneration by antitubulins and other drugs and isolation of resistant mutants. J. Bacteriol. 118, 59–69.

Fletcher, B.J., Brentnall, S.J., Anderson, C.W., Berner, R. a., Beerling, D.J., 2008. Atmospheric carbon dioxide linked with Mesozoic and early Cenozoic climate change. Nat. Geosci. 1, 43–48. https://doi.org/10.1038/ngeo.2007.29

Franklin, S., Ngo, B., Efuet, E., Mayfield, S.P., 2002. Development of a GFP reporter gene for *Chlamydomonas reinhardtii* chloroplast. Plant J. 30, 733–744. https://doi.org/10.1046/j.1365-313X.2002.01319.x

Freeman Rosenzweig, E.S., Xu, B., Kuhn Cuellar, L., Martinez-Sanchez, A., Schaffer, M., Strauss, M., Cartwright, H.N., Ronceray, P., Plitzko, J.M., Förster, F., Wingreen, N.S., Engel, B.D., Mackinder, L.C.M., Jonikas, M.C., 2017. The eukaryotic $CO_2$-concentrating organelle is liquid-like and exhibits dynamic reorganization. Cell 171, 148–162.e19. https://doi.org/10.1016/j.cell.2017.08.008

Freibauer, A., Mathijs, E., Brunori, G., Damianova, Z., Faroult, E., Gironagomis, J., O'Brien, L., Treyer, S., 2011. Sustainable food consumption and production in a resource-constrained world. [Report] 3rd SCAR Foresight Exercise. Brussels.

Fritze, C.E., Anderson, T.R., 2000. Epitope tagging: General method for tracking recombinant proteins BT - Applications of Chimeric Genes and Hybrid Proteins - Part B: Cell Biology and Physiology. Appl. Chimeric Genes Hybrid Proteins - Part B Cell Biol. Physiol. Volume 327, 3–16. https://doi.org/10.1016/S0076-6879(00)27263-7

Fuhrmann, M., Oertel, W., Hegemann, P., 1999. A synthetic gene coding for the green fluorescent protein (GFP) is a versatile reporter in *Chlamydomonas reinhardtii*. Plant J. 19, 353–361. https://doi.org/10.1046/j.1365-313X.1999.00526.x

Fujiwara, S., Fukuzawa, H., Tachiki, A., Miyachi, S., 1990. Structure and differential

expression of two genes encoding carbonic anhydrase in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. U. S. A. 87, 9779–9783.

Fujiwara, S., Ishida, N., Tsuzuki, M., 1996. Circadian expression of the carbonic anhydrase gene, Cah1, in *Chlamydomonas reinhardtii*. Plant Mol Biol 32, 745–749. https://doi.org/10.1007/BF00020215

Fukuzawa, H., Fujiwara, S., Tachiki, A., Miyachi, S., 1990. Nucleotide sequences of two genes CAH1 and CAH2 which encode carbonic anhydrase polypeptides in *Chlamydomonas reinhardtii*. Nucleic Acids Res. 18, 6441–6442.

Fukuzawa, H., Miura, K., Ishizaki, K., Kucho, K.I., Saito, T., Kohinata, T., Ohyama, K., 2001. Ccm1, a regulatory gene controlling the induction of a carbon-concentrating mechanism in *Chlamydomonas reinhardtii* by sensing $CO_2$ availability. Proc. Natl. Acad. Sci. U. S. A. 98, 5347–5352. https://doi.org/10.1073/pnas.081593498

Gao, H., Sage, T.L., Osteryoung, K.W., 2006. FZL, an FZO-like protein in plants, is a determinant of thylakoid and chloroplast morphology. Proc. Natl. Acad. Sci. 103, 6759–6764. https://doi.org/10.1073/pnas.0507287103

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein identification and analysis tools on the ExPASy server, in: Walker, J.M. (Ed.), The Proteomics Protocols Handbook. Humana Press Inc., Totowa, New Jersey, pp. 571–607. https://doi.org/10.1385/1592598900

Genkov, T., Meyer, M., Griffiths, H., Spreitzer, R.J., 2010. Functional hybrid Rubisco enzymes with plant small subunits and algal large subunits: Engineered RbcS cDNA for expression in *Chlamydomonas*. J. Biol. Chem. 285, 19833–41. https://doi.org/10.1074/jbc.M110.124230

Genkov, T., Spreitzer, R.J., 2009. Highly conserved small subunit residues influence Rubisco large subunit catalysis. J. Biol. Chem. 284, 30105–12. https://doi.org/10.1074/jbc.M109.044081

Geraghty, A.M., Spalding, M.H., 1996. Molecular and structural changes in *Chlamydomonas* under limiting $CO_2$ (A possible mitochondrial role in adaptation). Plant Physiol. 111, 1339–1347. https://doi.org/10.1104/PP.111.4.1339

Gibbs, S.P., 1962. The ultrastructure of the pyrenoids of green algae. J. Ultrastruct. Res. 7, 262–272. https://doi.org/10.1016/S0022-5320(62)90022-9

Giordano, M., Beardall, J., Raven, J.A., 2005. $CO_2$ concentrating mechanisms in algae:

Mechanisms, environmental modulation, and evolution. Annu. Rev. Plant Biol. 56, 99–131. https://doi.org/10.1146/annurev.arplant.56.032604.144052

Golberg, I., Ohad, I., 1970. Biogenesis of chloroplast membranes V: A Radioautographic Study of Membrane Growth in a Mutant of *Chlamydomonas reinhardi y-1*. J. Cell Biol. 44, 572–591.

Gong, H.Y., Li, Y., Fang, G., Hu, D.H., Jin, W. Bin, Wang, Z.H., Li, Y.S., 2015. Transgenic rice expressing Ictb and FBP/Sbpase derived from cyanobacteria exhibits enhanced photosynthesis and mesophyll conductance to $CO_2$. PLoS One 10, 1–23. https://doi.org/10.1371/journal.pone.0140928

González-Ballester, D., de Montaigu, A., Galván, A., Fernández, E., 2005. Restriction enzyme site-directed amplification PCR: A tool to identify regions flanking a marker DNA. Anal. Biochem. 340, 330–335. https://doi.org/10.1016/j.ab.2005.01.031

Goodenough, U.W., 1970. Chloroplast division and pyrenoid formation in *Chlamydomonas reinhardtii*. J. Phycol. 6, 1–6.

Goodenough, U.W., Levine, R.P., 1970. Chloroplast structure and function in *ac-20*, a mutant strain of *Chlamydomonas reinhardi*. III. Chloroplast ribosomes and membrane organization. J. Cell Biol. 44, 547–562. https://doi.org/10.1083/jcb.44.3.547

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., Rokhsar, D.S., 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40, D1178–D1186.

Goto, M., Semimaru, T., Furukawa, K., Hayashida, S., 1994. Analysis of the raw starch-binding domain by mutation of a glucoamylase from *Aspergillus awamori* var. *kawachi* expressed in *Saccharomyces cerevisiae*. Appl. Environ. Microbiol. 60, 3926–3930.

Griffiths, D.J., 1980. The pyrenoid and its role in algal metabolism. Sci. Prog. 66, 537–553.

Griffiths, D.J., 1970. The Pyrenoid. Bot. Rev. 36, 29–58.

Grossman, A.R., Harris, E.E., Hauser, C., Lefebvre, P.A., Martinez, D., Rokhsar, D., Shrager, J., Silflow, C.D., Stern, D., Vallon, O., Zhang, Z., 2003. *Chlamydomonas*

*reinhardtii* at the crossroads of genomics. Eukaryot. Cell 2, 1137–1150. https://doi.org/10.1128/EC.2.6.1137

Gruber, M., Söding, J., Lupas, A.N., 2006. Comparative analysis of coiled-coil prediction methods. J. Struct. Biol. 155, 140–145. https://doi.org/10.1016/j.jsb.2006.03.009

Gunning, B.E.S., Schwartz, O.M., 2000. Confocal microscopy of thylakoid autofluorescence in relation to origin of grana and phylogeny in the green algae. Aust. J. Plant Physiol. 27, 645. https://doi.org/10.1071/PP97167

Hall, L.M., Taylor, K.B., Jones, D.D., 1993. Expression of a foreign gene in *Chlamydomonas reinhardtii*. Gene 124, 75–81. https://doi.org/10.1016/0378-1119(93)90763-S

Hamel, V., Steib, E., Hamelin, R., Armand, F., Borgers, S., Flückiger, I., Busso, C., Olieric, N., Sorzano, C.O.S., Steinmetz, M.O., Guichard, P., Gönczy, P., 2017. Identification of *Chlamydomonas* central core centriolar proteins reveals a role for human WDR90 in ciliogenesis. Curr. Biol. 27, 2486–2498.e6. https://doi.org/10.1016/j.cub.2017.07.011

Han, D., Li, Y., Hu, Q., 2013. Biology and commercial aspects of *Haematococcus pluvialis*. Absorpt. Adsorpt. Heavy Met. by Microalgae 388–405.

Hay, W.T., Bihmidine, S., Mutlu, N., Hoang, K. Le, Awada, T., Weeks, D.P., Clemente, T.E., Long, S.P., 2017. Enhancing soybean photosynthetic $CO_2$ assimilation using a cyanobacterial membrane protein, *ictB*. J. Plant Physiol. 212, 58–68. https://doi.org/10.1016/j.jplph.2017.02.003

He, P., Zhang, D., Chen, G., Liu, Q., Wu, W., 2003. Gold immunolocalization of Rubisco and Rubisco activase in pyrenoid of *Chlamydomonas reinhardtii*. Algae 18, 121–127. https://doi.org/10.4490/ALGAE.2003.18.2.121

Hibberd, J.M., Sheehy, J.E., Langdale, J.A., 2008. Using $C_4$ photosynthesis to increase the yield of rice - rationale and feasibility. Curr. Opin. Plant Biol. 11, 228–231. https://doi.org/10.1016/j.pbi.2007.11.002

Holdsworth, R.H., 1971. The isolation and partial characterization of the pyrenoid protein of *Eremosphaera viridis*. J. Cell Biol. 51, 499–513. https://doi.org/10.1083/jcb.51.2.499

Holdsworth, R.H., 1968. The presence of a crystalline matrix in pyrenoids of the diatom,

*Achnanthes brevipes*. J. Cell Biol. 37, 831–837. https://doi.org/10.1083/jcb.37.3.831

Houtz, R.L., Magnani, R., Nayak, N.R., Dirk, L.M.A., 2008. Co- and post-translational modifications in Rubisco: Unanswered questions. J. Exp. Bot. 59, 1635–1645. https://doi.org/10.1093/jxb/erm360

Hu, H., Boisson-Dernier, A., Israelsson-Nordström, M., Böhmer, M., Xue, S., Ries, A., Godoski, J., Kuhn, J.M., Schroeder, J.I., 2010. Carbonic anhydrases are upstream regulators of $CO_2$-controlled stomatal movements in guard cells. Nat. Cell Biol. 12, 87–93. https://doi.org/10.1038/ncb2009

Huang, K., Beck, C.F., 2003. Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. 100, 6269–6274. https://doi.org/10.1073/pnas.0931459100

Huang, K., Merkle, T., Beck, C.F., 2002. Isolation and characterization of a *Chlamydomonas* gene that encodes a putative blue-light photoreceptor of the phototropin family. Physiol. Plant. 115, 613–622. https://doi.org/10.1034/j.1399-3054.2002.1150416.x

Hutner, S.H., Provasoli, L., Schatz, A., Haskins, C.P., 1950. Some approaches to the study of the role of metals in the metabolism of microorganisms. Proc. Am. Philos. Soc. 94, 152–170.

Im, C.S., Grossman, A.R., 2002. Identification and regulation of high light-induced genes in *Chlamydomonas reinhardtii*. Plant J. 30, 301–313. https://doi.org/10.1046/j.1365-313X.2001.01287.x

Inwood, W., Yoshihara, C., Zalpuri, R., Kim, K.-S., Kustu, S., 2008. The ultrastructure of a *Chlamydomonas reinhardtii* mutant strain lacking phytoene synthase resembles that of a colorless alga. Mol Plant 1, 925–937. https://doi.org/10.1093/mp/ssn046

Ishikawa, C., Hatanaka, T., Misoo, S., Miyake, C., Fukayama, H., 2011. Functional incorporation of sorghum small subunit increases the catalytic turnover rate of Rubisco in transgenic rice. Plant Physiol. 156, 1603–1611. https://doi.org/10.1104/pp.111.177030

Jefferson, R.A., 1989. The GUS reporter gene system. Nature 342, 837–838. https://doi.org/10.1038/342837a0

Jessberger, R., 2002. The many functions of SMC proteins in chromosome dynamics. Nat Rev Mol Cell Biol 3, 767–778. https://doi.org/10.1038/nrm930

Jin, S., Sun, J., Wunder, T., Tang, D., Cousins, A.B., Sze, S.K., Mueller-Cajar, O., Gao, Y.-G., 2016. Structural insights into the LCIB protein family reveals a new group of β-carbonic anhydrases. Proc. Natl. Acad. Sci. 113, 14716–14721. https://doi.org/10.1073/pnas.1616294113

Jinkerson, R.E., Jonikas, M.C., 2015. Molecular techniques to interrogate and edit the *Chlamydomonas* nuclear genome. Plant J. 82, 393–412. https://doi.org/10.1111/tpj.12801

Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., Nickelsen, J., Stern, D.B., Wollman, F.-A., Vallon, O., 2010. MRL1, a conserved pentatricopeptide repeat protein, is required for stabilization of rbcL mRNA in *Chlamydomonas* and *Arabidopsis*. Plant Cell 22, 234–248. https://doi.org/10.1105/tpc.109.066266

Jungnick, N., Ma, Y., Mukherjee, B., Cronan, J.C., Speed, D.J., Laborde, S.M., Longstreth, D.J., Moroney, J. V., 2014. The carbon concentrating mechanism in *Chlamydomonas reinhardtii*: Finding the missing pieces. Photosynth. Res. 121, 159–173. https://doi.org/10.1007/s11120-014-0004-x

Karlsson, J., Clarke, A.K., Chen, Z.Y., Hugghins, S.Y., Park, Y.I., Husic, H., Moroney, J. V, Samuelsson, G., 1998. Novel alpha-type carbonic anhydrase associated with the thylakoid membrane in *Chlamydomonas reinhardtii* is required for growth at ambient $CO_2$. EMBO J. 17, 1208–1216.

Karlsson, J., Hiltonen, T., Husic, H.D., Ramazanov, Z., Samuelsson, G., 1995. lntracellular carbonic anhydrase of *Chlamydomonas reinhardtii*. Plant Physiol. 109, 533–539.

Kasahara, M., Swartz, T.E., Olney, M.A., Onodera, A., Mochizuki, N., Fukuzawa, H., Asamizu, E., Tabata, S., Kanegae, H., Takano, M., Christie, J.M., Nagatani, A., Briggs, W.R., 2002. Photochemical properties of the flavin mononucleotide-binding domains of the phototropins from *Arabidopsi*s, Rice, and *Chlamydomonas reinhardtii*. Plant Physiol. 129, 762–773. https://doi.org/10.1104/pp.002410.762

Kathiresan, S., Chandrashekar, A., Ravishankar, G.A., Sarada, R., 2009. *Agrobacterium*-mediated transformation in the green alga *Haematococcus pluvialis*

(Chlorophyceae, Volvocales). J. Phycol. 45, 642–649. https://doi.org/10.1111/j.1529-8817.2009.00688.x

Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.J., Rosenkranz, R., Stäbler, N., Schönfeld, B., Kreuzaler, F., Peterhänsel, C., 2007. Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in *Arabidopsis thaliana*. Nat. Biotechnol. 25, 593–599. https://doi.org/10.1038/nbt1299

Keeley, J.E., Rundel, P.W., 2003. Evolution of CAM and $C_4$ Carbon-Concentrating Mechanisms. Int. J. Plant Sci. 164, S55–S77. https://doi.org/1058-5893/2003/16403S-0005$15.00

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modelling, prediction, and analysis. Nat. Protoc. 10, 845–858. https://doi.org/10.1038/nprot.2015-053

Kendall, H.W., Pimentel, D., 2000. Constraints on the expansion of the global food supply, in: Kendall, H.W. (Ed.), A Distant Light: Scientists and Public Policy. Springer-Verlag New York, Inc., New York, pp. 179–268. https://doi.org/10.1007/978-1-4419-8507-1_20

Kendall, H.W., Pimentel, D., 1994. Constraints on the expansion of the global food supply. Ambio 23, 198–205.

Khalil, A.S., Collins, J.J., 2010. Synthetic biology: Applications come of age. Nat. Rev. Genet. 11, 367–379. https://doi.org/10.1038/nrg2775

Kim, S., Lee, Y.C., Cho, D.H., Lee, H.U., Huh, Y.S., Kim, G.J., Kim, H.S., 2014. A simple and non-invasive method for nuclear transformation of intact-walled *Chlamydomonas reinhardtii*. PLoS One 9, 1–9. https://doi.org/10.1371/journal.pone.0101018

Kindle, K.L., 1990. High-frequency nuclear transformation of *Chlamydomonas reinhardtii*. Methods Enzymol. 297, 27–38. https://doi.org/10.1073/pnas.87.3.1228

Kobayashi, Y., Takusagawa, M., Harada, N., Fukao, Y., Yamaoka, S., Kohchi, T., Hori, K., Ohta, H., Shikanai, T., Nishimura, Y., 2015. Eukaryotic components remodeled chloroplast nucleoid organization during the Green Plant Evolution. Genome Biol. Evol. 8, 1–16. https://doi.org/10.1093/gbe/evv233

Kohinata, T., Nishino, H., Fukuzawa, H., 2008. Significance of zinc in a regulatory

protein, CCM1, which regulates the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. Plant Cell Physiol. 49, 273–283. https://doi.org/10.1093/pcp/pcn003

Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23, 1289–1291. https://doi.org/10.1093/bioinformatics/btm091

Kozminski, K.G., 2012. Intraflagellar transport--the 'new motility' 20 years later. Mol. Biol. Cell 23, 751–753. https://doi.org/10.1091/mbc.E11-11-0922

Kozminski, K.G., Johnson, K.A., Forscher, P., Rosenbaum, J.L., 1993. A motility in the eukaryotic flagellum unrelated to flagellar beating. Proc. Natl. Acad. Sci. 90, 5519–5523. https://doi.org/10.1073/pnas.90.12.5519

Kuchitsu, K., Tsuzuki, M., Miyachi, S., 1988. Characterization of the pyrenoid isolated from unicellular green alga *Chlamydomonas reinhardtii*: Particulate form of Rubisco protein. Protoplasma 144, 17–24. https://doi.org/10.1007/BF01320276

Kuchitsu, K., Tsuzuki, M., Miyachi, S., 1988. Changes of starch localization within the chloroplast induced by changes in $CO_2$ concentration during growth of *Chlamydomonas reinhardtii*: Independent regulation of pyrenoid starch and stroma starch. Plant Cell Physiol. 29, 1269–1278.

Kuhn, M., Hyman, A.A., Beyer, A., 2014. Coiled-coil proteins facilitated the functional expansion of the centrosome. PLoS Comput. Biol. 10. https://doi.org/10.1371/journal.pcbi.1003657

Kumar, A., Falcao, V.R., Sayre, R.T., 2013. Evaluating nuclear transgene expression systems in *Chlamydomonas reinhardtii*. Algal Res. 2, 321–332. https://doi.org/10.1016/j.algal.2013.09.002

Kurotani, A., Sakurai, T., 2015. *In silico* analysis of correlations between protein disorder and post-translational modifications in algae. Int. J. Mol. Sci. 16, 19812–19835. https://doi.org/10.3390/ijms160819812

Lacoste-Royal, G., Gibbs, S.P., 1987. Immunocytochemical localization of ribulose-1,5-bisphosphate carboxylase in the pyrenoid and thylakoid region of the chloroplast of *Chlamydomonas reinhardtii*. Plant Physiol. 83, 602–606. https://doi.org/10.1104/pp.83.3.602

Lee, G.J., Kostov, R. V., McFadden, B.A., 1993. A facile method to determine the

$CO_2/O_2$ specificity factor for ribulose bisphosphate carboxylase/oxygenase. Photosynth. Res. 37, 81–86. https://doi.org/10.1007/BF02185441

Lefebvre, S., Lawson, T., Fryer, M., Zakhleniuk, O.V., Lloyd, J.C., Raines, C.A., 2005. Increased sedoheptulose-1,7-bisphosphatase activity in transgenic tobacco plants stimulates photosynthesis and growth from an early stage in development. Plant Physiol. 138, 451–460. https://doi.org/10.1104/pp.104.055046.1

Lehtimäki, N., Koskela, M.M., Mulo, P., 2015. Posttranslational modifications of chloroplast proteins: An emerging field. Plant Physiol. 168, 768–775. https://doi.org/10.1104/pp.15.00117

Levine, R.P., 1960. Genetic control of photosynthesis in *Chlamydomonas reinhardi*. Proc. Natl. Acad. Sci. U. S. A. 46, 972–978.

Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N., Lopez, R., 2015. The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res. 43, W580–W584. https://doi.org/10.1093/nar/gkv279

Li, X., Moellering, E.R., Liu, B., Johnny, C., Fedewa, M., Sears, B.B., Kuo, M.-H., Benning, C., 2012. A galactoglycerolipid lipase is required for triacylglycerol accumulation and survival following nitrogen deprivation in *Chlamydomonas reinhardtii*. Plant Cell 24, 4670–86. https://doi.org/10.1105/tpc.112.105106

Li, X., Zhang, R., Patena, W., Gang, S.S., Blum, S.R., Ivanova, N., Yue, R., Robertson, J.M., Lefebvre, P.A., Fitz-Gibbon, S.T., Grossman, A.R., Jonikas, M.C., 2016. An indexed, mapped mutant library enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. Plant Cell 28, 367–387. https://doi.org/10.1105/tpc.15.00465

Lieman-Hurwitz, J., Rachmilevitch, S., Mittler, R., Marcus, Y., Kaplan, A., 2003. Enhanced photosynthesis and growth of transgenic plants that express *ictB*, a gene involved in $HCO_3^-$ accumulation in cyanobacteria. Plant Biotechnol. J. 1, 43–50. https://doi.org/10.1046/j.1467-7652.2003.00003.x

Life Technologies Corporation, 2013. GeneArt® MAX Efficiency® Transformation Reagent for Algae.

Lim, B.H., 2012. Phylogenetic engineering of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit in *Chlamydomonas reinhardtii*. University of

Nebraska-Lincoln.

Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A.J., Hanson, M.R., 2014. A faster Rubisco with potential to increase photosynthesis in crops. Nature 513, 547–550. https://doi.org/10.1038/nature13776

Liu, J., Hu, Q., 2013. *Chlorella*: Industrial production of cell mass and chemicals. Handb. Microalgal Cult. Appl. Phycol. Biotechnol. 327–338. https://doi.org/10.1002/9781118567166.ch16

Lloyd, N.D.H., Canvin, D.T., Culver, D.A., 1977. Photosynthesis and photorespiration in algae. Plant Physiol. 59, 936–940.

Long, S.P., 2014. We need winners in the race to increase photosynthesis in rice, whether from conventional breeding, biotechnology or both. Plant, Cell Environ. 37, 19–21. https://doi.org/10.1111/pce.12193

Long, S.P., Marshall-Colon, A., Zhu, X.G., 2015. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. Cell 161, 56–66. https://doi.org/10.1016/j.cell.2015.03.019

Lopez-Ruiz, A., Verbelen, J.P., Roldan, J.M., Diez, J., 1985. Nitrate reductase of green algae is located in the pyrenoid. Plant Physiol. 79, 1006–1010. https://doi.org/10.1104/pp.79.4.1006

Ma, Y., Pollock, S. V, Xiao, Y., Cunnusamy, K., Moroney, J. V, 2011. Identification of a novel gene, CIA6, required for normal pyrenoid formation in *Chlamydomonas reinhardtii*. Plant Physiol. 156, 884–896. https://doi.org/10.1104/pp.111.173922

Machingura, M.C., Bajsa-Hirschel, J., Laborde, S.M., Schwartzenburg, J.B., Mukherjee, B., Mukherjee, A., Pollock, S. V., Förster, B., Price, G.D., Moroney, J. V., 2017. Identification and characterization of a solute carrier, CIA8, involved in inorganic carbon acclimation in *Chlamydomonas reinhardtii*. J. Exp. Bot. https://doi.org/10.1093/jxb/erx189

Mackinder, L.C.M., 2017. The *Chlamydomonas* $CO_2$-concentrating mechanism and its potential for engineering photosynthesis in plants. New Phytol. https://doi.org/10.1111/nph.14749

Mackinder, L.C.M., Chen, C., Leib, R.D., Patena, W., Blum, S.R., Rodman, M., Ramundo, S., Adams, C.M., Jonikas, M.C., 2017. A spatial interactome reveals the protein organization of the algal $CO_2$-concentrating mechanism. Cell 171, 133–147.

https://doi.org/10.1016/j.cell.2017.08.044

Mackinder, L.C.M., Meyer, M.T., Mettler-Altmann, T., Chen, V.K., Mitchell, M.C., Caspari, O., Freeman Rosenzweig, E.S., Pallesen, L., Reeves, G., Itakura, A., Roth, R., Sommer, F., Geimer, S., Mühlhaus, T., Schroda, M., Goodenough, U., Stitt, M., Griffiths, H., Jonikas, M.C., 2016. A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. Proc. Natl. Acad. Sci. 113, 5958–5963. https://doi.org/10.1073/pnas.1522866113

Mahlow, S., Orzechowski, S., Fettke, J., 2016. Starch phosphorylation: insights and perspectives. Cell. Mol. Life Sci. 73, 2753–2764. https://doi.org/10.1007/s00018-016-2248-4

Maier, A., Fahnenstich, H., von Caemmerer, S., Engqvist, M.K.M., Weber, A.P.M., Flügge, U.-I., Maurino, V.G., 2012. Transgenic introduction of a glycolate oxidative cycle into *A. thaliana* chloroplasts leads to growth improvement. Front. Plant Sci. 3, 1–12. https://doi.org/10.3389/fpls.2012.00038

Marchler-Bauer, A., Bo, Y., Han, L., He, J., Lanczycki, C.J., Lu, S., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Geer, L.Y., Bryant, S.H., 2017. CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. Nucleic Acids Res. 45, D200–D203. https://doi.org/10.1093/nar/gkw1129

Matthews, B.F., Saunders, J. A., Gebhardt, J.S., Lin, J., Koehler, S.M., 1995. Reporter genes and transient assays for plants, in: Nickoloff, J.A. (Ed.), Plant Cell Electroporation and Electrofusion. Humana Press Inc., Totowa, New Jersey, pp. 147–162. https://doi.org/10.1385/0-89603-328-7:147

McKay, R.M.L., Gibbs, S.P., 1991. Composition and function of pyrenoids: cytochemical and immunocytochemical approaches. Can. J. Bot. 69, 1040–1052. https://doi.org/10.1139/b91-134

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., Lopez, R., 2013. Analysis tool web services from the EMBL-EBI. Nucleic Acids Res. 41, 597–600. https://doi.org/10.1093/nar/gkt376

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Maréchal-Drouard, L., Marshall, W.F.,

Qu, L.-H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., *Chlamydomonas* Annotation Team, JGI Annotation Team, Grigoriev, I. V., Rokhsar, D.S., Grossman, A.R., 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. Science 318, 245–251. https://doi.org/10.1126/science.1143609

Meslet-Cladière, L., Vallon, O., 2012. A new method to identify flanking sequence tags in *Chlamydomonas* using 3'-RACE. Plant Methods 8, 21. https://doi.org/10.1186/1746-4811-8-21

Meyer, M., Griffiths, H., 2015. The internal plumbing of algal chloroplasts. Elife 4, 1–3. https://doi.org/10.7554/eLife.04889

Meyer, M., Griffiths, H., 2013. Origins and diversity of eukaryotic $CO_2$-concentrating mechanisms: lessons for the future. J. Exp. Bot. 64, 769–86. https://doi.org/10.1093/jxb/ers390

Meyer, M.T., Genkov, T., Skepper, J.N., Jouhet, J., Mitchell, M.C., Spreitzer, R.J., Griffiths, H., 2012. Rubisco small-subunit α-helices control pyrenoid formation in *Chlamydomonas*. Proc. Natl. Acad. Sci. 109, 19474–19479.

Meyer, M.T., Whittaker, C., Griffiths, H., 2017. The algal pyrenoid: key unanswered questions. J. Exp. Bot. 68, 3739–3749. https://doi.org/10.1093/jxb/erx178

Mitchell, M., 2014. Regulation of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. University of Cambridge.

Mitchell, M.C., Meyer, M.T., Griffiths, H., 2014. Dynamics of carbon-concentrating mechanism induction and protein relocalization during the dark-to-light transition in synchronized *Chlamydomonas reinhardtii*. Plant Physiol. 166, 1073–82. https://doi.org/10.1104/pp.114.246918

Mitra, M., Lato, S.M., Ynalvez, R.A., Xiao, Y., Moroney, J. V., 2004. Identification of a new chloroplast carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Physiol. 135, 173–182. https://doi.org/10.1104/pp.103.037283

Miura, K., Yamano, T., Yoshioka, S., Kohinata, T., Inoue, Y., Taniguchi, F., Asamizu, E., Nakamura, Y., Tabata, S., Yamato, K.T., Ohyama, K., Fukuzawa, H., 2004. Expression profiling-based identification of $CO_2$-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas*

*reinhardtii*. Plant Physiol. 135, 1595–607. https://doi.org/10.1104/pp.104.041400

Miyachi, S., Tsuzuki, M., Maruyoma, I., Gantar, M., Miyachi, S., 1986. Effects of $CO_2$ concentration during growth on the intracellular structure of Chlorella and Scenedesmus (Chlorophyta). J. Phycol. 22, 313–320.

Miyagawa, Y., Tamoi, M., Shigeoka, S., 2001. Overexpression of a cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. Nat. Biotechnol. 19, 965–969. https://doi.org/10.1038/nbt1001-965

Morita, E., Kuroiwa, H., Kuoriwa, T., Hisayoshi, N., 1997. High localization of ribulose-1,5-bisphosphate carboxylase/oxygenase in the pyrenoids of *Chlamydomonas reinhardtii* (Chlorophyta), as revealed by cryofixation and immunogold electron microscopy. J. Phycol. 33, 68–72.

Moroney, J. V., Husic, H.D., Tolbert, N.E., Kitayama, M., Manuel, L.J., Togasaki, R.K., 1989. Isolation and characterization of a mutant of *Chlamydomonas reinhardtii* deficient in the $CO_2$ concentrating mechanism. Plant Physiol. 89, 897–903. https://doi.org/10.1104/pp.89.3.897

Moroney, J. V., Jungnick, N., DiMario, R.J., Longstreth, D.J., 2013. Photorespiration and carbon concentrating mechanisms: two adaptations to high O2, low $CO_2$ conditions. Photosynth. Res. 117, 121–131. https://doi.org/10.1007/s11120-013-9865-7

Moroney, J. V., Ynalvez, R.A., 2007. Proposed carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*. Eukaryot. Cell 6, 1251–9. https://doi.org/10.1128/EC.00064-07

Nakamura, Y., Gojobori, T., Ikemura, T., 2000. Codon usage tabulated from the international DNA sequence databases: status for the year 2000. Nucleic Acids Res. 28, 292. https://doi.org/10.1093/nar/27.1.292

Nelson, C.E., Hersh, B.M., Carroll, S.B., 2004. The regulatory content of intergenic DNA shapes genome architecture. Genome Biol. 5, R25. https://doi.org/10.1186/gb-2004-5-4-r25

Neupert, J., Karcher, D., Bock, R., 2009. Generation of *Chlamydomonas* strains that efficiently express nuclear transgenes. Plant J. 57, 1140–1150. https://doi.org/10.1111/j.1365-313X.2008.03746.x

Neupert, J., Shao, N., Lu, Y., Bock, R., 2012. Genetic transformation of the model green alga *Chlamydomonas reinhardtii*, in: Dunwell, J., Wetten, R. (Eds.), Transgenic Plants. Methods in Molecular Biology (Methods and Protocols). Humana Press, pp. 35–47. https://doi.org/10.1007/978-1-61779-558-9

Nölke, G., Houdelet, M., Kreuzaler, F., Peterhänsel, C., Schillberg, S., 2014. The expression of a recombinant glycolate dehydrogenase polyprotein in potato (*Solanum tuberosum*) plastids strongly enhances photosynthesis and tuber yield. Plant Biotechnol. J. 12, 734–742. https://doi.org/10.1111/pbi.12178

Ohad, I., Siekevitz, P., Palade, G.E., 1967a. Biogenesis of chloroplast membranes. I. Plastid differentiation during greening of a dark-grown algal mutant (*Chlamydomonas reinhardi*). J. Cell Biol. 35, 553–584. https://doi.org/10.1083/jcb.35.3.553

Ohad, I., Siekevitz, P., Palade, G.E., 1967b. Biogenesis of chloroplast membranes. II. Plastid differentiation during greening of a dark-grown algal mutant (*Chlamydomonas reinhardi*). J. Cell Biol. 35, 553–584. https://doi.org/10.1083/jcb.35.3.553

Ohnishi, N., Mukherjee, B., Tsujikawa, T., Yanase, M., Nakano, H., Moroney, J. V, Fukuzawa, H., 2010. Expression of a low $CO_2$-inducible protein, LCI1, increases inorganic carbon uptake in the green alga *Chlamydomonas reinhardtii*. Plant Cell 22, 3105–17. https://doi.org/10.1105/tpc.109.071811

Owen, N.A., Griffiths, H., 2013. A system dynamics model integrating physiology and biochemical regulation predicts extent of crassulacean acid metabolism (CAM) phases. New Phytol. 200, 1116–1131. https://doi.org/10.1111/nph.12461

Park, J.M., Cho, J.H., Kang, S.G., Jang, H.J., Pih, K.T., Piao, H.L., Cho, M.J., Hwang, I., 1998. A dynamin-like protein in *Arabidopsis thaliana* is involved in biogenesis of thylakoid membranes. EMBO J. 17, 859–867. https://doi.org/10.1093/emboj/17.4.859

Pazour, G.J., Agrin, N., Leszyk, J., Witman, G.B., 2005. Proteomic analysis of a eukaryotic cilium. J. Cell Biol. 170, 103–113. https://doi.org/10.1083/jcb.200504008

Peterhänsel, C., Krause, K., Braun, H.-P., Espie, G.S., Fernie, A.R., Hanson, D.T., Keech, O., Maurino, V.G., Mielewczik, M., Sage, R.F., 2013. Engineering

photorespiration: current state and future possibilities. Plant Biol. 15, 754–8. https://doi.org/10.1111/j.1438-8677.2012.00681.x

Pollock, S. V, Colombo, S.L., Jr, D.L.P., Godfrey, A.C., Moroney, J. V, 2003. Rubisco activase is required for optimal photosynthesis in the green alga *Chlamydomonas reinhardtii* in a Low-$CO_2$ Atmosphere. Plant Physiol. 133, 1854–1861. https://doi.org/10.1104/pp.103.032078.mutant

Porra, R.J., Thompson, W.A., Kriedemann, P.E., 1989. Determination of accurate extinction coefficients and simulteneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim. Biophys. Acta 975, 384–394.

Price, G.D., Badger, M.R., Woodger, F.J., Long, B.M., 2008. Advances in understanding the cyanobacterial $CO_2$-concentrating mechanism (CCM): Functional components, $C_i$ transporters, diversity, genetic regulation and prospects for engineering into plants. J. Exp. Bot. 59, 1441–1461. https://doi.org/10.1093/jxb/erm112

Price, G.D., Pengelly, J.J.L., Forster, B., Du, J., Whitney, S.M., von Caemmerer, S., Badger, M.R., Howitt, S.M., Evans, J.R., 2013. The cyanobacterial CCM as a source of genes for improving photosynthetic $CO_2$ fixation in crop species. J. Exp. Bot. 64, 753–768. https://doi.org/10.1093/jxb/err313

Qu, Z., Hartzell, H.C., 2008. Bestrophin $Cl^-$ channels are highly permeable to $HCO_3^-$. AJP Cell Physiol. 294, C1371–C1377. https://doi.org/10.1152/ajpcell.00398.2007

Rae, B.D., Long, B.M., Förster, B., Nguyen, N.D., Velanis, C.N., Atkinson, N., Hee, W.Y., Mukherjee, B., Price, G.D., McCormick, A.J., 2017. Progress and challenges of engineering a biophysical carbon dioxide-concentrating mechanism into higher plants. J. Exp. Bot. 68, 3717–3737. https://doi.org/10.1093/jxb/erx133

Ramazanov, Z., Rawat, M., Henk, M.C., Mason, C.B., Matthews, S.W., Moroney, J. V, 1994. The induction of the $CO_2$-concentrating mechanism is correlated with the formation of the starch sheath around the pyrenoid of *Chlamydomonas reinhardtii*. Planta 195, 210–216. https://doi.org/10.1007/BF00199681

Rasala, B.A., Barrera, D.J., Ng, J., Plucinak, T.M., Rosenberg, J.N., Weeks, D.P., Oyler, G.A., Peterson, T.C., Haerizadeh, F., Mayfield, S.P., 2013. Expanding the spectral

palette of fluorescent proteins for the green microalga *Chlamydomonas reinhardtii*. Plant J. 74, 545–556. https://doi.org/10.1111/tpj.12165

Rasala, B.A., Lee, P.A., Shen, Z., Briggs, S.P., Mendez, M., Mayfield, S.P., 2012. Robust expression and secretion of xylanase1 in *Chlamydomonas reinhardtii* by fusion to a selection gene and processing with the FMDV 2A peptide. PLoS One 7. https://doi.org/10.1371/journal.pone.0043349

Rast, A., Heinz, S., Nickelsen, J., 2015. Biogenesis of thylakoid membranes. Biochim. Biophys. Acta 1847, 821–830. https://doi.org/10.1016/j.bbabio.2015.01.007

Rathod, J.P., Prakash, G., Pandit, R., Lali, A.M., 2013. *Agrobacterium*-mediated transformation of promising oil-bearing marine algae *Parachlorella kessleri*. Photosynth. Res. 118, 141–146. https://doi.org/10.1007/s11120-013-9930-2

Raven, J.A., 1997. $CO_2$-concentrating mechanisms: A direct role for thylakoid lumen acidification? Plant Cell Environ. 20, 147–154. https://doi.org/10.1046/j.1365-3040.1997.d01-67.x

Raven, J.A., Beardall, J., Sánchez-Baracaldo, P., 2017. The possible evolution, and future, of $CO_2$-concentrating mechanisms. J. Exp. Bot. 68, 3701–3716. https://doi.org/10.1093/jxb/erx110

Raven, J.A., Cockell, C.S., De La Rocha, C.L., 2008. The evolution of inorganic carbon concentrating mechanisms in photosynthesis. Philos. Trans. R. Soc. B Biol. Sci. 363, 2641–2650. https://doi.org/10.1098/rstb.2008.0020

Raven, J.A., Giordano, M., Beardall, J., 2008. Insights into the evolution of CCMs from comparisons with other resource acquisition and assimilation processes. Physiol. Plant. 133, 4–14. https://doi.org/10.1111/j.1399-3054.2007.01024.x

Raven, J.A., Giordano, M., Beardall, J., Maberly, S.C., 2012. Algal evolution in relation to atmospheric $CO_2$: carboxylases, carbon-concentrating mechanisms and carbon oxidation cycles. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 367, 493–507. https://doi.org/10.1098/rstb.2011.0212

Rawat, M., Henk, M.C., Lavigne, L.L., Moroney, J. V., 1996. *Chlamydomonas reinhardtii* mutants without ribulose-1,5-bisphosphate carboxylase-oxygenase lack a detectable pyrenoid. Planta 198, 263–270. https://doi.org/10.1007/BF00206252

Rawat, M., Moroney, J. V., 1991. Partial characterization of a new isoenzyme of carbonic anhydrase. Biochemistry 266, 9719–9723.

Recuenco-Muñoz, L., Offre, P., Valledor, L., Lyon, D., Weckwerth, W., Wienkoop, S., 2015. Targeted quantitative analysis of a diurnal RuBisCO subunit expression and translation profile in *Chlamydomonas reinhardtii* introducing a novel Mass Western approach. J. Proteomics 113, 143–153. https://doi.org/10.1016/j.jprot.2014.09.026

Remacle, C., Cardol, P., Coosemans, N., Gaisne, M., Bonnefoy, N., 2006. High-efficiency biolistic transformation of *Chlamydomonas* mitochondria can be used to insert mutations in complex I genes. Proc. Natl. Acad. Sci. 103, 4771–4776. https://doi.org/10.1073/pnas.0509501103

Renberg, L., Johansson, A.I., Shutova, T., Stenlund, H., Aksmann, A., Raven, J. a, Gardeström, P., Moritz, T., Samuelsson, G., 2010. A metabolomic approach to study major metabolite changes during acclimation to limiting $CO_2$ in *Chlamydomonas reinhardtii*. Plant Physiol. 154, 187–196. https://doi.org/10.1104/pp.110.157651

Rochaix, J.-D., 2002. The three genomes of *Chlamydomonas*. Photosynth. Res. 73, 285–293.

Rochaix, J.-D., 1995. *Chlamydomonas reinhardtii* as the photosynthetic yeast. Annu. Rev. Genet. 29, 209–230. https://doi.org/10.1146/annurev.genet.29.1.209

Rochaix, J.D., van Dillewijn, J., 1982. Transformation of the green alga *Chlamydomonas* reinhardii with yeast DNA. Nature. https://doi.org/10.1038/296070a0

Romero-Campero, F.J., Perez-Hurtado, I., Lucas-Reina, E., Romero, J.M., Valverde, F., 2016. ChlamyNET: a *Chlamydomonas* gene co-expression network reveals global properties of the transcriptome and the early setup of key co-expression patterns in the green lineage. BMC Genomics 17, 227. https://doi.org/10.1186/s12864-016-2564-y

Rosenthal, D.M., Locke, A.M., Khozaei, M., Raines, C.A., Long, S.P., Ort, D.R., 2011. Over-expressing the C3 photosynthesis cycle enzyme Sedoheptulose-1-7 Bisphosphatase improves photosynthetic carbon gain and yield under fully open air $CO_2$ fumigation (FACE). BMC Plant Biol. 11, 123. https://doi.org/10.1186/1471-2229-11-123

Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M.C., Leonhardt,

H., 2008. A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. Mol. Cell. Proteomics 7, 282–289. https://doi.org/10.1074/mcp.M700342-MCP200

Rueden, C.T., Schindelin, J., Hiner, M.C., DeZonia, B.E., Walter, A.E., Arena, E.T., Eliceiri, K.W., 2017. ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinformatics 1, 1–26. https://doi.org/10.1186/s12859-017-1934-z

Sager, R., 1954. Mendelian and non-mendelian inheritance of streptomycin resistance in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. 40, 356–363.

Sager, R., Palade, G.E., 1957. Structure and development of the chloroplast in *Chlamydomonas*. J. Biophys. Biochem. Cytol. 3.

Sanford, J.C., 1990. Biolistic plant transformation. Physiol. Plant. 79, 206–209. https://doi.org/10.1111/j.1399-3054.1990.tb05888.x

Scaife, M.A., Nguyen, G.T., Rico, J., Lambert, D., Helliwell, K.E., Smith, A.G., 2015. Establishing *Chlamydomonas reinhardtii* as an industrial biotechnology host. Plant J. 82, 532–546. https://doi.org/10.1111/tpj.12781

Scaife, M.A., Smith, A.G., 2016. Towards developing algal synthetic biology. Biochem. Soc. Trans. 44, 716–722. https://doi.org/10.1042/BST20160061

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: An open-source platform for biological-image analysis. Nat. Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675. https://doi.org/10.1038/nmeth.2089

Schroda, M., Blöcker, D., Beck, C.F., 2000. The HSP70A promoter as a tool for the improved expression of transgenes in *Chlamydomonas*. Plant J. 21, 121–131. https://doi.org/10.1046/j.1365-313X.2000.00652.x

Schuler, M.L., Mantegazza, O., Weber, A.P.M., 2016. Engineering $C_4$ photosynthesis into $C_3$ chassis in the synthetic biology age. Plant J. 87, 51–65. https://doi.org/10.1111/tpj.13155

Scranton, M.A., Ostrand, J.T., Fields, F.J., Mayfield, S.P., 2015. *Chlamydomonas* as a

model for biofuels and bio-products production. Plant J. 82, 523–531. https://doi.org/10.1111/tpj.12780

Shah, M.M.R., Liang, Y., Cheng, J.J., Daroch, M., 2016. Astaxanthin-producing green microalga *Haematococcus pluvialis*: From single cell to high value commercial products. Front. Plant Sci. 7. https://doi.org/10.3389/fpls.2016.00531

Sharwood, R.E., 2017. Engineering chloroplasts to improve Rubisco catalysis: prospects for translating improvements into food and fiber crops. New Phytol. 213, 494–510. https://doi.org/10.1111/nph.14351

Shen, C., Dupont, C.L., Hopkinson, B.M., 2017. The diversity of carbon dioxide-concentrating mechanisms in marine diatoms as inferred from their genetic content. J. Exp. Bot. 68, 3937–3948. https://doi.org/10.1093/jxb/erx163

Shibata, M., Katoh, H., Sonoda, M., Ohkawa, H., Shimoyama, M., Fukuzawa, H., Kaplan, A., Ogawa, T., 2002. Genes essential to sodium-dependent bicarbonate transport in cyanobacteria: Function and phylogenetic analysis. J. Biol. Chem. 277, 18658–18664. https://doi.org/10.1074/jbc.M112468200

Shimogawara, K., Fujiwara, S., Grossman, A., Usuda, H., 1998. High-efficiency transformation of *Chlamydomonas reinhardtii* by electroporation. Genetics 148, 1821–1828.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7. https://doi.org/10.1038/msb.2011.75

Simkin, A.J., McAusland, L., Headland, L.R., Lawson, T., Raines, C.A., 2015. Multigene manipulation of photosynthetic carbon assimilation increases $CO_2$ fixation and biomass yield in tobacco. J. Exp. Bot. 66, 4075–4090. https://doi.org/10.1093/jxb/erv204

Sinetova, M.A., Kupriyanova, E. V, Markelova, A.G., Allakhverdiev, S.I., Pronina, N.A., 2012. Identification and functional role of the carbonic anhydrase Cah3 in thylakoid membranes of pyrenoid of *Chlamydomonas reinhardtii*. Biochim. Biophys. Acta 1817, 1248–55. https://doi.org/10.1016/j.bbabio.2012.02.014

Sizova, I., Fuhrmann, M., Hegemann, P., 2001. A *Streptomyces rimosus aphVIII* gene coding for a new type phosphotransferase provides stable antibiotic resistance to

*Chlamydomonas reinhardtii*. Gene 277, 221–229. https://doi.org/10.1016/S0378-1119(01)00616-3

Smith, E.C., Griffiths, H., 1996. A pyrenoid-based carbon-concentrating mechanism is present in terrestrial bryophytes of the class *Anthocerotae*. Planta 200, 203–212. https://doi.org/10.1007/BF00208310

Snapp, E., 2005. Design and Use of Fluorescent Fusion Proteins in Cell Biology. Curr. Protoc. Cell Biol. 27, 21.4.1-24.4.13. https://doi.org/https://doi.org/10.1002/0471143030.cb2104s27

Snapp, E.L., 2009. Fluorescent proteins: a cell biologist's user guide. Trends Cell Biol. 19, 649–655. https://doi.org/10.1016/j.tcb.2009.08.002

Sorimachi, K., Gal-Coëffet, M.-F. Le, Williamson, G., Archer, D.B., Williamson, M.P., 1997. Solution structure of the granular starch binding domain of *Aspergillus niger* glucoamylase bound to β-cyclodextrin. Structure 5, 647–661. https://doi.org/10.1016/S0969-2126(97)00220-7

Spalding, M.H., 2008. Microalgal carbon-dioxide-concentrating mechanisms: *Chlamydomonas* inorganic carbon transporters. J. Exp. Bot. 59, 1463–73. https://doi.org/10.1093/jxb/erm128

Spalding, M.H., Spreitzer, R.J., Ogren, W.L., 1983. Carbonic anhydrase-deficient mutant of *Chlamydomonas reinhardii* requires elevated carbon dioxide concentration for photoautotrophic growth. Plant Physiol. 73, 268–72. https://doi.org/10.1104/pp.73.2.268

Spreitzer, R.J., Goldschmidt-Clermont, M., Rahire, M., Rochaix, J.D., 1985. Nonsense mutations in the *Chlamydomonas* chloroplast gene that codes for the large subunit of ribulosebisphosphate carboxylase/oxygenase. Proc. Natl. Acad. Sci. U. S. A. 82, 5460–5464. https://doi.org/10.1073/pnas.82.16.5460

Spreitzer, R.J., Mets, L., 1981. Photosynthesis-deficient mutants of *Chlamydomonas* reinhardii with associated light-sensitive phenotypes. Plant Physiol. 67, 565–569. https://doi.org/10.1104/pp.67.3.565

Spreitzer, R.J., Salvucci, M.E., 2002. Rubisco: Structure, regulatory interactions, and possibilities for a better enzyme. Annu. Rev. Plant Biol. 53, 449–475. https://doi.org/10.1146/annurev.arplant.53.100301.135233

Strokopytov, B., Knegtel, R.M., Penninga, D., Rozeboom, H.J., Kalk, K.H., Dijkhuizen,

L., Dijkstra, B.W., 1996. Structure of cyclodextrin glycosyltransferase complexed with a maltononaose inhibitor at 2.6 angstrom resolution. Implications for product specificity. Biochemistry 35, 4241–4249. https://doi.org/10.1021/bi952339h

Süss, K.-H., Prokhorenko, I., Adler, K., 1995. *In situ* association of Calvin cycle enzymes, ribulose-1,5-bisphosphate carboxylase/oxygenase activase, ferredoxin-$NADP^+$ reductase, and nitrite reductase with thylakoid and pyrenoid membranes of *Chlamydomonas reinhardtii* chloroplasts as revealed by immunoelectron microscopy. Plant Physiol. 107, 1387–1397. https://doi.org/10.1104/pp.107.4.1387

Tachiki, A., Fukuzawa, H., Miyachi, S., 1992. Characterization of carbonic anhydrase isozyme CA2, which is the CAH2 gene product, in *Chlamydomonas reinhardtii*. Biosci. Biotechnol. Biochem. 56, 794–798.

Tam, L.W., Lefebvre, P. a., 1993. Cloning of flagellar genes in *Chlamydomonas reinhardtii* by DNA insertional mutagenesis. Genetics 135, 375–384.

Tardif, M., Atteia, A., Specht, M., Cogne, G., Rolland, N., Brugière, S., Hippler, M., Ferro, M., Bruley, C., Peltier, G., Vallon, O., Cournac, L., 2012. PredAlgo: A new subcellular localization prediction tool dedicated to green algae. Mol. Biol. Evol. 29, 3625–3639. https://doi.org/10.1093/molbev/mss178

Tenenboim, H., Smirnova, J., Willmitzer, L., Steup, M., Brotman, Y., 2014. VMP1-deficient *Chlamydomonas* exhibits severely aberrant cell morphology and disrupted cytokinesis. BMC Plant Biol. 14, 121. https://doi.org/10.1186/1471-2229-14-121

Terbush, A.D., MacCready, J.S., Chen, C., Ducat, D.C., Osteryoung, K.W., 2017. Conserved dynamics of chloroplast cytoskeletal FtsZ Proteins Across Photosynthetic Lineages. Plant Physiol. pp.00558.2017. https://doi.org/10.1104/pp.17.00558

TerBush, A.D., Yoshida, Y., Osteryoung, K.W., 2013. FtsZ in chloroplast division: Structure, function and evolution. Curr. Opin. Cell Biol. 25, 461–470. https://doi.org/10.1016/j.ceb.2013.04.006

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., Garg, K., John, S., Sandstrom, R., Bates, D., Boatman, L., Canfield, T.K., Diegel, M., Dunn, D., Ebersol, A.K., Frum, T., Giste, E., Johnson, A.K., Johnson, E.M., Kutyavin, T.,

Lajoie, B., Lee, B.-K., Lee, K., London, D., Lotakis, D., Neph, S., Neri, F., Nguyen, E.D., Qu, H., Reynolds, A.P., Roach, V., Safi, A., Sanchez, M.E., Sanyal, A., Shafer, A., Simon, J.M., Song, L., Vong, S., Weaver, M., Yan, Y., Zhang, Z., Zhang, Z., Lenhard, B., Tewari, M., Dorschner, M.O., Hansen, R.S., Navas, P.A., Stamatoyannopoulos, G., Iyer, V.R., Lieb, J.D., Sunyaev, S.R., Akey, J.M., Sabo, P.J., Kaul, R., Furey, T.S., Dekker, J., Crawford, G.E., Stamatoyannopoulos, J.A., 2012. The accessible chromatin landscape of the human genome. Nature 489, 75–82. https://doi.org/10.1038/nature11232

Thyssen, C., Schlichting, R., Giersch, C., 2001. The $CO_2$-concentrating mechanism in the physiological context: lowering the $CO_2$ supply diminishes culture growth and economises starch utilisation in *Chlamydomonas reinhardtii*. Planta 213, 629–639. https://doi.org/10.1007/s004250100534

Trippens, J., Greiner, A., Schellwat, J., Neukam, M., Rottmann, T., Lu, Y., Kateriya, S., Hegemann, P., Kreimer, G., 2012. Phototropin influence on eyespot development and regulation of phototactic behavior in *Chlamydomonas reinhardtii*. Plant Cell 24, 4687–4702. https://doi.org/10.1105/tpc.112.103523

Tsuzuki, M., Gantar, M., Aizawa, K., Miyachi, S., 1986. Ultrastructure of *Dunaliella tertiolecta* cells grown under low and high $CO_2$ concentrations. Plant Cell Physiol 27, 737–739.

Uniacke, J., Zerges, W., 2007. Photosystem II assembly and repair are differentially localized in *Chlamydomonas*. Plant Cell 19, 3640–3654. https://doi.org/10.1105/tpc.107.054882

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3 - new capabilities and interfaces. Nucleic Acids Res. 40, 1–12. https://doi.org/10.1093/nar/gks596

Urrutia, R., Henley, J.R., Cook, T., McNiven, M. a, 1997. The dynamins: redundant or distinct functions for an expanding family of related GTPases? Proc. Natl. Acad. Sci. U. S. A. 94, 377–84. https://doi.org/10.1073/pnas.94.2.377

Van, K., Spalding, M.H., 1999. Periplasmic carbonic anhydrase structural gene (Cah1) mutant in *Chlamydomonas reinhardtii*. Plant Physiol. 120, 757–764. https://doi.org/10.1104/pp.120.3.757

Van, K., Wang, Y., Nakamura, Y., Spalding, M.H., 2001. Insertional mutants of

*Chlamydomonas reinhardtii* that require elevated $CO_2$ for survival. Plant Physiol. 127, 607–614. https://doi.org/10.1104/pp.010333.1

van Lun, M., Hub, J.S., van der Spoel, D., Andersson, I., 2014. $CO_2$ and $O_2$ distribution in Rubisco suggests the small subunit functions as a $CO_2$ reservoir. J. Am. Chem. Soc. 136, 3165–3171. https://doi.org/10.1021/ja411579b

Villarejo, A., Martinez, F., Plumed, M.D.P., Ramazanov, Z., 1996. The induction of the $CO_2$ concentrating mechanism in a starch-less mutant of *Chlamydomonas reinhardtii*. Physiol. Plant. 98, 798–802. https://doi.org/10.1034/j.1399-3054.1996.980417.x

Wagner, V., Ullmann, K., Mollwo, A., Kaminski, M., Mittag, M., Kreimer, G., 2007. The phosphoproteome of a *Chlamydomonas reinhardtii* eyespot fraction includes key proteins of the light signaling pathway. Plant Physiol. 146, 772–788. https://doi.org/10.1104/pp.107.109645

Walne, P.L., 1967. The effects of colchicine on cellular organization in *Chlamydomonas* II. Ultrastructure. Am. J. Bot. 54, 564–577.

Walne, P.L., 1966. The effects of colchicine on cellular organization in *Chlamydomonas*. I. Light microscopy and cytochemistry. Am. J. Bot. 53, 908–916.

Wang, L., Yamano, T., Kajikawa, M., Hirono, M., Fukuzawa, H., 2014. Isolation and characterization of novel high-$CO_2$-requiring mutants of *Chlamydomonas reinhardtii*. Photosynth. Res. 121, 175–184. https://doi.org/10.1007/s11120-014-9983-x

Wang, L., Yamano, T., Takane, S., Niikawa, Y., Toyokawa, C., Ozawa, S., Tokutsu, R., Takahashi, Y., Minagawa, J., Kanesaki, Y., Yoshikawa, H., Fukuzawa, H., 2016. Chloroplast-mediated regulation of $CO_2$-concentrating mechanism by $Ca^{2+}$-binding protein CAS in the green alga *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. 113, 12586–12591. https://doi.org/10.1073/pnas.1606519113

Wang, Y., Duanmu, D., Spalding, M.H., 2011. Carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*: Inorganic carbon transport and $CO_2$ recapture. Photosynth. Res. 109, 115–122. https://doi.org/10.1007/s11120-011-9643-3

Wang, Y., Spalding, M.H., 2014. Acclimation to very low $CO_2$: Contribution of limiting $CO_2$ inducible proteins, LCIB and LCIA, to inorganic carbon uptake in

*Chlamydomonas reinhardtii*. Plant Physiol. 166, 2040–2050. https://doi.org/10.1104/pp.114.248294

Wang, Y., Spalding, M.H., 2014. LCIB in the *Chlamydomonas* $CO_2$-concentrating mechanism. Photosynth. Res. 121, 185–192. https://doi.org/10.1007/s11120-013-9956-5

Wang, Y., Stessman, D.J., Spalding, M.H., 2015. The $CO_2$ concentrating mechanism and photosynthetic carbon assimilation in limiting $CO_2$: How *Chlamydomonas* works against the gradient. Plant J. 82, 429–448. https://doi.org/10.1111/tpj.12829

Warr, J.R., 1968. A mutant of *Chlamydomonas reinhardii* with abnormal cell division. J. Gen. Microbiol. 52, 243–251.

Weber, E., Engler, C., Gruetzner, R., Werner, S., Marillonnet, S., 2011. A modular cloning system for standardized assembly of multigene constructs. PLoS One 6. https://doi.org/10.1371/journal.pone.0016765

Weber, K., Geisler, N., Plessmann, U., Bremerich, A., Lechtreck, K.F., Melkonian, M., 1993. SF-assemblin, the structural protein of the 2-nm filaments from striated microtubule associated fibers of algal flagellar roots, forms a segmented coiled coil. J. Cell Biol. 121, 837–845. https://doi.org/10.1083/jcb.121.4.837

Wen, Z., Johnson, M.B., 2009. Microalgae as a feedstock for biofuel production. Virginia Coop. Ext. Publ. 442-886 7.

Wheeler, G.L., Brownlee, C., 2008. $Ca^{2+}$ signalling in plants and green algae - changing channels. Trends Plant Sci. 13, 506–514. https://doi.org/10.1016/j.tplants.2008.06.004

Wichuk, K., Brynjólfsson, S., Fu, W., 2014. Biotechnological production of value-added carotenoids from microalgae: Emerging technology and prospects. Bioengineered 5, 204–208.

Winck, F.V., Arvidsson, S., Riaño-Pachón, D.M., Hempel, S., Koseska, A., Nikoloski, Z., Urbina Gomez, D.A., Rupprecht, J., Mueller-Roeber, B., 2013. Genome-wide identification of regulatory elements and reconstruction of gene regulatory networks of the green alga *Chlamydomonas reinhardtii* under carbon deprivation. PLoS One 8, e79909. https://doi.org/10.1371/journal.pone.0079909

Wloga, D., Gaertig, J., 2011. Post-translational modifications of microtubules. J. Cell Sci. 124, 154–154. https://doi.org/10.1242/jcs.083576

Xiang, Y., Zhang, J., Weeks, D.P., 2001. The Cia5 gene controls formation of the carbon concentrating mechanism in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. U. S. A. 98, 5341–5346. https://doi.org/10.1073/pnas.101534498

Xing, S., Wallmeroth, N., Berendzen, K.W., Grefen, C., 2016. Techniques for the analysis of protein-protein interactions in vivo. Plant Physiol. 171, pp.00470.2016. https://doi.org/10.1104/pp.16.00470

Xu, M., Bernát, G., Singh, A., Mi, H., Rögner, M., Pakrasi, H.B., Ogawa, T., 2008. Properties of mutants of *Synechocystis* sp. strain PCC 6803 lacking inorganic carbon sequestration systems. Plant Cell Physiol. 49, 1672–1677. https://doi.org/10.1093/pcp/pcn139

Yamano, T., Asada, A., Sato, E., Fukuzawa, H., 2014. Isolation and characterization of mutants defective in the localization of LCIB, an essential factor for the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. Photosynth. Res. 121, 193–200. https://doi.org/10.1007/s11120-013-9963-6

Yamano, T., Sato, E., Iguchi, H., Fukuda, Y., Fukuzawa, H., 2015. Characterization of cooperative bicarbonate uptake into chloroplast stroma in the green alga *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. 112, 7315–7320. https://doi.org/10.1073/pnas.1501659112

Yamano, T., Tsujikawa, T., Hatano, K., Ozawa, S.I., Takahashi, Y., Fukuzawa, H., 2010. Light and low-$CO_2$-dependent LCIB-LCIC complex localization in the chloroplast supports the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. Plant Cell Physiol. 51, 1453–1468. https://doi.org/10.1093/pcp/pcq105

Ynalvez, R.A., Moroney, J. V., 2008. Identification and characterisation of a novel inorganic carbon acquisition gene, CIA7, from an insertional mutant of *Chlamydomonas reinhardtii*. Funct. Plant Biol. 35, 373–381. https://doi.org/10.1071/FP08005

Yoshioka, S., Taniguchi, F., Miura, K., Inoue, T., Yamano, T., Fukuzawa, H., 2004. The novel Myb transcription factor LCR1 regulates the $CO_2$-responsive gene Cah1, encoding a periplasmic carbonic anhydrase in *Chlamydomonas reinhardtii*. Plant Cell 16, 1466–77. https://doi.org/10.1105/tpc.021162

Zhang, R., Patena, W., Armbruster, U., Gang, S.S., Blum, S.R., Jonikas, M.C., 2014. High-throughput genotyping of green algal mutants reveals random distribution of

mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. Plant Cell 26, 1398–1409.

Zhu, X.-G., Long, S.P., Ort, D.R., 2010. Improving photosynthetic efficiency for greater yield. Annu. Rev. Plant Biol. 61, 235–261. https://doi.org/10.1146/annurev-arplant-042809-112206

Zittelli, G.C., Biondi, N., Rodolfi, L., Tredici, M.R., 2013. Photobioreactors for mass production of microalgae. Handb. Microalgal Cult. Appl. Phycol. Biotechnol. 225–266. https://doi.org/10.1002/9781118567166.ch13

Zones, J.M., Blaby, I.K., Merchant, S.S., Umen, J.G., 2015. High-Resolution Profiling of a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic Differentiation. Plant Cell 27, 2743–2769. https://doi.org/10.1105/tpc.15.00498

# 8 APPENDICES

Appendix 1 Alignment of the whole genome sequencing data of the CCM defective mutants used in this study.

Appendix 2 Pyrenoid morphology of wild type *Chlamydomonas reinhardtii* cells. Scale bar 500 nm.

Appendix 3 The lack of visible thylakoid tubules network in the pyrenoids of *saga* cells. Scale bar 500 nm.

Appendix 4 The presence of traversing thylakoid tubules in the pyrenoids of *saga* cells. Scale bar 500 nm.

Appendix 5 The restoration of traversing thylakoid tubule network in the pyrenoids of *saga::SAGA-Venus-3xFLAG* A7 cells. Scale bar 500 nm.

Appendix 6 The restoration of traversing thylakoid tubule network in the pyrenoids of *saga::SAGA-Venus-3xFLAG* B2 cells. Scale bar 500 nm.

Appendix 7 Proteomic report displaying the protein identification probability of all the proteins detected by LC-MS/MS of the GFP-Trap_A coimmunoprecipitation assay.

**Appendix 1 Alignment of the whole genome sequencing data of the CCM defective mutants used in this study.**

| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| A5.sam - 14 | A05 | chloroplast | 193279 | | | | |
| | A05 | chromosome_11 | 1327257 | Cre11.g467712 | chromosome_11:1315112..1333013 forward | PTHR18937 - STRUCTURAL MAINTENANCE OF CHROMOSOMES SMC FAMILY MEMBER | g11642.t1 Cre18.g743450.t1.1 |
| | A05 | chromosome_11 | 1327257 | | | | |
| | A05 | chromosome_11 | 1327304 | | | | |
| | A05 | chromosome_11 | 1327304 | | | | |
| | A05 | chromosome_11 | 1327329 | | | | |
| | A05 | chromosome_11 | 1327329 | | | | |
| | A05 | chromosome_11 | 1327357 | | | | |
| | A05 | chromosome_11 | 1327456 | | | | |
| | A05 | chromosome_11 | 1327608 | | | | |
| | A05 | chromosome_12 | 5285369 | Cre12.g528750 | chromosome_12:5283273..5285510 forward | Ribosomal protein L12, component of cytosolic 80S ribosome and 60S large subunit | g13020.t1 Cre12.g528750.t1.1, RPL12 |
| | A05 | chromosome_16 | 3697340 | Cre16.g688850 | chromosome_16:3690780..3697605 reverse | Hexose/glycerol-3-phosphate transporter | g16171.t1, HPT2 |
| | A05 | chromosome_6 | 1716970 | Cre06.g261500 | chromosome_6:1714710..1719679 reverse | PTHR11260:SF136 - THIOREDOXIN FAMILY PROTEIN | g5877.t1, Cre06.g261500.t1.1, CPLD58 |
| | A05 | chromosome_8 | 2521924 | Cre08.g372100 | chromosome_8:2518227..2522077 reverse | Heat shock protein 70A | g8716.t1, Cre08.g372100.t1.1, HSP70A |
| A7.sam - 48 | A07 | chromosome_1 | 16865 | NA | | | |
| | A07 | chromosome_1 | 16997 | NA | | | |
| | A07 | chromosome_10 | 3606899 | NA | | | |
| | A07 | chromosome_11 | 1317016 | Cre11.g467712 | chromosome_11:1315112..1333013 forward | PTHR18937 - STRUCTURAL MAINTENANCE OF CHROMOSOMES SMC FAMILY | g11642.t1, Cre18.g743450.t1.1 |
| | A07 | chromosome_11 | 1317016 | | | | |
| | A07 | chromosome_11 | 1317126 | | | | |

| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| | A07 | chromosome_11 | 1317139 | | | MEMBER | |
| | A07 | chromosome_11 | 1317139 | | | | |
| | A07 | chromosome_11 | 1317139 | | | | |
| | A07 | chromosome_11 | 1317141 | | | | |
| | A07 | chromosome_11 | 1317141 | | | | |
| | A07 | chromosome_11 | 1317141 | | | | |
| | A07 | chromosome_11 | 1317143 | | | | |
| | A07 | chromosome_11 | 1317143 | | | | |
| | A07 | chromosome_11 | 1317144 | | | | |
| | A07 | chromosome_11 | 1317145 | | | | |
| | A07 | chromosome_11 | 1317163 | | | | |
| | A07 | chromosome_11 | 1317184 | | | | |
| | A07 | chromosome_11 | 2052537 | Cre11.g468600 | chromosome_11:2047506..2053373 | Unannotated | g11766.t1, Cre11.g468600.t1.1 |
| | A07 | chromosome_11 | 2052558 | | | | |
| | A07 | chromosome_11 | 2052558 | | | | |
| | A07 | chromosome_11 | 2052563 | | | | |
| | A07 | chromosome_11 | 2052600 | | | | |
| | A07 | chromosome_11 | 2052600 | | | | |
| | A07 | chromosome_12 | 5285472 | Cre12.g528750 | chromosome_12:5283273..5285510 forward | Ribosomal protein L12, component of cytosolic 80S ribosome and 60S large subunit | g13020.t1, Cre12.g528750.t1.1, RPL12 |
| | A07 | chromosome_14 | 4150454 | Cre14.g634322 | chromosome_14:4147306..4150687 forward | Differentially upregulated in photoheterotrophic highlight in MidLog vs. Early growth phase comparison (Gene Atlas Desc ) | g15294.t1 |
| | A07 | chromosome_14 | 4150657 | | | | |
| | A07 | chromosome_2 | 6501166 | Cre02.g116500 | chromosome_2:6499550..6505579 forward | Serine carboxypeptidase | g2411.t1, CPY3 |
| | A07 | chromosome_2 | 6501167 | | | | |

| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| | A07 | chromosome_2 | 6943511 | Cre02.g120200 | chromosome_2:6943486..6946581 reverse | 4.1.1.39 - Ribulose-bisphosphate carboxylase / RuBP carboxylase | g2496.t1, Cre02.g120200.t1.1 |
| | A07 | chromosome_5 | 2978392 | Cre05.g238332 | chromosome_5:2978242..2979235 reverse | Photosystem I reaction center subunit II, 20 kDa | g5492.t1, PSAD, Cre56.g791050.t1.1, Cre56.g791050.t1.2 |
| | A07 | chromosome_5 | 2978392 | | | | |
| | A07 | chromosome_5 | 2979858 | NA | | | |
| | A07 | chromosome_5 | 2979873 | NA | | | |
| | A07 | chromosome_5 | 2979881 | NA | | | |
| | A07 | chromosome_5 | 2979883 | NA | | | |
| | A07 | chromosome_5 | 2979944 | NA | | | |
| | A07 | chromosome_5 | 2987758 | NA | | | |
| | A07 | chromosome_5 | 2987786 | NA | | | |
| | A07 | chromosome_5 | 2987808 | NA | | | |
| | A07 | chromosome_5 | 2987808 | NA | | | |
| | A07 | chromosome_6 | 6088874 | Cre06.g289800 | chromosome_6:6083476..6089657 forward | 2.3.1.67 - 1-alkylglycerophosphocholine O-acetyltransferase / Platelet-activating factor-synthesizing enzyme | g6724.t1, Cre06.g289800.t1.1, CAM12 |
| | A07 | chromosome_6 | 6088874 | | | | |
| | A07 | chromosome_7 | 1066443 | Cre07.g319650 | chromosome_7:1049013..1069845 reverse | FixL-like PAS domain protein | g7393.t1, FXL4 |
| | A07 | chromosome_7 | 2384226 | Cre07.g328550 | chromosome_7:2383948..2397143 reverse | Subunit of the VPS-C complex | g7642.t1, VPS18 |
| | A07 | chromosome_9 | 5399714 | Cre09.g399960 | chromosome_9:5391745..5403687 reverse | 2.7.7.49 - RNA-directed DNA polymerase / Revertase | g9945.t1 |
| | A07 | scaffold_27 | 15418 | Cre27.g757197 | scaffold_27:5860..16535 reverse | 2.7.7.49 - RNA-directed DNA polymerase / Revertase | g18301.t1 |
| | A07 | scaffold_30 | 22965 | Cre30.g758147 | scaffold_30:20600..23925 reverse | Unannotated | g18328.t1 |

| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| A8.sam - 20 | A08 | chromosome_10 | 1247881 | Cre10.g426750 | chromosome_10:1246442..1252817 reverse | Cytochrome P450, CYP120 superfamily | g10545.t1, Cre10.g426750.t1.1, CYP29, CYP739A5 |
| | A08 | chromosome_11 | 1321442 | Cre11.g467712 | chromosome_11:1315112..1333013 forward | PTHR18937 - STRUCTURAL MAINTENANCE OF CHROMOSOMES SMC FAMILY MEMBER | g11642.t1, Cre18.g743450.t1.1 |
| | A08 | chromosome_11 | 1321442 | | | | |
| | A08 | chromosome_11 | 1321445 | | | | |
| | A08 | chromosome_11 | 1321451 | | | | |
| | A08 | chromosome_11 | 1321534 | | | | |
| | A08 | chromosome_11 | 1321534 | | | | |
| | A08 | chromosome_11 | 1321536 | | | | |
| | A08 | chromosome_12 | 5285483 | Cre12.g528750 | chromosome_12:5283273..5285510 forward | Ribosomal protein L12, component of cytosolic 80S ribosome and 60S large subunit | g13020.t1, Cre12.g528750.t1.1, RPL12 |
| | A08 | chromosome_12 | 6172792 | NA | | | |
| | A08 | chromosome_13 | 3070511 | Cre13.g584600 | chromosome_13:3066500..3070925 reverse | Unannotated | g14349.t1 |
| | A08 | chromosome_2 | 6943067 | Cre02.g120150 | chromosome_2:6941968..6943439 reverse | Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit 2 | g2495.t1, Cre02.g120150.t1.1, RBCS2 |
| | A08 | chromosome_7 | 4832694 | Cre07.g345900 | chromosome_7:4812506..4837787 forward | Dicer-like protein | g8048.t1, DCL3 |
| | A08 | chromosome_7 | 4832695 | | | | |
| | A08 | chromosome_7 | 4832699 | | | | |
| | A08 | chromosome_7 | 4832723 | | | | |
| | A08 | chromosome_7 | 4832729 | | | | |
| | A08 | chromosome_7 | 4832735 | | | | |
| | A08 | chromosome_7 | 4832766 | | | | |
| | A08 | chromosome_8 | 2521923 | Cre08.g372100 | chromosome_8:2518227..2522077 reverse | Heat shock protein 70A | g8716.t1, Cre08.g372100.t1.1, HSP70A |

| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| A11.sam - 18 | A11 | chloroplast | 100486 | | | | |
| | A11 | chloroplast | 119032 | | | | |
| | A11 | chromosome_11 | 1324877 | Cre11.g467712 | chromosome_11:1315112..1333013 forward | PTHR18937 - STRUCTURAL MAINTENANCE OF CHROMOSOMES SMC FAMILY MEMBER | g11642.t1, Cre18.g743450.t1.1 |
| | A11 | chromosome_11 | 1324960 | | | | |
| | A11 | chromosome_11 | 1324960 | | | | |
| | A11 | chromosome_11 | 1324967 | | | | |
| | A11 | chromosome_11 | 1324967 | | | | |
| | A11 | chromosome_11 | 1324969 | | | | |
| | A11 | chromosome_11 | 1324983 | | | | |
| | A11 | chromosome_11 | 2365925 | | | | |
| | A11 | chromosome_14 | 412742 | | | | |
| | A11 | chromosome_2 | 6943341 | Cre02.g120200 | chromosome_2:6943486..6946581 reverse | 4.1.1.39 - Ribulose-bisphosphate carboxylase / RuBP carboxylase | g2496.t1, Cre02.g120200.t1.1 |
| | A11 | chromosome_2 | 6946588 | | | | |
| | A11 | chromosome_5 | 2978498 | Cre05.g238332 | chromosome_5:2978242..2979235 reverse | Photosystem I reaction center subunit II, 20 kDa | g5492.t1, PSAD, Cre56.g791050.t1.1, Cre56.g791050.t1.2 |
| | A11 | chromosome_5 | 2987809 | NA | | | |
| | A11 | chromosome_6 | 7630960 | Cre06.g301350 | chromosome_6:7628850..7635904 forward | Differentially upregulated in heterotrophic ammonium in MidLog vs. Early growth phase comparison | g6972.t1 |
| | A11 | chromosome_7 | 5919864 | Cre07.g354100 | chromosome_7:5913020..5921352 reverse | K07739 - elongator complex protein 3 [EC:2.3.1.48] (ELP3, KAT9) | g8238.t1, Cre07.g354100.t1.1, HAT4 |
| | A11 | chromosome_9 | 7474428 | Cre09.g413700 | chromosome_9:7472736..7475677 forward | PTHR11132//PTHR11132:SF125 - SOLUTE CARRIER FAMILY 35 // SUBFAMILY NOT NAMED | g10248.t1, TPT13 |
| C4.sam - 22 | C04 | chromosome_1 | 6613293 | Cre01.g047350 | chromosome_1:6611336..6617929 forward | Unannotated | g1053.t1 |

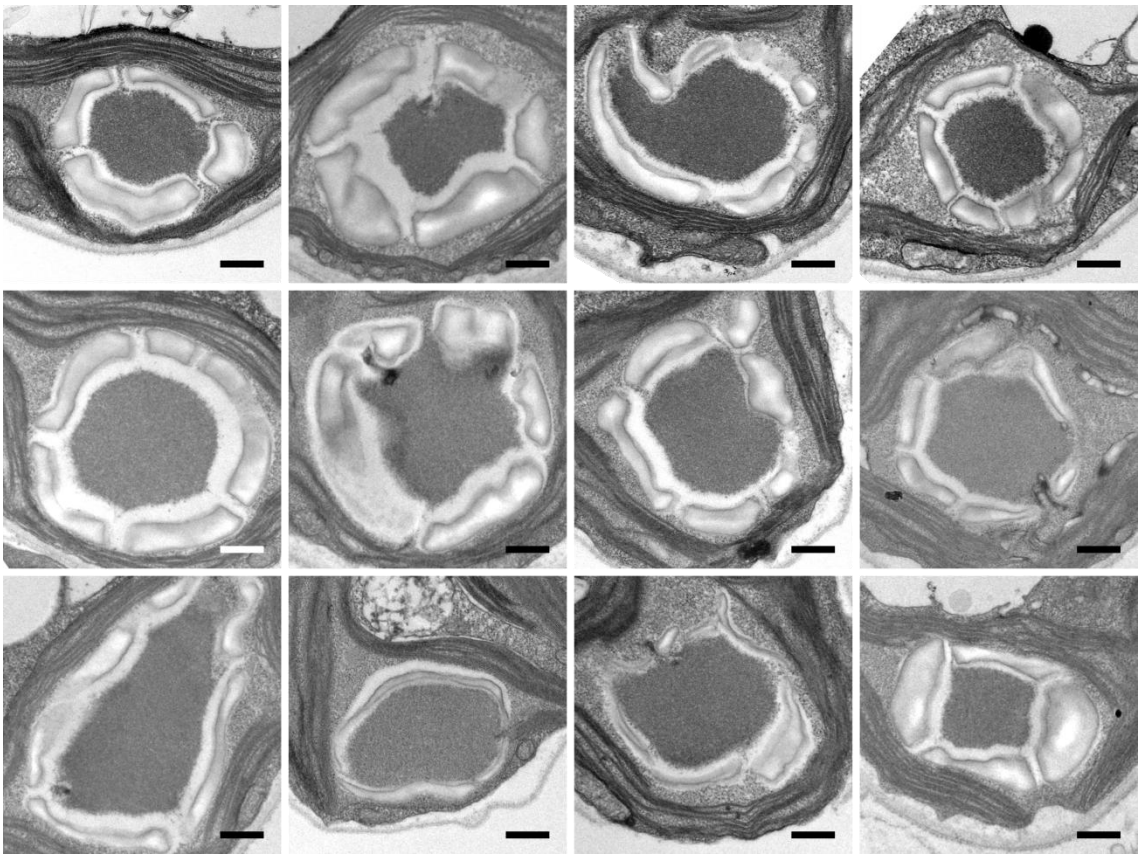
| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| | C04 | chromosome_11 | 1219982 | Cre11.g467697 | chromosome_11:1219972..1221943 forward | Unannotated | Cre11.g467697.t2.1 Cre11.g467697.t4.1, Cre11.g467697.t3.1 |
| | C04 | chromosome_11 | 1220028 | | | | |
| | C04 | chromosome_11 | 1220028 | | | | |
| | C04 | chromosome_16 | 5993959 | Cre16.g676600 | chromosome_16:5991432..5994829 forward | Unannotated | g16583.t1 |
| | C04 | chromosome_16 | 5994078 | | | | |
| | C04 | chromosome_16 | 5994085 | | | | |
| | C04 | chromosome_16 | 5994085 | | | | |
| | C04 | chromosome_16 | 5994085 | | | | |
| | C04 | chromosome_2 | 9115221 | Cre02.g143387 | chromosome_2:9114756..9117427 | Unannotated | g2829.t1 |
| | C04 | chromosome_2 | 9115221 | | | | |
| | C04 | chromosome_2 | 9115263 | | | | |
| | C04 | chromosome_2 | 9115427 | | | | |
| | C04 | chromosome_2 | 9115427 | | | | |
| | C04 | chromosome_2 | 9115427 | | | | |
| | C04 | chromosome_2 | 9115438 | | | | |
| | C04 | chromosome_2 | 9115438 | | | | |
| | C04 | chromosome_2 | 9115438 | | | | |
| | C04 | chromosome_5 | 2978173 | Cre05.g238332 | chromosome_5:2978242..2979235 reverse | Photosystem I reaction center subunit II, 20 kDa | g5492.t1, PSAD, Cre56.g791050.t1.1, Cre56.g791050.t1.2 |
| | C04 | chromosome_5 | 2987804 | NA | | | |
| | C04 | chromosome_6 | 8668788 | Cre06.g309550 | chromosome_6:8662125..8670223 reverse | Unannotated | g7162.t2 |
| | C04 | mitochondrial | 9665 | | | | |
| C8.sam - 22 | C08 | chromosome_11 | 1328224 | Cre11.g467712 | chromosome_11:1315112..1333013 forward | PTHR18937 - STRUCTURAL MAINTENANCE OF | g11642.t1, Cre18.g743450.t1.1 |
| | C08 | chromosome_11 | 1328237 | | | | |

| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| | C08 | chromosome_11 | 1328292 | | | CHROMOSOMES SMC FAMILY MEMBER | |
| | C08 | chromosome_11 | 1328292 | | | | |
| | C08 | chromosome_11 | 1328319 | | | | |
| | C08 | chromosome_11 | 1328322 | | | | |
| | C08 | chromosome_11 | 1328342 | | | | |
| | C08 | chromosome_11 | 1328342 | | | | |
| | C08 | chromosome_11 | 1328459 | | | | |
| | C08 | chromosome_11 | 1328473 | | | | |
| | C08 | chromosome_11 | 1328473 | | | | |
| | C08 | chromosome_11 | 1328473 | | | | |
| | C08 | chromosome_11 | 1328475 | | | | |
| | C08 | chromosome_11 | 1328479 | | | | |
| | C08 | chromosome_11 | 1328519 | | | | |
| | C08 | chromosome_11 | 1328526 | | | | |
| | C08 | chromosome_11 | 1328534 | | | | |
| | C08 | chromosome_12 | 2613474 | Cre12.g505000 | chromosome_12:2612227..2614512 reverse | Unannotated | g12468.t1, Cre12.g505000.t1.1 |
| | C08 | chromosome_12 | 5285434 | Cre12.g528750 | chromosome_12:5283273..5285510 forward | Ribosomal protein L12, component of cytosolic 80S ribosome and 60S large subunit | g13020.t1, Cre12.g528750.t1.1, RPL12 |
| | C08 | chromosome_14 | 4117329 | Cre14.g634113 | chromosome_14:4115936..4117720 forward | Unannotated | g15289.t1, Cre14.g634113.t1.1 |
| | C08 | chromosome_14 | 4117329 | | | | |
| | C08 | chromosome_5 | 2978507 | Cre05.g238332 | chromosome_5:2978242..2979235 reverse | Photosystem I reaction center subunit II, 20 kDa | g5492.t1, PSAD, Cre56.g791050.t1.1 , Cre56.g791050.t1.2 |
| D4.sam - | D04 | chromosome_12 | 6796973 | NA | | | |

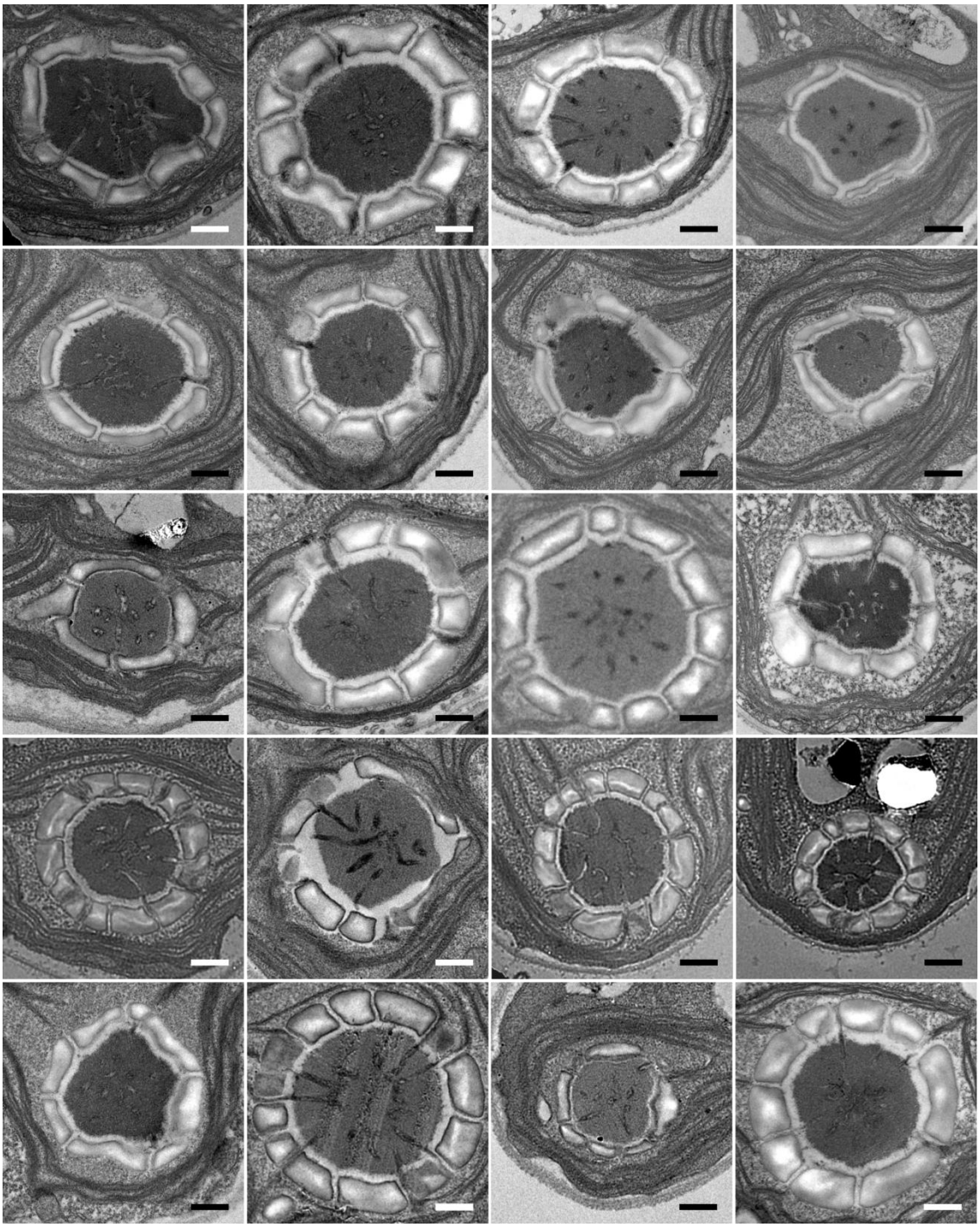
| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| 16 | D04 | chromosome_3 | 6842594 | Cre03.g198288 | chromosome_3:6839620..6848765 reverse | Unannotated | g4112.t4 |
| | D04 | chromosome_3 | 7131432 | NA | | | |
| | D04 | chromosome_5 | 2978010 | NA | | | |
| | D04 | chromosome_6 | 4414218 | NA | | | |
| | D04 | chromosome_6 | 4414253 | NA | | | |
| | D04 | chromosome_6 | 4414416 | NA | | | |
| | D04 | chromosome_6 | 4414421 | NA | | | |
| | D04 | chromosome_6 | 4414421 | NA | | | |
| | D04 | chromosome_6 | 4414429 | NA | | | |
| | D04 | chromosome_6 | 4414450 | NA | | | |
| | D04 | chromosome_6 | 4414450 | NA | | | |
| | D04 | chromosome_6 | 4414455 | NA | | | |
| | D04 | chromosome_6 | 4414467 | NA | | | |
| | D04 | chromosome_6 | 4614273 | NA | | | |
| | D04 | chromosome_6 | 4623799 | Cre06.g278286 | chromosome_6:4602520..4610791 forward | RNA-directed DNA polymerase / Revertase | g6443.t1 |
| F8.sam - 20 | F08 | chloroplast | 165961 | | | | |
| | F08 | chromosome_1 | 5290248 | Cre01.g036950 | chromosome_1:5285897..5290902 reverse | 2.7.8.26 - Adenosylcobinamide-GDP ribazoletransferase / Cobalamin-5'-phosphate synthase | g829.t1, Cre01.g036950.t1.1, CBA1 |
| | F08 | chromosome_14 | 4145975 | Cre14.g634279 | chromosome_14:4144686..4147259 forward | Unannotated | g15293.t1 |
| | F08 | chromosome_2 | 5004434 | Cre02.g104250 | chromosome_2:5002840..5015585 reverse | PTHR18945:SF559 - PROTEIN LGC-12 | g2133.t1 |
| | F08 | chromosome_2 | 6943035 | Cre02.g120150 | chromosome_2:6941968..6943439 reverse | Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit 2 | g2495.t1, Cre02.g120150.t1.1, RBCS2 |

| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| | F08 | chromosome_2 | 6943540 | Cre02.g120200 | chromosome_2:6943486..6946581 reverse | 4.1.1.39 - Ribulose-bisphosphate carboxylase / RuBP carboxylase | g2496.t1, Cre02.g120200.t1.1 |
| | F08 | chromosome_3 | 8308502 | Cre03.g199100 | chromosome_3:8304748..8308853 reverse | Unannotated | g4346.t1 Cre03.g199100.t1.1 |
| | F08 | chromosome_3 | 8326313 | Cre03.g199000 | chromosome_3:8321420..8329663 reverse | Phototropin | g4349.t1, Cre03.g199000.t1.1, PHT1 , PHOT, PHOT1 |
| | F08 | chromosome_3 | 8326315 | | | | |
| | F08 | chromosome_3 | 8326315 | | | | |
| | F08 | chromosome_3 | 8326315 | | | | |
| | F08 | chromosome_3 | 8326315 | | | | |
| | F08 | chromosome_3 | 8326315 | | | | |
| | F08 | chromosome_3 | 8326336 | | | | |
| | F08 | chromosome_3 | 8326427 | | | | |
| | F08 | chromosome_3 | 8326427 | | | | |
| | F08 | chromosome_3 | 8326431 | | | | |
| | F08 | chromosome_5 | 605828 | Cre05.g232050 | chromosome_5:602953..613048 reverse | Unannotated | g5138.t1 |
| | F08 | chromosome_5 | 2979843 | Cre05.g238332 | chromosome_5:2978242..2979235 reverse | Photosystem I reaction center subunit II, 20 kDa | g5492.t1, PSAD, Cre56.g791050.t1.1, Cre56.g791050.t1.2 |
| | F08 | chromosome_5 | 2987766 | NA | | | |
| G6.sam - 12 | G06 | chromosome_12 | 5285420 | Cre12.g528750 | chromosome_12:5283273..5285510 forward | Ribosomal protein L12, component of cytosolic 80S ribosome and 60S large subunit | g13020.t1, Cre12.g528750.t1.1, RPL12 |
| | G06 | chromosome_14 | 1279807 | Cre14.g616600 | chromosome_14:1274369..1283459 forward | PTHR11649:SF36 - FZL | g14855.t1, Cre14.g616600.t1.1 |
| | G06 | chromosome_14 | 1279850 | | | | |
| | G06 | chromosome_14 | 1279896 | | | | |
| | G06 | chromosome_14 | 1280010 | | | | |
| | G06 | chromosome_14 | 1280027 | | | | |

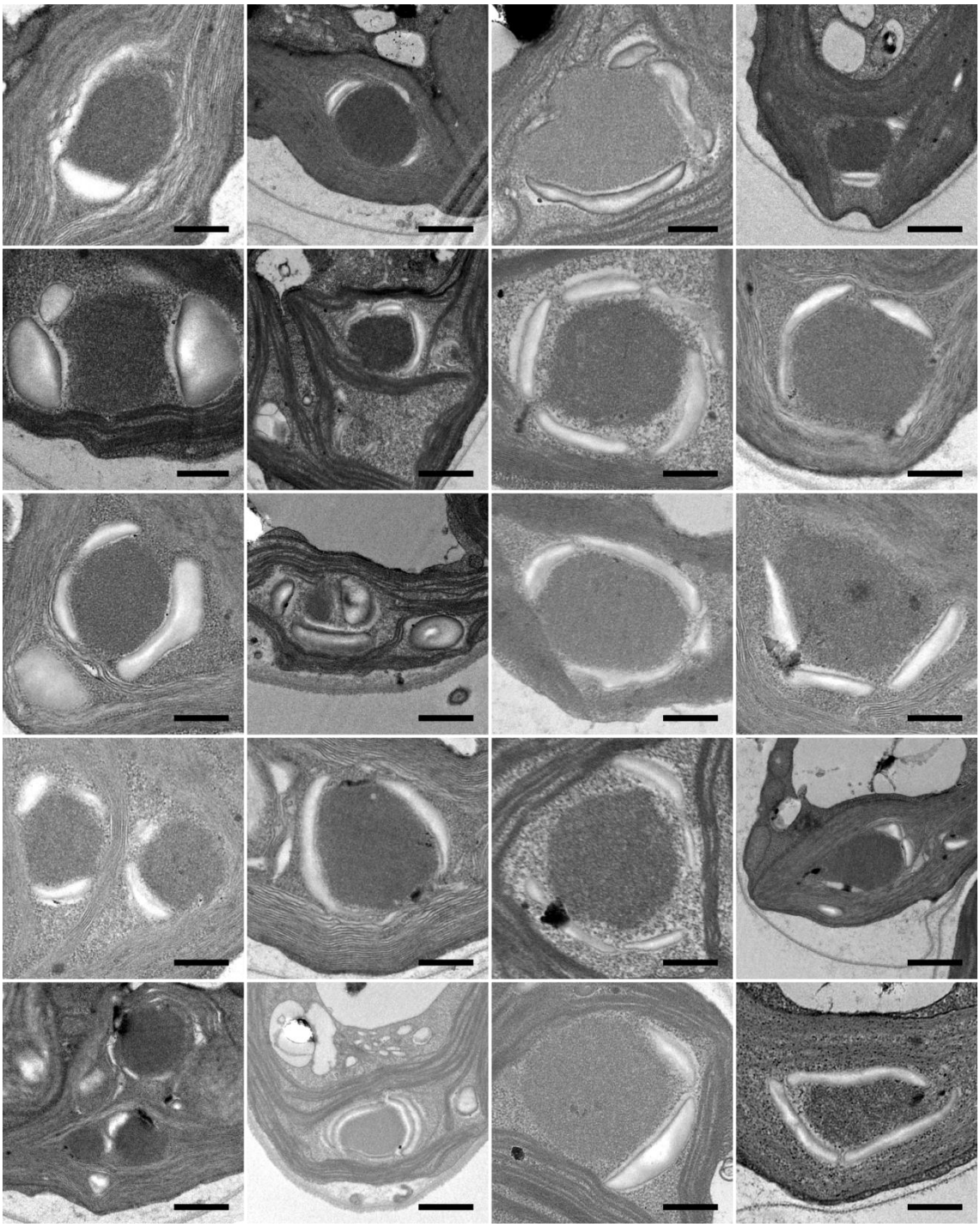
| SAM file name | Sample | Chromosome | Location | Gene | Location | Description | Alias |
|---|---|---|---|---|---|---|---|
| | G06 | chromosome_14 | 1280029 | | | | |
| | G06 | chromosome_14 | 1280032 | | | | |
| | G06 | chromosome_14 | 1280090 | | | | |
| | G06 | chromosome_14 | 1280215 | | | | |
| | G06 | chromosome_2 | 6943178 | Cre02.g120150 | chromosome_2:6941968..6943439 reverse | Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit 2 | g2495.t1, Cre02.g120150.t1.1, RBCS2 |
| | G06 | chromosome_2 | 6946371 | Cre02.g120200 | chromosome_2:6943486..6946581 reverse | 4.1.1.39 - Ribulose-bisphosphate carboxylase / RuBP carboxylase | g2496.t1, Cre02.g120200.t1.1 |
| H10.sam - 11 | H10 | chromosome_12 | 5285373 | Cre12.g528750 | chromosome_12:5283273..5285510 forward | Ribosomal protein L12, component of cytosolic 80S ribosome and 60S large subunit | g13020.t1, Cre12.g528750.t1.1, RPL12 |
| | H10 | chromosome_12 | 5285480 | | | | |
| | H10 | chromosome_2 | 6943212 | Cre02.g120200 | chromosome_2:6943486..6946581 reverse | 4.1.1.39 - Ribulose-bisphosphate carboxylase / RuBP carboxylase | g2496.t1, Cre02.g120200.t1.1 |
| | H10 | chromosome_5 | 2978499 | Cre05.g238332 | chromosome_5:2978242..2979235 reverse | Photosystem I reaction center subunit II, 20 kDa | g5492.t1, PSAD, Cre56.g791050.t1.1, Cre56.g791050.t1.2 |
| | H10 | chromosome_6 | 1428745 | Cre06.g259100 | chromosome_6:1418964..1430024 reverse | Unannotated | g5825.t1 |
| | H10 | chromosome_6 | 1428772 | | | | |
| | H10 | chromosome_6 | 1428805 | | | | |
| | H10 | chromosome_6 | 4289350 | NA | | | |
| | H10 | chromosome_6 | 4289350 | | | | |
| | H10 | mitochondrial | 4628 | | | | |
| | H10 | mitochondrial | 4628 | | | | |



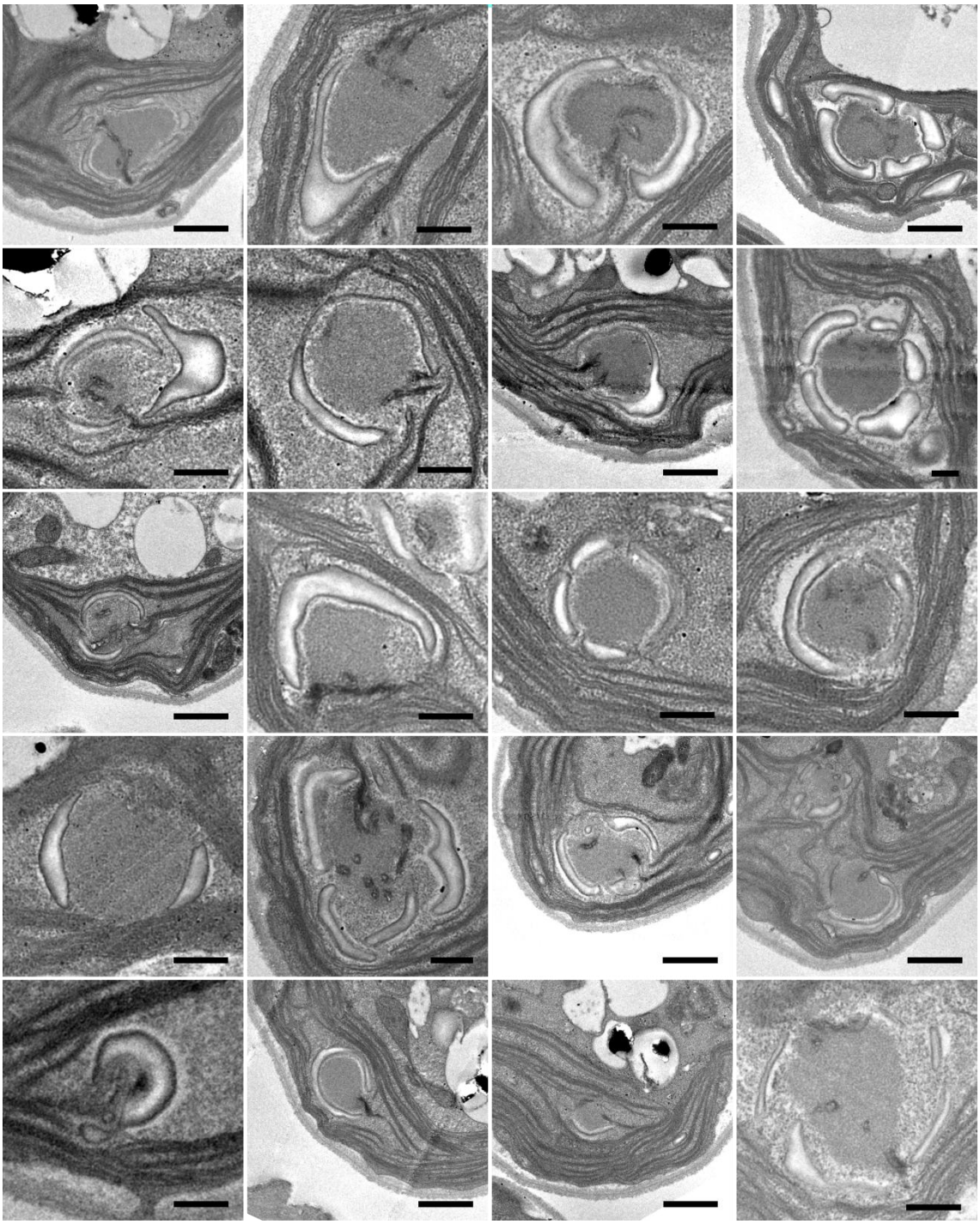
**Appendix 2 The lack of visible thylakoid tubules network in the pyrenoids of mutant H10 cells. Scale bar 500 nm.**



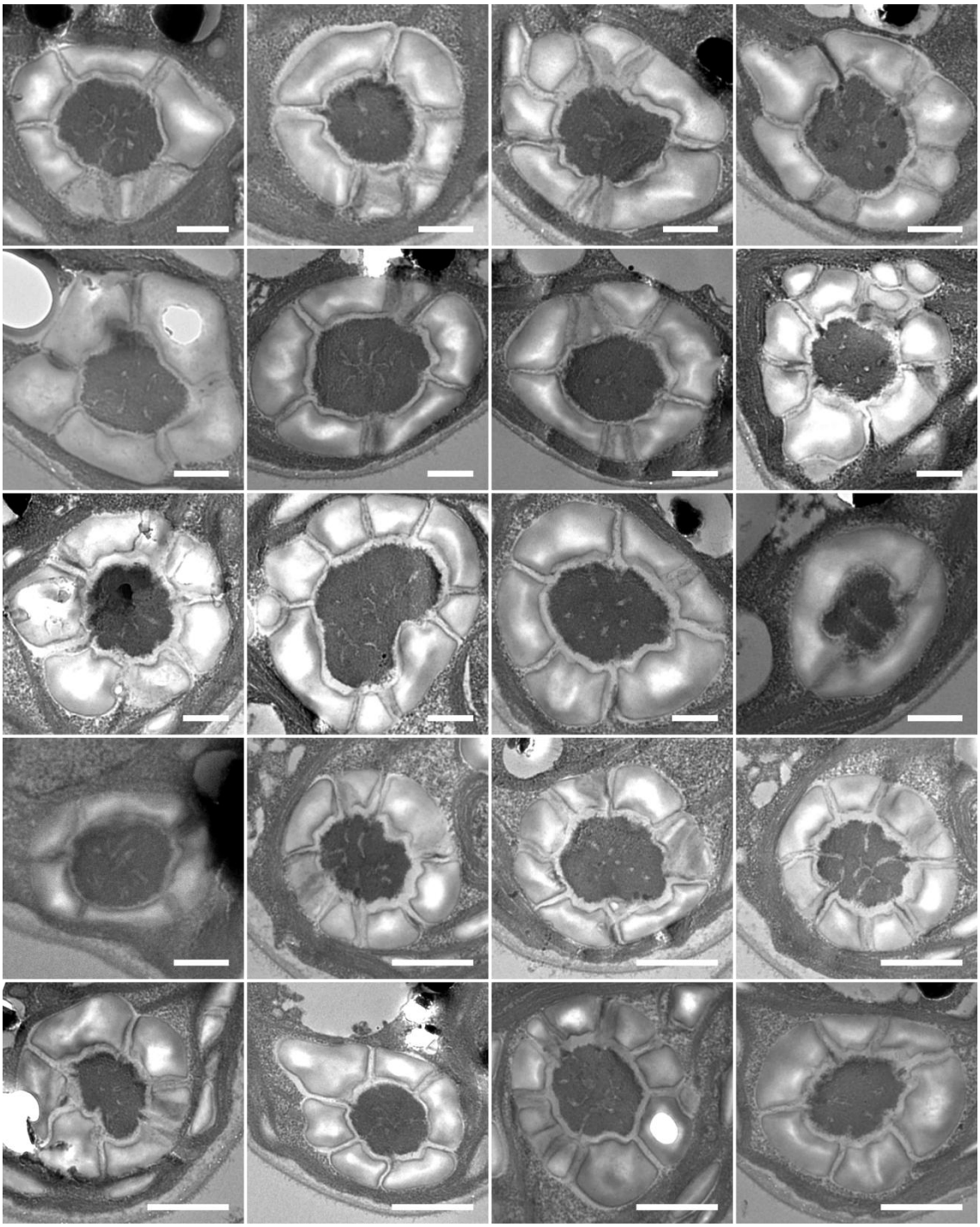
**Appendix 3 Pyrenoid morphology of wild type *Chlamydomonas reinhardtii* cells. Scale bar 500 nm.**



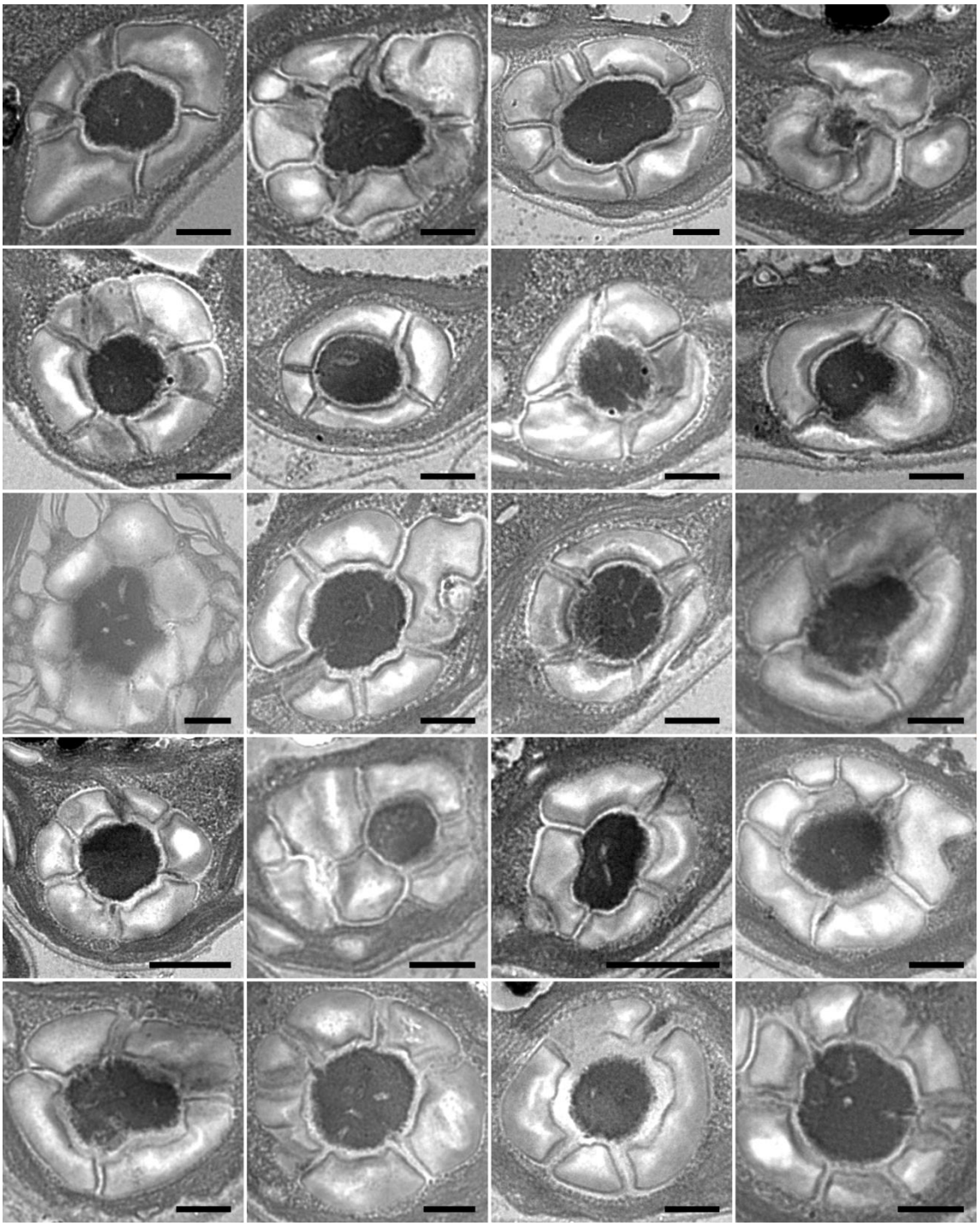
**Appendix 4 The lack of visible thylakoid tubules network in the pyrenoids of *saga* cells. Scale bar 500 nm.**



**Appendix 5 The presence of traversing thylakoid tubules in the pyrenoids of *saga* cells. Scale bar 500 nm.**



**Appendix 6 The restoration of traversing thylakoid tubule network in the pyrenoids of *saga::SAGA-Venus-3xFLAG* A7 cells. Scale bar 500 nm.**



**Appendix 7 The restoration of traversing thylakoid tubule network in the pyrenoids of *saga::SAGA-Venus-3xFLAG* B2 cells. Scale bar 500 nm.**

**Appendix 8 Proteomic report displaying the protein identification probability of all the proteins detected by LC-MS/MS of the GFP-Trap_A coimmunoprecipitation assay.**

**Three biological replicates of wild type (WT), *saga*, SVA7 (*saga::SAGA-Venus-3xFLAG A7*) and SVB2 (*saga::SAGA-Venus-3xFLAG B2*) were grown photoautotrophically under 5% $CO_2$ (high $CO_2$) for three days and then acclimated for 24 h under 0.04% $CO_2$ (low $CO_2$) for CCM induction. SAGA and green fluorescent protein were identified in all the replicates of the complemented lines but none was detected in the negative controls. 0, absent in all biological replicates; 1, present in one biological replicate; 2, present in two biological replicates; 3 present in all biological replicates.**

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre01.g000900.t1.2 | 51 kDa | 3 | 3 | 3 | 3 |
| Cre01.g001100.t1.2 | 70 kDa | 2 | 0 | 3 | 3 |
| Cre01.g002300.t1.2 | 18 kDa | 3 | 3 | 3 | 3 |
| Cre01.g002500.t1.2 | 27 kDa | 3 | 3 | 2 | 2 |
| Cre01.g004300.t1.2 | 65 kDa | 1 | 1 | 1 | 1 |
| Cre01.g004500.t1.2 | 53 kDa | 3 | 3 | 3 | 3 |
| Cre01.g005050.t1.1 | 97 kDa | 2 | 1 | 0 | 1 |
| Cre01.g005150.t1.1 | 49 kDa | 3 | 3 | 3 | 3 |
| Cre01.g007051.t1.2 | 15 kDa | 0 | 1 | 0 | 1 |
| Cre01.g007700.t1.1 | 62 kDa | 0 | 0 | 1 | 1 |
| Cre01.g010900.t1.2 | 40 kDa | 3 | 3 | 3 | 3 |
| Cre01.g011000.t1.2 | 24 kDa | 2 | 2 | 3 | 2 |
| Cre01.g012350.t1.2 | 154 kDa | 2 | 1 | 1 | 1 |
| Cre01.g013700.t1.2 (+1) | 28 kDa | 2 | 1 | 2 | 1 |
| Cre01.g014050.t1.2 | 46 kDa | 0 | 0 | 0 | 1 |
| Cre01.g015350.t1.1 | 42 kDa | 3 | 3 | 3 | 3 |
| Cre01.g015600.t1.1 | 282 kDa | 1 | 1 | 0 | 0 |
| Cre01.g016050.t1.2 | 278 kDa | 0 | 1 | 1 | 1 |
| Cre01.g016300.t1.1 | 36 kDa | 1 | 0 | 0 | 0 |
| Cre01.g016514.t1.1 | 61 kDa | 1 | 2 | 3 | 3 |
| Cre01.g020918.t1.1 | 117 kDa | 2 | 2 | 2 | 2 |
| Cre01.g021251.t1.1 | 56 kDa | 1 | 2 | 3 | 3 |
| Cre01.g023787.t1.1 | 46 kDa | 0 | 0 | 2 | 1 |
| Cre01.g025250.t1.1 | 42 kDa | 1 | 2 | 2 | 2 |
| Cre01.g026450.t1.2 | 35 kDa | 0 | 0 | 2 | 2 |
| Cre01.g026550.t1.1 | 61 kDa | 3 | 3 | 3 | 3 |
| Cre01.g027000.t1.2 | 20 kDa | 3 | 3 | 3 | 3 |
| Cre01.g027800.t1.2 | 53 kDa | 3 | 2 | 2 | 3 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre01.g029300.t1.2 | 30 kDa | 0 | 0 | 1 | 1 |
| Cre01.g029650.t1.1 | 100 kDa | 1 | 0 | 1 | 1 |
| Cre01.g032300.t1.2 | 24 kDa | 1 | 0 | 1 | 0 |
| Cre01.g032650.t1.2 (+1) | 44 kDa | 3 | 3 | 1 | 1 |
| Cre01.g033550.t1.2 | 27 kDa | 0 | 0 | 2 | 2 |
| Cre01.g034000.t1.2 | 97 kDa | 2 | 3 | 3 | 3 |
| Cre01.g034400.t1.1 | 39 kDa | 1 | 1 | 3 | 2 |
| Cre01.g034950.t1.2 | 82 kDa | 0 | 0 | 1 | 1 |
| Cre01.g039250.t1.1 (+1) | 32 kDa | 2 | 3 | 3 | 3 |
| Cre01.g039300.t1.2 | 70 kDa | 2 | 0 | 1 | 0 |
| Cre01.g040000.t1.2 | 16 kDa | 3 | 3 | 3 | 3 |
| Cre01.g042750.t1.2 | 89 kDa | 3 | 3 | 3 | 3 |
| Cre01.g042800.t1.2 | 45 kDa | 0 | 1 | 1 | 0 |
| Cre01.g044800.t1.2 | 91 kDa | 3 | 3 | 3 | 3 |
| Cre01.g045550.t1.2 | 43 kDa | 2 | 1 | 0 | 1 |
| Cre01.g047750.t1.2 | 21 kDa | 3 | 3 | 3 | 3 |
| Cre01.g048950.t1.2 | 55 kDa | 3 | 3 | 3 | 3 |
| Cre01.g050100.t1.2 | 52 kDa | 0 | 0 | 1 | 1 |
| Cre01.g050150.t1.1 | 44 kDa | 0 | 0 | 0 | 1 |
| Cre01.g050850.t1.2 (+1) | 96 kDa | 0 | 0 | 0 | 1 |
| Cre01.g050950.t1.2 | 56 kDa | 3 | 2 | 1 | 1 |
| Cre01.g052100.t1.2 | 16 kDa | 1 | 1 | 0 | 0 |
| Cre01.g053000.t1.2 | 78 kDa | 1 | 0 | 0 | 0 |
| Cre01.g053288.t1.2 (+1) | 46 kDa | 1 | 1 | 0 | 0 |
| Cre01.g054500.t1.2 | 112 kDa | 2 | 1 | 1 | 1 |
| Cre01.g054800.t1.2 | 135 kDa | 1 | 0 | 0 | 0 |
| Cre01.g055400.t1.2 | 40 kDa | 0 | 0 | 0 | 1 |
| Cre01.g055408.t1.1 | 77 kDa | 3 | 3 | 3 | 3 |
| Cre01.g055453.t1.1 | 53 kDa | 3 | 3 | 3 | 3 |
| Cre01.g057061.t1.1 | 95 kDa | 1 | 0 | 0 | 0 |
| Cre01.g061077.t1.1 | 29 kDa | 1 | 2 | 1 | 0 |
| Cre01.g065822.t1.1 | 139 kDa | 0 | 0 | 1 | 1 |
| Cre01.g066917.t1.1 | 28 kDa | 2 | 2 | 1 | 1 |
| Cre01.g069472.t1.1 | 67 kDa | 1 | 1 | 0 | 0 |
| Cre02.g075150.t1.2 (+1) | 181 kDa | 0 | 0 | 0 | 1 |
| Cre02.g075400.t1.1 (+1) | 24 kDa | 1 | 0 | 0 | 0 |
| Cre02.g075700.t1.2 | 23 kDa | 3 | 2 | 3 | 3 |
| Cre02.g076250.t1.1 | 85 kDa | 3 | 3 | 3 | 3 |
| Cre02.g076350.t1.2 | 56 kDa | 3 | 3 | 3 | 3 |
| Cre02.g076900.t1.1 | 115 kDa | 1 | 1 | 3 | 2 |
| Cre02.g077100.t1.2 | 54 kDa | 1 | 1 | 1 | 1 |
| Cre02.g077300.t1.2 | 33 kDa | 3 | 3 | 3 | 3 |
| Cre02.g077750.t1.2 | 82 kDa | 1 | 2 | 3 | 2 |
| Cre02.g077800.t1.2 | 86 kDa | 0 | 0 | 1 | 1 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre02.g079550.t1.2 | 91 kDa | 1 | 1 | 0 | 1 |
| Cre02.g079600.t1.2 | 49 kDa | 1 | 0 | 0 | 0 |
| Cre02.g079700.t1.2 | 41 kDa | 2 | 3 | 2 | 2 |
| Cre02.g080200.t1.2 | 78 kDa | 3 | 3 | 3 | 3 |
| Cre02.g080650.t1.2 | 93 kDa | 0 | 0 | 1 | 1 |
| Cre02.g080700.t1.2 | 72 kDa | 0 | 1 | 1 | 3 |
| Cre02.g081050.t1.2 | 56 kDa | 2 | 2 | 2 | 2 |
| Cre02.g082250.t1.1 | 36 kDa | 0 | 1 | 0 | 1 |
| Cre02.g083600.t1.2 (+1) | 45 kDa | 0 | 0 | 1 | 1 |
| Cre02.g083800.t1.2 (+1) | 41 kDa | 2 | 2 | 3 | 3 |
| Cre02.g083950.t1.1 | 33 kDa | 2 | 2 | 1 | 1 |
| Cre02.g085450.t1.2 | 41 kDa | 1 | 1 | 2 | 2 |
| Cre02.g087150.t1.2 | 58 kDa | 1 | 0 | 0 | 0 |
| Cre02.g087700.t1.2 | 36 kDa | 1 | 0 | 1 | 1 |
| Cre02.g087950.t1.2 | 104 kDa | 2 | 3 | 3 | 3 |
| Cre02.g088000.t1.2 | 34 kDa | 1 | 1 | 0 | 0 |
| Cre02.g088200.t1.2 | 58 kDa | 2 | 2 | 3 | 3 |
| Cre02.g088600.t1.2 | 65 kDa | 1 | 2 | 3 | 3 |
| Cre02.g088900.t1.2 | 32 kDa | 3 | 3 | 3 | 3 |
| Cre02.g089100.t1.2 | 56 kDa | 2 | 0 | 2 | 3 |
| Cre02.g090850.t1.1 | 110 kDa | 2 | 0 | 0 | 0 |
| Cre02.g091050.t1.2 | 43 kDa | 3 | 3 | 3 | 3 |
| Cre02.g091100.t1.2 | 27 kDa | 0 | 0 | 2 | 0 |
| Cre02.g091700.t1.2 | 72 kDa | 0 | 1 | 2 | 3 |
| Cre02.g092900.t1.2 | 59 kDa | 2 | 2 | 3 | 3 |
| Cre02.g093650.t1.2 | 15 kDa | 2 | 2 | 0 | 1 |
| Cre02.g095146.t1.1 | 31 kDa | 0 | 0 | 0 | 1 |
| Cre02.g095400.t1.2 (+1) | 39 kDa | 0 | 0 | 0 | 1 |
| Cre02.g096150.t1.2 | 24 kDa | 0 | 0 | 1 | 1 |
| Cre02.g097650.t1.2 | 46 kDa | 3 | 2 | 1 | 3 |
| Cre02.g097800.t1.1 | 147 kDa | 2 | 0 | 0 | 0 |
| Cre02.g097900.t1.2 | 47 kDa | 3 | 3 | 3 | 3 |
| Cre02.g098450.t1.2 | 50 kDa | 3 | 3 | 3 | 3 |
| Cre02.g099850.t1.1 | 52 kDa | 1 | 1 | 1 | 1 |
| Cre02.g101350.t1.2 | 24 kDa | 3 | 3 | 3 | 3 |
| Cre02.g101400.t1.2 | 190 kDa | 2 | 2 | 2 | 2 |
| Cre02.g101950.t1.2 | 85 kDa | 1 | 2 | 1 | 0 |
| Cre02.g102250.t1.2 | 21 kDa | 3 | 3 | 3 | 3 |
| Cre02.g103550.t1.2 | 16 kDa | 0 | 0 | 0 | 1 |
| Cre02.g104650.t1.1 (+1) | 126 kDa | 0 | 0 | 2 | 2 |
| Cre02.g104900.t1.2 | 44 kDa | 1 | 2 | 2 | 2 |
| Cre02.g106600.t1.2 | 17 kDa | 3 | 3 | 3 | 3 |
| Cre02.g106850.t1.2 | 34 kDa | 0 | 0 | 1 | 0 |
| Cre02.g107300.t1.2 | 38 kDa | 1 | 1 | 2 | 1 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre02.g108850.t1.2 | 19 kDa | 2 | 2 | 2 | 2 |
| Cre02.g111450.t1.2 | 27 kDa | 0 | 0 | 0 | 1 |
| Cre02.g113200.t1.1 | 55 kDa | 2 | 3 | 3 | 3 |
| Cre02.g114600.t1.2 | 22 kDa | 3 | 3 | 2 | 2 |
| Cre02.g115200.t1.2 | 16 kDa | 3 | 3 | 3 | 3 |
| Cre02.g116750.t1.1 (+1) | 62 kDa | 3 | 3 | 3 | 3 |
| Cre02.g118500.t1.1 | 108 kDa | 2 | 1 | 3 | 2 |
| Cre02.g120100.t1.2 | 21 kDa | 3 | 3 | 3 | 3 |
| Cre02.g141200.t1.2 | 34 kDa | 1 | 1 | 0 | 0 |
| Cre02.g141350.t1.2 | 41 kDa | 0 | 0 | 1 | 1 |
| Cre02.g141400.t1.2 | 68 kDa | 2 | 3 | 3 | 3 |
| Cre02.g141626.t1.1 | 126 kDa | 1 | 1 | 0 | 0 |
| Cre02.g142146.t1.1 | 41 kDa | 1 | 2 | 2 | 2 |
| Cre02.g142206.t1.1 | 32 kDa | 1 | 2 | 2 | 3 |
| Cre02.g142352.t1.1 | 92 kDa | 2 | 2 | 3 | 3 |
| Cre02.g142687.t1.1 (+2) | 21 kDa | 3 | 2 | 1 | 2 |
| Cre02.g143000.t1.2 | 46 kDa | 0 | 1 | 0 | 1 |
| Cre02.g143200.t1.1 | 122 kDa | 3 | 3 | 3 | 3 |
| Cre02.g143250.t1.2 | 37 kDa | 3 | 3 | 3 | 3 |
| Cre02.g143307.t1.1 | 6 kDa | 0 | 0 | 1 | 1 |
| Cre02.g145800.t1.2 | 38 kDa | 1 | 1 | 0 | 0 |
| Cre02.g147302.t1.1 | 45 kDa | 2 | 2 | 1 | 1 |
| Cre02.g147850.t1.2 | 70 kDa | 1 | 0 | 0 | 0 |
| Cre02.g147900.t1.1 (+3) | 65 kDa | 3 | 3 | 3 | 3 |
| Cre03.g143887.t1.1 | 65 kDa | 3 | 3 | 3 | 3 |
| Cre03.g144627.t1.1 | 51 kDa | 2 | 2 | 3 | 2 |
| Cre03.g144807.t1.1 | 61 kDa | 3 | 3 | 3 | 3 |
| Cre03.g144847.t1.1 | 270 kDa | 2 | 1 | 2 | 2 |
| Cre03.g145127.t1.1 | 504 kDa | 2 | 2 | 2 | 0 |
| Cre03.g145747.t1.1 | 44 kDa | 2 | 1 | 0 | 0 |
| Cre03.g145967.t1.1 | 46 kDa | 3 | 3 | 3 | 3 |
| Cre03.g146167.t1.1 | 25 kDa | 1 | 0 | 0 | 0 |
| Cre03.g146187.t1.1 | 41 kDa | 2 | 2 | 3 | 3 |
| Cre03.g146487.t1.1 | 122 kDa | 1 | 1 | 0 | 0 |
| Cre03.g146527.t1.1 | 69 kDa | 3 | 2 | 2 | 1 |
| Cre03.g149100.t1.2 | 55 kDa | 3 | 3 | 3 | 3 |
| Cre03.g154550.t1.1 | 36 kDa | 2 | 2 | 2 | 2 |
| Cre03.g156750.t1.2 | 59 kDa | 3 | 3 | 3 | 3 |
| Cre03.g156900.t1.2 | 29 kDa | 2 | 2 | 2 | 2 |
| Cre03.g158000.t1.2 | 49 kDa | 3 | 3 | 2 | 2 |
| Cre03.g158750.t1.1 | 62 kDa | 3 | 2 | 3 | 3 |
| Cre03.g158900.t1.2 | 50 kDa | 3 | 3 | 3 | 3 |
| Cre03.g160500.t1.2 | 67 kDa | 3 | 3 | 3 | 3 |
| Cre03.g165000.t1.2 | 70 kDa | 1 | 0 | 0 | 0 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre03.g166050.t1.2 | 52 kDa | 2 | 3 | 3 | 3 |
| Cre03.g168450.t1.2 | 58 kDa | 3 | 2 | 3 | 3 |
| Cre03.g169400.t1.2 | 37 kDa | 3 | 3 | 3 | 3 |
| Cre03.g169850.t1.2 | 94 kDa | 2 | 2 | 2 | 1 |
| Cre03.g171950.t1.1 | 131 kDa | 3 | 3 | 3 | 3 |
| Cre03.g172000.t1.2 | 33 kDa | 2 | 1 | 2 | 1 |
| Cre03.g172300.t1.2 | 37 kDa | 2 | 1 | 1 | 1 |
| Cre03.g172950.t1.2 | 58 kDa | 3 | 3 | 3 | 3 |
| Cre03.g175200.t1.2 | 87 kDa | 1 | 1 | 0 | 0 |
| Cre03.g175400.t1.2 (+1) | 72 kDa | 3 | 3 | 3 | 3 |
| Cre03.g176250.t1.2 | 41 kDa | 1 | 2 | 2 | 2 |
| Cre03.g177053.t1.1 | 137 kDa | 1 | 1 | 0 | 1 |
| Cre03.g178600.t1.1 | 77 kDa | 0 | 0 | 1 | 1 |
| Cre03.g179860.t1.1 | 97 kDa | 1 | 1 | 0 | 0 |
| Cre03.g180750.t1.2 | 87 kDa | 3 | 3 | 3 | 3 |
| Cre03.g180800.t1.1 | 133 kDa | 0 | 0 | 1 | 1 |
| Cre03.g181200.t1.2 | 64 kDa | 2 | 1 | 1 | 1 |
| Cre03.g181300.t1.2 | 54 kDa | 1 | 2 | 0 | 0 |
| Cre03.g181500.t1.2 | 66 kDa | 0 | 1 | 2 | 1 |
| Cre03.g182050.t1.2 | 72 kDa | 0 | 1 | 0 | 0 |
| Cre03.g182150.t1.2 | 28 kDa | 1 | 1 | 0 | 0 |
| Cre03.g182450.t1.2 | 36 kDa | 3 | 2 | 2 | 1 |
| Cre03.g182500.t1.2 | 77 kDa | 0 | 0 | 0 | 1 |
| Cre03.g182950.t1.2 | 79 kDa | 1 | 3 | 3 | 3 |
| Cre03.g185500.t1.1 | 168 kDa | 2 | 0 | 0 | 1 |
| Cre03.g185550.t1.2 | 41 kDa | 2 | 2 | 2 | 2 |
| Cre03.g188250.t1.2 | 55 kDa | 3 | 3 | 3 | 3 |
| Cre03.g189250.t1.2 | 24 kDa | 1 | 0 | 1 | 1 |
| Cre03.g189400.t1.2 | 53 kDa | 1 | 1 | 3 | 2 |
| Cre03.g189800.t1.2 | 45 kDa | 1 | 1 | 0 | 0 |
| Cre03.g190100.t1.1 | 83 kDa | 3 | 3 | 3 | 3 |
| Cre03.g190500.t1.2 | 101 kDa | 0 | 1 | 2 | 2 |
| Cre03.g190950.t1.2 (+1) | 50 kDa | 3 | 3 | 3 | 3 |
| Cre03.g191050.t1.2 | 25 kDa | 3 | 3 | 3 | 3 |
| Cre03.g192850.t1.2 | 118 kDa | 1 | 1 | 1 | 1 |
| Cre03.g193750.t1.1 | 45 kDa | 3 | 3 | 3 | 3 |
| Cre03.g193800.t1.1 | 56 kDa | 1 | 1 | 2 | 2 |
| Cre03.g193850.t1.2 | 34 kDa | 1 | 1 | 3 | 3 |
| Cre03.g194200.t1.2 | 40 kDa | 1 | 1 | 1 | 1 |
| Cre03.g194350.t1.2 | 72 kDa | 2 | 1 | 0 | 1 |
| Cre03.g194400.t1.2 | 37 kDa | 3 | 3 | 3 | 2 |
| Cre03.g194850.t1.2 | 37 kDa | 3 | 3 | 3 | 3 |
| Cre03.g196850.t1.2 | 57 kDa | 1 | 0 | 0 | 1 |
| Cre03.g197000.t1.2 | 37 kDa | 2 | 1 | 2 | 2 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre03.g197500.t1.2 (+1) | 47 kDa | 0 | 0 | 0 | 2 |
| Cre03.g199000.t1.2 | 81 kDa | 3 | 3 | 2 | 2 |
| Cre03.g199535.t1.1 | 27 kDa | 1 | 1 | 0 | 0 |
| Cre03.g199647.t1.1 | 46 kDa | 3 | 3 | 3 | 3 |
| Cre03.g199900.t1.2 | 23 kDa | 0 | 1 | 1 | 2 |
| Cre03.g200250.t1.2 | 37 kDa | 3 | 3 | 3 | 2 |
| Cre03.g204250.t1.2 | 53 kDa | 3 | 3 | 3 | 3 |
| Cre03.g204350.t1.2 | 31 kDa | 0 | 1 | 0 | 1 |
| Cre03.g205800.t1.2 | 61 kDa | 1 | 0 | 1 | 1 |
| Cre03.g206600.t1.2 | 64 kDa | 3 | 3 | 3 | 3 |
| Cre03.g207713.t1.1 | 83 kDa | 1 | 1 | 1 | 1 |
| Cre03.g207825.t1.1 (+2) | 53 kDa | 1 | 0 | 0 | 0 |
| Cre04.g211800.t1.2 | 15 kDa | 2 | 0 | 1 | 2 |
| Cre04.g214250.t1.1 | 113 kDa | 1 | 0 | 0 | 0 |
| Cre04.g214500.t1.1 | 49 kDa | 3 | 3 | 3 | 3 |
| Cre04.g214501.t1.1 | 81 kDa | 1 | 1 | 0 | 1 |
| Cre04.g214502.t1.1 | 39 kDa | 2 | 2 | 3 | 3 |
| Cre04.g215850.t1.2 | 61 kDa | 1 | 2 | 2 | 3 |
| Cre04.g216600.t1.2 | 44 kDa | 1 | 1 | 3 | 2 |
| Cre04.g216950.t1.2 | 39 kDa | 2 | 2 | 0 | 0 |
| Cre04.g217550.t1.1 | 130 kDa | 3 | 3 | 3 | 3 |
| Cre04.g217921.t1.1 | 105 kDa | 2 | 1 | 2 | 1 |
| Cre04.g218250.t1.2 | 20 kDa | 2 | 3 | 3 | 3 |
| Cre04.g220200.t1.2 (+2) | 109 kDa | 1 | 1 | 0 | 0 |
| Cre04.g222700.t1.2 | 115 kDa | 3 | 3 | 3 | 3 |
| Cre04.g223100.t1.2 | 42 kDa | 0 | 0 | 1 | 0 |
| Cre04.g223300.t1.2 | 38 kDa | 1 | 1 | 0 | 1 |
| Cre04.g223850.t1.2 | 46 kDa | 1 | 3 | 1 | 2 |
| Cre04.g229300.t1.1 | 45 kDa | 3 | 2 | 3 | 3 |
| Cre04.g229700.t1.2 | 50 kDa | 2 | 2 | 2 | 3 |
| Cre04.g231222.t2.1 | 99 kDa | 3 | 3 | 3 | 3 |
| Cre04.g232104.t1.1 | 27 kDa | 2 | 2 | 2 | 2 |
| Cre05.g232150.t1.2 | 49 kDa | 3 | 2 | 3 | 2 |
| Cre05.g232300.t1.1 | 176 kDa | 1 | 0 | 0 | 0 |
| Cre05.g232750.t1.2 | 58 kDa | 1 | 2 | 2 | 3 |
| Cre05.g233300.t1.2 | 32 kDa | 0 | 0 | 0 | 1 |
| Cre05.g233303.t1.1 (+1) | 189 kDa | 0 | 0 | 0 | 1 |
| Cre05.g233305.t1.1 (+1) | 41 kDa | 3 | 2 | 0 | 2 |
| Cre05.g233800.t1.2 | 76 kDa | 3 | 3 | 3 | 3 |
| Cre05.g234550.t1.2 (+1) | 41 kDa | 3 | 3 | 3 | 3 |
| Cre05.g234637.t1.1 | 15 kDa | 1 | 2 | 2 | 2 |
| Cre05.g234638.t1.1 | 66 kDa | 3 | 3 | 3 | 3 |
| Cre05.g234639.t1.1 | 57 kDa | 2 | 3 | 3 | 3 |
| Cre05.g234652.t1.1 | 48 kDa | 0 | 0 | 2 | 1 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre05.g234700.t1.1 | 248 kDa | 1 | 1 | 1 | 1 |
| Cre05.g237450.t1.2 | 32 kDa | 3 | 3 | 3 | 3 |
| Cre05.g238322.t1.1 | 45 kDa | 1 | 0 | 0 | 0 |
| Cre05.g238332.t1.1 | 21 kDa | 2 | 2 | 3 | 2 |
| Cre05.g238650.t1.1 | 58 kDa | 0 | 0 | 1 | 0 |
| Cre05.g238687.t1.1 | 83 kDa | 0 | 0 | 0 | 1 |
| Cre05.g241750.t1.2 | 45 kDa | 0 | 0 | 1 | 0 |
| Cre05.g241950.t1.2 | 29 kDa | 2 | 1 | 2 | 1 |
| Cre05.g242000.t1.2 | 83 kDa | 3 | 3 | 2 | 2 |
| Cre05.g242300.t1.2 | 62 kDa | 2 | 2 | 2 | 3 |
| Cre05.g245158.t1.1 | 77 kDa | 3 | 3 | 2 | 3 |
| Cre05.g245550.t1.1 | 238 kDa | 1 | 2 | 1 | 1 |
| Cre05.g245950.t1.1 | 70 kDa | 2 | 2 | 2 | 1 |
| Cre05.g247600.t1.2 | 16 kDa | 0 | 0 | 1 | 1 |
| Cre05.g248400.t1.2 (+1) | 28 kDa | 3 | 1 | 3 | 3 |
| Cre06.g250100.t1.2 | 72 kDa | 3 | 3 | 3 | 3 |
| Cre06.g250200.t1.2 | 43 kDa | 3 | 3 | 3 | 3 |
| Cre06.g250250.t1.2 | 42 kDa | 2 | 3 | 3 | 3 |
| Cre06.g250300.t1.2 | 481 kDa | 2 | 1 | 1 | 1 |
| Cre06.g250902.t1.1 | 268 kDa | 1 | 0 | 0 | 0 |
| Cre06.g251450.t1.1 | 77 kDa | 3 | 2 | 0 | 1 |
| Cre06.g251500.t1.2 | 35 kDa | 0 | 0 | 1 | 1 |
| Cre06.g254400.t1.1 | 63 kDa | 3 | 3 | 3 | 3 |
| Cre06.g257150.t1.2 | 10 kDa | 1 | 0 | 1 | 1 |
| Cre06.g257500.t1.2 | 28 kDa | 2 | 3 | 3 | 3 |
| Cre06.g257601.t1.2 | 26 kDa | 3 | 3 | 3 | 3 |
| Cre06.g257950.t1.2 | 48 kDa | 3 | 2 | 1 | 1 |
| Cre06.g258700.t1.1 | 103 kDa | 3 | 3 | 3 | 3 |
| Cre06.g258733.t1.1 (+1) | 74 kDa | 3 | 3 | 3 | 3 |
| Cre06.g258800.t1.1 | 132 kDa | 3 | 3 | 3 | 3 |
| Cre06.g259150.t1.2 | 48 kDa | 2 | 2 | 3 | 2 |
| Cre06.g259900.t1.2 | 39 kDa | 3 | 3 | 3 | 3 |
| Cre06.g260350.t1.2 | 60 kDa | 1 | 0 | 0 | 0 |
| Cre06.g261700.t1.2 | 30 kDa | 0 | 2 | 3 | 3 |
| Cre06.g261800.t1.2 | 41 kDa | 0 | 0 | 1 | 1 |
| Cre06.g263450.t1.2 (+1) | 51 kDa | 3 | 3 | 3 | 3 |
| Cre06.g264200.t1.2 | 32 kDa | 3 | 3 | 3 | 3 |
| Cre06.g264350.t1.2 | 25 kDa | 3 | 3 | 3 | 3 |
| Cre06.g264600.t1.2 (+29) | 11 kDa | 3 | 3 | 3 | 3 |
| Cre06.g264800.t1.2 | 17 kDa | 2 | 3 | 3 | 3 |
| Cre06.g265800.t1.2 | 22 kDa | 3 | 3 | 3 | 3 |
| Cre06.g267100.t1.2 | 73 kDa | 0 | 0 | 0 | 1 |
| Cre06.g268800.t1.2 | 152 kDa | 1 | 1 | 1 | 1 |
| Cre06.g269050.t1.2 | 91 kDa | 3 | 2 | 2 | 2 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre06.g269950.t1.2 | 90 kDa | 3 | 3 | 3 | 3 |
| Cre06.g271200.t1.2 | 63 kDa | 3 | 3 | 3 | 3 |
| Cre06.g272050.t1.2 | 67 kDa | 1 | 1 | 2 | 1 |
| Cre06.g272650.t1.2 | 26 kDa | 2 | 1 | 0 | 1 |
| Cre06.g272800.t1.2 | 24 kDa | 3 | 3 | 3 | 3 |
| Cre06.g272850.t1.2 | 26 kDa | 3 | 3 | 3 | 3 |
| Cre06.g272900.t1.2 | 56 kDa | 1 | 1 | 1 | 1 |
| Cre06.g272950.t1.1 | 18 kDa | 3 | 3 | 3 | 3 |
| Cre06.g273300.t1.2 | 79 kDa | 3 | 3 | 3 | 3 |
| Cre06.g273413.t1.1 | 94 kDa | 1 | 0 | 1 | 1 |
| Cre06.g273600.t1.2 | 20 kDa | 3 | 2 | 3 | 2 |
| Cre06.g273700.t1.2 | 43 kDa | 0 | 1 | 0 | 0 |
| Cre06.g275100.t1.2 | 54 kDa | 1 | 1 | 3 | 3 |
| Cre06.g275650.t1.2 | 55 kDa | 0 | 1 | 0 | 0 |
| Cre06.g278087.t1.1 | 154 kDa | 3 | 2 | 2 | 1 |
| Cre06.g278097.t1.1 | 39 kDa | 0 | 0 | 0 | 1 |
| Cre06.g278098.t1.1 | 82 kDa | 3 | 2 | 3 | 2 |
| Cre06.g278125.t1.1 | 139 kDa | 3 | 3 | 3 | 3 |
| Cre06.g278135.t1.1 | 26 kDa | 3 | 3 | 3 | 3 |
| Cre06.g278148.t1.1 | 39 kDa | 2 | 0 | 0 | 0 |
| Cre06.g278162.t1.1 | 47 kDa | 3 | 3 | 3 | 3 |
| Cre06.g278163.t1.1 | 49 kDa | 1 | 2 | 0 | 0 |
| Cre06.g278168.t1.1 | 34 kDa | 2 | 1 | 0 | 0 |
| Cre06.g278170.t1.1 | 133 kDa | 2 | 1 | 0 | 1 |
| Cre06.g278185.t1.1 | 53 kDa | 3 | 3 | 3 | 3 |
| Cre06.g278210.t1.1 | 65 kDa | 3 | 3 | 3 | 3 |
| Cre06.g278213.t1.1 | 28 kDa | 1 | 1 | 1 | 0 |
| Cre06.g278215.t1.1 | 43 kDa | 0 | 0 | 0 | 1 |
| Cre06.g278221.t1.1 | 100 kDa | 1 | 3 | 2 | 1 |
| Cre06.g278222.t1.1 | 35 kDa | 3 | 3 | 3 | 3 |
| Cre06.g278225.t1.1 | 100 kDa | 3 | 2 | 1 | 1 |
| Cre06.g278256.t1.1 | 37 kDa | 0 | 0 | 1 | 1 |
| Cre06.g278750.t1.2 | 31 kDa | 1 | 1 | 0 | 0 |
| Cre06.g279150.t1.2 | 60 kDa | 3 | 3 | 3 | 3 |
| Cre06.g280650.t1.2 | 29 kDa | 1 | 0 | 0 | 0 |
| Cre06.g280950.t1.2 | 78 kDa | 3 | 3 | 3 | 3 |
| Cre06.g282000.t1.1 | 138 kDa | 1 | 3 | 2 | 2 |
| Cre06.g282500.t1.2 | 16 kDa | 3 | 3 | 3 | 3 |
| Cre06.g282800.t1.2 | 46 kDa | 3 | 3 | 3 | 2 |
| Cre06.g283050.t1.2 | 24 kDa | 1 | 0 | 0 | 0 |
| Cre06.g283800.t1.2 | 101 kDa | 0 | 0 | 2 | 1 |
| Cre06.g283950.t1.2 | 27 kDa | 2 | 2 | 2 | 2 |
| Cre06.g284750.t1.2 | 63 kDa | 3 | 2 | 3 | 3 |
| Cre06.g284900.t1.2 (+1) | 24 kDa | 3 | 3 | 3 | 3 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre06.g286250.t1.2 | 32 kDa | 0 | 1 | 0 | 0 |
| Cre06.g288700.t1.2 | 119 kDa | 3 | 3 | 2 | 2 |
| Cre06.g289550.t1.2 | 15 kDa | 0 | 0 | 1 | 2 |
| Cre06.g290950.t1.2 | 22 kDa | 3 | 3 | 3 | 3 |
| Cre06.g291850.t1.2 (+2) | 92 kDa | 1 | 0 | 0 | 0 |
| Cre06.g292550.t1.2 | 35 kDa | 1 | 1 | 3 | 3 |
| Cre06.g293582.t1.1 | 52 kDa | 3 | 3 | 3 | 3 |
| Cre06.g293950.t1.2 | 53 kDa | 3 | 3 | 3 | 3 |
| Cre06.g294650.t1.2 | 46 kDa | 0 | 1 | 1 | 1 |
| Cre06.g294950.t1.1 | 39 kDa | 3 | 3 | 3 | 3 |
| Cre06.g295450.t1.2 | 45 kDa | 1 | 1 | 1 | 1 |
| Cre06.g296250.t1.1 | 68 kDa | 2 | 1 | 0 | 0 |
| Cre06.g296400.t1.2 | 45 kDa | 0 | 0 | 0 | 1 |
| Cre06.g298350.t1.2 | 47 kDa | 3 | 3 | 3 | 3 |
| Cre06.g298650.t1.2 | 53 kDa | 3 | 3 | 3 | 3 |
| Cre06.g299650.t1.2 | 31 kDa | 0 | 0 | 1 | 1 |
| Cre06.g300550.t1.2 | 106 kDa | 1 | 0 | 0 | 0 |
| Cre06.g300650.t1.2 | 87 kDa | 1 | 0 | 0 | 0 |
| Cre06.g300700.t1.1 | 29 kDa | 1 | 1 | 0 | 0 |
| Cre06.g300800.t1.2 | 17 kDa | 0 | 0 | 1 | 2 |
| Cre06.g306300.t1.2 | 45 kDa | 3 | 3 | 3 | 3 |
| Cre06.g306400.t1.2 | 46 kDa | 1 | 0 | 0 | 1 |
| Cre06.g306601.t1.1 | 67 kDa | 0 | 0 | 0 | 1 |
| Cre06.g306850.t1.2 | 56 kDa | 0 | 0 | 0 | 1 |
| Cre06.g307500.t1.1 | 48 kDa | 3 | 3 | 3 | 3 |
| Cre06.g308250.t1.2 | 29 kDa | 3 | 3 | 3 | 3 |
| Cre06.g308500.t1.2 | 50 kDa | 3 | 3 | 3 | 3 |
| Cre06.g309100.t1.2 | 60 kDa | 3 | 3 | 3 | 3 |
| Cre06.g310750.t1.2 | 99 kDa | 2 | 2 | 2 | 1 |
| Cre07.g313700.t1.1 | 61 kDa | 2 | 0 | 0 | 0 |
| Cre07.g314150.t1.2 | 63 kDa | 3 | 3 | 3 | 2 |
| Cre07.g314900.t1.1 | 80 kDa | 0 | 0 | 0 | 1 |
| Cre07.g318750.t1.2 | 39 kDa | 2 | 2 | 2 | 2 |
| Cre07.g319300.t1.1 | 166 kDa | 1 | 1 | 1 | 1 |
| Cre07.g321400.t1.1 | 199 kDa | 2 | 1 | 2 | 2 |
| Cre07.g321750.t1.2 | 41 kDa | 1 | 0 | 0 | 0 |
| Cre07.g322000.t1.2 | 50 kDa | 1 | 0 | 0 | 0 |
| Cre07.g322176.t1.1 | 114 kDa | 0 | 0 | 1 | 1 |
| Cre07.g324000.t1.2 | 188 kDa | 1 | 0 | 0 | 0 |
| Cre07.g324550.t1.1 | 46 kDa | 0 | 0 | 0 | 1 |
| Cre07.g325400.t1.2 | 43 kDa | 3 | 3 | 3 | 3 |
| Cre07.g325500.t1.1 | 154 kDa | 2 | 2 | 0 | 0 |
| Cre07.g325736.t1.1 | 56 kDa | 2 | 1 | 1 | 0 |
| Cre07.g325746.t1.1 | 8 kDa | 0 | 0 | 1 | 0 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre07.g325748.t1.1 | 149 kDa | 2 | 1 | 0 | 0 |
| Cre07.g325762.t1.1 | 95 kDa | 0 | 0 | 1 | 2 |
| Cre07.g326950.t1.1 | 25 kDa | 0 | 0 | 1 | 1 |
| Cre07.g328900.t1.2 | 54 kDa | 1 | 2 | 2 | 2 |
| Cre07.g329700.t1.2 | 49 kDa | 0 | 0 | 1 | 0 |
| Cre07.g330250.t1.2 | 14 kDa | 2 | 1 | 0 | 0 |
| Cre07.g330750.t1.1 | 99 kDa | 0 | 0 | 0 | 1 |
| Cre07.g331500.t1.1 | 47 kDa | 2 | 1 | 2 | 2 |
| Cre07.g331550.t1.2 | 44 kDa | 1 | 2 | 3 | 3 |
| Cre07.g331900.t1.2 | 17 kDa | 3 | 2 | 3 | 3 |
| Cre07.g332300.t1.2 | 124 kDa | 3 | 3 | 2 | 1 |
| Cre07.g332800.t1.2 | 149 kDa | 1 | 0 | 0 | 0 |
| Cre07.g335300.t1.2 | 49 kDa | 1 | 1 | 0 | 0 |
| Cre07.g335400.t1.1 | 70 kDa | 2 | 2 | 2 | 1 |
| Cre07.g336950.t1.1 | 98 kDa | 3 | 3 | 3 | 3 |
| Cre07.g337650.t1.2 | 44 kDa | 2 | 2 | 2 | 2 |
| Cre07.g338050.t1.2 | 40 kDa | 2 | 1 | 0 | 1 |
| Cre07.g338451.t1.1 | 42 kDa | 1 | 0 | 0 | 1 |
| Cre07.g339150.t1.1 | 62 kDa | 3 | 3 | 3 | 3 |
| Cre07.g340200.t1.1 | 36 kDa | 2 | 1 | 0 | 0 |
| Cre07.g340350.t1.1 | 63 kDa | 3 | 2 | 3 | 2 |
| Cre07.g342350.t1.2 | 89 kDa | 0 | 1 | 1 | 2 |
| Cre07.g342800.t1.1 | 41 kDa | 1 | 0 | 0 | 0 |
| Cre07.g343700.t1.2 | 48 kDa | 3 | 3 | 3 | 3 |
| Cre07.g344400.t1.2 | 54 kDa | 3 | 3 | 3 | 3 |
| Cre07.g344550.t1.2 | 54 kDa | 3 | 3 | 3 | 3 |
| Cre07.g344600.t1.2 | 64 kDa | 2 | 1 | 0 | 0 |
| Cre07.g344900.t1.2 (+1) | 44 kDa | 1 | 0 | 0 | 0 |
| Cre07.g344950.t1.2 | 23 kDa | 1 | 1 | 1 | 1 |
| Cre07.g349200.t1.2 | 16 kDa | 2 | 1 | 0 | 0 |
| Cre07.g350500.t1.2 | 108 kDa | 1 | 2 | 0 | 1 |
| Cre07.g352800.t1.2 | 144 kDa | 3 | 1 | 1 | 1 |
| Cre07.g353450.t1.2 | 73 kDa | 3 | 3 | 3 | 3 |
| Cre07.g356350.t1.1 | 79 kDa | 3 | 3 | 2 | 2 |
| Cre07.g356400.t1.2 | 21 kDa | 0 | 0 | 1 | 1 |
| Cre07.g356600.t1.2 | 85 kDa | 0 | 0 | 1 | 2 |
| Cre07.g357850.t1.2 | 14 kDa | 3 | 3 | 3 | 3 |
| Cre08.g358526.t1.1 | 62 kDa | 2 | 1 | 1 | 1 |
| Cre08.g358562.t1.1 | 23 kDa | 0 | 0 | 1 | 0 |
| Cre08.g358580.t1.1 | 128 kDa | 3 | 3 | 3 | 3 |
| Cre08.g358900.t1.2 | 26 kDa | 0 | 1 | 0 | 1 |
| Cre08.g359350.t1.2 | 56 kDa | 3 | 3 | 3 | 3 |
| Cre08.g359750.t1.2 (+1) | 22 kDa | 3 | 3 | 3 | 3 |
| Cre08.g360900.t1.2 | 17 kDa | 3 | 1 | 0 | 0 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre08.g362450.t1.2 | 78 kDa | 3 | 3 | 3 | 3 |
| Cre08.g362650.t1.1 | 157 kDa | 0 | 0 | 1 | 1 |
| Cre08.g364800.t1.2 | 151 kDa | 2 | 2 | 2 | 2 |
| Cre08.g365692.t1.1 | 54 kDa | 2 | 3 | 3 | 3 |
| Cre08.g367400.t1.1 (+1) | 28 kDa | 1 | 0 | 0 | 0 |
| Cre08.g368900.t1.2 (+1) | 101 kDa | 1 | 0 | 0 | 0 |
| Cre08.g370550.t1.1 | 63 kDa | 1 | 0 | 0 | 0 |
| Cre08.g372100.t1.2 | 71 kDa | 3 | 3 | 3 | 3 |
| Cre08.g372450.t1.2 | 22 kDa | 2 | 2 | 1 | 2 |
| Cre08.g372800.t1.1 | 132 kDa | 0 | 2 | 3 | 2 |
| Cre08.g373100.t1.1 | 63 kDa | 2 | 0 | 0 | 0 |
| Cre08.g374800.t1.2 | 63 kDa | 0 | 0 | 0 | 1 |
| Cre08.g375500.t1.2 | 78 kDa | 0 | 0 | 1 | 1 |
| Cre08.g376100.t1.2 | 34 kDa | 0 | 0 | 1 | 1 |
| Cre08.g384000.t1.1 | 179 kDa | 1 | 0 | 0 | 0 |
| Cre08.g384864.t1.2 | 153 kDa | 1 | 0 | 0 | 0 |
| Cre09.g386200.t1.1 | 72 kDa | 3 | 1 | 0 | 0 |
| Cre09.g386400.t1.2 | 114 kDa | 1 | 1 | 2 | 2 |
| Cre09.g386650.t1.2 (+1) | 34 kDa | 3 | 2 | 2 | 2 |
| Cre09.g386750.t1.2 | 81 kDa | 3 | 3 | 3 | 3 |
| Cre09.g386758.t1.1 | 74 kDa | 3 | 3 | 3 | 2 |
| Cre09.g387171.t1.1 (+1) | 35 kDa | 1 | 1 | 2 | 1 |
| Cre09.g387726.t1.1 | 47 kDa | 3 | 3 | 3 | 3 |
| Cre09.g387875.t1.1 | 22 kDa | 0 | 1 | 1 | 1 |
| Cre09.g388200.t1.1 | 26 kDa | 3 | 3 | 3 | 3 |
| Cre09.g388467.t1.1 | 50 kDa | 1 | 2 | 3 | 2 |
| Cre09.g388500.t1.2 | 89 kDa | 0 | 1 | 2 | 2 |
| Cre09.g388800.t1.2 | 49 kDa | 0 | 1 | 1 | 1 |
| Cre09.g389430.t1.1 | 56 kDa | 1 | 1 | 0 | 1 |
| Cre09.g389689.t1.1 | 40 kDa | 1 | 1 | 1 | 1 |
| Cre09.g390615.t1.1 | 52 kDa | 0 | 2 | 0 | 0 |
| Cre09.g390763.t1.1 | 24 kDa | 0 | 1 | 0 | 0 |
| Cre09.g390986.t1.1 | 70 kDa | 2 | 1 | 0 | 0 |
| Cre09.g391023.t1.1 | 36 kDa | 1 | 3 | 3 | 3 |
| Cre09.g391097.t1.1 (+1) | 20 kDa | 3 | 3 | 3 | 3 |
| Cre09.g391652.t1.1 | 111 kDa | 0 | 0 | 0 | 1 |
| Cre09.g392350.t1.2 (+1) | 16 kDa | 3 | 3 | 3 | 3 |
| Cre09.g392500.t1.2 | 37 kDa | 0 | 0 | 2 | 1 |
| Cre09.g392867.t1.1 | 433 kDa | 3 | 3 | 2 | 3 |
| Cre09.g393200.t1.2 | 71 kDa | 0 | 2 | 1 | 2 |
| Cre09.g393543.t1.1 | 70 kDa | 3 | 3 | 3 | 3 |
| Cre09.g393728.t1.1 | 75 kDa | 0 | 0 | 1 | 1 |
| Cre09.g394102.t1.1 | 62 kDa | 1 | 1 | 0 | 0 |
| Cre09.g394200.t1.1 | 156 kDa | 3 | 3 | 3 | 3 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre09.g394350.t1.2 | 176 kDa | 2 | 0 | 0 | 0 |
| Cre09.g394436.t1.1 | 80 kDa | 2 | 2 | 2 | 1 |
| Cre09.g394450.t1.2 | 26 kDa | 0 | 0 | 3 | 3 |
| Cre09.g394750.t1.2 | 48 kDa | 2 | 2 | 3 | 2 |
| Cre09.g396100.t1.2 | 18 kDa | 0 | 0 | 0 | 1 |
| Cre09.g396213.t1.1 | 31 kDa | 2 | 3 | 3 | 3 |
| Cre09.g396252.t1.1 | 55 kDa | 0 | 0 | 1 | 0 |
| Cre09.g396438.t1.1 | 209 kDa | 3 | 3 | 0 | 0 |
| Cre09.g396650.t1.2 | 85 kDa | 3 | 3 | 3 | 3 |
| Cre09.g397200.t1.2 | 60 kDa | 3 | 3 | 3 | 3 |
| Cre09.g397697.t1.1 | 45 kDa | 3 | 3 | 3 | 3 |
| Cre09.g400650.t1.2 | 28 kDa | 3 | 3 | 3 | 3 |
| Cre09.g401515.t1.1 | 41 kDa | 2 | 3 | 3 | 1 |
| Cre09.g401886.t1.1 | 39 kDa | 0 | 1 | 0 | 0 |
| Cre09.g402775.t1.1 | 85 kDa | 1 | 1 | 0 | 0 |
| Cre09.g402950.t1.1 | 47 kDa | 1 | 0 | 0 | 0 |
| Cre09.g403800.t1.2 | 520 kDa | 2 | 1 | 1 | 0 |
| Cre09.g405050.t1.2 | 30 kDa | 2 | 2 | 2 | 2 |
| Cre09.g406050.t1.2 | 65 kDa | 3 | 3 | 3 | 3 |
| Cre09.g406600.t1.1 | 38 kDa | 1 | 1 | 1 | 2 |
| Cre09.g406983.t1.1 | 32 kDa | 0 | 0 | 0 | 1 |
| Cre09.g407700.t1.2 | 54 kDa | 1 | 2 | 1 | 1 |
| Cre09.g408200.t1.2 | 22 kDa | 1 | 1 | 3 | 2 |
| Cre09.g409350.t1.2 | 70 kDa | 3 | 3 | 3 | 2 |
| Cre09.g409901.t1.1 | 33 kDa | 0 | 1 | 1 | 1 |
| Cre09.g410600.t1.1 | 39 kDa | 1 | 0 | 0 | 1 |
| Cre09.g410650.t1.2 | 47 kDa | 0 | 0 | 1 | 1 |
| Cre09.g410700.t1.2 | 45 kDa | 3 | 3 | 3 | 3 |
| Cre09.g411100.t1.2 (+1) | 19 kDa | 2 | 3 | 3 | 3 |
| Cre09.g411900.t1.1 | 53 kDa | 3 | 3 | 3 | 3 |
| Cre09.g412100.t1.2 | 24 kDa | 3 | 3 | 3 | 3 |
| Cre09.g413750.t1.1 | 50 kDa | 1 | 0 | 0 | 0 |
| Cre09.g414050.t1.2 | 50 kDa | 0 | 0 | 1 | 0 |
| Cre09.g414200.t1.2 | 37 kDa | 0 | 2 | 3 | 3 |
| Cre09.g415550.t1.2 | 48 kDa | 2 | 2 | 1 | 1 |
| Cre09.g415700.t1.2 | 33 kDa | 0 | 0 | 1 | 1 |
| Cre09.g415800.t1.2 | 79 kDa | 3 | 3 | 2 | 1 |
| Cre09.g415950.t1.2 | 22 kDa | 3 | 2 | 3 | 3 |
| Cre09.g416050.t1.2 | 49 kDa | 3 | 3 | 3 | 3 |
| Cre09.g416150.t1.2 | 22 kDa | 2 | 2 | 1 | 2 |
| Cre09.g416750.t1.2 | 57 kDa | 2 | 1 | 3 | 3 |
| Cre09.g417150.t1.2 | 56 kDa | 3 | 3 | 3 | 3 |
| Cre09.g417200.t1.2 | 43 kDa | 2 | 2 | 2 | 2 |
| Cre10.g417700.t1.2 | 44 kDa | 3 | 3 | 3 | 3 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre10.g420750.t1.2 | 12 kDa | 3 | 3 | 3 | 3 |
| Cre10.g421600.t1.2 | 70 kDa | 3 | 3 | 1 | 1 |
| Cre10.g422000.t1.2 | 64 kDa | 2 | 2 | 1 | 2 |
| Cre10.g423250.t1.2 | 36 kDa | 2 | 0 | 0 | 1 |
| Cre10.g423500.t1.2 | 29 kDa | 0 | 0 | 1 | 1 |
| Cre10.g424450.t1.2 | 37 kDa | 1 | 0 | 0 | 0 |
| Cre10.g425900.t1.2 | 28 kDa | 2 | 1 | 1 | 1 |
| Cre10.g427700.t1.2 | 67 kDa | 2 | 2 | 3 | 3 |
| Cre10.g428750.t1.1 | 224 kDa | 2 | 2 | 2 | 2 |
| Cre10.g429100.t1.2 | 61 kDa | 3 | 2 | 3 | 3 |
| Cre10.g430150.t1.2 | 50 kDa | 1 | 0 | 0 | 0 |
| Cre10.g430501.t1.1 | 73 kDa | 2 | 2 | 1 | 2 |
| Cre10.g432800.t1.2 | 31 kDa | 3 | 3 | 3 | 3 |
| Cre10.g433000.t1.1 | 118 kDa | 2 | 1 | 0 | 0 |
| Cre10.g433600.t1.2 | 66 kDa | 3 | 2 | 3 | 2 |
| Cre10.g434750.t1.2 | 60 kDa | 3 | 3 | 3 | 3 |
| Cre10.g435300.t1.2 (+1) | 57 kDa | 1 | 1 | 1 | 1 |
| Cre10.g435800.t1.2 | 45 kDa | 3 | 2 | 2 | 2 |
| Cre10.g436550.t1.2 | 32 kDa | 3 | 0 | 3 | 2 |
| Cre10.g439100.t1.2 | 59 kDa | 3 | 3 | 2 | 2 |
| Cre10.g439150.t1.2 | 48 kDa | 0 | 0 | 1 | 1 |
| Cre10.g439850.t1.2 | 73 kDa | 2 | 1 | 0 | 0 |
| Cre10.g440050.t1.2 | 48 kDa | 1 | 0 | 1 | 0 |
| Cre10.g441200.t1.2 | 106 kDa | 1 | 0 | 0 | 0 |
| Cre10.g441400.t1.2 | 55 kDa | 3 | 3 | 3 | 3 |
| Cre10.g442700.t1.2 | 52 kDa | 2 | 2 | 3 | 3 |
| Cre10.g443250.t1.2 | 61 kDa | 3 | 3 | 3 | 3 |
| Cre10.g443350.t1.2 | 125 kDa | 1 | 0 | 0 | 0 |
| Cre10.g444700.t1.1 | 98 kDa | 3 | 3 | 3 | 3 |
| Cre10.g445600.t1.1 | 151 kDa | 0 | 0 | 0 | 1 |
| Cre10.g447950.t1.2 | 22 kDa | 0 | 0 | 1 | 0 |
| Cre10.g448051.t1.1 | 19 kDa | 1 | 0 | 0 | 0 |
| Cre10.g451752.t1.1 | 58 kDa | 0 | 0 | 2 | 1 |
| Cre10.g451900.t1.1 | 54 kDa | 1 | 2 | 1 | 1 |
| Cre10.g451950.t1.2 | 58 kDa | 1 | 2 | 3 | 2 |
| Cre10.g452050.t1.2 | 29 kDa | 2 | 1 | 0 | 0 |
| Cre10.g452250.t1.1 | 68 kDa | 3 | 3 | 0 | 0 |
| Cre10.g452450.t1.2 | 115 kDa | 3 | 2 | 0 | 0 |
| Cre10.g452800.t1.2 | 48 kDa | 3 | 3 | 3 | 3 |
| Cre10.g456050.t1.2 (+2) | 21 kDa | 0 | 1 | 0 | 0 |
| Cre10.g456200.t1.2 | 15 kDa | 0 | 1 | 2 | 2 |
| Cre10.g457700.t1.2 | 40 kDa | 1 | 0 | 1 | 1 |
| Cre10.g458350.t1.1 (+2) | 932 kDa | 0 | 0 | 0 | 1 |
| Cre10.g459200.t1.2 | 117 kDa | 1 | 0 | 0 | 0 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre10.g459250.t1.2 | 12 kDa | 1 | 0 | 1 | 0 |
| Cre10.g461050.t1.2 | 68 kDa | 3 | 3 | 3 | 3 |
| Cre10.g461250.t1.1 | 79 kDa | 1 | 2 | 2 | 2 |
| Cre10.g465550.t1.1 | 110 kDa | 0 | 0 | 1 | 0 |
| Cre11.g467350.t1.2 | 73 kDa | 3 | 3 | 3 | 3 |
| Cre11.g467531.t1.1 | 44 kDa | 1 | 0 | 0 | 1 |
| Cre11.g467535.t1.1 | 44 kDa | 2 | 1 | 1 | 1 |
| Cre11.g467547.t1.1 | 83 kDa | 3 | 3 | 3 | 3 |
| Cre11.g467550.t1.2 | 42 kDa | 1 | 1 | 1 | 1 |
| Cre11.g467556.t1.1 | 145 kDa | 2 | 2 | 2 | 0 |
| Cre11.g467569.t1.1 | 24 kDa | 2 | 2 | 1 | 1 |
| Cre11.g467578.t1.1 | 14 kDa | 3 | 3 | 3 | 3 |
| Cre11.g467617.t1.1 | 30 kDa | 1 | 0 | 0 | 0 |
| Cre11.g467699.t1.1 | 68 kDa | 2 | 2 | 2 | 2 |
| Cre11.g467700.t1.1 | 43 kDa | 1 | 1 | 1 | 2 |
| **Cre11.g467712.t1.1** | **168 kDa** | **0** | **0** | **3** | **3** |
| Cre11.g467723.t1.1 (+1) | 49 kDa | 3 | 3 | 3 | 3 |
| Cre11.g467770.t1.1 | 49 kDa | 3 | 3 | 3 | 3 |
| Cre11.g467778.t1.1 | 29 kDa | 0 | 0 | 1 | 0 |
| Cre11.g468300.t1.2 | 22 kDa | 1 | 0 | 1 | 1 |
| Cre11.g468359.t1.1 | 75 kDa | 0 | 1 | 3 | 2 |
| Cre11.g476050.t1.2 | 339 kDa | 1 | 1 | 0 | 0 |
| Cre11.g476550.t1.2 | 37 kDa | 0 | 2 | 1 | 2 |
| Cre11.g476750.t1.2 | 38 kDa | 1 | 2 | 0 | 0 |
| Cre11.g477300.t1.1 | 65 kDa | 3 | 3 | 3 | 3 |
| Cre11.g477625.t1.1 | 190 kDa | 1 | 0 | 0 | 0 |
| Cre11.g477850.t1.1 | 111 kDa | 0 | 0 | 1 | 1 |
| Cre11.g477950.t1.1 | 94 kDa | 1 | 1 | 1 | 0 |
| Cre11.g478800.t1.2 | 75 kDa | 2 | 0 | 0 | 0 |
| Cre11.g479250.t1.2 | 54 kDa | 1 | 0 | 0 | 1 |
| Cre11.g479500.t1.2 | 26 kDa | 3 | 3 | 3 | 3 |
| Cre11.g480150.t1.2 | 16 kDa | 3 | 3 | 3 | 3 |
| Cre11.g480800.t1.2 | 36 kDa | 0 | 0 | 0 | 1 |
| Cre11.g481000.t1.2 | 41 kDa | 0 | 0 | 1 | 1 |
| Cre11.g481500.t1.2 | 50 kDa | 3 | 1 | 1 | 1 |
| Cre12.g483950.t1.2 | 35 kDa | 1 | 0 | 0 | 1 |
| Cre12.g484000.t1.2 | 49 kDa | 3 | 3 | 2 | 2 |
| Cre12.g484050.t1.2 | 11 kDa | 3 | 3 | 3 | 3 |
| Cre12.g485150.t1.2 | 40 kDa | 3 | 3 | 3 | 3 |
| Cre12.g485800.t1.2 | 78 kDa | 2 | 1 | 1 | 1 |
| Cre12.g487450.t1.1 | 33 kDa | 0 | 0 | 0 | 1 |
| Cre12.g489153.t1.1 | 19 kDa | 3 | 3 | 3 | 3 |
| Cre12.g489750.t1.2 | 68 kDa | 0 | 2 | 2 | 2 |
| Cre12.g490000.t1.2 | 37 kDa | 2 | 3 | 3 | 3 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre12.g490350.t1.1 | 75 kDa | 3 | 3 | 3 | 3 |
| Cre12.g493250.t1.2 | 81 kDa | 3 | 3 | 3 | 3 |
| Cre12.g493550.t1.2 | 229 kDa | 1 | 0 | 0 | 1 |
| Cre12.g493950.t1.2 | 19 kDa | 1 | 1 | 3 | 3 |
| Cre12.g494050.t1.2 | 22 kDa | 3 | 3 | 3 | 3 |
| Cre12.g494750.t1.2 | 18 kDa | 3 | 3 | 3 | 3 |
| Cre12.g494850.t1.2 | 29 kDa | 1 | 0 | 0 | 0 |
| Cre12.g495850.t1.2 | 28 kDa | 3 | 3 | 3 | 3 |
| Cre12.g495951.t1.1 (+1) | 21 kDa | 2 | 3 | 3 | 3 |
| Cre12.g497300.t1.2 (+1) | 39 kDa | 2 | 1 | 0 | 0 |
| Cre12.g497350.t1.1 | 190 kDa | 1 | 1 | 2 | 2 |
| Cre12.g498100.t1.2 | 49 kDa | 2 | 2 | 1 | 2 |
| Cre12.g498250.t1.2 | 16 kDa | 3 | 3 | 3 | 3 |
| Cre12.g498500.t1.2 | 48 kDa | 2 | 1 | 1 | 2 |
| Cre12.g498550.t1.1 | 35 kDa | 1 | 0 | 0 | 0 |
| Cre12.g498900.t1.2 | 22 kDa | 3 | 3 | 3 | 3 |
| Cre12.g500150.t1.1 | 58 kDa | 0 | 0 | 1 | 0 |
| Cre12.g500950.t1.2 | 38 kDa | 3 | 3 | 3 | 3 |
| Cre12.g502800.t1.2 | 61 kDa | 0 | 0 | 1 | 1 |
| Cre12.g503300.t1.2 | 68 kDa | 3 | 2 | 0 | 0 |
| Cre12.g504200.t1.2 | 16 kDa | 0 | 0 | 0 | 1 |
| Cre12.g505850.t1.2 | 26 kDa | 1 | 2 | 2 | 2 |
| Cre12.g507150.t1.2 | 32 kDa | 0 | 0 | 0 | 1 |
| Cre12.g507400.t1.2 | 77 kDa | 0 | 1 | 0 | 0 |
| Cre12.g507650.t1.2 (+1) | 45 kDa | 0 | 0 | 1 | 0 |
| Cre12.g508050.t1.2 | 183 kDa | 0 | 0 | 2 | 1 |
| Cre12.g508150.t1.2 | 124 kDa | 1 | 2 | 3 | 2 |
| Cre12.g508200.t1.2 | 47 kDa | 0 | 0 | 1 | 0 |
| Cre12.g508250.t1.1 | 273 kDa | 2 | 0 | 0 | 1 |
| Cre12.g508750.t1.2 | 27 kDa | 0 | 1 | 0 | 0 |
| Cre12.g508900.t1.2 | 43 kDa | 2 | 3 | 3 | 3 |
| Cre12.g509000.t1.1 | 40 kDa | 1 | 0 | 0 | 1 |
| Cre12.g509650.t1.2 | 63 kDa | 3 | 3 | 0 | 0 |
| Cre12.g509750.t1.2 | 50 kDa | 1 | 1 | 0 | 0 |
| Cre12.g510050.t1.2 | 47 kDa | 1 | 0 | 0 | 0 |
| Cre12.g510650.t1.2 | 44 kDa | 2 | 2 | 3 | 3 |
| Cre12.g510800.t1.2 | 46 kDa | 2 | 0 | 0 | 0 |
| Cre12.g511050.t1.2 | 39 kDa | 0 | 0 | 0 | 2 |
| Cre12.g511850.t1.2 | 55 kDa | 0 | 1 | 1 | 2 |
| Cre12.g511900.t1.2 | 28 kDa | 3 | 3 | 1 | 1 |
| Cre12.g512000.t1.1 | 120 kDa | 2 | 1 | 1 | 1 |
| Cre12.g512300.t1.2 | 118 kDa | 1 | 2 | 2 | 2 |
| Cre12.g512600.t1.2 | 21 kDa | 3 | 3 | 3 | 3 |
| Cre12.g513150.t1.2 | 54 kDa | 0 | 0 | 0 | 1 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre12.g513200.t1.2 | 52 kDa | 2 | 2 | 3 | 2 |
| Cre12.g514050.t1.2 | 175 kDa | 3 | 2 | 3 | 3 |
| Cre12.g514200.t1.2 | 64 kDa | 3 | 2 | 2 | 2 |
| Cre12.g514500.t1.2 | 18 kDa | 2 | 3 | 3 | 3 |
| Cre12.g514750.t1.2 | 51 kDa | 3 | 3 | 3 | 3 |
| Cre12.g514850.t1.2 | 89 kDa | 2 | 2 | 2 | 2 |
| Cre12.g515550.t1.2 | 77 kDa | 1 | 0 | 0 | 0 |
| Cre12.g515650.t1.2 | 24 kDa | 3 | 2 | 2 | 2 |
| Cre12.g515850.t1.2 | 30 kDa | 1 | 1 | 0 | 0 |
| Cre12.g516200.t1.2 | 94 kDa | 3 | 3 | 3 | 3 |
| Cre12.g517150.t1.1 | 46 kDa | 3 | 3 | 3 | 3 |
| Cre12.g517900.t1.1 | 114 kDa | 3 | 3 | 1 | 1 |
| Cre12.g519100.t1.2 | 54 kDa | 3 | 3 | 2 | 2 |
| Cre12.g519180.t1.1 (+1) | 118 kDa | 3 | 3 | 2 | 1 |
| Cre12.g519900.t1.2 | 86 kDa | 0 | 1 | 1 | 1 |
| Cre12.g520500.t1.1 | 35 kDa | 3 | 3 | 3 | 3 |
| Cre12.g522450.t1.2 | 138 kDa | 0 | 1 | 0 | 1 |
| Cre12.g522550.t1.2 | 82 kDa | 0 | 0 | 2 | 2 |
| Cre12.g522600.t1.2 | 12 kDa | 2 | 3 | 3 | 3 |
| Cre12.g523300.t1.2 | 66 kDa | 1 | 1 | 0 | 0 |
| Cre12.g523850.t1.2 | 55 kDa | 2 | 1 | 0 | 2 |
| Cre12.g525200.t1.2 | 56 kDa | 3 | 3 | 3 | 3 |
| Cre12.g526800.t1.2 | 61 kDa | 2 | 3 | 3 | 3 |
| Cre12.g526850.t1.2 | 75 kDa | 1 | 0 | 0 | 0 |
| Cre12.g528000.t1.2 | 44 kDa | 2 | 1 | 1 | 1 |
| Cre12.g528750.t1.2 | 18 kDa | 2 | 2 | 3 | 3 |
| Cre12.g529400.t1.2 | 10 kDa | 2 | 1 | 2 | 2 |
| Cre12.g529500.t1.2 | 96 kDa | 0 | 0 | 0 | 1 |
| Cre12.g529600.t1.2 | 204 kDa | 0 | 0 | 1 | 1 |
| Cre12.g529651.t1.1 | 22 kDa | 3 | 3 | 3 | 3 |
| Cre12.g529950.t1.2 | 70 kDa | 1 | 1 | 1 | 2 |
| Cre12.g530650.t1.1 (+1) | 41 kDa | 3 | 3 | 3 | 3 |
| Cre12.g531550.t1.2 | 31 kDa | 0 | 1 | 2 | 1 |
| Cre12.g531900.t1.2 | 76 kDa | 1 | 0 | 0 | 1 |
| Cre12.g532550.t1.1 | 19 kDa | 3 | 3 | 3 | 3 |
| Cre12.g533050.t1.2 | 73 kDa | 0 | 0 | 1 | 1 |
| Cre12.g533550.t1.1 | 55 kDa | 3 | 3 | 3 | 3 |
| Cre12.g534800.t1.1 | 111 kDa | 1 | 2 | 2 | 2 |
| Cre12.g535851.t1.1 | 31 kDa | 3 | 3 | 3 | 3 |
| Cre12.g537100.t1.2 | 142 kDa | 3 | 3 | 3 | 3 |
| Cre12.g537200.t1.2 | 117 kDa | 3 | 3 | 3 | 3 |
| Cre12.g537581.t1.1 | 56 kDa | 3 | 3 | 3 | 3 |
| Cre12.g537800.t1.2 | 28 kDa | 3 | 3 | 3 | 3 |
| Cre12.g538100.t1.2 | 39 kDa | 0 | 0 | 1 | 1 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre12.g538500.t1.1 | 74 kDa | 0 | 0 | 0 | 1 |
| Cre12.g539100.t1.1 | 68 kDa | 2 | 2 | 2 | 2 |
| Cre12.g540200.t1.2 | 113 kDa | 0 | 0 | 1 | 1 |
| Cre12.g542250.t1.1 (+1) | 50 kDa | 3 | 3 | 3 | 3 |
| Cre12.g543350.t1.1 (+2) | 41 kDa | 2 | 1 | 3 | 2 |
| Cre12.g544114.t1.1 (+1) | 26 kDa | 1 | 1 | 2 | 2 |
| Cre12.g544450.t1.2 | 81 kDa | 0 | 0 | 3 | 2 |
| Cre12.g546050.t1.2 | 50 kDa | 0 | 1 | 0 | 0 |
| Cre12.g548000.t1.2 | 256 kDa | 1 | 0 | 0 | 0 |
| Cre12.g549000.t1.2 | 53 kDa | 0 | 1 | 1 | 1 |
| Cre12.g549750.t1.2 | 38 kDa | 0 | 0 | 1 | 2 |
| Cre12.g550850.t1.2 | 26 kDa | 3 | 3 | 3 | 3 |
| Cre12.g552200.t1.2 | 114 kDa | 3 | 3 | 3 | 3 |
| Cre12.g554800.t1.2 | 42 kDa | 1 | 1 | 3 | 3 |
| Cre12.g554953.t1.1 | 29 kDa | 2 | 0 | 0 | 0 |
| Cre12.g556050.t1.2 | 22 kDa | 3 | 3 | 3 | 3 |
| Cre12.g556600.t1.2 | 53 kDa | 0 | 1 | 0 | 1 |
| Cre12.g557600.t1.1 | 26 kDa | 2 | 2 | 2 | 2 |
| Cre12.g558100.t1.2 | 42 kDa | 0 | 0 | 0 | 1 |
| Cre12.g558450.t1.2 | 35 kDa | 1 | 0 | 0 | 0 |
| Cre12.g558700.t1.2 | 42 kDa | 0 | 0 | 3 | 3 |
| Cre12.g559250.t1.2 | 30 kDa | 3 | 3 | 3 | 3 |
| Cre12.g560900.t1.2 | 49 kDa | 0 | 1 | 0 | 0 |
| Cre12.g560950.t1.2 | 13 kDa | 3 | 2 | 1 | 2 |
| Cre13.g562450.t1.2 | 114 kDa | 1 | 1 | 0 | 1 |
| Cre13.g562550.t1.1 | 56 kDa | 1 | 2 | 2 | 1 |
| Cre13.g564250.t1.2 | 116 kDa | 2 | 2 | 2 | 2 |
| Cre13.g565450.t1.2 | 58 kDa | 2 | 2 | 2 | 2 |
| Cre13.g565800.t1.2 | 39 kDa | 0 | 2 | 3 | 2 |
| Cre13.g565850.t1.1 | 137 kDa | 2 | 1 | 1 | 1 |
| Cre13.g566000.t1.2 | 68 kDa | 3 | 3 | 3 | 3 |
| Cre13.g566650.t1.2 (+1) | 75 kDa | 1 | 1 | 0 | 0 |
| Cre13.g567800.t1.2 | 58 kDa | 2 | 2 | 2 | 2 |
| Cre13.g567850.t1.2 | 22 kDa | 1 | 1 | 2 | 1 |
| Cre13.g567950.t1.2 | 56 kDa | 3 | 3 | 3 | 3 |
| Cre13.g568650.t1.2 | 29 kDa | 3 | 3 | 3 | 3 |
| Cre13.g568900.t1.2 | 21 kDa | 3 | 3 | 3 | 3 |
| Cre13.g572700.t1.2 | 151 kDa | 0 | 1 | 1 | 1 |
| Cre13.g572900.t1.1 (+1) | 86 kDa | 0 | 0 | 2 | 2 |
| Cre13.g573250.t1.2 | 26 kDa | 1 | 0 | 1 | 0 |
| Cre13.g573351.t1.2 | 16 kDa | 3 | 3 | 3 | 3 |
| Cre13.g578100.t1.1 | 233 kDa | 1 | 0 | 0 | 0 |
| Cre13.g578950.t1.1 | 260 kDa | 1 | 0 | 0 | 0 |
| Cre13.g580850.t1.2 (+1) | 19 kDa | 2 | 2 | 2 | 3 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre13.g581450.t1.2 | 43 kDa | 2 | 3 | 3 | 3 |
| Cre13.g581600.t1.2 | 34 kDa | 2 | 2 | 2 | 2 |
| Cre13.g582800.t1.2 | 67 kDa | 1 | 1 | 0 | 1 |
| Cre13.g584200.t1.2 | 105 kDa | 1 | 1 | 0 | 0 |
| Cre13.g584850.t1.1 | 26 kDa | 1 | 0 | 0 | 0 |
| Cre13.g585900.t1.2 | 70 kDa | 0 | 1 | 0 | 0 |
| Cre13.g587050.t1.2 | 49 kDa | 2 | 2 | 3 | 3 |
| Cre13.g589450.t1.2 | 107 kDa | 1 | 0 | 0 | 1 |
| Cre13.g592200.t1.2 | 242 kDa | 2 | 3 | 3 | 3 |
| Cre13.g603225.t1.1 | 45 kDa | 3 | 3 | 2 | 1 |
| Cre13.g603700.t1.2 | 42 kDa | 3 | 3 | 3 | 3 |
| Cre13.g603900.t1.2 | 69 kDa | 3 | 2 | 2 | 3 |
| Cre13.g604650.t1.2 (+1) | 41 kDa | 3 | 3 | 3 | 3 |
| Cre13.g605150.t1.2 | 25 kDa | 0 | 0 | 1 | 1 |
| Cre14.g608700.t1.2 | 64 kDa | 1 | 0 | 0 | 0 |
| Cre14.g608800.t1.2 | 40 kDa | 1 | 0 | 0 | 0 |
| Cre14.g612450.t1.2 | 26 kDa | 3 | 2 | 3 | 3 |
| Cre14.g613800.t1.1 | 48 kDa | 1 | 1 | 0 | 0 |
| Cre14.g614300.t1.2 | 54 kDa | 2 | 2 | 3 | 3 |
| Cre14.g614850.t1.1 | 214 kDa | 0 | 1 | 0 | 0 |
| Cre14.g614900.t2.1 | 89 kDa | 3 | 3 | 3 | 3 |
| Cre14.g615950.t1.1 | 116 kDa | 1 | 2 | 2 | 1 |
| Cre14.g616100.t1.1 | 46 kDa | 3 | 2 | 3 | 3 |
| Cre14.g616600.t1.2 | 108 kDa | 1 | 1 | 0 | 0 |
| Cre14.g617900.t1.2 | 15 kDa | 0 | 0 | 1 | 1 |
| Cre14.g619100.t1.2 | 46 kDa | 1 | 1 | 1 | 1 |
| Cre14.g619133.t1.1 | 69 kDa | 3 | 3 | 3 | 3 |
| Cre14.g619166.t1.1 | 36 kDa | 0 | 0 | 1 | 1 |
| Cre14.g620702.t1.1 | 52 kDa | 3 | 2 | 3 | 3 |
| Cre14.g621450.t1.2 | 34 kDa | 3 | 3 | 3 | 3 |
| Cre14.g621650.t1.1 | 37 kDa | 3 | 3 | 3 | 2 |
| Cre14.g625400.t1.1 | 48 kDa | 0 | 1 | 2 | 2 |
| Cre14.g625802.t1.1 | 98 kDa | 0 | 0 | 0 | 1 |
| Cre14.g626900.t1.2 | 28 kDa | 0 | 0 | 1 | 1 |
| Cre14.g627488.t1.1 | 25 kDa | 1 | 0 | 0 | 0 |
| Cre14.g629550.t1.2 | 40 kDa | 0 | 1 | 2 | 3 |
| Cre14.g630100.t1.2 | 24 kDa | 3 | 3 | 3 | 3 |
| Cre14.g630847.t1.1 | 42 kDa | 3 | 3 | 1 | 1 |
| Cre14.g632775.t1.1 | 54 kDa | 2 | 2 | 2 | 1 |
| Cre14.g633750.t1.1 | 122 kDa | 3 | 2 | 2 | 2 |
| Cre15.g635067.t1.1 | 87 kDa | 1 | 1 | 0 | 0 |
| Cre15.g635600.t1.2 | 21 kDa | 0 | 0 | 2 | 2 |
| Cre15.g635650.t1.2 (+1) | 281 kDa | 2 | 0 | 1 | 1 |
| Cre15.g635850.t1.2 (+1) | 37 kDa | 3 | 3 | 3 | 3 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre15.g637249.t1.1 | 219 kDa | 1 | 0 | 0 | 0 |
| Cre15.g638500.t1.2 | 34 kDa | 0 | 1 | 0 | 0 |
| Cre15.g643600.t1.2 | 59 kDa | 1 | 0 | 0 | 0 |
| Cre15.g644051.t1.1 | 72 kDa | 1 | 0 | 0 | 0 |
| Cre16.g650600.t1.2 | 205 kDa | 2 | 3 | 3 | 3 |
| Cre16.g651750.t1.2 | 130 kDa | 3 | 3 | 3 | 3 |
| Cre16.g651923.t1.2 | 67 kDa | 1 | 0 | 0 | 0 |
| Cre16.g652550.t1.2 | 18 kDa | 3 | 2 | 3 | 2 |
| Cre16.g653350.t1.2 | 66 kDa | 1 | 2 | 2 | 2 |
| Cre16.g654500.t1.1 | 32 kDa | 1 | 1 | 2 | 1 |
| Cre16.g656400.t1.2 | 53 kDa | 2 | 2 | 3 | 3 |
| Cre16.g657650.t1.2 | 13 kDa | 0 | 0 | 2 | 2 |
| Cre16.g658650.t1.1 | 183 kDa | 1 | 0 | 0 | 0 |
| Cre16.g659050.t1.1 | 281 kDa | 1 | 0 | 0 | 0 |
| Cre16.g659950.t1.1 | 72 kDa | 3 | 3 | 3 | 3 |
| Cre16.g660150.t1.2 | 13 kDa | 3 | 3 | 3 | 3 |
| Cre16.g661050.t1.2 | 14 kDa | 0 | 0 | 1 | 1 |
| Cre16.g663200.t1.1 | 80 kDa | 1 | 1 | 0 | 1 |
| Cre16.g663900.t1.2 | 37 kDa | 3 | 2 | 3 | 2 |
| Cre16.g664550.t1.2 | 57 kDa | 3 | 3 | 3 | 3 |
| Cre16.g665364.t1.1 | 49 kDa | 1 | 1 | 2 | 2 |
| Cre16.g670973.t1.1 | 23 kDa | 0 | 0 | 1 | 0 |
| Cre16.g672650.t1.2 | 29 kDa | 2 | 1 | 1 | 2 |
| Cre16.g672750.t1.2 | 35 kDa | 2 | 2 | 3 | 3 |
| Cre16.g672800.t1.2 | 40 kDa | 3 | 3 | 3 | 3 |
| Cre16.g673550.t1.1 | 92 kDa | 0 | 0 | 0 | 1 |
| Cre16.g673650.t1.1 | 31 kDa | 2 | 2 | 2 | 2 |
| Cre16.g673700.t1.2 | 94 kDa | 0 | 0 | 0 | 1 |
| Cre16.g673852.t1.1 | 109 kDa | 3 | 3 | 3 | 2 |
| Cre16.g674050.t1.2 | 36 kDa | 2 | 0 | 0 | 0 |
| Cre16.g674950.t1.2 | 60 kDa | 1 | 1 | 0 | 0 |
| Cre16.g675350.t1.1 | 155 kDa | 1 | 1 | 0 | 1 |
| Cre16.g675650.t1.2 | 58 kDa | 1 | 2 | 2 | 2 |
| Cre16.g676197.t1.1 | 96 kDa | 3 | 3 | 3 | 2 |
| Cre16.g677000.t1.2 | 84 kDa | 3 | 3 | 3 | 3 |
| Cre16.g677026.t1.1 | 38 kDa | 0 | 1 | 0 | 0 |
| Cre16.g677850.t1.1 | 29 kDa | 1 | 0 | 0 | 0 |
| Cre16.g680000.t1.2 | 26 kDa | 0 | 1 | 2 | 2 |
| Cre16.g680700.t1.2 | 31 kDa | 1 | 0 | 0 | 0 |
| Cre16.g680900.t1.2 | 199 kDa | 1 | 0 | 0 | 1 |
| Cre16.g682100.t1.2 | 28 kDa | 1 | 0 | 0 | 1 |
| Cre16.g682300.t1.2 | 13 kDa | 3 | 3 | 2 | 2 |
| Cre16.g682900.t1.2 (+1) | 32 kDa | 0 | 0 | 0 | 1 |
| Cre16.g683050.t1.1 | 98 kDa | 0 | 0 | 0 | 1 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre16.g683793.t1.1 | 90 kDa | 1 | 1 | 2 | 1 |
| Cre16.g685550.t1.2 | 38 kDa | 1 | 2 | 2 | 2 |
| Cre16.g685901.t1.1 | 103 kDa | 1 | 0 | 0 | 0 |
| Cre16.g687350.t1.2 | 68 kDa | 3 | 3 | 3 | 3 |
| Cre16.g687518.t1.1 | 36 kDa | 0 | 0 | 0 | 1 |
| Cre16.g687900.t1.2 | 26 kDa | 2 | 2 | 2 | 2 |
| Cre16.g688550.t1.2 | 24 kDa | 1 | 0 | 1 | 0 |
| Cre16.g689050.t1.1 | 78 kDa | 0 | 0 | 1 | 1 |
| Cre16.g689647.t1.1 | 108 kDa | 2 | 2 | 3 | 2 |
| Cre16.g689871.t1.1 | 115 kDa | 1 | 0 | 0 | 0 |
| Cre16.g690431.t1.1 (+1) | 53 kDa | 0 | 1 | 0 | 0 |
| Cre16.g690500.t1.2 | 18 kDa | 3 | 3 | 3 | 3 |
| Cre16.g691000.t1.2 | 24 kDa | 0 | 0 | 1 | 0 |
| Cre16.g691450.t1.2 | 73 kDa | 0 | 1 | 0 | 0 |
| Cre16.g692901.t1.1 | 57 kDa | 2 | 0 | 0 | 0 |
| Cre16.g693500.t1.1 | 42 kDa | 0 | 1 | 3 | 1 |
| Cre16.g693600.t1.2 (+1) | 39 kDa | 1 | 3 | 3 | 3 |
| Cre16.g695050.t1.2 | 77 kDa | 3 | 3 | 3 | 3 |
| Cre16.g695100.t1.1 | 52 kDa | 0 | 0 | 2 | 3 |
| Cre16.g696000.t1.2 | 156 kDa | 3 | 1 | 0 | 1 |
| Cre17.g697450.t1.2 | 61 kDa | 2 | 1 | 2 | 2 |
| Cre17.g697650.t1.1 | 92 kDa | 1 | 1 | 0 | 1 |
| Cre17.g698000.t1.2 | 62 kDa | 3 | 3 | 3 | 3 |
| Cre17.g698450.t1.2 | 38 kDa | 1 | 0 | 0 | 0 |
| Cre17.g699000.t1.2 (+1) | 86 kDa | 1 | 1 | 2 | 2 |
| Cre17.g701200.t1.2 (+1) | 15 kDa | 2 | 3 | 3 | 3 |
| Cre17.g701650.t1.2 | 15 kDa | 2 | 3 | 3 | 3 |
| Cre17.g701700.t1.2 (+1) | 45 kDa | 3 | 3 | 1 | 1 |
| Cre17.g703700.t1.1 | 44 kDa | 2 | 2 | 3 | 3 |
| Cre17.g703900.t1.2 | 198 kDa | 0 | 1 | 1 | 1 |
| Cre17.g704850.t1.2 | 20 kDa | 0 | 0 | 1 | 1 |
| Cre17.g705000.t1.2 | 67 kDa | 3 | 3 | 3 | 3 |
| Cre17.g705500.t1.1 | 95 kDa | 0 | 0 | 0 | 1 |
| Cre17.g705850.t1.2 | 49 kDa | 0 | 0 | 0 | 1 |
| Cre17.g708750.t1.2 | 34 kDa | 0 | 1 | 3 | 1 |
| Cre17.g708800.t1.1 | 53 kDa | 3 | 3 | 3 | 3 |
| Cre17.g709850.t1.2 | 43 kDa | 1 | 1 | 2 | 2 |
| Cre17.g710150.t1.2 | 45 kDa | 0 | 0 | 2 | 1 |
| Cre17.g712100.t1.2 | 45 kDa | 1 | 2 | 2 | 2 |
| Cre17.g712300.t1.1 | 100 kDa | 2 | 0 | 0 | 0 |
| Cre17.g713200.t1.1 (+1) | 54 kDa | 2 | 1 | 0 | 0 |
| Cre17.g718468.t1.2 | 68 kDa | 0 | 1 | 1 | 1 |
| Cre17.g719000.t1.2 | 49 kDa | 3 | 3 | 3 | 3 |
| Cre17.g719900.t1.2 | 131 kDa | 1 | 2 | 2 | 1 |

| Accession Number | Molecular Weight | WT | *saga* | SVA7 | SVB2 |
|---|---|---|---|---|---|
| Cre17.g720050.t1.2 | 74 kDa | 3 | 2 | 2 | 2 |
| Cre17.g720250.t1.2 | 30 kDa | 3 | 2 | 2 | 2 |
| Cre17.g721500.t1.2 | 77 kDa | 1 | 0 | 1 | 1 |
| Cre17.g723650.t1.2 | 47 kDa | 3 | 3 | 3 | 3 |
| Cre17.g724850.t1.2 | 53 kDa | 0 | 0 | 0 | 1 |
| Cre17.g726750.t1.2 | 56 kDa | 0 | 0 | 1 | 0 |
| Cre17.g726850.t1.2 | 83 kDa | 3 | 1 | 2 | 3 |
| Cre17.g727300.t1.2 | 42 kDa | 0 | 0 | 1 | 1 |
| Cre17.g727700.t1.1 | 87 kDa | 0 | 0 | 0 | 1 |
| Cre17.g727950.t1.1 | 108 kDa | 2 | 3 | 3 | 3 |
| Cre17.g728800.t1.2 | 41 kDa | 2 | 2 | 3 | 3 |
| Cre17.g729800.t1.2 | 45 kDa | 1 | 0 | 0 | 0 |
| Cre17.g730950.t1.2 | 87 kDa | 0 | 0 | 2 | 1 |
| Cre17.g734100.t1.2 | 50 kDa | 2 | 3 | 3 | 3 |
| Cre17.g734200.t1.2 | 48 kDa | 2 | 2 | 1 | 1 |
| Cre17.g734300.t1.1 (+1) | 102 kDa | 1 | 1 | 0 | 0 |
| Cre17.g734400.t1.1 | 88 kDa | 0 | 1 | 0 | 0 |
| Cre17.g734450.t1.2 | 17 kDa | 3 | 2 | 3 | 3 |
| Cre17.g734500.t1.2 | 26 kDa | 3 | 2 | 2 | 2 |
| Cre17.g734516.t1.1 | 44 kDa | 1 | 0 | 0 | 1 |
| Cre17.g734548.t1.1 | 104 kDa | 3 | 2 | 3 | 2 |
| Cre17.g734900.t1.2 (+1) | 61 kDa | 1 | 0 | 0 | 0 |
| Cre17.g737250.t1.2 | 109 kDa | 0 | 1 | 1 | 1 |
| Cre17.g738050.t1.2 | 14 kDa | 1 | 0 | 0 | 0 |
| Cre17.g741450.t1.2 | 62 kDa | 3 | 3 | 3 | 2 |
| Cre17.g742200.t1.1 | 49 kDa | 0 | 1 | 1 | 1 |
| Cre17.g742450.t1.1 | 23 kDa | 1 | 1 | 1 | 1 |
| Cre17.g743897.t1.1 | 42 kDa | 2 | 0 | 1 | 1 |
| Cre17.g744147.t1.1 | 240 kDa | 3 | 1 | 2 | 2 |
| Cre17.g746997.t1.1 | 102 kDa | 3 | 3 | 3 | 3 |
| Cre18.g749447.t1.1 (+1) | 99 kDa | 3 | 3 | 3 | 3 |
| Cre18.g749497.t1.1 | 48 kDa | 3 | 3 | 3 | 3 |
| Cre18.g749847.t1.1 (+1) | 52 kDa | 3 | 3 | 3 | 3 |
| Cre18.g750047.t1.1 | 447 kDa | 0 | 0 | 3 | 3 |
| Cre19.g750547.t1.1 (+3) | 69 kDa | 3 | 1 | 2 | 2 |
| Cre19.g751197.t1.1 | 156 kDa | 1 | 1 | 0 | 0 |
| Cre24.g755997.t1.1 | 69 kDa | 3 | 3 | 0 | 0 |
| Cre43.g760497.t1.1 | 29 kDa | 3 | 3 | 3 | 3 |
| Cre48.g761197.t1.1 | 28 kDa | 3 | 3 | 3 | 3 |
| **sp\|cRAP032\|P42212\|GFP_AEQVI** | **?** | **0** | **0** | **3** | **3** |